Lecture Notes for Analysis II
MA131

Xue-Mei Li *with revisions by* David Mond

January 8, 2013

Contents

The module

0.1 Introduction

Analysis I and Analysis II together make up a 24 CATS core module for first year students.
Assessment 7.5% Term 1 assignments, 7.5% Term 2 assignments, 25% January exam (on Analysis 1) and 60% June exam (on Analysis 1 and 2).

Analysis II Assignments: Given out on Thursday and Fridays in lectures, to hand in the following Thursday in supervisor's pigeonhole by 14:00. Each assignment is divided into part A, B and C. Part B is for handing in.

Take the exercises very seriously. They help you to understand the theory much better than you do by simply memorising it.

Useful books to consult:
E. Hairer & G. Wanner: *Analysis by its History*, Springer-Verlag, 1996. M.H. Protter and C.B. Morrey Jr., *A first course in real analysis*, 2nd edition, Springer-Verlag, 1991
M. Spivack, *Calculus*, 3rd edition, Cambridge University Press, 1994

Feedback Ask questions in lectures! Talk to the lecturer before or after the lectures. He or she will welcome your interest.

Plan your study! Shortly after each lecture, spend an hour going over lecture notes and working on the assignments. Few students understand everything during the lecture. There is plentiful evidence that putting in this effort shortly after the lecture pays ample dividends in terms of understanding. And if you begin a lecture having understood the previous one, you learn much more from it, so the process is cumulative.

Please allow at least two hours per week for the exercises.

Topics by Lecture (approximate guide)

1. Introduction. Continuity, $\epsilon - \delta$ formulation
2. Properties of continuous functions
3. Algebra of continuity
4. Composition of continuous functions, examples
5. The Intermediate Value Theorem (IVT)
6. The Intermediate Value Theorem (continued)
7. Continuous Limits, $\epsilon - \delta$ formulation, relation with to sequential limits and continuity
8. One sided limits, left and right continuity
9. The Extreme Value Theorem
10. The Inverse Function Theorem (continuous version)
11. Differentiation
12. Derivatives of sums, products and composites
13. The Inverse Function Theorem (Differentiable version)
14. Local Extrema, critical points, Rolle's Theorem
15. The Mean Value Theorem
16. Inequalities and behaviour of $f(x)$ as $x \to \pm\infty$
17. Higher order derivatives, C^k functions, convexity and graphing
18. Formal power series, radius of convergence
19. Limit superior
20. Hadamard's Test for the radius of convergence, functions defined by power series
21. Term by Term Differentiation
22. Classical Functions of Analysis

23. Polynomial approximation, Taylor series, Taylor's formula
24. Taylor's Theorem
25. Taylor's Theorem
26. Techniques for evaluating limits
27. L'Hôpital's rule
28. L'Hôpital's rule
29. Question Time

Chapter 1

Continuity of Functions of One Real Variable

Let $\mathbf{R}$ be the set of real numbers. We will often use the letter E to denote a subset of $\mathbf{R}$. Here are some examples of the kind of subsets we will be considering: $E = \mathbf{R}$, $E = (a, b)$ (open interval), $E = [a, b]$ (closed interval), $E = (a, b]$ (semi-closed interval), $E = [a, b)$, $E = (a, \infty)$, $E = (-\infty, b)$, and $E = (1, 2) \cup (2, 3)$. The set of rational numbers $\mathbf{Q}$ is also a subset of $\mathbf{R}$.

1.1 Functions

Definition 1.1.1 *By a* **function** $f : E \to \mathbf{R}$ *we mean a rule which to every number in* E *assigns a number from* $\mathbf{R}$*. This correspondence is denoted by*

$$y = f(x), \qquad \textit{or} \qquad x \mapsto f(x).$$

- *We say that* y *is the* **image** *of* x *and* x *is a* **pre-image** *of* y*.*
- *The set* E *is the* **domain** *of* f*.*
- *The* **range***, or image, of* f *consists of the images of all points of* E*. It is often denoted* $f(E)$*.*

We denote by $\mathbf{N}$ the set of natural numbers, $\mathbf{Z}$ the set of integers and $\mathbf{Q}$ the set of rational numbers:

$$\begin{aligned} \mathbf{N} &= \{1, 2, \dots\} \\ \mathbf{Z} &= \{0, \pm 1, \pm 2, \dots\} \\ \mathbf{Q} &= \{\frac{p}{q} : p, q \in \mathbf{Z}, q \neq 0\}. \end{aligned}$$

Example 1.1.2 *1.* $E = [-1, 3]$.

$$f(x) = \begin{cases} 2x, & -1 \le x \le 1 \\ 3 - x, & 1 < x \le 3. \end{cases}$$

The range of f *is* $[-2, 2]$.

2. $E = \mathbf{R}$.

$$f(x) = \begin{cases} 1, & \textit{if } x \in \mathbf{Q} \\ 0, & \textit{if } x \notin \mathbf{Q} \end{cases}$$

Range $f = \{0, 1\}$.

3. $E = \mathbf{R}$.

$$f(x) = \begin{cases} 1/q, & \textit{if } x = \frac{p}{q} \textit{ where } p \in \mathbf{Z}, q \in \mathbf{N} \textit{ have no common divisor} \\ 0, & \textit{if } x \notin \mathbf{Q} \end{cases}$$

Range $f = \{0\} \cup \{\frac{1}{q}, q = \pm 1, \pm 2, \dots\}$.

4. $E = (-\infty, 0) \cup (0, \infty)$.

$$f(x) = \begin{cases} \frac{\sin x}{x} & \textit{if } x \ne 0 \\ 1 & \textit{if } x = 0 \end{cases}.$$

Range $f = (c, 1]$, *where* $c = -\cos x_0$, x_0 *the smallest positive solution of* $x = \tan x$. *Can you see why?*

5. $E = \mathbf{R}$.

$$f(x) = \begin{cases} \frac{\sin x}{x}, & x \ne 0 \\ 2, & \textit{if } x = 0 \end{cases}$$

Range $f = (c, 1) \cup \{2\}$

6. $E = (-1, 1)$.

$$f(x) = \sum_{n=1}^{\infty} (-1)^n \frac{x^n}{n}.$$

This function is a representation of $-\log(1+x)$, *see chapter on Taylor series. Range* $f = (-\log 2, \infty)$.

1.2 Useful Inequalities and Identities

$$\begin{aligned} a^2+b^2 &\geq 2ab \\ |a+b| &\leq |a|+|b| \\ |b-a| &\geq \max(|b|-|a|, |a|-|b|). \end{aligned}$$

Proof The first follows from $(a-b)^2 > 0$. The second inequality is proved by squaring $|a+b|$ and noting that $ab \leq |a||b|$. The third inequality follows from the fact that

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

and by symmetry $|a| \leq |b| + |b-a|$.

□

Pythagorean theorem : $\sin^2 x + \cos^2 x = 1$.

$$\cos^2 x - \sin^2 x = \cos(2x), \qquad \cos(x) = 1 - 2\sin^2(\frac{x}{2}).$$

1.3 Continuous Functions

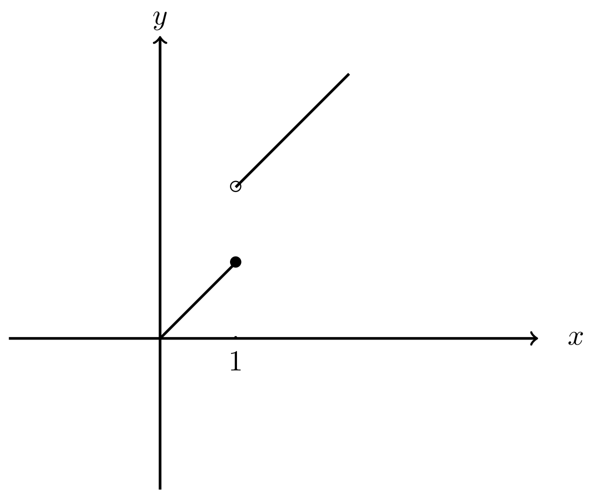
What do we mean by saying that "$f(x)$ varies continuously with x"?

It is reasonable to say f is continuous if the graph of f is an "unbroken continuous curve". The concept of an unbroken continuous curve seems easy to understand. However we may need to pay attention.

For example we look at the graph of

$$f(x) = \begin{cases} x, & x \leq 1 \\ x+1, & \text{if } x \geq 1 \end{cases}$$

It is easy to see that the curve is continuous everywhere except at $x = 1$. The function is not continuous at $x = 1$ since there is a gap of 1 between the values of $f(1)$ and $f(x)$ for x close to 1. It is continuous everywhere else.



Now take the function

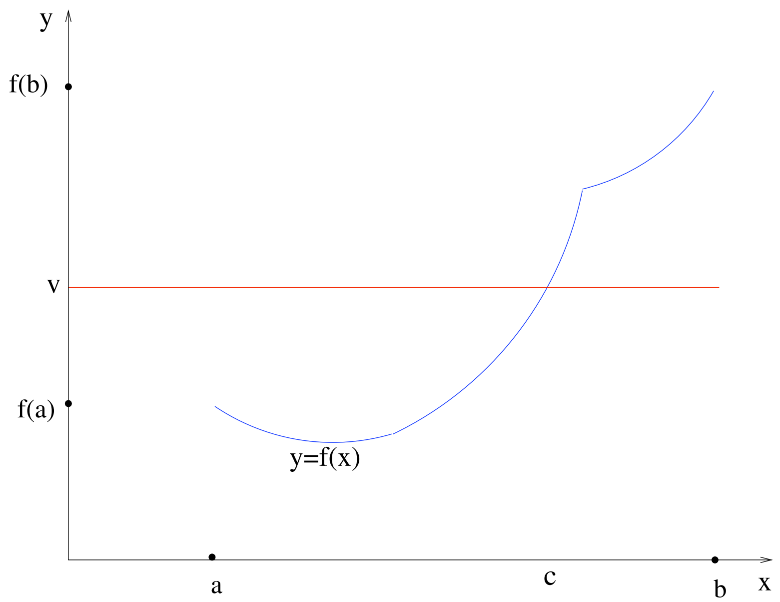
$$F(x) = \begin{cases} x, & x \leq 1 \\ x + 10^{-30}, & \text{if } x \geq 1 \end{cases}$$

The function is not continuous at $x = 1$ since there is a gap of 10^{-30}. However can we see this gap on a graph with our naked eyes? No, unless you have exceptional eyesight!

