
Calculus 101

Release

Sep 15, 2017

Contents

1	History and Motivation	1
2	Preliminaries	3
3	Building the Integers	5
4	Building the Rationals	9
5	Irrationals – the Secret Pythagoras Kept	13
6	Axiomatizing the Real Numbers	15
7	Sequences and Convergence	19
8	Limit Properties	21
9	Cantor’s Lemma	23
10	Subsequences and convergence	25
11	Completeness	27
12	Building the Real Numbers	29
13	Series	31
14	Real functions	33
15	Limit	35
16	Continuity	37
17	Derivative	39
18	Roll’s Theorem	41
19	Riemann Integral	43
20	Exponential	45

CHAPTER 1

History and Motivation

Once upon a time, Isaac Newton wanted to define the speed of an object, whose speed changes continuously. He erected a weird tower of concepts like “infinitesimals” to analyze this concept, and came up with something that, surprisingly enough, worked. However, Newton’s notation and concepts were hard to understand and, for anyone who was not Newton, to convince themselves that this was not a tower built on sand. When mathematicians started to figure out how to properly define these things so anyone can understand them, they realized that before even delving into the question of speed, they needed to understand the things that measure speed – numbers.

Natural numbers harken back to the prehistory of man, but a proper understanding of them in Europe, where calculus was invented, was a 13th century discovery. Still, even then, a proper understanding of zero, negative numbers and fractions would need a while. However, to properly understand those numbers Pythagoras and his students hated so much they denigrated them as “irrational” (how much vitriol does someone need to call a number, literally, a “crazy” number?) required still more machinery. But proper calculus could not be done without those numbers – those “crazy” numbers kept cropping up anywhere where instantaneous change was hiding – even in calculating the answers like “if your bank decided to charge interest based on APR-to-EAR calculation based on very small periods, what is the maximum they can charge?” the “crazy” number e makes an appearance!

In this text, we try to work through the work of these 18th and 19th century mathematicians, to understand what we mean by numbers, limits and, finally, speeds. Tying all this together is the quest to tell a story using the language of math. Imagine a 100 meter dash course. At the beginning, you are at the starting position. You want to spend some time at the end position, say a minute, chatting with someone, and then get back to the starting position. The catch? the points in between are boring, and you want to spend as little time as possible in them. Our quest will be to find a function that captures that story while being amenable to questions such as “what is the speed? what is the acceleration?” at each point.

For this, we will need to understand what numbers are, how to define speed and how to relate the distance travelled to speed. It’s a long journey, with many N , ϵ and δ – but at the end, the concepts of calculus will seem so obvious that it will be hard to believe it took a Newton to invent them.

CHAPTER 2

Preliminaries

This text covers things from the basics. However, there are a few things that will be assumed without deep justification – perhaps to be filled by a “Set Theory 101” text.

The first is sets. A set is a collection of elements. A set can be indicated by its elements: $2, 5, 9$. A set can also be indicated by a property of its elements: $x : x > 5$.

The second is functions. The intuitive idea of a “function” is “something that assigns each element in its domain a value”. Making this notion more formal requires a surprise – we do not have to have any rules of assignment or a way of calculating the result. The formal definition of a function is “a set of pairs (x, y) such that there are no two pairs (x, y) and (z, w) where $x = z$ but $y \neq w$ ”. The intuition should be not the rules of assignment, but the graph of a function as its true meaning. When we specify a function like $f(x) = x + 5$, what we are saying is “the set of all pairs (x, y) where, say, x is a rational number and $y = x + 5$, or $(x, y) : y = x + 5$ ”.

The third is the natural numbers. For mathematicians, the natural numbers always include $0 - 0, 1, 2, \dots$. The natural numbers come with addition and multiplication, which obey the following rules:

- Neutrality of 0: $x + 0 = x$ for every x
- Neutrality of 1: $x1 = x$ for every x
- Uniqueness of addition: $x + y = x + z$ implies $y = z$
- Almost-uniqueness of multiplication: $xy = xz$ implies $y = z$ or $x = 0$
- Associativity of addition: $x + (y + z) = (x + y) + z$
- Associativity of multiplication: $x(yz) = (xy)z$
- Commutativity of addition: $x + y = y + x$
- Commutativity of multiplication: $xy = yx$
- Distributive law: $x(y + z) = xy + xz$
- Almost inverses: if $x \neq y$, then either there is a z such that $x + z = y$ or $y + z = x$, but not both.

Definition: $x \leq y$ if there is a z such that $x + z = y$

Claim: If $x \leq y$ and $y \leq z$ then $x \leq z$

Before proving this, note that even though it seems utterly obvious, in math nothing is taken on faith. Even things like this must be formally proven.

How can a proof proceed? Clearly, this would not be true of any relationship, so the definition of \leq needs to be used. When substituting the definition, the results are:

$$x + t = y$$

$$y + s = z$$

Note that since we use y and z in the theorem statement, different letters were used instead.

If $x + t = y$, then $x + t$ can be substituted in the second equation for y , leading to $(x + t) + s = z$. Applying the associative law,

And so $x \leq z$

We write $x < y$ for $x \leq y$ and $x \neq y$.

The last thing taken for granted for the natural numbers is the law of induction. The law of induction says that if a property holds for 0 and whenever it holds for x it holds for $x + 1$, then it holds for all integers. This is a powerful axiom: it can actually build addition, multiplication and order by itself.

Here is a simple example: “no numbers between 0 and 1”: there is no number x such that $0 < x < 1$.

Proof: Since $0 = 0$, for $x = 0$ it is not the case that $0 < x < 1$.

Let $y = x + 1$. We prove that it is not the case that $y < 1$. Indeed, since $y = x + 1$, we know that $1 \leq y$. By the axioms on order, either $y = 1$ or it is not the case that $y \leq 1$, and in either case, it is not the case that $y < 1$. QED

Notice that this was a simple application – the fact that the inductive statement holds for x was not even used.

TODO [Issue #1]: Write hint-along-exercises for:

- Prove $x \leq y$ and $y \leq x$ implies $x = y$
- Prove $0 \leq x$ for every x
- Prove order respects addition and multiplication, as an exercise
- Prove “no numbers between x and $x + 1$ ”

Here is a more interesting example: every non-empty set of integers has a least element.

Now, this does not seem to be a statement where we can use the law of induction. Instead, we transform it into a statement where we can use induction: for every set A , if $x \in A$ (read as “ x is in A ”), then A has a least element. Induction on this statement, however, will fail. As is often the case, the hardest part is figuring out the correct inductive statement.

Proof: The inductive statement is: “for every x , if $x \leq y$ and $x \in A$, then A has a least element”.

First, verify for 0: if $x \leq 0$, and $x \in A$, then $x = 0$, and since $0 \leq z$ for all natural numbers z , 0 is the least element.

Assume the statement holds for x . Assume $y = x + 1 \in A$. If there is a $z < y$ and $z \in A$, then $z \leq x$ and by the assumption, A has a least element. Otherwise, for every $z \in A$, $y \leq z$ and so y is the least element. Either way, A has a least element, and so the induction is proven. QED

We write \mathbb{N} for the set of natural numbers.

TODO [Issue #1]: write hint-a-long exercise for:

- If A is a set, and $p \in A$ is a member, B is the set of all pairs (a, n) for $a \in A$ and n a natural number, and $f : B \rightarrow A$ is a function, then there is a unique function, $s : \mathbb{N} \rightarrow A$ such that $s(0) = p$ and $s(n + 1) = f(s(n), n)$

We call function from the natural numbers is also called a “sequence”, and the technique above is called “defining a sequence by recursion”. It is a powerful technique for generating sequences.

TODO [Issue #2]: Show a few examples of defining functions recursively

Building the Integers

We are already one leg up on the ancient Greeks, as we have a 0 in our natural numbers. However, it turns out that we need some more numbers to be able to answer questions like “If I have \$3, and I spend \$4, how much money do I have?” (answer: “\$1 in debt”, of course).

There are several ways to construct the integers. The technique used here will also come in handy later: imagine they exist, figure out how they would look, and then formalize. So, assume that there is an answer for “I have X things, and I spend Y things. How much do I have left?” Call this answer $A(X, Y)$. Intuitively, imagine it to be $X - Y$. What properties will it have?

- If $X + W = Y + Z$, then $X - Y = Z - W$, or $A(X, Y) = A(Z, W)$.
- $(X - Y) + (Z - W) = (X + Z) - (Y + W)$ or $A(X, Y) + A(Z, W) = A(X + Z, Y + W)$
- $(X - Y)(Z - W) = (XZ + YW) - (XW + YZ)$ or $A(X, Y)A(Z, W) = A(XZ + YW, XW + YZ)$

In order to formalize this, another powerful technique is helpful – equivalence classes. Take the collection of all pairs of natural numbers (X, Y) .

Definition: $(X, Y) \sim (Z, W)$ (read as “ (X, Y) is equivalent to (Z, W) ”) if $X + W = Z + Y$.

In order for a definition of a relation to be a proper equivalence, three things need to be proven: reflectivity, symmetry and transitivity.

Reflectivity: this means each thing should be equivalent to itself. Indeed, $(X, Y) \sim (X, Y)$ because $X + Y = X + Y$.

