MA2108S (Mathematical Analysis I)

Jia Cheng

January 2021

1 The real numbers

1.1 p-th powers and roots are order preserving ($p > 0$)

It is very easy to see that for $n \in \mathbb{Z}^+$, taking n-th roots on $r > 0$ preserves order. That is, $0 < r_1 < r_2 \implies r_1^{\frac{1}{n}} < r_2^{\frac{1}{n}}$. This can be easily proven by contradiction.

However, later in chapter 3, theorem 3.20(a) uses the idea that given $p > 0$, $\frac{1}{\epsilon} < n^p \iff (\frac{1}{\epsilon})^{\frac{1}{p}} < n$. Here, p is a positive real number, not just a positive integer.

PMA Ex 1.6 gives a definition on rational powers, then real powers of positive real numbers.

Ex 1.6 shows that it makes sense to define rational powers. Now, we show that positive rational powers $q > 0$ are order preserving. Since $q > 0$, $q = \frac{m}{n}$ for some $m, n \in \mathbb{N}^+$. Suppose $0 < r_1 < r_2$. Then $r_1^{\frac{1}{n}} < r_2^{\frac{1}{n}}$. Hence, $(r_1^{\frac{1}{n}})^m < (r_2^{\frac{1}{n}})^m$. By definition of rational powers, $r_1^q < r_2^q$. So now we have proven that taking rational powers is order preserving. This implies that taking rational roots is also order preserving. To define rational roots, we need to prove that $(r^a)^b = r^{ab}$. Then, $(r^{\frac{1}{q}})^q = r^1 = r$ will be a consequence of this.

To show that real powers are also order preserving, I believe it suffices to consider supremums of sets and show that one supremum is larger than the other. (see definition of real powers in Ex 1.6)

I believe we might also need to prove that $(r^{\frac{1}{p}})^p$ indeed recovers r for real p, or more generally, $(r^a)^b = r^{ab}$.

1.2 Methods of bounding $b^{\frac{1}{n}}, b > 1$

In both proofs, we assume $M > 1$

Method 1 (Ex 1.7) In general, $x^n - 1 = (x - 1)(x^{n-1} + \cdots + 1)$
Hence, $b - 1 \geq n(b^{\frac{1}{n}} - 1)$

$b^{\frac{1}{n}} < M \iff b^{\frac{1}{n}} - 1 < M - 1$
It suffices to show $b^{\frac{1}{n}} - 1 \leq \frac{b-1}{n} < M - 1$
The strict inequality will hold iff $n > \frac{b-1}{M-1}$.

Method 2 This uses binomial expansion and Bernoulli's inequality. $b^{\frac{1}{n}} < M \iff b < M^n = (1 + (M-1))^n$
It suffices to show $b < 1 + n(M-1) \leq M^n$.
The strict inequality will hold iff $n > \frac{b-1}{M-1}$.

1.3 Complex Numbers and $\mathbb{R}^2$

Both $\mathbb{C}$ and $\mathbb{R}^2$ are normed spaces. With respect to metrics, these 2 spaces are essentially equivalent.

The only difference between $\mathbb{C}$ and $\mathbb{R}^2$ is the inner product defined on them (and also the multiplication operation on complex numbers). For $\mathbb{R}^2$ (and $\mathbb{R}^n$ in general), it is the dot/scalar product. For $\mathbb{C}$ (and $\mathbb{C}^n$), it is the Hermitian product.
Regardless, the normed induced by these inner products with respect to $\mathbb{C}$ and $\mathbb{R}^2$ are effectively the same. A complex number can be represented by a 2-tuple (a, b), and it's norm is given by $a^2 + b^2$, which the same for a tuple in $\mathbb{R}^2$.

Hence, many theorems that apply to sequences in $\mathbb{R}^2$ later on also apply to complex number sequences.

1.4 Dense sets

(Theorem 1.20)
The notion of denseness in Chapter 1 claims that, to prove S is dense in $\mathbb{R}$, it is equivalent to show that $\forall p, q \in \mathbb{R}, p < q, \exists x \in S, p < x < q$.
Note that in general, this is not always true for any set.

For instance, consider $\mathbb{Z}$. $\mathbb{Z}$ is dense with respect to itself. However, between 0 and 1, there exists no element of $\mathbb{Z}$.

Of course with respect to $\mathbb{R}$, topological denseness of a set S is the same as saying that we can always find another element of S between two elements p, q of $\mathbb{R}$.