Here is a theorem we will prove, once we have the definition of "continuous function".

Theorem 1.3.1 *(Intermediate value theorem): Let* $f : [a, b] \to \mathbf{R}$ *be a continuous function. Suppose that* $f(a) \neq f(b)$*, and that the real number* v *lies between* $f(a)$ *and* $f(b)$*. Then there is a point* $c \in [a, b]$ *such that* $f(c) = v$*.*

This looks "obvious", no? In the picture shown here, it says that if the graph of the continuous function $y = f(x)$ starts, at $(a, f(a))$, *below* the straight line $y = v$ and ends, at $(b, f(b))$, *above* it, then at some point between these two points it must *cross* this line.



But how can we *prove* this? Notice that its truth uses some subtle facts about the real numbers. If, instead of the domain of f being an interval in $\mathbf{R}$, it is an interval in $\mathbf{Q}$, then the statement is no longer true. For example, we would probably agree that the function $F(x) = x^2$ is "continuous" (soon we will see that it is). If we now consider the function $f : \mathbf{Q} \to \mathbf{Q}$ defined by the same formula, then the rational number $v = 2$ lies between $0 = f(0)$ and $9 = f(3)$, but even so there is no c between 0 and 3 (or anywhere else, for that matter) such that $f(c) = 2$.

Sometimes what seems obvious becomes a little less obvious if you widen your perspective.

These examples call for a proper definition for the notion of continuous function.

In Analysis, the letter ε is often used to denote a distance, and generally we want to find way to make some quantity "smaller than ε":

- The sequence $(x_n)_{n\in\mathbf{N}}$ converges to ℓ if for every $\varepsilon > 0$, there exists $N \in \mathbf{N}$ such that if $n > N$, then $|x_n - \ell| < \varepsilon$.

- The sequence $(x_n)_{n\in\mathbf{N}}$ is Cauchy if for every $\varepsilon > 0$, there exists $N \in \mathbf{N}$ such that if $m, n > N$ then $|x_m - x_n| < \varepsilon$.

In each case, by taking n, or n and m, sufficiently big, we make the distance between x_n and ℓ, or the distance between x_n and x_m, smaller than ε.

Important: The distance between a and b is $|a-b|$. The set of x such that $|x-a|<\delta$ is the same as the set of x with $a-\delta<x<a-\delta$.

The definition of continuity is similar in spirit to the definition of convergence. It too begins with an ε, but instead of meeting the challenge by finiding a suitable $N \in \mathbf{N}$, we have to find a positive real number δ, as follows:

Definition 1.3.2 *Let E be subset of $\mathbf{R}$ and c a point of E.*

1. *A function $f: E \to \mathbf{R}$ is* **continuous at** *c if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that*

$$\text{if } |x-c|<\delta \text{ and } x \in E$$

$$\text{then} \quad |f(x)-f(c)|<\varepsilon.$$

2. *If f is continuous at every point c of E, we say f* **is continuous on** *E or simply that f is continuous.*

The reason that we require $x \in E$ is that $f(x)$ is only defined if $x \in E$! If f is a function with domain $\mathbf{R}$, we generally drop E in the formulation above.

Example 1.3.3 The function $f: \mathbf{R} \to \mathbf{R}$ given by $f(x)=3x$ is continuous at every point x_0. It is very easy to see this. Let $\varepsilon > 0$. If $\delta = \varepsilon/3$ then

$$|x-x_0|<\delta \implies |x-x_0|<\varepsilon/3 \implies |3x-3x_0|<\varepsilon \implies |f(x)-f(x_0)|<\varepsilon$$

as required.

Example 1.3.4 The function $f: \mathbf{R} \to \mathbf{R}$ given by $f(x)=x^2+x$ is continuous at $x=2$. Here is a proof: for any $\varepsilon>0$ take $\delta=\min(\frac{\varepsilon}{6},1)$. Then if $|x-2|<\delta$,

$$|x+2| \le |x-2|+4 \le \delta+4 \le 1+4=5.$$

And

$$\begin{aligned} |f(x)-f(2)| &= |x^2+x-(4+2)| \le | x^2-4|+|x-2| \\ &= |x+2||x-2|+|x-2| \le 5|x-2|+|x-2| \\ &= 6|x-2| < 6\delta \le \varepsilon \end{aligned}$$

I do not like this proof! One of the apparently difficult parts about the notion of continuity is figuring out how to choose δ to ensure that *if* $|x - c| < \delta$ *then* $|f(x) - f(c)| < \varepsilon$. The proof simply produces δ without explanation. It goes on to show very clearly that this δ works, but we are left wondering how on earth we would choose δ if asked a similar question.

There are two things which make this easier than it looks at first sight:

1. If δ works, then so does any $\delta' > 0$ which is smaller than δ. There is no single right answer. Sometimes you can make a guess which may not be the biggest possible δ which works, but still works. In the last example, we could have taken

$$\delta = \min\{\frac{\varepsilon}{10}, 1\}$$

   instead of $\min\{\frac{\varepsilon}{6}, 1\}$, just to be on the safe side. It would still have been right.

2. Fortunately, we don't need to find δ very often. In fact, it turns out that sometimes it is easier to use general theorems than to prove continuity directly from the definition. Sometimes it's easier to *prove* general theorems than to prove continuity directly from the definition in an example. My preferred proof of the continuity of the function $f(x) = x^2 + x$ goes like this:

   (a) First, prove a general theorem showing that if f and g are functions which are continuous at c, then so are the functions $f + g$ and fg, defined by

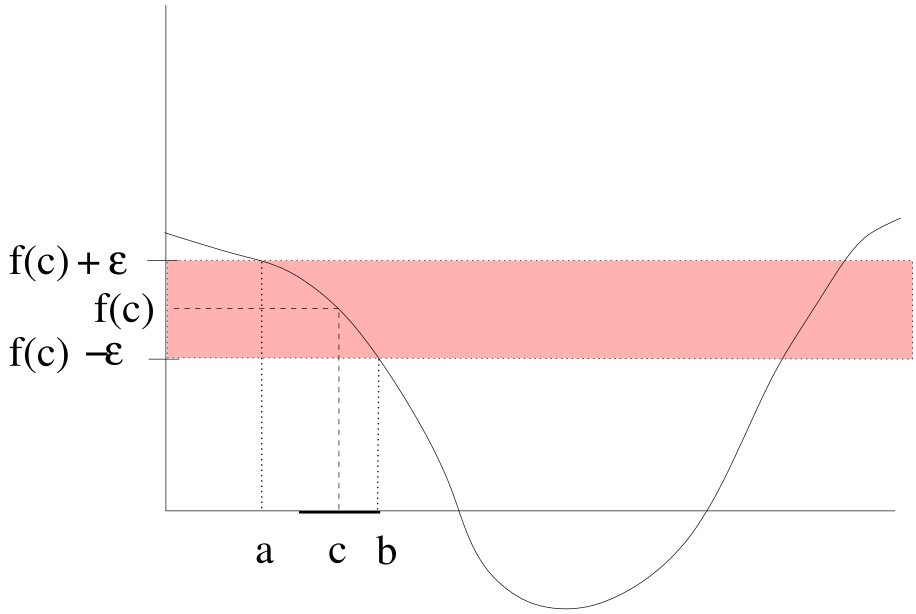
$$\begin{cases} (f+g)(x) = f(x) + g(x) \\ (fg)(x) = f(x) \times g(x) \end{cases}$$

   (b) Second, show how to assemble the function $h(x) = x^2 + x$ out of simpler functions, whose continuity we can prove effortlessly. In this case, h is the sum of the functions $x \mapsto x^2$ and $x \mapsto x$. The first of these is the product of the functions $x \mapsto x$ and $x \mapsto x$. Continuity of the function $f(x) = x$ is very easy to show!

   As you can imagine, the general theorem in the first step will be used on many occasions, as will the continuity of the simple functions in the second step. So general theorems save work.

   But that will come later. For now, you have to learn how to apply the definition directly in some simple examples.

First, a visual example with no calculation at all:



Here, as you can see, for all x in (a, b), we have $f(c)-\varepsilon < f(x) < f(c)+\varepsilon$. What could be the value of δ? The largest possible interval *centred* on c is shown in the picture; so the biggest possible value of δ here is $b - c$. If we took as δ the larger value $c - a$, there would be points within a distance δ of c, but for which $f(x)$ is *not* within a distance ε of $f(c)$.

Exercise 1.3.5 Suppose that $f(x) = x^2$, $x_0 = 7$ and $\varepsilon = 5$. What is the largest value of δ such that if $|x - x_0| < \delta$ then $|f(x) - f(x_0)| < \varepsilon$? What if $\varepsilon = 50$? What if $\varepsilon = 0.1$? Drawing a picture like the last one will almost certainly help.

We end this section with a small result which will be very useful later. The following lemma says that if f is continuous at c and $f(c) \neq 0$ then f does not vanish close to c. In fact f has the same sign as $f(c)$ in a neighbourhood of c.

We introduce a new notation: if c is a point in $\mathbf{R}$ and r is also a real number, then $B_r(c) = \{x \in \mathbf{R} : |x - c| < r\}$. It is sometimes called "the ball of radius r centred on c". The same definition makes sense in higher dimensions ($B_r(c)$ is the set of points whose distance from c is less than r); in $\mathbf{R}^2$ and $\mathbf{R}^3$, $B_r(c)$ looks more round than it does in $\mathbf{R}$.

Lemma 1.3.6 Non-vanishing lemma *Suppose that $f : E \to \mathbf{R}$ is continuous at c.*

1. *If $f(c) > 0$, then there exists $r > 0$ such that $f(x) > 0$ for $x \in B_r(c)$.*

2. *If $f(c) < 0$, then there exists $r > 0$ such that $f(x) < 0$ for $x \in B_r(c)$.*

In both cases there is a neighbourhood of c on which f does not vanish.

Proof Suppose that $f(c) > 0$. Take $\varepsilon = \frac{f(c)}{2} > 0$. Let $r > 0$ be a number such that if $|x - c| < r$ and $x \in E$ then

$$|f(x) - f(c)| < \varepsilon.$$

For such x, $f(x) > f(c) - \varepsilon = \frac{f(c)}{2} > 0$.