Symmetry: if one thing is equivalent to another, then the other thing should be equivalent to the first thing. Indeed, if $(X, Y) \sim (Z, W)$ then $X + W = Z + Y$, but then $Z + Y = X + W$, and so $(Z, W) \sim (X, Y)$.

Transitivity: Assume $(X, Y) \sim (Z, W)$ and $(Z, W) \sim (U, V)$. Then $(X, Y) \sim (U, V)$, or $X + V = Y + U$. We know that $X + W = Y + Z$ and $Z + V = U + W$. Adding both sides, we get $X + W + Z + V = Y + Z + U + W$. Using the commutative law, $X + V + W + Z = Y + U + W + Z$. Now, use the eliminative property of addition, and get $X + V = Y + U$.

The other part of the technique is moving from the set of “things” to the set of “equivalence classes of things”.

Define an equivalence class as “all elements that are equivalent to a given element”

Formally: $[(X, Y)] = (Z, W) : (Z, W) \sim (X, Y)$.

TODO [Issue #4]: Show that $(X, Y) \sim (Z, W)$ iff $[(X, Y)] = [(Z, W)]$.

The set of the integers: $\mathbb{Z} = [(X, Y)] : X, Y \in \mathbb{N}$. Just like the natural numbers, the integers need operations.

In order to inspire the definition, again informally view (X, Y) as $X - Y$. So $X - Y + Z - W = (X + Z) - (Y + W)$, and $(X - Y)(Z - W) = (XZ + YW) - (XW + YZ)$.

Formally:

$$[(X, Y)] + [(Z, W)] = [(X + Z, Y + W)]$$

For this definition to make sense, the following theorem is needed:

Theorem: If $(X, Y) \sim (X', Y')$ and $(Z, W) \sim (Z', W')$ then $(X + Y, Z + W) \sim (X' + Y', Z' + W')$.

Proof:

$$X + Y' = X' + Y$$

$$Z + W' = Z' + W$$

Adding the equalities:

$$X + Y' + Z + W' = Z' + W + X' + Y$$

Rearrange:

$$X + Z + Y' + W' = X' + Z' + Y + W$$

Which is the definition of $(X + Y, Z + W) \sim (X' + Y', Z' + W')$.

Because of this theorem, it does not matter which elements of the equivalence class are used to define addition, and so the definition of addition makes sense.

Define multiplication by $[(X, Y)][(Z, W)] = [(XZ + YW, XW + YZ)]$

Exercise: Prove that this definition makes sense. Hint [Issue #5]: Follow the same steps as the proof above – write down the equivalences. This time, how about multiplying them?

Note that $[(X, 0)] + [(Y, 0)] = [(X + Y, 0)]$ and $[(X, 0)][(Y, 0)] = [(XY, 0)]$. Also note that if $X \neq Y$, it is not the case that $(X, 0) \sim (Y, 0)$, and so $[(X, 0)] \neq [(Y, 0)]$.

Therefore, the function $i : \mathbb{N} \rightarrow \mathbb{Z}$ preserves uniqueness, addition and multiplication. Functions that preserve uniqueness and structure are called “embeddings”. It is often convenient, and perfectly safe, to consider the embedding as an identity, and confuse n and $[(n, 0)]$.

Note $[(X, Y)] + [(Y, X)] = [(X + Y, Y + X)] = [(0, 0)] = e(0)$. Define $-(X, Y) = [(Y, X)]$, and the equality above can be written as:

$$a + (-a) = 0$$

The integers with addition and multiplication are called a ring for satisfying the following axioms:

- $a + 0 = a$
- $a + (b + c) = (a + b) + c$
- $a + b = b + a$
- For every a there is a $-a$ such that $a + (-a) = 0$
- $a \cdot 1 = a$
- $a(bc) = (ab)c$

- $ab = ba$
- $a(b + c) = ab + ac$

(Notice that above we wrote 0 and 1 for $e(0)$ and $e(1)$).

Define $a - b = a + (-b)$, and any two integers can be subtracted.

Note that for every X and Y , either $Y > X$ in which case there is a number n such that $X = Y + n$ or $X + 0 = Y + n$, which means $[(X, Y)] = [(n, 0)]$, either $Y < X$ in which case there is an integer n such that $Y = X + n$ or $n + X = 0 + Y$, which means $[(X, Y)] = [(0, n)] = -[(n, 0)]$, $X = Y$ which means $[(X, Y)] = [(0, 0)]$.

In other words, for every integer, z , either $z = n$, $z = -n$ or $z = 0$, with n a non-zero natural number.

This is our structure theorem for the integers: every integer is either a natural number or the inverse of a negative number.

Building the Rationals

The integers allow addition and subtraction – however, division is still tricky. For example, there is no x such that, $2x = 1$. If $2x = 1$, then $x + x = 1$, and so $x \leq 1$. Since $1 + 1 = 2 \neq 1$, $x < 1$. However, $0 + 0 = 0$, so $0 < x < 1$.

Much like the integers, first consider how X/Y should behave and then formalize. Much like we did with the rationals, we want to think how X/Y behave, and then formalize it.

- $X/Y = Z/W$ if and only if $XW = ZY$
- $X/Y + Z/W = (XW + YZ)/YW$
- $(X/Y)(Z/W) = (XZ)/(YW)$

Formally define: for two pairs of integers, $(X, Y) (Z, W)$ if $XW = ZY$.

Exercise:

- Prove this relationship is reflective, symmetric and transitive.

TODO: Add Hints (Issue #6)

Let \mathbb{Q} be the set of equivalence classes under this relationship.

Define multiplication and addition based on the earlier inspiration as:

$$[(X, Y)] + [(Z, W)] = [(XW + YZ, YW)]$$

$$[(X, Y)][(Z, W)] = [(XZ, YW)]$$

Exercise:

- Prove these definitions depend only on the equivalence classes.

TODO: Add Hints (Issue #6)

Note that $[(X, 1)] + [(Y, 1)] = [(X + Y, 1)]$ and $[(X, 1)][(Y, 1)] = [(XY, 1)]$ and so $e : \mathbb{Z} \rightarrow \mathbb{Q}$ defined by $e(X) = [(X, 1)]$ is an embedding. Much like earlier, it is sometimes useful to confuse $e(X)$ and X .

Note that if $[(X, Y)] \neq 0$ (or, really, $e(0)$), then $X \neq 0$. In this case, $[(Y, X)] \in \mathbb{Q}$ and $[(X, Y)][(Y, X)] = [(XY, YX)] = [(1, 1)] = e(1)$ So every element different from 0 has a multiplicative inverse.

In conclusion, the field axioms are satisfied:

- $a + 0 = a$
- $a + (b + c) = (a + b) + c$
- $a + b = b + a$
- For every a there is a $-a$ such that $a + (-a) = 0$
- $a1 = a$
- $a(bc) = (ab)c$
- $ab = ba$
- For every $a \neq 0$ that is a a^{-1} such that $aa^{-1} = 1$
- $a(b + c) = ab + ac$

The field of rationals is called \mathbb{Q} .

The embedding of the natural numbers into the integers, combined with the embedding of the integers into the rationals, gives an embedding of the natural numbers into the rationals. Define a rational number, r to be “positive” if there are natural numbers that are not 0, n and m , such that $n = rm$.

Theorem: for every rational, r , exactly one of those is true:

- $r = 0$
- r is positive
- $-r$ is positive

Proof: Assume $r = [(X, Y)]$ for X and Y integers. Since $Y \neq 0$, by the structure of integers, either Y is natural or $-Y$ is natural. If $-Y$ is natural, by the definition of equivalence, $r = [(-X, -Y)]$, so we can assume Y is natural.

If $X = 0$, then $r = 0$, and $rm = 0$ for every natural number m .

If X is natural, then $rY = X$.

If $-X$ is natural, then $-rY = -X$.

QED

Theorem: If r and s are positive, then so are $r + s$ and rs .

Proof: If $rm = n$ and $sl = k$, then $(rs)ml = nk$ and

$$(r + s)ml = rml + sml = rml + slm = nl + km$$

And ml and $nl + km$ are natural numbers, since the natural numbers are closed under addition and multiplication.

QED

Define $r \leq s$ if $r = s$ or $s - r$ is positive.

It is straightforward to verify that

- $x \leq y$ or $y \leq x$ and if both are true, $x = y$
- If $x \leq y$ then $(x + z) \leq (y + z)$
- If $x \leq y$: and : math : $0 \leq z$ then $xz \leq yz$

Together with the field axioms, those are the axioms of an “ordered field”.

Exercise: Assume F is an ordered field, there is a function $e : \mathbb{Q} \rightarrow F$, which is one-to-one (i.e., $e(x) = e(y)$ implies $x = y$), preserves addition, multiplication, zero, one and order. (Such a function is called an “embedding”).

Hint-a-long: Define a sequence by recursion:

- $e(0) = 0_F$
- $e(n+1) = e(n) + 1_F$

Prove that $n < m$ implies $e(n) < e(m)$

Define $e(-n) = -e(n)$, and $e(n/m) = e(n)/e(m)$.

The rationals, therefore, are the smallest ordered field.