1.5 Linear Algebra

The following theorem can be developed without the notion of basis, linear independence or dimension. We simply need definition of orthogonality. It is useful when dealing with orthogonality.

Given subspace U of Euclidean vector space V and vector $v \in V$. Also assume that U is spanned by orthogonal vectors $\{u_1, \ldots, u_m\}$. Then there exists vectors $p \in U$, $n \in V$, where n is orthogonal to U. In particular,

$$p = \sum_{i=1}^{m} \frac{v \cdot u_i}{|u_i|^2} u_i, \; n = v - p$$

We call p the projection of v on U.

This is used in ex 1.18.

1.6 Strategies

1.6.1 To show $\sup A \leq \sup B$

One of the ways to show this is the following:
For each $a \in A$, there exists $b \in B, a \leq b$.
Take supremum of RHS to get $a \leq \sup B$.
Use definition of supremum to get $\sup A \leq \sup B$.

1.6.2 To show $\inf A \leq \inf B$

This mirrors the previous section. But we'll show it, for clarity.
For each $b \in B$, there exists $a \in A, a \leq b$.

Take infimum of LHS to get inf $A \leq b$.
Use definition of infimum to get inf $A \leq$ inf B.

2 Point Set Topology

2.1 Countability

Proposition If there is an injection from A to countable set B, then A is at most countable.

An alternative way of saying this is that if there is a 1-1 correspondence between A and some subset of B, then A is at most countable.

2.2 Definitions

- Metric Space M

- Open ball $B(p, r) = \{q \in M : d(p, q) < r\}$

- Boundary of S: $bd(S) = \{p \in S : \forall r > 0, \exists q \in S, \exists q' \in S^c, q, q' \in B(p, r)\}$

- Limit points of S: $lim(S) = \{p \in S : \forall r > 0, \exists q \in S, q \neq p, q \in B(p, r)\}$

- Interior of S: $int(S) = \{p \in S : \exists r > 0, B(p, r) \subseteq S\}$

2.3 Equivalence of definitions of closed set

The following definitions of a closed set are equivalent.

1. $bd(S) \subseteq S$

2. $lim(S) \subseteq S$

Proof: Suppose $bd(S) \subseteq S$. Let $x \in lim(S)$ and fix some arbitrary $r \in \mathbb{R}^+$. If $x \in S$, we are done. Otherwise, $x \notin S$. As x is a limit point, $\exists q \in S, q \neq x$ such that $q \in B(x, r)$. Now, notice that x is a boundary point of S, since x itself is not in S, and q is in S, and both $x, q \in B(x, r)$.
But by our initial assumption, we have $x \in bd(S) \subseteq S$. Contradiction.
Hence, $x \in S$ and $\lim(S) \subseteq S$.

Conversely, suppose $lim(S) \subseteq S$.
Let $x \in bd(S)$ and fix some arbitrary $r \in \mathbb{R}^+$. If $x \in S$, we are done.
Otherwise, $x \notin S$. As x is a boundary point, $\exists q \in S, q' \in S^c$ such that both are in $B(x, r)$. In particular, $q \neq x$ since $x \notin S$. This says that x is a limit point of S.
Our initial supposition says that $x \in lim(S) \subseteq S$. Again, we have a contradiction.

2.4 Sequential Compactness

The following 2 notions are equivalent.

1. Every infinite subset E of a set X has a limit point in X.

2. X is sequentially compact, that is, for every sequence $\{p_n\}$ in X, there exists a subsequence converging to some point of X.

To prove (2) using (1), consider 2 cases: a sequence with finite range, and a sequence with infinite range (this case is proven with (1)).

3 Sequences

3.1 Infinite subsets vs sequences

There is a subtle difference between these 2 concepts. Of course, infinite sets are unordered, whereas sequences are ordered.
Additionally, it must be noted that sequences can have repeat elements (with different indices). The range of an infinite sequence may very well be finite. An example would be the sequence $\{i^n\}_{n \in \mathbb{N}}$, with range $\{\pm 1, \pm i\}$.

Also, with regard to infinite sets, we speak of limit points, and with regard to sequences, we speak of subsequence limits. Again, there is a difference between limit points and subsequential limits (also known as accumulation points in some texts).

For example, a point can be a subsequential limit without being a limit point. Consider the earlier defined set $\{i^n\}$. Every element in the range $\{\pm 1, \pm i\}$ is a subsequential limit, but since the range is finite, none of these subsequential limits are limit points of the range.