If $f(c) < 0$, take $\varepsilon = \frac{-f(c)}{2} > 0$. There is $r > 0$ such that on $B_r(c)$, $f(x) < f(c) + \varepsilon = \frac{f(c)}{2} < 0$. □

1.4 New continuous functions from old

Proposition 1.4.1 The algebra of continuous functions *Suppose f and g are both defined on a subset E of $\mathbf{R}$ and are both continuous at c, then*

1. *$f + g$ is continuous at c.*

2. *λf is continuous at c for any constant λ.*

3. *fg is continuous at c*

4. *$1/g$ is well-defined in a neighbourhood of c, and is continuous at c if $g(c) \neq 0$.*

5. *f/g is continuous at c if $g(c) \neq 0$.*

Proof

1. Given $\varepsilon > 0$ we have to find $\delta > 0$ such that if $|x - c| < \delta$ then $|(f(x) + g(x)) - (f(c) + g(c))| < \varepsilon$. Now

$$|(f(x) + g(x)) - (f(c) + g(c))| = |f(x) - f(c) + g(x) - g(c)|;$$

since, by the triangle inequality, we have

$$|f(x) - f(c) + g(x) - g(c)| \leq| f(x) - f(c) + g(x) - g(c)|$$

it is enough to show that $|f(x) - f(c)| < \varepsilon/2$ and $|g(x) - g(c)| < \varepsilon/2$. This is easily achieved: $\varepsilon/2$ is a positive number, so by the continuity of f and g at c, there exist $\delta_1 > 0$ and $\delta_2 > 0$ such that

$$\begin{aligned} |x - c| < \delta_1 &\implies |f(x) - f(c)| < \varepsilon/2 \\ |x - c| < \delta_2 &\implies |g(x) - g(c)| < \varepsilon/2. \end{aligned}$$

Now take $\delta = \min\{\delta_1, \delta_2\}$.

2. Suppose $\lambda \neq 0$. Given $\varepsilon > 0$, choose $\delta > 0$ such that if $|x-c| < \delta$ then $|f(x) - f(c)| < \varepsilon/\lambda$. The last inequality implies $|\lambda f(x) - \lambda f(c)| < \varepsilon$. The case where $\lambda = 0$ is easier!

3. We have

$$\begin{aligned} |f(x)g(x) - f(c)g(c)| &= |f(x)g(x) - f(c)g(x) + f(c)g(c) - f(c)g(c)| \\ &\leq |f(x)g(x) - f(c)g(x)| + |f(c)g(x) - f(c)g(c)|. \\ &\leq |g(x)||f(x) - f(c)| + |f(c)||g(x) - g(c)|. \end{aligned}$$

   Provided each of the two terms on the right hand side here is less than $\varepsilon/2$, the left hand side will be less that ε. The second of the two terms on the right hand side is easy to deal with: provided $f(c) \neq 0$ then it is enough to make $|g(x) - g(c)|$ less than $\varepsilon/2f(c)$. So we choose $\delta_1 > 0$ such that if $|x - c| < \delta_1$ then $|g(x) - g(c)| < \varepsilon/2f(c)$. If $f(c) = 0$, on the other hand, the second term on the RHS term is equal to 0, so causes us no problem at all.

   The first of the two terms on the right hand side is a little harder. We argue as follows: taking $\varepsilon = 1$ in the definition of continuity, there exists $\delta_2 > 0$ such that if $|x-c| < \delta_2$ then $|g(x) - g(c)| < 1$, and hence $|g(x)| < |g(c)| + 1$. Now $\varepsilon/2(|g(c) + 1)$ is a positive number; so by the continuity of f at c, we can choose $\delta_3 > 0$ such that if $|x-c| < \delta_3$ then $|f(x) - f(c)| < \varepsilon/2(|g(c)| + 1)$. Finally, take $\delta = \min\{\delta_1, \delta_2, \delta_3\}$. Then if $|x - c| < \delta$, we have

$$|g(x) - g(c)| < \varepsilon/2f(c), \quad |g(x)| < |g(c)| + 1 \quad \text{and } |f(x) - f(c)| < \varepsilon/2|g(x)|$$

   from which it follows that $|f(x)g(x) - f(c)g(c)| < \varepsilon$.

4. This is the most complicated! We have to show that given $\varepsilon > 0$, there exists $\delta > 0$ such that if $|x - c| < \delta$ then

$$\left| \frac{1}{g(x)} - \frac{1}{g(c)} \right| < \varepsilon. \tag{1.4.1}$$

Now

$$\left|\frac{1}{g(x)} - \frac{1}{g(c)}\right| = \left|\frac{g(c) - g(x)}{g(c)g(x)}\right|.$$

We have to make the numerator small while keeping the denominator big.

First step: Choose $\delta_1 > 0$ such that if $|x - c| < \delta_1$ then $|g(x)| > |g(c)|/2$. (Why is this possible)? Note that if $|g(x)| > |g(c)|/2$ then $|g(x)g(c)| > |g(c)|^2/2$.

Second step: Choose $\delta_2 > 0$ such that if $|x-c| < \delta_2$ then $|g(x)-g(c)| < \varepsilon|g(c)|^2/2$.

Third step: Take $\delta = \min\{\delta_1, \delta_2\}$.

5. If f and g are continuous at c with $g(c) \neq 0$ then by (3), $1/g$ is continuous at c, and now by (2), f/g is continuous at c.

□ The continuity of the function $f(x) = x^2 + x$, which we proved from first principles in Example 1.3.4, can be proved more easily using Proposition 1.4.1. For the function $g(x) = x$ is obviously continuous (take $\delta = \varepsilon$), the function $h(x) = x^2$ is equal to $g \times g$ and is therefore continuous by 1.4.1(2), and finally $f = h + g$ and is therefore continuous by 1.4.1(1).

Proposition 1.4.2 composition of continuous functions *Suppose $f : D \to \mathbf{R}$ and $g : E \to \mathbf{R}$ are functions, and that $f(D) \subset E$ (so that the composite $g \circ f$ is defined). If f is continuous at x_0, and g is continuous at $f(x_0)$, then $g \circ f$ is continuous at x_0.*

Proof Denote the domain of G by E. If $x \in D$, write $y_0 = f(x_0)$. By hypothesis g is continuous at y_0. So given $\varepsilon > 0$, there exists $\delta_0 > 0$ such that

$$y \in E \text{ and } |y - y_0| < \delta_0 \quad \Longrightarrow \quad |g(y) - g(y_0)| < \varepsilon \tag{1.4.2}$$

As f is continuous at x_0, there exists $\delta > 0$ such that

$$x \in D \text{ and } |x - x_0| < \delta \quad \Longrightarrow \quad |f(x) - y_0| < \delta_0 \tag{1.4.3}$$

Putting (1.4.2) and (1.4.3) together we see that

$$x \in D \text{ and } |x-x_0| < \delta \quad \Longrightarrow \quad |f(x)-y_0| < \delta_0 \Longrightarrow |g(f(x))-g(y_0)| < \varepsilon \tag{1.4.4}$$

That is,

$$x \in D \text{ and } |x - x_0| < \delta \quad \Longrightarrow \quad |g(f(x)) - g(f(x_0))| < \varepsilon.$$

□

Example 1.4.3 *1. A polynomial of degree n is a function of the form $P_n(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0$. Polynomials are continuous, by repeated application of Proposition 1.4.1(1) and (2).*

2. A rational function is a function of the form: $\frac{P(x)}{Q(x)}$ where P and Q are polynomials. The rational function $\frac{P}{Q}$ is continuous at x_0 provided that $Q(x_0) \neq 0$, by 1.4.1(3).

Example 1.4.4 *1. The exponential function $\exp(x) = e^x$ is continuous. This we will prove later in the course.*

2. Given that $\exp$ is continuous, the function g defined by $g(x) = \exp(x^{2n+1} + x)$ is continuous (use continuity of $\exp$, 1.4.3(1) and 1.4.2).

Example 1.4.5 *The function $x \mapsto \sin x$ is continuous. This will be proved shortly. From the continuity of $\sin x$ we can deduce that $\cos x$, $\tan x$ and $\cot x$ are continuous: the function $x \mapsto \cos x$ is continuous by 1.4.2, since $\cos x = \sin(x + \frac{\pi}{2})$ and is thus the composite of the continuous function $\cos$ and the continuous function $x \mapsto x + \pi/2$; the functions $x \mapsto \tan x$ and $x \mapsto \cot x$ are continuous on all of their domains, by 1.4.1(3), because $\tan x = \frac{\sin x}{\cos x}$ and $\cot x = \frac{\cos x}{\sin x}$.*

Discussion of continuity of $\tan x$. If we restrict the domain of $\tan x$ to the region $(-\frac{\pi}{2}, \frac{\pi}{2})$, its graph is a continuous curve and we believe that $\tan x$ is continuous. On a larger range, $\tan x$ is made of disjoint pieces of continuous curves. How could it be a continuous function ? Surely the graph looks discontinuous at $\frac{\pi}{2}$!!! The trick is that the domain of the function does not contain these points where the graph looks broken. By the definition of continuity we only consider x with values in the domain.

The largest domain for $\tan x$ is

$$\mathbf{R}/\{\frac{\pi}{2} + k\pi, k \in Z\} = \cup_{k \in Z}(k\pi - \frac{\pi}{2}, k\pi + \frac{\pi}{2}).$$

For each c in the domain, we locate the piece of continuous curve where it belongs. We can find a small neighbourhood on which the graph is is part of this single piece of continuous curve.

For example if $c \in (-\frac{\pi}{2}, \frac{\pi}{2})$, make sure that δ is smaller than $\min(\frac{\pi}{2} - c, c + \frac{\pi}{2})$.

1.5 Discontinuity

Let us now consider a problem of logic. We gave a definition of what is meant by 'f is continuous at c':

$$\forall \varepsilon > 0 \; \exists \delta > 0 \text{ such that if } |x - c| < \delta \text{ then } |f(x) - f(c) < \varepsilon.$$

How to express the meaning of 'f is *not* continuous at c' in these same terms? Before answering, we think for a moment about the meaning of negation.