Irrationals – the Secret Pythagoras Kept

The integers allow solving equations like $a + x = b$. The rationals allow solving equations like $ax = b$. What other numbers are needed?

In Ancient Greece, Pythagoras and his followers did two things: they built the rationals to understand music and they taught each other what is now known in Pythagoras's theorem: if a right-angled triangle has sides a, b and a hypotenuse c then $a^2 + b^2 = c^2$. Students were taught to keep those secrets, on pain of death. Why the secrets and the threats?

Consider a right-angled triangle with sides of length 1. Call the hypotenuse c . $1^2 + 1^2 = c^2$, or $2 = c^2$.

Can c be rational? Assume it is. We can assume c is positive (because $(-c)^2 = c^2$), so $cm = n$, and $2m^2 = c^2m^2 = n^2$. If we write $n = (2^k)p$, with p odd, then $n^2 = 2^{2k}p^2$, and p^2 is odd. If we write $m = (2^l)q$, with q odd, then $m^2 = 2^{2l}q^2$, and q^2 is odd. In other words: $(2^{2k})p^2 = 2 * 2^{2l}q^2 = 2^{2l+1}q^2$. Since $2k \neq 2l+1$: then $2^{2k} \neq 2^{2l+1}$. If $2^{2k} < 2^{2l+1}$, then $p^2 = 2^{2l+1-2k}q^2$, and so p^2 is even, and so p is even. If $2^{2l+1} < 2^{2k}$, for similar reasons, q is even.

Since both p and q are odd, neither can be the case, so the assumption is wrong: c is not "rational".

That is why the Pythagoreans kept those secrets – they knew the side of the hypotenuse could not be rational. Indeed the word "irrational" comes from that era, where it was "illogical" that this would be the case.

It would be nice to be able to solve equations like c^2 so that triangles' sides will have length. How can the answers be calculated?

If $0 < x < y$, then $x^2 < y^2$.

Define two sequences: a_n and b_n so that $a_n^2 < 2 < b_n^2$.

- $a_0 = 1$ and $b_0 = 2$
- After having a_n and b_n , calculate $t = (a_n + b_n)/2$. If $t^2 < 2$ set $a_{n+1} = t, b_{n+1} = b_n$. Otherwise, set $a_{n+1} = a_n, b_{n+1} = t$.

Note:

- $a_n \leq a_{n+1}$
- $b_{n+1} \leq b_n$
- $a_n < b_n$.

- $b_{n+1} - a_{n+1} = 1/2(b_n - a_n)$

So a and b seem to be growing closer, trying to find the square root of 2.

Axiomatizing the Real Numbers

Building the real numbers is a bit more difficult. Instead, a different route will be taken – first, the real numbers will be treated axiomatically. On the assumption that the real numbers exist, and obey certain axioms, theorems will be proven. After getting experience with the real numbers, and some related techniques, it will be more easier to understand the construction.

The real numbers are assumed to be a *complete ordered field*.

The axioms for ordered field have been given already, but here they are repeated:

Addition is an abelian group

- $a + 0 = a$
- $a + (b + c) = (a + b) + c$
- For every a , there is a $-a$ such that $a + (-a) = 0$
- $a + b = b + a$

Multiplication is almost an abelian group

- $a1 = a$
- $a(bc) = (ab)c$
- For every $a \neq 0$, there is a a^{-1} such that $aa^{-1} = 1$
- $ab = ba$

Multiplication distributes over addition

- $a(b + c) = ab + ac$

Ordered field

There exist a subset of the field, P (the positive numbers), such that:

- For every a , exactly one of $a \in P$, $-a \in P$ or $a = 0$ is true
- If $a, b \in P$ then $ab, a + b \in P$

Before defining the axioms of completeness, the concept of bounds needs to be introduced. If A is a subset of an ordered field, q is an upper bound if $a \leq q$ for every $a \in A$. Sometimes, an upper bound will not exist. For example, in the field of rational numbers, the set of embedded natural numbers has no upper bound. Sometimes, an upper bound will exist. For example, the set $1, 2, 3$ has 5 as an upper bound.

An ordered field will be called complete if whenever a set A does have an upper bound, it has a least upper bound.

- For every A , if there exists a q upper bound, then there exists an upper bound p such that if r is also an upper bound, $p \leq r$

If p is the least upper bound of a set, then it is the only least upper bound of the set. Because of that, we call such a number “the supremum” of the set, and denote $p = \sup A$

The rational numbers, as we have seen, are not complete. Specifically, the set of values that a_n , defined previously, does not have an least upper bound. Here is a proof: assume there was a least upper bound a . Then because every b_n is an upper bound, $0 < a_n \leq a \leq b_n < 2$ for every n . Because of that, $a_n^2 \leq a \leq b_n^2$. Recall that $b_n - a_n = 2^{-n}$, and so $a_n^2 \leq a^2 \leq b_n^2$ and $b_n^2 - a_n^2 = (b_n - a_n)(b_n + a_n) \leq 2^{-n} * 4$, and so $b_n^2 - a^2 \leq 2^{-n} * 4$ and $a^2 - a_n^2 \leq 2^{-n} * 4$ and If $a^2 > 2$, then write $d = a^2 - 2 > 0$. Since d is a rational number, $d = m/n \geq 1/n > 2^{-n} = 2^{-(n+3)} * 4$ and now $a^2 - a_n^2 < d$, so $a^2 < d + a_n^2 < d + 2 = a^2$ which is a contradiction. A similar argument shows $a^2 < 2$ cannot be true, using b_n . So $a^2 = 2$, but such an a cannot be a rational number. This shows that a_n does not have a least upper bound in the field of rational numbers.

So it turns out the rationals are *not* a complete ordered field. Maybe there is no complete ordered field? It would not be surprising to know that some axioms have no structures that satisfy them. However, maybe there is a complete ordered field. Just as the natural numbers have been taken up on faith, the consequences of the axiom “there exists a complete ordered field”.

It has already been shown that the natural numbers embed in every ordered field. However, complete ordered fields have a stronger property.

Theorem: (Archimedean property) If \mathbb{F} is a complete ordered field, then for every $a \in \mathbb{F}$ there is a natural number n , such that $a < n$.

Proof: Suppose there is an a without such a natural number. Then $n \leq a$ for every n , and so the natural numbers have an upper bound. By completeness, they have a least upper bound, b . Since $n + 1$ is a natural number if n is, $n + 1 \leq b$ for every natural number n . Therefore, $n \leq b - 1 < b$ for every n , so $b - 1$ is also an upper bound. Since b was a least upper bound, this cannot be.

If a field is Archimedean, and $x > 0$, then for any $y > 0$ there is a natural number $n > y/x$ or $y < nx$. This is known as the “hair grows at some amount of miles per hour” consequence: no matter how small some number is, and how big another number is, they are commensurable: if we add the small one to itself enough times, we will overtake the bigger one.

Assume $a < b$ in an Archimedean field. Then there is a rational number $a < q < b$. First the case $0 < a < b$. Pick a natural number n such that $n > (b - a)^{-1}$. Therefore, $1/n < (b - a)$. Using the Archimedean property again, there is a natural number k such that $k/n > a$. Any set of natural numbers has a first element, so there is a smallest number m such that $m/n > a$. Now, if $m/n \geq b$, then $(m - 1)/n \geq b - 1/n > b - (b - a) = a$, and so m is not the smallest number. Therefore $q = m/n$ satisfies the condition. If $a < 0 < b$, then 0 satisfies the condition. Lastly, if $a < b < 0$, find a q such that $-b < q < -a$ (by the first part), and so $a < -q < b$ and $-q$ satisfies the condition.

This result can be summarized as “in an Archimedean field, there is a rational number between any two elements”, or “the rationals are dense in any Archimedean field”.

Assume \mathbb{F} and \mathbb{G} are complete ordered fields. If A is the set of all rationals $q < f$ for some $f \in \mathbb{F}$. There is a natural number $n > f$ and so A is bounded from above by n in \mathbb{G} as well. Therefore, there is a least upper bound in \mathbb{G} . Define $g(f)$ to be this upper bound. This defines a function g from \mathbb{F} to \mathbb{G} .

It is straight-forward to see that this function preserves addition, multiplication and order. It is also straight-forward to verify that the same function defined from \mathbb{G} to \mathbb{F} . Therefore, any two complete ordered fields are one and the same.

So given that a complete ordered field does exist, it is Archimedean and unique. Therefore, under the axiom that it exists, it is fine to talk about “the complete ordered field”, and for short, this will be referred to as “the real number field”.

Sequences and Convergence

Recall the definition of a sequence: a sequence of X is a function from the natural numbers to X . Usually, instead of $f(n)$ it is written as a_n . In this chapter, unless mentioned otherwise, sequences will be sequences in an ordered field.

Let a and b real numbers, and $\epsilon > 0$ be a real number (informally, think of it as “small”). If $x < y + \epsilon$, define it as x is ϵ -almost smaller than y . If a is ϵ -smaller than b and b is ϵ -smaller than a , then $a - \epsilon < b < a + \epsilon$, or $-\epsilon < b - a < \epsilon$.