As a consequence, these 2 propositions say different things:

- The set of limit points E' of a set E is closed.

- The set of subsequential limits of a sequence $\{p_n\}$ is closed.

Similarly, supremum of a finite subset of $\mathbb{R}$ is clearly not a limit point of the set.

3.1.1 Supremum of closed subset of $\mathbb{R}$ belongs to set

Like before, we need to consider 2 cases, whether the set if infinite or not.

If the set is finite, the supremum of a set is the maximum of the set, and is clearly an element of the set.
If the set is infinite, the supremum is then a limit point of the set, and by definition of "closed", belongs to the set.

3.2 Subsequential limits in $\mathbb{R}$ and $\pm\infty$

This discussion is inspired by Rudin 3.15 to 3.17.

Notice that subsequential limits of a real sequence lie in $(-\infty, +\infty)$, i.e. they are finite. $\pm\infty$ are **not** considered limits.

This is because the definition of a limit p of a sequence $\{p_n\}$ (or a subsequence) involved the notion of distance $d(p_n, p)$ approaching 0. In particular, $d(p_n, p)$ is defined.
Clearly, $d(p_n, +\infty)$ or $d(p_n, -\infty)$ are not defined, since metric d is always a real-valued function.

3.3 The Extended Real Line

While this was introduced in Chapter 1, it is in Chapter 3 that we see the application of $\mathbb{R} \cup \{\pm\infty\} = [-\infty, +\infty]$.

Properties of the extended reals

- The extended real line is no longer a field.

- $\mathbb{R}$ has the completeness property, such that for any upper bounded subset S of $\mathbb{R}$, $\sup S$ exists in $\mathbb{R}$. However, $\sup S$ does not exist in $\mathbb{R}$ if S is unbounded.
  In contrast, all subsets S of $[-\infty, +\infty]$ have supremums and infimums. For an upper unbounded

set S, we would have $\sup S = +\infty$, and for a lower unbounded set S, we would have $\inf S = -\infty$. For a bounded set, we can then apply the completeness property of $\mathbb{R}$ to find a supremum/infimum in $\mathbb{R} \subset [-\infty, +\infty]$

The fact that all subsets of $[-\infty, +\infty]$ have well-defined supremum and infimum makes Theorem 3.17 in PMA very easy to state. Given the set E of all subsequential limits of a sequence $\{p_n\}$ **which additionally includes $\pm\infty$ if the sequence is unbounded**, we are guaranteed to find $\sup E, \inf E$, without having to talk about special cases (i.e. bounded vs unbounded sequences), which can be troublesome to write.

Of course, if $\sup E$ turns out to be $+\infty$ or $-\infty$, we must be careful. We can't treat the infinities like subsequential limits. For e.g., we can't use a theorem like *the set of subsequential limits is closed*, since $\pm\infty$ do not belong in the set of subsequential limits.
To emphasise: E in theorem 3.17 is the union of subsequential limits (which are finite) and possibly $+\infty$ if the sequence is unbounded above, $-\infty$ if the sequence is unbounded below.

3.4 Two proofs of the Squeeze Theorem

We are given functions f, g, h where $f(x) \le g(x) \le h(x)$ in some neighborhood of x_0, such that $f(x) \to L$ and $h(x) \to L$. Fix $\epsilon > 0$ and we can obtain the following bound for all $0 < |x - x_0| < \delta$ for some $\delta > 0$.

$$L - \epsilon < f(x) \le g(x) \le h(x) < L + \epsilon$$

This says that $g(x) \to L$ as well.

This can easily be adapted to the proof of squeeze theorem for sequences as well.

A second proof of the squeeze theorem that is more in line with Rudin Chapter 3:

Given sequences $\{a_n\}, \{b_n\}, \{c_n\}$ such that $a_n \le b_n \le c_n$ for all but finitely many n, and $a_n, c_n \to L$.

Then we have

$$L = \lim a_n = \liminf a_n \le \liminf b_n \le \limsup b_n \le \limsup c_n = \lim c_n = L$$

which implies

$$\liminf b_n = \limsup b_n = L \implies \lim b_n = L$$

The last implies is due to the following: For any $\epsilon > 0$, there exists $N \in \mathbb{N}$ s.t. $n \ge N \implies L - \epsilon = \liminf b_n - \epsilon < b_n < \limsup b_n + \epsilon = L + \epsilon$.