Definition 1.5.1 *The negation of a statement A is a statement which is true whenever statement A is false, and false whenever statement A is true.*

This is very easy to apply when the statement is simple: the negation of

The student passed the exam

is

The student did not pass the exam

or, equivalently,

The student failed the exam

If the statement involves one or more quantifiers, the negation may be less obvious: the negation of

All the students passed the exam (1.5.1)

is *not*

All the students failed the exam (1.5.2)

but, in view of the definition of negation, 1.5.1, is instead

At least one student failed the exam (1.5.3)

If we avoid the correct but lazy option of simply writing "not" at the start of the sentence, and instead look for something with the opposite meaning, then the negation of

At least one student stayed awake in my lecture

is

Every student fell asleep in my lecture

and the negation of

Every student stayed awake in my lecture

is

At least one student fell asleep in my lecture.

Negation turns statements involving "∀" into statements involving "∃", and vice versa.

Exercise 1.5.2 *Give the negations of*

1. *Somebody failed the exam.*
2. *Everybody failed the exam.*
3. *Everybody needs somebody to love.*

Now let us get to the point: what is the negation of the three-quantifier statement

> for all $\varepsilon > 0$, there exists $\delta > 0$ such that for all x satisfying $|x - c| < \delta$, we have $|f(x) - f(c)| < \varepsilon$?

First attempt

> *there exists an* $\varepsilon > 0$ *such that there does* **not** *exist a* $\delta > 0$ *such that for all* x *satisfying* $|x - c| < \delta$*, we have* $|f(x) - f(c)| < \varepsilon$

This is correct, but "there does not exist" is the lazy option. We avoid it with

Second attempt

> *there exists an* $\varepsilon > 0$ *such that for all* $\delta > 0$*, it is* **not** *true that for all* x *satisfying* $|x - c| < \delta$*, we have* $|f(x) - f(c)| < \varepsilon$

But this still uses the lazy option with the word *not*: we improve it again, to

Third attempt

there exists an $\varepsilon > 0$ such that for all $\delta > 0$, there exists an x satisfying $|x - c| < \delta$ but $|f(x) - f(c)| \geq \varepsilon$.

Now finally we have a negation which does not use the word "not"! It is the most informative, and the version we will make use of. Incidentally, the word "but" means the same as "and", except that it warns the reader that he or she will be surprised by the statement which follows. It has the same purely logical contect as "and".

Example 1.5.3 Consider $f : \mathbf{R} \to \mathbf{R}$,

$$f(x) = \begin{cases} 0, & \text{if } x \in \mathbf{Q} \\ 1, & \text{if } x \notin \mathbf{Q}. \end{cases}$$

Let $c \in \mathbf{Q}$. The function f is not continuous at c. The key point is that between any two numbers there is an irrational number.

Take $\varepsilon = 0.5$. For any $\delta > 0$, there is an $x \notin \mathbf{Q}$ with $|x - c| < \delta$; because $x \notin \mathbf{Q}$, $f(x) = 1$. Thus $|f(x) - f(c)| = |1 - 0| > 0.5$. We have shown that *there exists* an $\varepsilon > 0$ (in this case $\varepsilon = 0.5$) such that *for every* $\delta > 0$, *there exists* an x such that $|x - c| < \delta$ but $|f(x) - f(c)| \geq \varepsilon$. So f is not continuous at c.

If $c \notin \mathbf{Q}$ then once again f is not continuous at c. The key point now is that between any two numbers there is a rational number. Take $\varepsilon = 0.5$. For any $\delta > 0$, take $x \in \mathbf{Q}$ with $|x - c| < \delta$; then $|f(x) - f(c)| = |0 - 1| > 0.5 = \varepsilon$.

Example 1.5.4 Consider $f : \mathbf{R} \to \mathbf{R}$,

$$f(x) = \begin{cases} 0, & \text{if } x \in \mathbf{Q} \\ x, & \text{if } x \notin \mathbf{Q} \end{cases}$$

Claim: This function is continuous at $c = 0$ and discontinuous everywhere else.

Let $c = 0$. For any $\varepsilon > 0$ take $\delta = \varepsilon$. If $|x-c| < \delta$, then $|f(x)-f(0)| = |x|$ in case $c \notin \mathbf{Q}$, and $|f(x) - f(0) = |0 - 0| = 0$ in case $c \in \mathbf{Q}$. In both cases, $|f(x) - f(0)| \leq |x| < \delta = \varepsilon$. Hence f is continuous at 0.

Exercise 1.5.5 Show that the function f of the last example is not continuous at $c \neq 0$. Hint: take $\varepsilon = |c|/2$.

Example 1.5.6 Suppose that $g : \mathbf{R} \to \mathbf{R}$ and $h : \mathbf{R} \to \mathbf{R}$ are continuous, and let c be some fixed real number. Define a new function $f : \mathbf{R} \to \mathbf{R}$ by

$$f(x) = \begin{cases} g(x), & x < c \\ h(x), & x \geq c. \end{cases}$$

Then the function f is continuous at c if and only if $g(c) = h(c)$.
Proof: Since g and h are continuous at c, for any $\varepsilon > 0$ there are $\delta_1 > 0$ and $\delta_2 > 0$ such that

$$|g(x) - g(c)| < \varepsilon \text{ when } |x - c| < \delta_1$$

and such that

$$|h(x) - h(c)| < \varepsilon \text{ when } |x - c| < \delta_2.$$

Define $\delta = \min(\delta_1, \delta_2)$. Then if $|x - c| < \delta$,

$$|g(x) - g(c)| < \varepsilon \quad \text{and} \quad |h(x) - h(c)| < \varepsilon.$$

- Case 1: Suppose $g(c) = h(c)$. For any $\varepsilon > 0$ and the above δ, if $|x - c| < \delta$,

  $$|f(x) - f(c)| = \begin{cases} |g(x) - g(c)| < \varepsilon & \text{if } x < c \\ |h(x) - h(c)| < \varepsilon & \text{if } x \geq c \end{cases}$$

  So f is continuous at c.

- Case 2: Suppose $g(c) \neq h(c)$.

  Take

  $$\varepsilon = \frac{1}{2}|g(c) - h(c)|.$$

  By the continuity of g at c, there exists δ_0 such that if $|x - c| < \delta_0$, then $|g(x) - g(c)| < \varepsilon$. If $x < c$, then

  $$|f(c) - f(x)| = |h(c) - g(x)|;$$

  moreover

  $$|h(c) - g(x)| = |h(c) - g(c) + g(c) - g(x)| \geq |\, h(c) - g(c)| - |g(c) - g(x)|.$$

  We have $|h(c) - g(c)| = 2\varepsilon$, and if $|c - x| < \delta_0$ then $|g(c) - g(x)| < \varepsilon$. It follows that if $c - \delta_0 < x < c$,

  $$|f(c) - f(x)| = |h(c) - h(x)| > \varepsilon.$$

  So if $c - \delta_0 < x < c$, then no matter how close x is to c we have

  $$|f(x) - f(c)| > \varepsilon.$$

  This shows f is not continuous at c.

Example 1.5.7 *Let $E = [-1, 3]$. Consider $f : [-1, 3] \to \mathbf{R}$,*

$$f(x) = \begin{cases} 2x, & -1 \le x \le 1 \\ 3 - x, & 1 < x \le 3. \end{cases}$$

The function is continuous everywhere.

Example 1.5.8 *Let $f : \mathbf{R} \to \mathbf{R}$.*

$$f(x) = \begin{cases} 1/q, & \text{if } x = \frac{p}{q}, q > 0 \quad p, q \text{ coprime integers} \\ 0, & \text{if } x \notin \mathbf{Q} \end{cases}$$

Then f is continuous at all irrational points, and discontinuous at all rational points.

Proof *Every rational number p/q can be written with positive denominator q, and in this proof we will always use this choice.*

- *Case 1. Let $c = p/q \in \mathbf{Q}$. We show that f is not continuous at c. Take $\varepsilon = \frac{1}{2q}$. No matter how small is δ there is an irrational number x such that $|x - c| < \delta$. And $|f(x) - f(c)| = |\frac{1}{q}| > \varepsilon$.*

- *Case 2. Let $c \notin \mathbf{Q}$. We show that f is continuous at c. Let $\varepsilon > 0$. If $x = \frac{p}{q} \in \mathbf{Q}$, $|f(x) - f(c)| = \frac{1}{q}$. So the only rational numbers p/q for which $|f(x) - f(p/q)| \ge \varepsilon$ are those with denominator q less than or equal to $1/\varepsilon$.*

  Let $A = \{x \in \mathbf{Q} \cap (c-1, c+1) : q \le \frac{1}{\varepsilon}\}$. Clearly $c \notin A$. The crucial point is that if it is not empty, A contains only finitely many elements. To see this, observe that its members have only finitely many possible denominators q, since q must be a natural number $\le 1/\varepsilon$. For each possible q, we have

  $$p/q \in (c-1, c+1) \iff p \in (qc - q, qc + q),$$

  so no more than $2q$ different values of p are possible.

  It now follows that the set $B := \{|x - c| : x \in A\}$, if not empty, has finitely many members, all strictly positive. Therefore if B is not empty, it has a least element, and this element is strictly positive. Take δ to be this element. If B is empty, take $\delta = \infty$. In either case, $(c - \delta, c + \delta)$ does not contain any number from A.

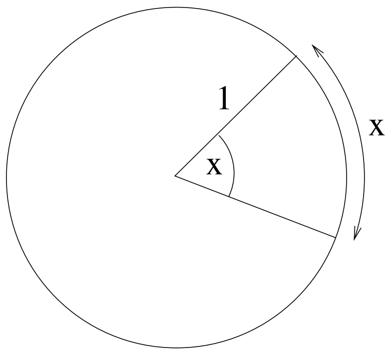
  Suppose $|x - c| < \delta$. If $x \notin \mathbf{Q}$, then $f(x) = f(x) = 0$, so $|f(x) - f(c)| = 0 < \varepsilon$. If $x = p/q \in \mathbf{Q}$ then since $x \notin A$, $|f(x) - f(c)| = |\frac{1}{q}| < \varepsilon$.

□

1.6 Continuity of Trigonometric Functions

Measurement of angles

Two different units are commonly used for measuring angles: Babylonian degrees and radians. We will use radians. The radian measure of an angle x is the length of the arc of a unit circle subtended by the angle x.