Definition: In an ordered field, the absolute value $|x|$ is defined as

$$|x| = \max x, -x$$

Since

$$b - a < \epsilon$$

and

$$a - b < \epsilon$$

then

$$|b - a| < \epsilon$$

Call this “ b and a are ϵ -close”. Note that since $|b - a| = |a - b|$, this relationship is symmetric (though not, in general, transitive). An alternative way of saying it is “within ϵ of each other”. The reason this relationship has no many ways is because this, above all else, defines calculus.

If a_n and b_n are sequences, and there exists N such that for all $n > N$, $a_n = b_n$, then they “share a tail”. Sharing a tail is obviously a reflexive and symmetric condition. It is also transitive: if a_n shares a tail with b_n , starting at N , and b_n shares a tail with c_n , starting at M , then if $n > \max N, M$, $a_n = b_n = c_n$, and so a_n shares a tail with c_n .

A property of sequences is called a tail property if it is the same for sequences sharing a tail.

Here is an example: the property of having all the elements in a sequence bounded from above is a tail property, though it is not immediately obvious that it is.

Lemma: Every finite set has an upper bound

Proof: By induction on the size of the set. If the size is 0, the set is empty and so any number (say, 1) is an upper bound. Assume every set of size n is bounded. If A is a set of $n + 1$ elements, let $x \in A$. $A - x$ is a set of n elements, so it has an upper bound, b . If $b \geq x$, then b is an upper bound for A . Otherwise $b < x$, and so x is an upper bound for A . Either way, A is bounded from above.

Now, if a_n is a sequence which shares a tail with b_n , and the latter is bounded from above by b , then for $n > N$, $a_n = b_n \leq b$. If a is an upper bound for a_1, \dots, a_{N-1} , then $\max a, b$ is an upper bound for all of a_n .

Being bounded from above by a particular number is not, of course, a tail property. It can be explicitly transformed to a tail property by looking only at tails: a sequence a_n will be said to be “eventually bounded from above” by a if there exists a N such that for all $n > N$, $a_n \leq a$. It is straightforward to verify that this is a tail property. Similarly, a sequence is defined to be “eventually bounded from below” by a if there exists a N such that for all $n > N$, $a_n \geq a$.

A sequence, a_n is eventually ϵ -close to a if it is both eventually bounded from above by $a + \epsilon$ and eventually bounded from below by $a - \epsilon$. If a_n is ϵ -close to a for all $\epsilon > 0$, then a_n converges to a .

Note that any sequence that converges has a bounded tail, and therefore it is bounded. In other words, an unbounded sequence will not converge to anything.

A sequence is said to be increasing if for all n , $a_n \leq a_{n+1}$. A sequence is said to be decreasing if for all n , $a_n \geq a_{n+1}$. A sequence is said to be monotonic if it is either increasing or decreasing.

Theorem: in a complete ordered field, every monotonic sequence converges.

Proof: First assume that a_n is increasing. Let a be a least upper bound. Then $a_n \leq a < a + \epsilon$ for every n and ϵ . Since a is a least upper bound, and $a - \epsilon < a$, then $a - \epsilon$ is not an upper bound. So there is a N , $a - \epsilon < a_N$. If $n > N$, $a - \epsilon < a_n < a + \epsilon$.

Now if a_n is decreasing, then $-a_n$ is increasing and converges to a . Note that $|-a_n - a| = |a_n - (-a)|$, and so a_n converges to $-a$.

QED

Note that this theorem is only true in complete ordered fields.

Limit Properties

As before, “sequences” here implicitly mean “in an ordered field”. Where a complete ordered field is assume, it will be explicitly noted.

Assume that a_n and b_n are sequences which converge to a and b respectively.

$$-a_n - (-a) = -a_n + a = a - a_n = -(a_n - a)$$

and so $|a_n - a| = |(-a_n) - (-a)|$, so $-a_n$ converges to a .

$$|(a_n + b_n) - (a + b)| = |(a_n - a) + (b_n - b)| \leq |a_n - a| + |b_n - b|$$

If we choose N such that the appropriate tails of a_n and b_n are $\epsilon/2$ -close to a and b respectively, then the sum above will be $< \epsilon/2 + \epsilon/2 = \epsilon$. Therefore, $a_n + b_n$ converges to $a + b$ (or in other words, the limit of a sum is the sum of the limits). This technique is known as the “half- ϵ ”, although in many applications, dividing by more than two is necessary.

What about multiplication?

$$|a_n b_n - ab| = |a_n b_n - a_n b + a_n b - ab| \leq |a_n(b_n - b)| + |(a_n - a)b| = |a_n||b_n - b| + |b||a_n - a|$$

Note that since a_n converges, it is bounded from both above and below, and so $|a_n|$ is bounded by a number, M .

Define N such that the N -tail has $|b_n - b| < \epsilon/2(M + 1)$ and $|a_n - a| < \epsilon/2(|b| + 1)$ (we add 1 to avoid dividing by zero). Then for $n > N$

$$|a_n||b_n - b| + |b||a_n - a| \leq M|b_n - b| + |b||a_n - a| < M\epsilon/2(M + 1) + |b|\epsilon/2(|b| + 1) < \epsilon$$

So the limit of a product is the product of the limits.

If a_n converges to $a > 0$, then a tail will be bounded from below by $a - a/2 = a/2$. Since the following only deals with tail properties, a simplifying assumption is that $a_n > a/2$ for all n . In particular, $a_n \neq 0$ and so a_n^{-1} is well defined.

$$|a_n^{-1} - a^{-1}| = |a - a_n|/(|a_n|a) < |a - a_n|/(2|a|^2)$$

If we take N for a_n to be $\epsilon(2|a|^2)$ -close to a , then the expression above is $< \epsilon$.

If $a < 0$, then $-a_n$ converges to $-a > 0$, $1/-a_n = -1/a_n$ converges to $1/-a = -1/a$ and so $1/a_n$ converges to $1/a$. In summary, the limit of the inverse is the inverse of the limit.

If $b > a$, let $\epsilon = (b - a)/2$ and find N where for $n > N$ both a_n is ϵ -close to a and b_n is ϵ -close to b , then $b_n - a_n > b - \epsilon - (a + \epsilon) = b - a - 2\epsilon = d - d = 0$, so $b_n > a_n$ at the tail. In particular, if $a_n \geq b_n$ on a tail, then it cannot be the case that $b > a$ so $a \geq b$.

The summary of these results is that in an ordered field, limits preserve the ordered field structure (weakly, for the order).

Assume a_n and c_n both converge to a , and $a_n \leq b_n \leq c_n$. For a tail, $a - \epsilon < a_n \leq b_n \leq c_n < a + \epsilon$, so b_n also converges to a . This result is called “the sandwich theorem”, and is often a powerful aid to finding and proving limits.

Cantor's Lemma

A (closed) interval, $[a, b]$ is a set of numbers: $\{x : a \leq x \leq b\}$. Intervals can contain other intervals: $[a, b] \subset [c, d]$ if and only if $a \leq c$ and $d \leq b$.

The next lemma allows defining limits without knowing what they are, or showing that a sequence of intervals has a point in common.

Assume $a_n \leq b_n$ are two sequences, a_n monotonically increasing and b_n monotonically decreasing. Also assume that the sequence $b_n - a_n$ converges to 0.

These sequences can be thought of a sequence of intervals. The monotonicity requirements mean that each interval contains all following ones. The convergence requirement says that the limit of the lengths of the intervals is 0.

Since b_n is decreasing, $b_n \leq b_1$, and so $a_n \leq b_1$. Therefore, a_n is bounded, and so has a limit, a . For similar reasons, b_n is bounded and so has a limit, b .

By the results on the limits of addition and negation, $b - a = 0$, so $b = a$. Call the limit c . Limits for monotonic sequences are the bounds, so $a_n \leq c \leq b_n$ for every n .

If $d < c$ then because c is the least upper bound on a_n , for some n , $d < a_n$, and therefore $d \notin [a_n, b_n]$. Similarly, if $d > c$, some b_n is smaller than d , and therefore $d \notin [a_n, b_n]$.

The only number that is common to all intervals is c .

This result is known as “Cantor’s Lemma”.

Here is an application of this lemma: if $f : N \rightarrow [0, 1]$ is a function there is a number x , $f(n) \neq x$ for every n .

Define a sequence by induction.

$a_0 = 0$ and $b_0 = 1$.

look at $f(n)$. For $n + 1$, define $c_n = b_n - a_n$. $[a_n, b_n]$ is the union of three intervals:

- $[a_n, a_n + c_n/3]$
- $[a_n + c_n/3, a_n + 2c_n/3]$
- $[a_n + 2c_n/3, b_n]$

Not all of them can contain $f(n)$, because the intersection of all three is empty. For $[a_{n+1}, b_{n+1}]$, choose the smallest that does not contain $f(n)$

Note that by induction, $0 < b_n - a_n = 3^{-n} < 1/n$ and so the conditions for Cantor's Lemma are fulfilled.

Let c be the common limit. By the $n + 1$ 'st stage, $c \neq f(n)$. This proves the result.

This result is often summarized as "there are more real numbers between 0 and 1 than

Subsequences and convergence

If a_n is a sequence of real numbers, and n_k is a sequence of natural numbers that is increasing (that is, $n_{k+1} > n_k$), then we can define a sequence a_{n_k} . This is called a subsequence of a_n .