Note: These 2 proofs really do the same thing, but are expressed differently.

3.5 Some common limits

Theorem 3.20 is what I find to be one of the hardest in Chapter 3 of PMA.
The core ideas are as follows.

The independent variable here is n.

- It is easier to show the convergence of monomials (n^p) than exponentials (x^n) because we can take roots($\frac{1}{p}$).

- For integer powers (such as in the case of x^n), we can use binomial theorem to expand out an exponential expression, e.g. $(1 + p)^n$ and extract out a monomial from the binomial expression, so as to prove properties of convergence/divergence.

3.6   Comparison Test

PMA Theorem 3.25

It is important to note that part (a) of this theorem holds not just for real sequences, but for complex (and $\mathbb{R}^k$) sequences as well.

See Ex 3.15 for more extensions to $\mathbb{R}^k$.

3.7   Series of Nonnegative terms

This will cover PMA 3.26 to 3.29.

3.7.1   Values of monomial n^p

Assume $n > 1$. We will prove explicitly that $a, b \in \mathbb{R}, a < 0 < b$ implies $n^a < 1 < n^b$. While this may be trivial for integers, this actually needs to be proven for rationals, then for reals. We will only prove the right inequality, since the left inequality is analogous.
Note that $n^0 = 1$ by definition.

Suppose $0 < b = \frac{p}{q} \in \mathbb{Q}$. Then $n^p \geq n = n^1 > 1$ since p is a positive integer. Hence $n^b = (n^p)^{\frac{1}{q}} > 1$.
(Easily proven by contradiction).
The result for positive real b then follows by considering supremums.

Given the above result, together with the theorem that $\forall x, y \in \mathbb{R}, n^{x+y} = n^x n^y$, we can then prove a more general result where $a < b \implies n^a < n^b$.
As a start, $a < b \implies \exists c > 0, b = a + c$ and $n^c > 1$.

Remark: With this understanding, we can now rest assured when we claim that $2^{(1-p)} < 1 \iff (1 - p) < 0$ in the proof of theorem 3.28.

3.7.2   Theorem 3.29

A supplement to the proof of theorem 3.29. (*I am not sure why Rudin did not write this detail in. Either it is too trivial, or I am missing something here.*)

First of all, it is not obvious whether $\frac{1}{n(\log n)^p}$ is monotonically decreasing for $p < 0$. Because one of the conditions for applying theorem 3.27 is for the terms of the sequence to be monotonically decreasing. Hence, I would just temporarily assume that $p \geq 0$, and work from there.

So assuming $p \geq 0$, following Rudin's proof, we will arrive at the conclusion that the series converges if $p > 1$ and diverges is $0 \leq p \leq 1$. Then, we make the observation that $p < 0$ implies that the terms of the series $\frac{1}{n(\log n)^p} > \frac{1}{n(\log n)^0} = \frac{1}{n}$, and by the comparison test, the series also diverges for $p < 0$.

3.8   Power Series

3.8.1   Theorem 3.39

The best way to remember this result is by the equality

$$\limsup_{n\to\infty} \sqrt[n]{|c_n z^n|} = |z| \limsup_{n\to\infty} \sqrt[n]{|c_n|}$$

Notice that this result holds even if $\limsup_{n\to\infty} \sqrt[n]{|c_n|} = +\infty$, i.e. this result holds for all values of $\limsup_{n\to\infty} \sqrt[n]{|c_n|} \in [0, +\infty]$. It simply says that the subsequential limits of LHS and RHS are unbounded.

Hence we let $\alpha = \limsup_{n\to\infty} \sqrt[n]{|c_n|}$, then $\limsup_{n\to\infty} \sqrt[n]{|c_n z^n|} < 1 \iff |z|\alpha < 1$.
For finite α, divide both sides by α to get $|z| < \frac{1}{\alpha}$.
For $\alpha = 0$, clearly $\forall z \in \mathbb{C}, |z|\alpha = 0 < 1$, hence $R = \infty$, where R is radius of convergence.
For $\alpha = \infty$, then $\forall z \in \mathbb{C}, |z| > 0 \implies |z|\alpha = \infty > 1$, hence $R = 0$.

By defining $\frac{1}{0} = +\infty$ and $\frac{1}{+\infty} = 0$, we then have $R = \frac{1}{\alpha}$ for all values of $\alpha \in [0, +\infty]$.

3.9 Summation by parts

This covers theorem 3.41.