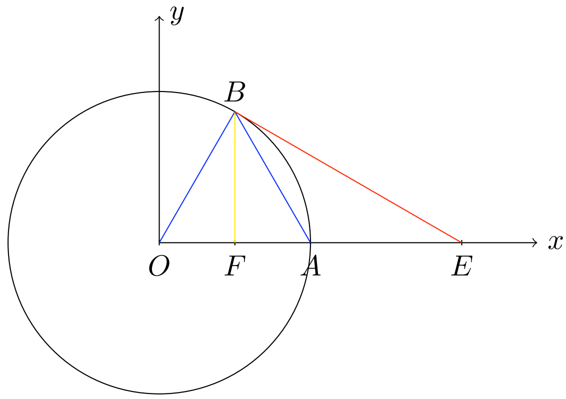
The following explains how to measure the length of an arc. Take a polygon of n segments of equal length inscribed in the unit circle. Let ℓ_n be its length. The length increases with n. As $n \to \infty$, it has a limit. The limit is called the *circumference* of the unit circle. The circumference of the unit circle was measured by Archimedes, Euler, Liu Hui etc.. It can now be shown to be an irrational number.

Historically $\sin x$ and $\cos x$ are defined in the following way. Later we define them by power series. The two definitions agree. Take a right-angled triangle, with the hypotenuse of length 1 and an angle x. Define $\sin x$ to be the length of the side facing the angle and $\cos x$ the length of the side adjacent to it. Extend this definition to $[0, 2\pi]$. For example, if $x \in [\frac{\pi}{2}, \pi]$, define $\sin x = \cos(x - \frac{\pi}{2})$. Extend to the rest of $\mathbf{R}$ by decreeing that $\sin(x + 2\pi) = \sin x, \cos(x + 2\pi) = \cos x$.

Lemma 1.6.1 *If* $0 < x < \frac{\pi}{2}$, *then*

$$\sin x < x < \tan x.$$

Proof Take a unit disk centred at 0, and consider a sector of the disk of angle x which we denote by OBA.



The area of the sector OBA is $x/2\pi$ times the area of the disk, and is therefore $\pi\frac{x}{2\pi} = \frac{x}{2}$. The area of the triangle OBA is $\frac{1}{2}\sin x$. So since the triangle is strictly smaller than the sector, we have $\sin x < x$.

Consider the right-angled triangle OBE, with one side tangent to the circle at B. Because $BE = \tan x$, the area of the triangle is $\frac{1}{2}\tan x$. This triangle is strictly bigger than the sector OBA. So

$$\text{Area(Sector } OBA) \;<\; \text{Area(Triangle } OBE),$$

and therefore $x < \tan x$.

□

Theorem 1.6.2 *The function $x \mapsto \sin x$ is continuous.*

Proof

$$\begin{aligned}
|\sin(x+h) - \sin x| &= |\sin x \cos h + \cos x \sin h - \sin x| \\
&= |\sin x(\cos h - 1) + \cos x \sin h| \\
&= 2|\sin x \sin^2 \frac{h}{2} + \cos x \sin h| \\
&\le 2|\sin x||\sin^2 \frac{h}{2}| + |\cos x||\sin h| \\
&\le \frac{|h^2|}{2} + |h| = |h|(\frac{|h|}{2} + 1).
\end{aligned}$$

If $|h| \le 1$, then $(\frac{|h|}{2} + 1) < \frac{3}{2}$. For any $\varepsilon > 0$, choose $\delta = \min(\frac{2}{3}\varepsilon, 1)$. If $|h| < \delta$ then

$$|\sin(x+h) - \sin x| \le |h|(\frac{|h|}{2} + 1) \le \frac{3}{2}|h| < \varepsilon.$$

□

Chapter 2

Continuous Functions on Closed Intervals

We recall the definition of the least upper bound or the supremum of a set A.

Definition 2.0.3 *A number c is an* **upper bound** *of a set A if for all $x \in A$ we have $x \leq c$. A number c is the* **least upper bound** *of a set A if*

- *c is an upper bound for A.*
- *if c' is an upper bound for A then $c \leq c'$.*

The least upper bound of the set A is denoted by $\sup A$. If $\sup A$ belongs to A, we call it the maximum of A.

Recall the **Completeness of R** If $A \subset \mathbf{R}$ is non empty and bounded above, it has a least upper bound in $\mathbf{R}$.

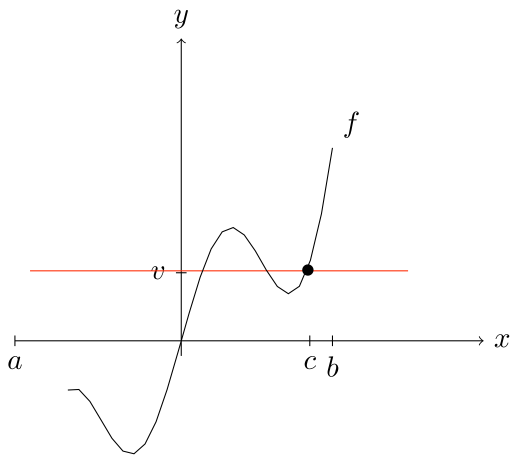
Exercise 2.0.4 *For any nonempty set $A \subset \mathbf{R}$ which is bounded above, there is a sequence $x_n \in A$ such that $c - \frac{1}{n} \leq x_n < c$. This sequence x_n evidently converges to c. (If $\sup A \in A$, then one can take $x_n = \sup A$ for all n.)*

2.1 The Intermediate Value Theorem

Recall Theorem 1.3.1, which we state again:

Theorem 2.1.1 Intermediate Value Theorem (IVT) *Let $f : [a, b] \to \mathbf{R}$ be continuous. Suppose that $f(a) \neq f(b)$. Then for any v strictly between $f(a)$ and $f(b)$, there exists $c \in (a, b)$ such that $f(c) = v$.*

A rigorous proof was first given by Bolzano in 1817.
Proof We give the proof for the case $f(a) < v < f(b)$. Consider the set $A = \{x \in [a, b] : f(x) \le v\}$. Note that $a \in A$ and A is bounded above by b, so it has a least upper bound, which we denote by c. We will show that $f(c) = v$.



- Since $c = \sup A$, there exists $x_n \in A$ with $x_n \to c$. Then $f(x_n) \to f(c)$ by continuity of f. Since $f(x_n) \le v$ then $f(c) \le v$.
- Suppose $f(c) < v$. Then by the Non-vanishing Lemma 1.3.6 applied to the function $x \mapsto v - f(x)$, there exists $r > 0$ such that for all $x \in B_r(c)$, $f(x) < v$. But then $c + r/2 \in A$. This contradicts the fact that c is an upper bound for A.

□

The idea of the proof is to identify the greatest number in (a, b) such that $f(c) = v$. The existence of the least upper bound (i.e. the completeness axiom) is crucial here, as it is on practically every occasion on which one wants to prove that there exists a point with a certain property, without knowing at the start where this point is.

Exercise 2.1.2 *To test your understanding of this proof, write out the (very similar) proof for the case* $f(a) > v > f(b)$.

Example 2.1.3 *Let* $f : [0, 1] \to \mathbf{R}$ *be given by* $f(x) = x^7 + 6x + 1$. *Does there exist* $x \in [0, 1]$ *such that* $f(x) = 2$*?*

Proof *The answer is yes. Since* $f(0) = 1, f(1) = 8$ *and* $2 \in (1, 8)$*, there is a number* $c \in (0, 1)$ *such that* $f(c) = 2$ *by the IVT.* □

Remark*: There is an alternative proof for the IVT, using the "bisection method". It is not covered in lectures.

Suppose that $f(a) < 0$, $f(b) > 0$. We construct nested intervals $[a_n, b_n]$ with length decreasing to zero. We then show that a_n and b_n have common limit c which satisfy $f(c) = 0$.

Divide the interval $[a, b]$ into two equal halves: $[a, c_1], [c_1, b]$. If $f(c_1) = 0$, done. Otherwise on (at least) one of the two sub-intervals, the value of f must change from negative to positive. Call this subinterval $[a_1, b_1]$. More precisely, if $f(c_1) > 0$ write $a = a_1$, $c_1 = b_1$; if $f(c_1) < 0$, let $a_1 = c_1$, $b_1 = b$. Iterate this process. Either at the k'th stage $f(c_k) = 0$, or we obtain a sequence of intervals $[a_k, b_k]$ with $[a_{k+1}, b_{k+1}] \subset [a_k, b_k]$, a_k increasing, b_k decreasing, and $f(a_k) < 0$, $f(b_k) > 0$ for all k. Because the sequences a_k and b_k are monotone and bounded, both converge, say to c' and c'' respectively. We must have $c' \le c''$, and $f(c') \le 0 \le f(c'')$. If we can show that $c' = c''$ then since $f(c') \le 0$, $f(c'') \ge 0$ then we must have $f(c') = 0$, and we have won. It therefore remains only to show that $c' = c''$. I leave this as an (easy!) exercise.

Example 2.1.4 *The polynomial* $p(x) = 3x^5 + 5x + 7 = 0$ *has a real root.*

Proof Let $f(x) = 3x^5 + 5x + 7$. Consider f as a function on $[-1, 0]$. Then f is continuous and $f(-1) = -3 - 5 + 7 = -1 < 0$ and $f(0) = 7 > 0$. By the IVT there exists $c \in (-1, 0)$ such that $f(c) = 0$. □

Discussion of assumptions. In the following examples the statement fails. For each one, which condition required in the IVT is not satisfied?

Example 2.1.5

1. *Let* $f : [-1, 1] \to \mathbf{R}$*,*

$$f(x) = \begin{cases} x + 1, & x > 0 \\ x, & x < 0 \end{cases}$$

   Then $f(-1) = -1 < f(1) = 2$*. Can we solve* $f(x) = 1/2 \in (-1, 2)$*? No: that the function is not continuous on* $[-1, 1]$*.*

2. *Define* $f : \mathbf{Q} \cap [0, 2] \to \mathbf{R}$ *by* $f(x) = x^2$*. Then* $f(0) = 0$ *and* $f(2) = 4$*. Is it true that for each* v *with* $0 < v < 4$*, there exists* $c \in \mathbf{Q} \cap [0, 2]$ *such that* $f(c) = v$*? No! If, for example,* $v = 2$*, then there is no rational number* c *such that* $f(c) = v$*.*

   Here, the domain of f *is* $\mathbf{Q} \cap [0, 2]$*, not an interval in the reals, as required by the IVT.*

Example 2.1.6 *Let $f, g : [a, b] \to \mathbf{R}$ be continuous functions. Suppose that $f(a) < g(a)$ and $f(b) > g(b)$. Then there exists $c \in (a, b)$ such that $f(c) = g(c)$.*

Proof Define $h = f - g$. Then $h(a) < 0$ and $h(b) > 0$ and h is continuous. Apply the IVT to h: there exists $c \in (a, b)$ with $h(c) = 0$, which means $f(c) = g(c)$. □

Let $f : E \to \mathbf{R}$ be a function. Any point $x \in E$ such that $f(x) = x$ is called a *fixed point* of f.