Subsequences interact with convergence in a few interesting ways. First, if a_n converges to a , then every subsequence does as well.

To prove this, we will first show a useful lemma: if n_k is an increasing sequence of natural numbers, then $n_k \geq k$. We show this by induction: n_0 is a natural number, so by definition, $n_0 \geq 0$. Assume $n_k \geq k$, then $n_{k+1} > n_k \geq k$. Since $n_{k+1} > k$, then $n_{k+1} \geq k + 1$.

Now, let $\epsilon > 0$. Since a_n converges to a , there is an N such that $n > N$ implies $|a_n - a| < \epsilon$. Now, if $k > N$, then $n_k > N$, so $|a_{n_k} - a| < \epsilon$. Since this is true for every ϵ , we have proved the subsequence converges.

Since a_n is a subsequence of itself (via $n_k = k$), if every subsequence converges to a , then so does a_n . Note that if a_n converges to a , then so does every subsequence as shown above, and therefore, so does every subsubsequence. In fact, a weaker result is true: if every subsequence has a subsubsequence that converges to a , then so does a_n .

In order to prove this, we need to master a technique that is mechanical yet subtle: reversing complicated statements. We need to specify what it means for a sequence to *not* converge to a . If a_n does not converge to a , then there exists an $\epsilon > 0$ such that for every N there is an $n > N$ such that $|a_n - a| \geq \epsilon$.

The easiest way to approach the theorem is to prove the logical converse: if a_n does not converge to a , then there is a subsequence with no subsubsequence that converges to a .

Let a_n be a sequence, and let us assume a_n does not converge to a . Let $N = 0$. Then we can find, as above, n_0 , so that $|a_{n_0} - a| \geq \epsilon$. Now we define a subsequence recursively: let $N = n_k$, find $n_{k+1} > N$ such that $|a_{n_{k+1}} - a| \geq \epsilon$. Note that since every element of this subsequence is ϵ -away from a , then no subsubsequence can converge to a .

We note that here we did not take advantage of the properties of real numbers, and indeed, the above results are true for rational numbers as well. The following result is not (and, indeed, it is possible to show that it is unique to the real numbers).

Let a_n be a bounded sequence, $-M < a_n < M$ for every n . Then it has a converging subsequence. Since this result is so important, both for understanding real numbers and, practically, to prove results about real numbers, there will be two proofs given.

We define two sequences, c_k increasing and d_k decreasing, by induction, with the property that there is an infinite number of a_n such that $c_k \leq a_n \leq d_k$ for every k . We set $c_0 = -M, d_0 = M$. Since all a_n satisfy that, an infinite number of them do. Assume we have $c_k < d_k$. Set $m = (c_k + d_k)/2$. Assume a finite number of a_n have $c_k \leq a_n \leq m$ [1] and a finite number of a_n have $m \leq a_n \leq d_k$ [2]. However every a_n is either $a_n \leq m$ or $a_n \geq m$, so only a finite number of a_n have $c_k \leq a_n \leq d_k$. This contradicts the inductive assumption, so one of [1] or [2] is infinite. If [1] is infinite, set $c_{k+1} = m, d_{k+1} = d_k$ otherwise $c_{k+1} = c_k, d_{k+1} = m$. We see that $d_k - c_k = 1/2^k$ and so converges to 0.

We apply Cantor's lemma, and get c_k and d_k converge to a . Set $n_0 = 0$. Then $c_0 = -M < a_{n_0} < M = d_0$. Recursively, find $n_{k+1} > n_k$ such that $c_{k+1} \leq a_{n_{k+1}} \leq d_{k+1}$. One such must exist, because there are infinite members of a_n between c_{k+1} and d_{k+1} , but only a finite number $\leq n_k$.

By the Sandwich theorem, we have a_{n_k} converging to a .

Here is another proof for this: let a_n be any sequence. We define a subsequence by induction: let n_0 be such that a_{n_0} is less than any element in the sequence. Let n_{k+1} be such that it is $a_{n_{k+1}} > a_{n_k}$ and the smallest element that is $a_{n_{k+1}} \geq a_{n_k}$. If this construction succeeds, we have an increasing subsequence. This construction might fail: at any stage, the "smallest" element might not exist. However, if the smallest element does not exist, we define a new subsequence, m_k . Without loss of generality, we assume that the smallest element does not exist in the first stage: therefore, for every n , a_n is not the smallest element. We take m_0 to be 0. If we have m_k , we notice there must be an infinite number of n , $a_n < a_{m_k}$.

This means for every sequence, we can either construct an increasing subsequence or a decreasing subsequence. Since bounded increasing or decreasing sequences, we constructed a converging subsequence.

Completeness

We already know that every bounded monotonic sequence converges. This is powerful – we know it converges, even though we cannot, necessarily point to the limit. Clearly, however, this is not necessary for convergence – we have already seen sequences which are not monotonic converge.

If we try to find a universal criterion for convergence, we need to start with things of which we are sure. We know, for example, we need to limit ourselves to tail properties. We also know that for any given ϵ , from some point on $|a_n - a| < \epsilon$. However, that mentions the limit, and we would like to avoid that. One way of doing it is noting that in that tail $|a_n - a_m| < 2\epsilon$, for any n, m . As we have already seen in our proof for convergence of addition, things like 2ϵ are not interesting: after all, if we had taken the tail corresponding to $\epsilon/2$, we would have gotten $|a_n - a_m| < \epsilon$.

A sequence with the property that for every ϵ , there is an N such that for all $n, m > N$, $|a_n - a_m| < \epsilon$ is called a Cauchy sequence. The discussion above shows that every converging sequence is a Cauchy sequence.

We know that in the rational numbers, not every Cauchy sequence converges: for example, the sequences we defined as a_n and b_n in the chapter about irrational numbers are Cauchy, but do not converge in the rational numbers. Is it possible for sequences of real numbers to be Cauchy and not to converge in the real numbers? It turns out the answer is negative.

Assume a_n is Cauchy. We remember that despite not looking it, being bounded is a tail property. If we take $\epsilon = 1$, we have that for some N , if $n > N$ then $|a_n - a_0| < 1$, or $a_0 - 1 < a_n < a_0 + 1$. Since the sequence has a bounded tail, it is bounded. Therefore, it has a converging subsequence. Call the subsequence a_{n_k} . Now let $\epsilon > 0$. Find N from the Cauchy property such that if $n, m > N$ $|a_n - a_m| < \epsilon/2$. Find M from the convergence of the subsequence such that if $k > M$, $|a_{n_k} - a| < \epsilon/2$. Now let O be the biggest of N and M . Then if $n > O$ then $n_{O+1} \geq O+1 > O \geq N$, so $|a_n - a_{n_{O+1}}| < \epsilon/2$ and $|a_{n_{O+1}} - a| < \epsilon/2$. Adding the inequalities, and using the triangle inequality, we get that if $n > O$ then $|a_n - a| = |a_n - a_{n_{O+1}} + a_{n_{O+1}} - a| \leq |a_n - a_{n_{O+1}}| + |a_{n_{O+1}} - a| < \epsilon/2 + \epsilon/2 = \epsilon$. Therefore, a_n converges to a .

This gives us a general machine to produce limits: define a sequence, show it is Cauchy, and then take its limit.

As an example of how to use the machine, we define here the decimal expansion.

Let d_n be a sequence of integers between 0 and 9. We will define a sequence by induction: $a_0 = 0$, $a_{n+1} = a_n + d_n/10^n$. In order to prove it is Cauchy, we will first note $|a_{n+1} - a_n| \leq 9/10^n$. We will prove by induction $|a_{n+k} - a_n| \leq 9/10^n [1 - 1/10^k]/[1 - 1/10]$. The case for 0 is trivial, since the difference between an element and itself is 0. Assume it holds for k , and now $|a_{n+k+1} - a_n| \leq |a_{n+k+1} - a_{n+k}| + |a_{n+k} - a_n| \leq 9/10^{n+k} + 9/10^n [1 - 1/10^k]/[1 - 1/10] = 9/10^n [1/10^k + [1 - 1/10^k]/[1 - 1/10]] = 9/10^n [1 - 1/10^{k+1}]/[1 - 1/10]$.

We also note that $9/10^n \leq 9/n$. Thus, if $N > 2 * 9/\epsilon$, $|a_{N+k} - a_N| < \epsilon/2$, and thus $|a_{N+k} - a_{N+l}| < \epsilon$, then $n = N + k, m = N + l$ for some k , we see that the sequence we defined is Cauchy. We call its limit $0.d_0d_1\dots$ the “decimal number” corresponding to the sequence of digits. Note that we have not used any special properties of 10 here, so that we can have expansions by any base.

(A proper treatment of decimal expansions would show that every real number between 0 and 1 has a decimal expansion, and would analyze the circumstances under which it is unique, but this is beyond the scope of showing the Cauchy machinery in use.)

Building the Real Numbers

At this point, with some understanding on how sequences and convergence work, we would like to take a step back and pay a debt so far, we have been taking the real numbers' existence as an assumption. Now, we would like to build them.