The following is what I recall from the book Concrete Mathematics Chapter 2. Some terminology may be different from in the book.

We introduce the following notions of finite calculus.
For a discrete function $f : \mathbb{Z} \to \mathbb{R}$, define the difference operator $(\Delta f)(x) = f(x+1) - f(x)$.
Define the summation operator (the discrete analog of the antiderivative) as $\sum \Delta f(x)\delta x = f(x) + C$.
i.e. the summation operator the reverse operation to the difference operator. Call this the antidifference. Here, C is a constant, just like antiderivatives, where particular antiderivatives belonging to the same equivalence class/family of antiderivatives differ by a constant.
Define the definite summation (the discrete analog of the definite integral) as $\sum_p^q \Delta f(x)\delta x = f(q) - f(p)$.

We make the observation that $\sum_p^q f(x)\delta x = \sum_{p \le n < q} f(n)$. Note that LHS is the definite summation operator, whereas the RHS is a regular sum.

3.9.1 Discrete summation by parts

With these definitions, we can derive some results. Recall that in calculus, for differentiable functions f, g, the product rule states $(D(fg))(x) = (Df)(x)g(x) + f(x)(Dg)(x)$ for all x in domain of definition, where D is the derivative operator.

Similarly, there is a product rule for finite calculus.

$$
\begin{aligned}
(\Delta(fg))(x) &= f(x+1)g(x+1) - f(x)g(x) \\
&= f(x+1)g(x+1) - f(x+1)g(x) + f(x+1)g(x) - f(x)g(x) \\
&= f(x+1)(g(x+1) - g(x)) + (f(x+1) - f(x))g(x) \\
&= f(x+1)(\Delta g)(x) + (\Delta f)(x)g(x)
\end{aligned}
$$

Hence, applying the summation operator to both sides,

$$(fg)(x) + C = \sum(f(x+1)(\Delta g)(x) + (\Delta f)(x)g(x))\delta x \,, C \in \mathbb{R}$$

Rearranging then recovers the rule for discrete summation by parts.

$$\sum(\Delta f)(x)g(x))\delta x = (fg)(x) - \sum f(x+1)(\Delta g)(x)\delta x$$

Again, similar to antiderivatives, the constant C is absorbed by the summation expression.

Notice that here, we split a single antidifference into 2. Why does this work? In regular calculus, this is because f, g are assumed to be continuously differentiable, such that Df, Dg are continuous, and thus $Df \cdot g$, $f \cdot Dg$ are continuous, and hence antiderivatives $\int (Df)(x)g(x)\,dx, \int f(x)(Dg)(x)\,dx$ both exist.
In finite calculus, the definite summation of a function always exists, since regular summation is always

well-defined. And consequently, we can define $F(p) = \sum_0^p f(x)\delta x = \sum_{0 \le x < p} f(x)$. Call this F the indefinite summation of f (*see remark below*). Then $(\Delta F)(x) = F(x+1) - F(x) = f(x)$. This tells us that the indefinite summation is a particular antidifference. Hence, the antidifference always exists. Notice the contrast between the indefinite integral and the indefinite summation. The indefinite integral is only guaranteed to be a particular antiderivative for continuous functions, whereas indefinite summation is always a particular antidifference for any discrete function defined on $\mathbb{Z}$.

Applying limits of summation, we obtain,

$$\sum_p^q (\Delta f)(x)g(x))\delta x = (fg)(x)|_p^q - \sum_p^q f(x+1)(\Delta g)(x)\delta x$$

Writing this in terms of a regular sum, we have,

$$\sum_{p \le x < q} (\Delta f)(x)g(x)) = (fg)(x)|_p^q - \sum_{p \le x < q} f(x+1)(\Delta g)(x)$$

This lends to a second proof of theorem 3.41. Define a_n, b_n, A_n as in the book. Also let $A(n) = A_n$. Then $(\Delta A)(n-1) = A_n - A_{n-1} = a_n$.
Theorem 3.41 can then be proven by using $p, q+1$ as limits of definite summation.

Remark: Notice that I am quite hesitant to speak of indefinite integrals. This is because, inspired by Prof Chin Chee Whye in MA2104 Multivariable Calculus, indefinite integrals are defined as $F(p) = \int_0^p f(x)dx$ and should not be viewed as the antiderivative for continuous functions f. At the very most, indefinite integrals are **particular** antiderivatives for continuous functions.
Hence, when speaking about the reverse process of differentiation, it is better to use terminology like antiderivatives. Similarly, the reverse of difference should be referred to as antidifference. Indefinite summation refers to a specific function, $\sum_0^p f(x)\delta x$.