Theorem 2.1.7 (Fixed Point Theorem) *Suppose $g : [a, b] \to [a, b]$ is a continuous function. Then there exists $c \in [a, b]$ such that $g(c) = c$.*

Proof The notation that $g : [a, b] \to [a, b]$ implies that the range of f is contained in $[a, b]$.

Set $f(x) = g(x) - x$. Then $f(a) = g(a) - a \geq a - a = 0$, $f(b) = g(b) - b \leq b - b = 0$.

- If $f(a) = 0$ then a is the sought after point.
- If $f(b) = 0$ then b is the sought after point.
- If $f(a) > 0$ and $f(b) < 0$, apply the intermediate value theorem to f to see that there is a point $c \in (a, b)$ such that $f(c) = 0$. This means $g(c) = c$.

□

Remark This theorem has a remarkable generalisation to higher dimensions, known as Brouwer's Fixed Point Theorem. Its statement requires the notion of continuity for functions whose domain and range are of higher dimension - in this case a function from a product of intervals $[a, b]^n$ to itself. Note that $[a, b]^2$ is a square, and $[a, b]^3$ is a cube. I invite you to adapt the definition of continuity we have given, to this higher-dimensional case.

Theorem 2.1.8 *(Brouwer, 1912) Let $f : [a, b]^n \to [a, b]^n$ be a continuous function. Then f has a fixed point.*

The proof of Brouwer's theorem when $n > 1$ is harder than when $n = 1$. It uses the techniques of Algebraic Topology.

Chapter 3

Continuous Limits

3.1 Continuous Limits

We wish to give a precise meaning to the statement "f approaches ℓ as x approaches c".

To make our definition, we require that f should be defined on some set $(a, b) \setminus \{c\}$, where $a < c < b$.

Definition 3.1.1 *Let $c \in (a, b)$ and let $f : (a, b) \setminus \{c\} \to \mathbf{R}$. Let ℓ be a real number. We say that f* **tends to** ℓ **as** x **approaches** c*, and write*

$$\lim_{x \to c} f(x) = \ell,$$

if for any $\varepsilon > 0$ there exists $\delta > 0$, such that for all $x \in (a, b)$ satisfying

$$0 < |x - c| < \delta \tag{3.1.1}$$

we have

$$|f(x) - \ell| < \varepsilon.$$

Remark 3.1.2

1. *The condition $|x - c| > 0$ in (3.1.1) means we do not care what happens when x equals c. For that matter, f does not even need to be defined at c. That is why the the definition of limit makes no reference to the value of f at c.*
2. *In the definition, we may assume that the domain of f is a subset E of $\mathbf{R}$ containing $(c - r, c) \cup (c, c + r)$ for some $r > 0$.*

Remark 3.1.3 *If a function has a limit at c, this limit is unique.*

Proof Suppose f has two limits ℓ_1 and ℓ_2. Take $\varepsilon = \frac{1}{4}|\ell_1 - \ell_2|$. If $\ell_1 \neq \ell_2$ then $\varepsilon > 0$, so by definition of limit there exists $\delta > 0$ such that if $0 < |x - c| < \delta$, $|f(x) - \ell_1| < \varepsilon$ and $|f(x) - \ell_2| < \varepsilon$. By the triangle inequality,

$$|\ell_1 - \ell_2| \leq |f(x) - \ell_1| + |f(x) - \ell_2| < 2\varepsilon = \frac{1}{2}|l_1 - l_2|.$$

This cannot happen. We must have $\ell_1 = \ell_2$. □

Theorem 3.1.4 *Let $c \in (a, b)$. Let $f : (a, b) \to \mathbf{R}$. The following are equivalent:*

1. *f is continuous at c.*
2. $\lim_{x \to c} f(x) = f(c)$.

Proof

- Assume that f is continuous at c.

  $\forall \varepsilon > 0$ there is a $\delta > 0$ such that for all x with

  $$|x - c| < \delta \quad \text{and} \quad x \in (a, b), \tag{3.1.2}$$

  we have

  $$|f(x) - f(c)| < \varepsilon.$$

  Condition (3.1.2) holds if condition (3.1.1) holds, so $\lim_{x \to c} f(x) = f(c)$.

- On the other hand suppose that $\lim_{x \to c} f(x) = f(c)$. For any $\varepsilon > 0$ there exists $\delta > 0$, such that for all $x \in (a, b)$ satisfying

  $$0 < |x - c| < \delta$$

  we have

  $$|f(x) - f(c)| < \varepsilon.$$

  If $|x - c| = 0$ then $x = c$. In this case $|f(c) - f(c)| = 0 < \varepsilon$. Hence for all $x \in (a, b)$ with $|x - c| < \delta$, we have $|f(x) - f(c)| < \varepsilon$. Thus f is continuous at c.

□

Example 3.1.5 *Does*

$$\lim_{x\to 1} \frac{x^2+3x+2}{\sin(\pi x)+2}$$

exist?
Yes: let

$$f(x) = \frac{x^2+3x+2}{\sin(\pi x)+2}.$$

Then f is continuous on all of $\mathbf{R}$, for both the numerator and the denominator are continuous on all of $\mathbf{R}$, and the numerator is never zero. Hence $\lim_{x\to 1} f(x) = f(1) = \frac{6}{2} = 3$.

Example 3.1.6 *Does*

$$\lim_{x\to 0} \frac{\sqrt{1+x}-\sqrt{1-x}}{x}$$

exist? Find its value if it does exist.

For $x \neq 0$ and x close to zero (e.g. for $|x| < 1/2$), the function

$$f(x) = \frac{\sqrt{1+x}-\sqrt{1-x}}{x}$$

is well defined. When $x \neq 0$,

$$f(x) = \frac{\sqrt{1+x}-\sqrt{1-x}}{x} = \frac{2x}{x(\sqrt{1+x}+\sqrt{1-x})} = \frac{2}{\sqrt{1+x}+\sqrt{1-x}}.$$

Since $\sqrt{x}$ is continuous at $x = 1$, (prove it by $\varepsilon - \delta$ argument!), the functions $x \mapsto \sqrt{1+x}$ and $x \mapsto \sqrt{1-x}$ are both continuous at $x = 0$; as their sum is non-zero when $x = 0$, it follows, by the algebra of continuous functions, that the function

$$f : [-\frac{1}{2}, \frac{1}{2}] \to \mathbf{R}, \qquad f(x) = \frac{2}{\sqrt{1+x}+\sqrt{1-x}}$$

is also continuous at $x = 0$. Hence

$$\lim_{x\to 0} \frac{\sqrt{1+x}-\sqrt{1-x}}{x} = \lim_{x\to 0} \frac{2}{\sqrt{1+x}+\sqrt{1-x}} = f(0) = 1.$$

Example 3.1.7 *The statement that $\lim_{x\to 0} f(x) = 0$ is equivalent to the statement that $\lim_{x\to 0} |f(x)| = 0$.*

Define $g(x) = |f(x)|$. Then $|g(x) - 0| = |f(x)|$. The statement $|g(x) - 0| < \varepsilon$ is the same as the statement $|f(x) - 0| < \varepsilon$.

Remark 3.1.8 (use of negation) *The statement "it is not true that* $\lim_{x \to c} f(x) = \ell$*" means precisely that there exists a number* $\varepsilon > 0$ *such that for all* $\delta > 0$ *there exists* $x \in (a, b)$ *with* $0 < |x - c| < \delta$ *and* $|f(x) - \ell| \ge \varepsilon$*. This can occur in two ways:*

1. *the limit does not exist, or*
2. *the limit exists, but differs from* ℓ*.*

In the second case, but not in the first case, we write

$$\lim_{x \to c} f(x) \neq \ell.$$

In the first case it would be wrong to write this since it suggests that the limit exists.

Theorem 3.1.9 *Let* $c \in (a, b)$ *and* $f : (a, b) \smallsetminus \{c\} \to \mathbf{R}$*. The following are equivalent.*

1. $\lim_{x \to c} f(x) = \ell$
2. *For every sequence* $x_n \in (a, b) \smallsetminus \{c\}$ *with* $\lim_{n \to \infty} x_n = c$ *we have*

$$\lim_{n \to \infty} f(x_n) = \ell.$$

Proof

- Step 1. If $\lim_{x \to c} f(x) = \ell$, define

$$g(x) = \begin{cases} f(x), & x \neq c \\ \ell, & x = c \end{cases}$$

  Then $\lim_{x \to c} g(x) = \lim_{x \to c} f(x) = \ell = g(c)$. This means that g is continuous at c. If x_n is a sequence with $x_n \in (a, b) \smallsetminus \{c\}$ and $\lim_{n \to \infty} x_n = c$, then

$$\lim_{n \to \infty} f(x_n) = \lim_{n \to \infty} g(x_n) = \ell$$

  by sequential continuity of g at c. We have shown that statement 1 implies statement 2.

- Step 2. Suppose that $\lim_{x\to c} f(x) = \ell$ does not hold.

  There is a number $\varepsilon > 0$ such that for all $\delta > 0$, there is a number $x \in (a,b) \setminus \{c\}$ with $0 < |x-c| < \delta$, but $|f(x)-\ell| \geq \varepsilon$.

  Taking $\delta = \frac{1}{n}$ we obtain a sequence $x_n \in (a,b)/\{c\}$ with $|x_n - c| < \frac{1}{n}$ with
  $$|f(x_n) - \ell| \geq \varepsilon$$
  for all n. Thus the statement $\lim_{n\to\infty} f(x_n) = \ell$ cannot hold. But x_n is a sequence with $\lim_{n\to\infty} x_n = c$. Hence statement 2 fails.

$\square$

Example 3.1.10 *Let* $f : \mathbf{R} \setminus \{0\} \to \mathbf{R}$, $f(x) = \sin(\frac{1}{x})$. *Then*

$$\lim_{x\to 0} \sin(\frac{1}{x})$$

does not exist.