We already have the rational numbers. We look at Cauchy sequences of rational numbers. Let a_n and b_n be Cauchy sequences of rational numbers. If $b_n - a_n$ converges to 0, we will call them equivalent. It is straightforward to see that this is an equivalence relation. We define addition and multiplication as $a_n + b_n$ and $a_n b_n$. It is straightforward to see that these definitions respect the equivalence relation. We also note that if we identify a rational number, q with the sequence q, q, \dots , then different rational numbers are not equivalent and a sequence a_n is equivalent to q if and only if a_n converges to q .

We define the real numbers as the equivalence classes of Cauchy sequences of rational numbers. In order to prove it is a field, we need to show that if a_n is not equivalent to 0, it has an inverse. Not being equivalent to 0 means there is an ϵ such that for every N there is an $n > N$ such that $|a_n| > \epsilon$. Since a_n is Cauchy, let N be such that $|a_n - a_m| < \epsilon/2$. Since there is an $n > N$, $|a_n| > \epsilon$. Assume $a_n > \epsilon$, then if $m > N$, $a_m > a_n - \epsilon/2 > \epsilon/2$. In particular, if $m > N$, $a_m \neq 0$. Let's define a sequence, b_n to be $1/a_n$ if $n > N$ and 0 otherwise. Then $a_n b_n = 1$ if $n > N$, and so $a_n b_n$ is equivalent to 1 – so b_n is an inverse of a_n .

We define $a_n \leq b_n$ if $b_n - a_n$ has a tail bounded from below by ϵ for every ϵ . The definition is subtle because we do want $(1/n) \leq 0$. Checking that with this ordering relationship, the field becomes ordered is straightforward.

Let a_n be a Cauchy sequence of rationals. Choose N that corresponds to 1 for this sequence. Then a_{N+1} is a rational number, and if $n > N$, $a_n < a_{N+1} + 1$. Since for every rational number there is a greater natural number M , $a_n N$ and $M > a_{N+1}$. Therefore, $a_n \leq M$ according to the definition above.

Now assume that m is a natural number, and a is a real number. There exists a natural number, $n > ma$ and so $a < 0$. For every natural number n there is a natural number m such that $m/2^n > b$. Since b is an upper bound, that means that for every a in A , $a \leq m/2^n$. For every n , we take the smallest m that satisfies the property (since any set of natural numbers has a minimal element). Now we note that if $m/2^n$ satisfies that property, then for $n + 1$, either $2m/2^{n+1}$ satisfies the property, or $(2m-1)/2^{n+1}$ because if $(2m-2)/2^{n+1}$ is an upper bound, then so is $(m-1)/2^{n+1}$ in contradiction to minimality. In particular, if we define this sequence as c_n , then $|c_n - c_{n+1}| < 1/2^{n+1}$. By induction, we can see that $|c_n - c_m| < 1/2^n$. Therefore, c_n is a Cauchy sequence of rationals. We can see that $c_n > a$ for every a in A , so c is an upper bound for A . We also note that $c \leq c_n$ for every n , since c is a decreasing sequence.

Is it a minimal upper bound? Assume $d < c$. In that case, $1/(c-d) > 0$. Let us take a natural number, $N > 1/(c-d)$, and so $2^N > N > 1/(c-d)$. Therefore, $c-d > 1/2^N$, or $c > d + 1/2^N$. But since d is an upper bound, $c - 1/2^N \geq a$

for all a in A . $c_N - 1/2^N \geq c - 1/2^N \geq a$, and so for stage N , m was not minimally chosen, in contradiction. Therefore c is a least upper bound.

For a general A , let a_0 be a member in A , and let A' be the set all numbers expressible as $a - a_0 + 1$ for a in A . if A is bounded from above, so is A' , and since 1 is in A' , A' has a least upper bound c . But if $a - a_0 + 1 \leq c$, $a \leq c + a_0 - 1$ and so $c + a_0 - 1$ is a least upper bound for A .

This shows that rational Cauchy sequences under the right equivalence relation is a complete ordered field, and so a complete ordered field exists.

Let a_n be a sequence. We define another sequence, called the partial sums, S_n , by induction: $S_0 = a_0$ and $S_{n+1} = S_n + a_{n+1}$. Intuitively, we think of those sums as approaching summing up all elements in a_n . If the sequence of partial sums converges, we call this the sum of a_n , and denote it by $\sum a_n$.

As a simple example, if $a_n = 1/2^n$, then $S_n = 2 - 1/2^n$ (by induction). Then S_n converges to 2 and $\sum 1/2^n = 2$.

A “series” is the partial sums of a sequence, and so we say that the series a_n converges if S_n converges.

We note that although the sum is not a tail property, the convergence of the series is: if S_n converges, after all, so does S_{n+N} , since sequence convergence is a tail property, and then so does $S_{n+N} - S_N$ – but this is the sequence of sums of the N -tail.

We also note that since for a sequence to converge it must be Cauchy, if S_n is Cauchy then for every ϵ there is an N such that if $n, m > N$ $|S_n - S_m| < \epsilon$. In particular, if $m = n + 1$, $|a_n| < \epsilon$ and therefore a_n converges to 0. This is our first negative criterion of series convergence: if a sequence does not converge to 0, its series does not converge.

Assume a_n is a sequence such that the series $|a_n|$ converges. In that case, the partial sum sequence is therefore Cauchy, and so converges. If we call the partial sum sequence for $|a_n|$ by the name of T_n , we can show by induction on k that $|S_{n+k} - S_n| \leq |T_{n+k} - T_n|$ – the induction step is the triangle inequality. Therefore S_n is Cauchy as well, and so converges. When the series a_n converges, we say that a_n converges absolutely.

Assume $0 \leq a_n \leq b_n$ and the series b_n converges. Therefore the sequence of b partial sums is bounded. Since the sequence of b partial sums, at each point, is greater than the sequence of a partial sums, we get that the partial sum sequence for a is bounded. We also see that the sequence of partial sums for a is monotonically increasing since $0 \leq a_n$. Therefore, the series a_n converges. In particular, if $|a_n| \leq |b_n|$, and b_n converges absolutely, so does a_n .

This is our first criteria of positive convergence. As an application, note that if $g(g + 2g + g)/4 > g$, but since $h < 1$, $h^2 > 0$ and $h = 1/(1 + a)$. We note that $(1 + a)^n \geq 1 + na$ (by induction, for example) and so $h^n < 1/na$, and so $nh^n < 1/a$. In particular, $ng^n < nh^{2n} = nh^n h^n < h^n/a$. Since the series h^n converges, so does the series h^n/a and by the criteria above, $\sum ng^n$ also converges.

This is an example of a so-called power series: let a_n be a sequence. For any x , the series $a_n x^n$ is called the power series of a_n at x . We have shown, above, that the power series $\sum nx^n$ converges for every $|x| < 1$. Then for any natural n , $|a_{n+N} x^{n+N}| < |a_N x^N| (|x|/N)^n$. Since series convergence is a tail property, we see that the series $a_n x^n$ converges, so that the power series of the inverse factorial converges at every point!

For further analysis of power series, it is useful to have the following result: if $\sum a_n$ converges absolutely, and k_n is a sequence of natural numbers such that every natural number appears exactly once, then $\sum a_{k_n}$ also converges, and to the same sum. Since k_n includes every natural number, for every natural number, we can ask at which place does it appear. We call that sequence l_n . Since a_n converges absolutely, given an ϵ , there is a N , such that $|S_{n+j} - S_n| < \epsilon$, $j > 0$ where S is the sequence of partial sums of $|a_n|$. Treating that a sequence of k , and looking at the limit, we can see that $\sum |a_{n+j}| \leq \epsilon$. If we take an M large enough that $n > M$ implies both $k_n > N$ and $l_n > M$, we have $|S[a_{k_n}] - S[a_n]| < 2\epsilon$, and therefore $|\sum a_{k_n} - \sum a_n| < 2\epsilon$. Since this is true for every ϵ , $|\sum a_{k_n} - \sum a_n| = 0$.

This is useful for power series in the following way: if we have $\sum a_n x^n$ and $\sum b_n x^n$ power series, what happens when we multiply them? Inside the common region of convergence, we know that the multiplication is well defined. So we have that $S[a_n x^n] S[b_n x^n]$ converges, and so if we use the right k_n , and move to a subsequence, we can have that $\sum [a_0 b_n + a_1 b_{n-1} + \dots + a_{n-1} b_1 + a_n b_0] x^n$ is the multiplication.

CHAPTER 14

Real functions

Up until now, we have dealt with sequences – functions from the natural numbers to real numbers. Now we turn our attention to functions from real numbers to real numbers, so-called “real functions”.

A real function is a function from a set, A to the real numbers, \mathbb{R} . Usually A will be an “interval” – all numbers between two numbers, all numbers less than a number, all numbers greater than a number or all numbers.

We will usually keep A implicit, and assume it contains anything we are interested in.

If f and g are functions, $f + g$ is the function defined by $(f + g)(x) = f(x) + g(x)$. Similarly, we define fg , $-f$ and $f - g$. If f is a function, everywhere that $f(x) \neq 0$, we can define $1/f$.