3.10 Limits of truncated sums

Proposition: For all $n \in \mathbb{N}$, $\sum_{k=1}^n a_k + \sum_{k=n+1}^\infty a_k = \sum_{k=1}^\infty a_k$

Proof: Fix $n \in \mathbb{N}$. For all $N \in \mathbb{N}, N > n$, we have $\sum_{k=1}^n a_k + \sum_{k=n+1}^N a_k = \sum_{k=1}^N a_k$.
Rearranging, we get $\sum_{k=n+1}^N a_k = \sum_{k=1}^N a_k - \sum_{k=1}^n a_k$. As $N \to \infty$, RHS converges to $\sum_{k=1}^\infty a_k - \sum_{k=1}^n a_k$. Hence the LHS converges, and $\sum_{k=n+1}^\infty a_k = \sum_{k=1}^\infty a_k - \sum_{k=1}^n a_k$.

This leads to the following result.

Suppose $\sum_{k=1}^\infty a_k = L$. Then fix $\epsilon > 0$ and choose $N \in \mathbb{N}$ such that $n \ge N \implies |\sum_{k=1}^n a_k - L| < \epsilon$.

By the previous proposition, $|\sum_{k=1}^n a_k - L| = |\sum_{k=n+1}^\infty a_k| < \epsilon$.

Alternatively, we can prove this using Cauchy property of convergent sequences, finding $N \in \mathbb{N}$ such that $m > n \ge N \implies |\sum_{k=n}^m a_k| < \frac{\epsilon}{2}$. Letting $m \to \infty$, we get $|\sum_{k=n}^\infty a_k| \le \frac{\epsilon}{2} < \epsilon$.

3.11 Rearrangements

This will cover PMA 3.52 to 3.55.

3.11.1 Theorem 3.55

See the first answer of this math.stackexchange post for how we get the ϵ upper bound for $|s_n - s'_n|$.

The consequence of this bound is that $s_n - \epsilon < s'_n < s_n + \epsilon$, and by taking the limsup and liminf, we get

$$L - \epsilon \leq \liminf_{n \to \infty} s'_n \leq \limsup_{n \to \infty} s'_n \leq L + \epsilon$$

and hence $L = \lim_{n \to \infty} s'_n$. (Here, L is defined to be the limit of s_n)

However, note that it is also possible to prove directly that $\lim_{n \to \infty} s'_n = L$.

Alternative Proof

Before we begin, define the permutation function $\sigma : \mathbb{N} \to \mathbb{N}$ such that $\forall N \in \mathbb{N}, a_{\sigma(n)} = a'_n$.

Fix $\epsilon > 0$.
Choose $N_0 \in \mathbb{N}$ such that $n \geq N_0 \implies |s_n - L| < \frac{\epsilon}{2}$.
Choose $N_1 \in \mathbb{N}$ such that $\{1, 2, \ldots, N_0\} \subset \{\sigma(1), \ldots, \sigma(N_1)\}$
Choose $N_2 \in \mathbb{N}$ such that $n \geq N_2 \implies \sum_{k=n}^{\infty} |a_k| < \frac{\epsilon}{2}$. This is possible since $\{s_n\}$ converges absolutely.

Let $N = \max(N_1, N_2)$. Then $n \geq N \implies$

$$
\begin{aligned}
|s_n - L| = |\sum_{k=1}^{n} a'_k - L| & \\
\leq |\sum_{k=1}^{N_1} a_k - L| + |\sum_{1 \leq k \leq n; \sigma^{-1}(k) > N_1} a'_k| & \qquad \text{by triangle inequality} \\
\leq |\sum_{k=1}^{N_1} a_k - L| + \sum_{1 \leq k \leq n; \sigma^{-1}(k) > N_1} |a'_k| & \qquad \text{by triangle inequality} \\
\leq |\sum_{k=1}^{N_1} a_k - L| + \sum_{k > N_1} |a_k| & \\
< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon &
\end{aligned}
$$

And we are done.

Compared to the book's proof, this involves the choice of N_0, which isn't needed in the book.

4 Functions and Continuity

Check out this [math.stackexchange post](#) on some of the subtleties of defining limits of a function and continuity.

4.1 Subtleties in definitions