Proof Take two sequences of points,

$$\begin{aligned} x_n &= \frac{1}{\frac{\pi}{2} + 2n\pi} \\ z_n &= \frac{1}{-\frac{\pi}{2} + 2n\pi}. \end{aligned}$$

Both sequences tend to 0 as $n \to \infty$. But $f(x_n) = 1$ and $f(z_n) = -1$ so $\lim_{n\to\infty} f(x_n) \neq \lim_{n\to\infty} f(z_n)$. By the sequential formulation for limits, Theorem 3.1.9, $\lim_{x\to 0} f(x)$ cannot exist. $\square$

Example 3.1.11 *Let* $f : \mathbf{R} \setminus \{0\} \to \mathbf{R}$, $f(x) = \frac{1}{|x|}$. *Then there is no number* $\ell \in \mathbf{R}$ *s.t.* $\lim_{x\to 0} f(x) = \ell$.

Proof Suppose that there exists ℓ such that $\lim_{x\to 0} f(x) = \ell$. Take $x_n = \frac{1}{n}$. Then $f(x_n) = n$ and $\lim_{n\to\infty} f(x_n)$ does not converges to any finite number! This contradicts that $\lim_{n\to\infty} f(x_n) = \ell$. $\square$

Example 3.1.12 *Denote by* $[x]$ *the integer part of* x. *Let* $f(x) = [x]$. *Then* $\lim_{x\to 1} f(x)$ *does not exist.*

Proof We only need to consider the function f near 1. Let us consider f on $(0, 2)$. Then

$$f(x) = \begin{cases} 1, & \text{if } 1 \le x < 2 \\ 0, & \text{if } 0 \le x < 1 \end{cases}.$$

Let us take a sequence $x_n = 1 + \frac{1}{n}$ converging to 1 from the right and a sequence $y_n = 1 - \frac{1}{n}$ converging to 1 from the left . Then

$$\lim_{n\to\infty} x_n = 1, \qquad \lim_{n\to\infty} y_n = 1.$$

Since $f(x_n) = 1$ and $f(y_n) = 0$ for all n, the two sequences $f(x_n)$ and $f(y_n)$ have different limits and hence $\lim_{x\to 1} f(x)$ does not exist. $\square$

The following follows from the sandwich theorem for sequential limits.

Proposition 3.1.13 (Sandwich Theorem/Squeezing Theorem) *Let* $c \in (a, b)$ *and* $f, g, h : (a, b) \smallsetminus \{c\} \to \mathbf{R}$. *If* $h(x) \le f(x) \le g(x)$ *on* $(a, b) \smallsetminus \{c\}$, *and*

$$\lim_{x\to c} h(x) = \lim_{x\to c} g(x) = \ell$$

then

$$\lim_{x\to c} f(x) = \ell.$$

Proof Let $x_n \in (a, b)\smallsetminus\{c\}$ be a sequence converging to c. Then by Theorem 3.1.9, $\lim_{n\to\infty} g(x_n) = \lim_{n\to\infty} h(x_n) = \ell$. Since $h(x_n) \le f(x_n) \le g(x_n)$,

$$\lim_{n\to\infty} f(x_n) = \ell.$$

By Theorem 3.1.9 again, we see that $\lim_{x\to c} f(x) = \ell$. $\square$

Example 3.1.14

$$\lim_{x\to 0} x \sin(\frac{1}{x}) = 0.$$

Proof Note that $0 \le |x \sin(\frac{1}{x})| \le |x|$. Now $\lim_{x\to 0} |x| = 0$ by the the continuity of the function $f(x) = |x|$. By the Sandwich theorem the conclusion $\lim_{x\to 0} x \sin(\frac{1}{x}) = 0$ follows. $\square$

Example 3.1.15

$$f(x) = \begin{cases} x \sin(1/x) & x \ne 0 \\ 0, & x = 0 \end{cases}$$

is continuous at 0.

This follows from Example 3.1.14, $\lim_{x\to 0} f(x) = 0 = f(0)$.

The following example will be revisited in Example 11.2.2 (as an application of L'Hôpital's rule).

Example 3.1.16 Important limit to remember

$$\lim_{x\to 0}\frac{\sin x}{x}=1.$$

Proof *Recall if* $0<x<\frac{\pi}{2}$*, then* $\sin x\le x\le \tan x$ *and*

$$\frac{\sin x}{\tan x}\le\frac{\sin x}{x}\le 1.$$

$$\cos x\le\frac{\sin x}{x}\le 1 \qquad (3.1.3)$$

The relation (3.1.3) also holds if $-\frac{\pi}{2}<x<0$*: Letting* $y=-x$ *then* $0<y<\frac{\pi}{2}$*. All three terms in (3.1.3) are even functions:* $\frac{\sin x}{x}=\frac{\sin y}{y}$ *and* $\cos y=\cos(-x)$*.*

Since

$$\lim_{x\to 0}\cos x=1$$

by the Sandwich Theorem and (3.1.3), $\lim_{x\to 0}\frac{\sin x}{x}=1$*.*

□

From the algebra of continuity and the continuity of composites of continuous functions we deduce the following for continuous limits, with the help of Theorem 3.1.9.

Proposition 3.1.17 (Algebra of limits) *Let* $c\in(a,b)$ *and* $f,g(a,b)/\{c\}$*. Suppose that*

$$\lim_{x\to c}f(x)=\ell_1,\qquad \lim_{x\to c}g(x)=\ell_2.$$

Then

1. $$\lim_{x\to c}(f+g)(x)=\lim_{x\to c}f(x)+\lim_{x\to c}g(x)=\ell_1+\ell_2.$$

2. $$\lim_{x\to c}(fg)(x)=\lim_{x\to xc}f(x)\lim_{x\to c}g(x)=\ell_1\ell_2.$$

3. *If* $\ell_2\neq 0$,
$$\lim_{x\to c}\frac{f(x)}{g(x)}=\frac{\lim_{x\to c}f(x)}{\lim_{x\to c}g(x)}=\frac{\ell_1}{\ell_2}.$$

Proposition 3.1.18 Limit of the composition of two functions *Let $c \in (a,b)$ and $f : (a,b) \smallsetminus \{c\} \to \mathbf{R}$. Suppose that*

$$\lim_{x\to c} f(x) = \ell. \tag{3.1.4}$$

Suppose that the range of f is contained in $(a_1,b_1) \smallsetminus \{\ell\}$, that $g : (a_1,b_1) \smallsetminus \{\ell\} \to \mathbf{R}$, and that

$$\lim_{y\to \ell} g(y) = L \tag{3.1.5}$$

Then

$$\lim_{x\to c} g(f(x)) = L. \tag{3.1.6}$$

Proof Given $\varepsilon > 0$, by (3.1.5) we can choose $\delta_1 > 0$ such that

$$0 < |y-\ell| < \delta_1 \implies \quad |g(y) - L| < \varepsilon. \tag{3.1.7}$$

By (3.1.4) we can choose $\delta_2 > 0$ such that

$$0 < |x-c| < \delta_2 \quad \implies \quad |f(x) - \ell| < \delta_1. \tag{3.1.8}$$

Since $f(x) \neq \ell$ for all x in the domain of f, it follows from (3.1.8) that if $0 < |x-c| < \delta_2$ then $0 < |f(x) - \ell| < \delta_1$; it therefore follows, by (3.1.7), that $|g(f(x)) - L| < \varepsilon$. □

Exercise 3.1.19 *Does the conclusion of Proposition 3.1.18 hold if the hypotheses on the range of f and the domain of g are replaced by the assumptions that the range of f is contained in (a_1,b_1) and that g is defined on all of (a_1,b_1)?*

The following example will be re-visited in Example 11.2.6 (an application of L'Hôpital's rule).

Example 3.1.20 (Important limit to remember)

$$\lim_{x\to 0} \frac{1-\cos x}{x} = 0.$$

Proof

$$\begin{aligned}
\lim_{x\to 0} \frac{1-\cos x}{x} &= \lim_{x\to 0} \frac{2\sin^2(\frac{x}{2})}{x} \\
&= \lim_{x\to 0} \sin(\frac{x}{2}) \frac{\sin(\frac{x}{2})}{\frac{x}{2}} \\
&= \lim_{x\to 0} \sin(\frac{x}{2}) \times \lim_{x\to 0} \frac{\sin(\frac{x}{2})}{\frac{x}{2}} \\
&= 0 \times 1 = 0.
\end{aligned}$$

The following lemma should be compared to Lemma 1.3.6. The existence of the $\lim_{x\to c} f(x)$ and its value depend only on the behaviour of f near to c. It is in this sense we say that 'limit' is a local property. Recall that $B_r(c) = (c-r, c+r)$; let $\mathring{B}_r(c) = (c-r, c+r) \setminus \{c\}$. If a property holds on $B_r(c)$ for some $r > 0$, we often say that the property holds "close to c".

Lemma 3.1.21 *Let $g : (a, b) \setminus \{c\} \to \mathbf{R}$, and $c \in (a, b)$. Suppose that*

$$\lim_{x\to c} g(x) = \ell.$$

1. *If $\ell > 0$, then $g(x) > 0$ on $\mathring{B}_r(c)$ for some $r > 0$.*
2. *If $\ell < 0$, then $g(x) < 0$ on $\mathring{B}_r(c)$ for some $r > 0$.*

In both cases $g(x) \neq 0$ for $x \in \mathring{B}_r(c)$.

Proof Define

$$\tilde{g}(x) = \begin{cases} g(x), & x \neq c \\ \ell, & x = c. \end{cases}$$

Then $\tilde{g}$ is continuous at c and so by Lemma 1.3.6, $\tilde{g}$ is not zero on $(c-r, c+r)$ for some r. The conclusion for g follows. □

Proof [Direct Proof] Suppose that $l > 0$. Let $\varepsilon = \frac{\ell}{2} > 0$. There exists $r > 0$ such that if $x \in (\mathring{B}_r(c)$ and $x \in (a, b)$, then.

$$g(x) > \ell - \varepsilon = \frac{\ell}{2} > 0.$$

We choose $r < \min(b-c, c-a)$ so that $(c-r, c+r) \subset (a, b)$.