We note that with these definitions, functions on a given set A obey the so-called ring axioms.

The simplest example of a function is $f(x) = a$, the constant function. The next-simplest example of a function is $f(x) = x$, the identity function.

With these two functions,

Examples

Assume $f(x)$ is a real function defined on some interval containing a point x_0 . We want to have the equivalent of a “tail property” for sequences – but this time for functions near x_0 . A property will be said to be a “local property” if whenever there exists a $\delta > 0$ such that $f(x) = g(x)$ when $|x - x_0| < \delta$, $f(x) = g(x)$ for every x such that $0 < |x - x_0| < \delta$ such that for all x such that $0 < |x - x_0| < \delta$ such that for the appropriate x , $f(x) < A$ ” is an lpf property. Similarly, $B < f(x)$ is an lpf property, and so is $B < f(x) < A$. In that case, we will say that f is locally between B and A .

Let C be a number, and assume that for every ϵ , f is between $C - \epsilon$ and $C + \epsilon$. As a synonym, we will also say that f has C as a limit at x_0 . In that case, we say that f tends to C as x tends to x_0 . Making the definition more explicit: if for every ϵ there exists a δ such that if $0 < |x - x_0| < \delta$ then $C - \epsilon < f(x) < C + \epsilon$ (or, equivalently, $|f(x) - C| < \epsilon$). We know that there is a $\delta > 0$, such that if $0 < |x - x_0| < \delta$, $|f(x) - C| < \epsilon$, $0 < |x_n - x_0| < \delta$, $|f(x_n) - C| < \epsilon$, we see that $f(x_n)$ is a sequence converging to C . So we have: if f has a limit at x_0 , $f(x_n)$ will converge to the same limit if for every sequence x_n converging to x_0 .

Now, assume that for every sequence converging to x_0 , $f(x_n)$ converges to C . Does that mean that f has a C as a limit at x_0 ? Assume the opposite. Then there is some ϵ_0 such that for every δ , we have some x such that $0 < |x - x_0| < \delta$ and $|f(x) - C| \geq \epsilon_0$. For $\delta = 1/n$, take x_n to be one such x . Then $0 < |x_n - x_0| < 1/n$, so that $f(x_n)$ does not converge to C , in contradiction to the original assumption. Therefore, if applying f to every sequence converging to x_0 yields a sequence converging to C , then f has a C as a limit at x_0 .

Because of that, we can use much of what we know of sequence convergence for function limits. For example, this implies that $f + g$ will have as its limit the sum of the limits, and similarly for fg and f/g .

Let f be a function. If f has $f(x_0)$ as a limit at x_0 , we say that f is continuous at x_0 . If A is a set of real numbers and f is continuous at every point of the set, we say that f is continuous on A . This is especially appropriate when A is a segment: all points between a and b .

Assume that f is continuous on $[a, b]$, all points between a and b , inclusive. We want to show that there is some M such that $|f(x)| \leq M$, $a \leq x \leq b$. Since x_n is a bounded sequence of real numbers, there is a convergent subsequence x_{n_k} . We note that since $n_k \geq k$, $|f(x_{n_k})| > n_k \geq k$. x_{n_k} converges to x_0 , and so $a \leq x_0 \leq b$. Therefore, $f(x_{n_k})$ converges to $f(x_0)$. But $f(x_{n_k})$ is not a bounded sequence, which is a contradiction. Therefore, f is bounded on A .

Let us again assume that A is a segment, and f is continuous there. Assume that there is a ϵ_0 such that for every δ there is an x, y such that $|x - y| < \delta$ and $|f(x) - f(y)| \geq \epsilon_0$. Now let us take x_n to be such an x for $\delta = 1/n$. x_n has a convergent subsequence, x_{n_k} converging to x_0 . We know that f is continuous at x_0 . Let us take a δ corresponding to $\epsilon_0/2$. We have that if $|x - x_0| < \delta$ then $|f(x) - f(x_0)| < \epsilon_0/2$. Now we know that for every x_{n_k} there is a y_k such that $|x_{n_k} - y_k| < \delta$ and such that $|f(x_{n_k}) - f(y_k)| \geq \epsilon_0/2$. We have $|y_k - x_0| < \delta$, so $|f(y_k) - f(x_0)| < \epsilon_0/2$, and therefore we have $|f(x_{n_k}) - f(y_k)| = |f(x_{n_k}) - f(x_0) + f(x_0) - f(y_k)| \leq |f(x_{n_k}) - f(x_0)| + |f(x_0) - f(y_k)| < \epsilon_0/2 + \epsilon_0/2 = \epsilon_0$ in contradiction to the definition of k . Therefore, there is no such ϵ_0 , and so for every ϵ there is a δ such that if $|x - y| < \delta$ then $|f(x) - f(y)| < \epsilon$. In this case, we call f uniformly continuous on A . As we have seen, every continuous function on a segment is uniformly continuous.

Assume f is a continuous function on a segment A . We know that f is bounded from above. In other words, the set $B = \{f(x) : x \in A\}$ has an upper bound, and therefore a least upper bound, say c . In other words, if $d < c$, there is an x , $d < f(x) \leq c$. Define x_n to be such an x for $d = c - 1/n$. Now we have x_{n_k} a converging subsequence, and $c - 1/k \leq c - 1/n_k < f(x_{n_k}) \leq c$. Assume that x_{n_k} converges to x_0 . Then $f(x_{n_k})$ converges, by continuity, to $f(x_0)$, and by sandwich, to c . Therefore, $f(x_0) = c$ and so the least upper bound is actually attained (in this case, we call it a maximum). For similar reasons, f has a minimum on A .

Let f be a continuous function on a segment $A = [a, b]$. If c is such that $f(a) < c < f(b)$, then there is an x , $f(c) = x$. Let a_n, b_n be defined by $a_0 = a, b_0 = b$. Let $t = (a_n + b_n)/2$. If $f(t) < c$, then $a_{n+1} = t, b_{n+1} = b_n$. If $f(t) > c$, then $a_{n+1} = a_n, b_{n+1} = t$. If $f(t) = c$, then $a_{n+1} = t, b_{n+1} = t$. Define δ to be the smaller of the two. If $|x - y| < \delta$, then $|f(x) - f(y)| < \epsilon$. Applying f^{-1} to the inequality (since f^{-1} is increasing) gives us $t - \epsilon < f^{-1}(y) < t + \epsilon$, or $|f^{-1}(y) - f^{-1}(x)| < \epsilon$. If $a > 0$ is a real number, then if we find $M > a, M > 1, M^n > a$. $0^n = 0$, x^q is a well defined real function on the non-negative numbers (and therefore, if we define $x^{-q} = 1/x^q$, we have x^q is a well defined function on the positive numbers for any rational q !

Let f be a function. For x_0 , and x , we can define a function on h as $[f(x) - f(x_0)]/[x - x_0]$. That measures the average change in f between x_0 and x . It is not well-defined, of course, at x_0 itself. However, that does not stop us from asking about what limit it has at x_0 . There might, of course, not be a limit. If there is one, however, we call that limit “the derivative of f at x_0 ”, and write $f'(x_0)$.

We note that if $f'(x_0)$ exists, then f must be continuous at x_0 .

From the definition, and the fact that limits of sums are sums of limits, we see that $(f + g)' = f' + g'$. With a bit of algebra, it is easy to see that $(fg)' = f'g + fg'$. If f and g are functions, we define $(f \circ g)(x) = f(g(x))$. With this definition, and the kind of ϵ/δ games we are used to, it is straightforward to show that $(f \circ g)'(x_0) = f'(g(x_0))g'(x_0)$.

Assume $f'(x_0) > 0$. Let $\epsilon = f'(x_0)/2$. Then by definition of limit, there is a δ , $0 < |x - x_0| < \delta$ implies that $|[f(x) - f(x_0)]/[x - x_0] - f'(x_0)| < \epsilon$. If $0 < x - x_0 < \delta$, then $[f(x) - f(x_0)]/[x - x_0] > f'(x_0) - \epsilon = f'(x_0)/2 > 0$. If $0 < x_0 - x < \delta$ then $[f(x) - f(x_0)]/[x - x_0] < f'(x_0) - \epsilon = f'(x_0)/2 < 0$. Therefore, on any segment that contains x_0 (except as an endpoint) $f(x_0)$ is not a minimum or a maximum. For analogous reasons, if $f'(x_0) < 0$, it is neither a minimum or a maximum. In other words, if $f'(x_0)$ is a minimum or a maximum, even if only on a short segment that includes x_0 not as an endpoint, then $f'(x_0) = 0$.

Roll's Theorem

We note here that if $f(x) = ax + b$, then $f(x) - f(x_0) = a(x - x_0)$ and so $[f(x) - f(x_0)]/[x - x_0] = a$, and so $f'(x) = a$ for every x .

Let f be a derivable function on a segment $A = [a, b]$, and assume that $f(a) = f(b)$, then there is a number c such that $a < c < b$ and $f'(c) = 0$. Why?

Since f is derivable, it is continuous. So it attains its minimum m and its maximum, M . If $m = M$ then all values of f are the same, and a constant function has derivative 0 at all points. Otherwise, one of those is not equal to $f(a)$. The point where the one not equal to $f(a)$ is attained, is not b , and so it is a c such that $a < c < b$. But at this point, a minimum or a maximum is attained not on an endpoint, so that $f'(c) = 0$.

Now let f be any derivable function on $[a, b]$. Define $S = [f(b) - f(a)]/[b - a]$, the slope of f . Define g by $g(x) = f(x) - [Sx - Sa]$. Then $g(a) = f(a)$, and $g(b) = f(b) - Sb = f(b) - S[b - a] = f(b) - [f(b) - f(a)]/[b - a][b - a] = f(b) - [f(b) - f(a)] = f(b) - f(b) + f(a) = f(a)$. Therefore, by the previous discussion, there is c , $g'(c) = 0$. But $g'(x) = f'(x) - S$, and so $f'(c) = S = [f(b) - f(a)]/[b - a]$. So we see that if a function is derivable on an interval there is a point where the derivative is equal to the slope.

Here is a simple use: if $f'(x) \geq 0$ for every x if $a < x < b$, then f is increasing. In order to see why, let $a \leq c < b$. For similar reasons, if the derivative is always negative, the function is decreasing.

Another way to write the same equation, by the way, is $f(b) = f(a) + f'(c)[b - a]$ for some c , $a < c < b$.

Assume f is a bounded function on $[a, b]$. Also assume that $f \geq 0$ on this interval. If a_i is a finite increasing sequence such that $a_0 = a$ and $a_n = b$, we call it a division of $[a, b]$. If a_i is a division, for $[a_i, a_{i+1}]$, we know that f is bounded there, and so it has a least upper bound. We will call that $f^u[a_i, a_{i+1}]$. Now we can define the upper step-sum of f on a division as $f^u[a_0, a_1](a_1 - a_0) + \dots + f^u[a_{n-1}, a_n](a_n - a_{n-1})$. If f is bounded from above by M , this sum is between $M[b - a]$ and 0.

If a_i is a division, and a'_j is a division such that for every i there is a j such that $a'_j = a_i$, then we say that a'_j is a subdivision of a_i . If a'_j is a subdivision of a_i , then we can prove that the step-sum of f on a' is less or equal to than the step-sum of f on a . Since both are finite, in particular there are only a finite number of a'_j not in a_i . Therefore, we can prove this by induction on the number of elements in a' which are not in a . If there are 0, the divisions are the same. Assume it is true when there are n elements in a' which are not in a , and assume that a' has $n + 1$ such elements. Let j_0 be an element of a' not in a . We define $a''_i = a'_i$ if $i < j_0$ and $a''_i = a'_{i+1}$ otherwise. Then a'' only has n items, and so the step sum on a'' is less than that on a' . However, the step sum on a' and a'' only differ in that in the latter, the term $f^u[a'_{j_0-1}, a'_{j_0+1}](a'_{j_0+1} - a'_{j_0-1})$ appears while in the former, $f^u[a'_{j_0-1}, a'_{j_0}](a'_{j_0} - a'_{j_0-1}) + f^u[a'_{j_0}, a'_{j_0+1}](a'_{j_0+1} - a'_{j_0})$. But since $f^u[a'_{j_0-1}, a'_{j_0+1}] \geq f^u[a'_{j_0-1}, a'_{j_0}]$, $f^u[a'_{j_0}, a'_{j_0+1}]$, the former is less or equal to the latter, and so the step-sum on a' is less or equal to the step-sum on a'' , as required.

If a_i is a division and a'_i is a different division, we can define a common subdivision of both, by induction: $a''_0 = a$, and if we have a''_i , we define a''_{i+1} as the least of a' and a that is greater than a''_i .

We can define the lower step-sum analogously by taking the infimum of f on the intervals. The analogous results hold, with inequality signs reversed.

We define the integral of f from a to b to be the least upper step-sum of f and the highest lower step-sum of f . If they are different, f does not have an integral (or, more properly, a Riemann integral).

Any continuous function has an integral, as the difference between the upper step-sum and the lower step sum of a division is no more than ϵ is any two elements in the division are no more than δ apart, and δ corresponds to ϵ for uniform continuity, and there are always such divisions – since if $1/n < \delta/(b - a)$, then $a_i = a + (b - a)i/n$ is such a division.

Assume that f is bounded. We define $f^+ = (|f| + f)/2$, $f^- = (|f| - f)/2$. If f is bounded, so are f^+ and f^- , which are both non-negative. $f = f^+ - f^-$. We define the integral of f as the integral of f^+ minus the integral of f^- . We will note that if f is continuous, then so is $|f|$ and then so are f^+ and f^- , and therefore any continuous function, not necessarily non-negative, still has an integral.

It is trivial to prove that $\int_a^b f = \int_a^c f + \int_c^b f$ – for any division, have a subdivision that includes c .

It is also trivial to prove that if $m < f < M$ then $m(b-a) \leq \int_a^b f \leq M(b-a)$.

Assume f is continuous. $\int_a^{x+h} f - \int_a^x f = \int_x^{x+h} f$. If δ is such that $0 < h < \delta$ such that $|f(x+h) - f(x)| < \epsilon$, then for such h , $0 \leq f \leq h$ implies $f(x) - \epsilon < f(x+h) < f(x) + \epsilon$, and so $f(x) - \epsilon < (\int_x^{x+h} f)/h < f(x) + \epsilon$. Analogous proof shows the same for $h < 0$, and so the derivative of $\int_a^x f$ is $f(x)$ when f is continuous.

Now assume that f is a function, and f' is continuous. Then the derivative of the integral of f' and the derivative of f are the same. Therefore, the derivative of their difference is always 0. As a consequence of Roll's theorem, if the derivative is identically 0, the difference is constant. So, the integral of the derivative of a function differs from the function by a constant. This is the fundamental theorem of calculus.

Let us define $e(x) = \sum x^n/n!$. We have already seen that $e(x)$ is defined everywhere. We also note that since $(x+y)^n = \sum_{i=0}^n n!/(i!(n-i)!) x^i y^{n-i}$ and so $(x+y)^n/n! = \sum_{i=0}^n (x^i y^{n-i})/(i!(n-i)!)$ then by the fact that it converges absolutely and by the permutation theorem, we have $e(x+y) = e(x)e(y)$. In particular, we have: $e(q) = e(1)^q$ for any rational q . By definition, we extend this to all numbers, and putting $e = e(1)$ we have $e(x) = e^x$. This is the so-called “natural exponent” function.

We note that $(e(x+h) - e(x))/h = e(x)[e(h) - 1]/h = e(x)[1 + h \sum_{n=1}^{\infty} h^{n-1}/(n-1)!]$. If $|h| < 1$, we have $|\sum h^{n-1}/(n-1)!| < (e-1)$, so for any $t > 1$, there is a number $e^s = t$. If $t > 1$, then we can write $b^2 = a$, $b = 1 + c$. $n/a^n = n/(b^2)^n = n/(b^n)^2 = n/((1+c)^n)^2 > n/(cn)^2 = 1/(c^2)n$ converges to 0. By sandwich, n/a^{n^2} also converges to 0. By continuity, $n^k/(a^k)^{n^2}$ converges to 0 for any real $a > 1$. If $c > 1$, let a be such that $a^k = c$, then n^k/c^{n^2} converges to 0 for any real $c > 1$. Therefore, the limit of $c^{-1/x^2}/x^k$ is 0 at 0. Away from 0, e^{-1/x^2} has a derivative. For any n th order derivative, it is: $\frac{d^n}{dx^n} e^{-1/x^2} = (a_n x^n + \dots + a_0) e^{-1/x^2}$. So we have that we can complete e^{-1/x^2} to 0 with 0, and it has all derivatives at 0 be 0.

For this reason, if we define $f(x) = 0$ if $x \leq 0$, $f(x) = e^{-1/x^2}$ if $x > 0$, we have f have all derivatives. If we define $g(x) = f(x)f(1-x)$ then g has all derivatives, and g is 0 outside of $[0, 1]$. $h = \int_0^x g$ is bounded by $\int_0^1 g$, and has all derivatives. If we assume the bound is M , $t = h/M$ is 0 for $x \leq 0$ and increasing between 0 and 1. If we put $s(x) = t(x/\epsilon)$, then we have $s(x) = 0$ if $x \leq 0$ and $s(x) = 1$ for $x \geq \epsilon$. If we take $q(x) = s(x)s(1-x)$ then we have the function we have promised at the beginning: it is 0 for $x < 0$, it is 1 for $x > \epsilon$. Therefore, if q defines your position on a hundred-meter dash course, then at the beginning, you’re at the beginning. After a minute, you are back at the beginning – and for most of the intervening minute (all but ϵ of it, say a microsecond) you are chatting with your friend at the other end. Note that q is a “nice” function: it has all derivatives, so at any given time, you will have a well-defined speed and acceleration.

CHAPTER 21

A Story

Analyze $z = e^{-1/x^2}$ (infinite derivatives, all 0 at 0)

Construct $w=0+z$ function

Construct $h=w*w(1-x)$

Construct $s = \int w$

Construct n , normalized s

Construct f_ϵ , fast step function

Construct t_ϵ , roughly $f*f(1-x)$

Tell the story