If $\ell < 0$ let $\varepsilon = -\frac{\ell}{2} > 0$, then there exists $r > 0$, with $r < \min(b-c, c-a)$, such that on $\mathring{B}_r(c)$,

$$g(x) < \ell + \varepsilon = \ell - \frac{\ell}{2} = \frac{\ell}{2} < 0.$$

□

3.2 One Sided Limits

How do we define "the limit of f as x approaches c from the right " or "the limit of f as x approaches c from the left "?

Definition 3.2.1 *A function $f : (a, c) \to \mathbf{R}$ has a left limit ℓ at c if for any $\varepsilon > 0$ there exists $\delta > 0$, such that*

$$\text{for all } x \in (a, c) \quad \text{ with } \; c - \delta < x < c \tag{3.2.1}$$

we have

$$|f(x) - \ell| < \varepsilon.$$

In this case we write

$$\lim_{x \to c-} f(x) = \ell.$$

Right limits are defined in a similar way:

Definition 3.2.2 *A function $f : (c, b) \to \mathbf{R}$ has a right limit ℓ at c and we write $\lim_{x \to c+} f(x) = \ell$ if for any $\varepsilon > 0$ there exists $\delta > 0$, such that for all $x \in (c, b) \quad$ with $\; c < x < c + \delta$ we have $|f(x) - \ell| < \varepsilon$.*

Remark 3.2.3

- *(3.2.1) is equivalent to*

$$x \in (a, c), \qquad \& \quad 0 < |x - c| < \delta.$$

- *$\lim_{x \to c} f(x) = \ell$ if and only if both $\lim_{x \to c+} f(x)$ and $\lim_{x \to c-} f(x)$ exist and are equal to ℓ.*

Definition 3.2.4

1. *A function $f : (a, c] \to \mathbf{R}$ is said to be left continuous at c if for any $\varepsilon > 0$ there exists a $\delta > 0$ such that if $c - \delta < x < c$ then*

$$|f(x) - f(c)| < \varepsilon.$$

2. *A function $f : [c, b) \to \mathbf{R}$ is said to be right continuous at c if for any $\varepsilon > 0$ there exists a $\delta > 0$ such that if $c < x < c + \delta$*

$$|f(x) - f(c)| < \varepsilon.$$

Theorem 3.2.5 *Let $f : (a, b) \to \mathbf{R}$ and for $c \in (a, b)$. The following are equivalent:*

(a) f is continuous at c

(b) f is right continuous at c and f is left continuous at c.

(c) Both $\lim_{x \to c+} f(x)$ and $\lim_{x \to c-} f(x)$ exist and are equal to $f(c)$.

(d) $\lim_{x \to c} f(x) = f(c)$

Example 3.2.6 *Denote by $[x]$ the integer part of x. Let $f(x) = [x]$. Show that for $k \in \mathbf{Z}$, $\lim_{x\to k+} f(x)$ and $\lim_{x\to k-} f(x)$ exist. Show that $\lim_{x\to k} f(x)$ does not exist. Show that f is discontinuous at all points $k \in \mathbf{Z}$.*

Proof Let $c = k$, we consider f near k, say on the open interval $(k-1, k+1)$. Note that

$$f(x) = \begin{cases} k, & \text{if } k \le x < k+1 \\ k-1, & \text{if } k-1 \le x < k \end{cases}.$$

It follows that

$$\begin{aligned} \lim_{x\to k+} f(x) &= \lim_{x\to 1+} k = k \\ \lim_{x\to k-} f(x) &= \lim_{x\to 1-} (k-1) = k-1. \end{aligned}$$

Since the left limit does not agree with the right limit, $\lim_{x\to k} f(x)$ does not exist. By the limit formulation of continuity, ('A function continuous at k must have a limit as x approaches k'), f is not continuous at k. □

Example 3.2.7 *For which number a is f defined below continuous at $x = 0$?*

$$f(x) = \begin{cases} a, & x \neq 0 \\ 2, & x = 0 \end{cases}.$$

Since $\lim_{x\to 0} f(x) = 2$, f is continuous if and only if $a = 2$.

3.3 Limits to ∞

What do we mean by

$$\lim_{x\to c} f(x) = \infty?$$

As ∞ is not a real number, the definition of limit given up to now does not apply.

Definition 3.3.1 *Let $c \in (a, b)$ and $f : (a, b)/\{c\} \to \mathbf{R}$. We say that*

$$\lim_{x\to c} f(x) = \infty,$$

if for all $M > 0$, there exists $\delta > 0$, such that for all $x \in (a, b)$ with $0 < |x - c| < \delta$ we have $f(x) > M$.

Example 3.3.2 $\lim_{x\to 0} \frac{1}{|x|} = \infty$. *For let $M > 0$. Define $\delta = \frac{1}{M}$. Then if $0 < |x - 0| < \delta$, we have $f(x) = \frac{1}{|x|} > \frac{1}{\delta} = M$.*

Definition 3.3.3 *Let $c \in (a, b)$ and $f : (a, b)/\{c\} \to \mathbf{R}$. We say that*

$$\lim_{x \to c} f(x) = -\infty,$$

if for all M, there exists $\delta > 0$, such that for all $x \in (a, b)$ with $0 < |x - c| < \delta$ we have $f(x) < M$.

One sided limits can be similarly defined. For example,

Definition 3.3.4 *Let $f : (c, b) \to \mathbf{R}$. We say that*

$$\lim_{x \to c+} f(x) = \infty,$$

if $\forall M > 0$, $\exists \delta > 0$, s.t. $f(x) > M$ for all $x \in (c, b) \cap (c, c + \delta)$.

Definition 3.3.5 *Let $f : (a, c) \to \mathbf{R}$. We say that*

$$\lim_{x \to c-} f(x) = -\infty,$$

if $\forall M > 0$, $\exists \delta > 0$, such that $f(x) < -M$ for all $x \in (a, c)$ with $c - \delta < x < c$.

Example 3.3.6 *Show that $\lim_{x \to 0-} \frac{1}{\sin x} = -\infty$.*

Proof Let $M > 0$. Since $\lim_{x \to 0} \sin x = 0$, there is a $\delta > 0$ such that if $0 < |x| < \delta$, then $|\sin x| < \frac{1}{M}$.

Since we are interested in the left limit, we may consider $x \in (-\frac{\pi}{2}, 0)$. In this case $|\sin x| = -\sin x$. So we have $-\sin x < \frac{1}{M}$ which is the same as

$$\sin x > -\frac{1}{M}.$$

In conclusion $\lim_{x \to 0-} \frac{1}{\sin x} = -\infty$. $\square$

3.4 Limits at ∞

What do we mean by

$$\lim_{x \to \infty} f(x) = \ell, \quad \text{or} \quad \lim_{x \to \infty} f(x) = \infty?$$

Definition 3.4.1 *1. Consider $f : (a, \infty) \to \mathbf{R}$. We say that*

$$\lim_{x \to \infty} f(x) = \ell,$$

if for any $\varepsilon > 0$ there is an M such that if $x > M$ we have $|f(x) - \ell| < \varepsilon$.

2. *Consider* $f : (-\infty, b) \to \mathbf{R}$. *We say that*

$$\lim_{x \to -\infty} f(x) = \ell,$$

if for any $\varepsilon > 0$ *there is an* $M > 0$ *such that for all* $x < -M$ *we have* $|f(x) - \ell| < \varepsilon$.

Definition 3.4.2 *Consider* $f : (a, \infty) \to \mathbf{R}$. *We say*

$$\lim_{x \to \infty} f(x) = \infty,$$

if for all $M > 0$, *there exists* X, *such that* $f(x) > M$ *for all* $x > X$.

Example 3.4.3 *Show that* $\lim_{x \to +\infty}(x^2 + 1) = +\infty$.

Proof For any $M > 0$, we look for x with the property that

$$x^2 + 1 > M.$$

Take $A = \sqrt{M}$ then if $x > A$, $x^2+1 > M+1$. Hence $\lim_{x \to +\infty}(x^2+1) = +\infty$.

□

Remark 3.4.4 *In all cases,*

1. *There is a unique limit if it exists.*
2. *There is a sequential formulation. For example,* $\lim_{x \to \infty} f(x) = \ell$ *if and only if* $\lim_{n \to \infty} f(x_n) = \ell$ *for all sequences* $\{x_n\} \subset E$ *with* $\lim_{n \to \infty} x_n = \infty$. *Here* E *is the domain of* f.
3. *Algebra of limits hold.*
4. *The Sandwich Theorem holds.*

Chapter 4

The Extreme Value Theorem

By now we understand quite well what is a continuous function. Let us look at the landscape drawn by a continuous function. There are peaks and valleys. Is there a highest peak or a lowest valley?

4.1 Bounded Functions

Consider the range of $f : E \to \mathbf{R}$:

$$A = \{f(x)|x \in E\}.$$

Is A bounded from above and from below? If so it has a greatest lower bound m_0 and a least upper bound M_0. And

$$m_0 \le f(x) \le M_0, \qquad \forall x \in E.$$

Do m_0 and M_0 belong to A? That they belong to A means that there exist $\underline{x} \in E$ and $\bar{x} \in E$ such that $f(\underline{x}) = m_0$ and $f(\bar{x}) = M_0$.

Definition 4.1.1 *We say that $f : E \to \mathbf{R}$ is bounded above if there is a number M such that for all $x \in E$, $f(x) \le M$. We say that f attains its maximum if there is a number $c \in E$ such that for all $x \in E$, $f(x) \le f(c)$*

That a function f is bounded above means that its range $f(E)$ is bounded above.

Definition 4.1.2 *We say that $f : E \to \mathbf{R}$ is bounded below if there is a number m such that for all $x \in E$, $f(x) \ge m$.*

We say that f attains its minimum if there is a number $c \in E$ such that for all $x \in E$, $f(x) \ge f(c)$.

That a function f is bounded below means that its range $f(E)$ is bounded below.

Definition 4.1.3 *A function which is bounded above and below is* **bounded**.

4.2 The Bolzano-Weierstrass Theorem

This theorem was introduced and proved in Analysis I.

Lemma 4.2.1 (Bolzano-Weierstrass Theorem) *A bounded sequence has at least one convergent sub-sequence.*

4.3 The Extreme Value Theorem

Theorem 4.3.1 (The Extreme Value theorem) *Let $f : [a,b] \to \mathbf{R}$ be a continuous function. Then*

1. *f is bounded above and below, i.e. there exist numbers m and M such that*
$$m \le f(x) \le M, \qquad \forall\, x \in [a,b].$$

2. *There exist $\underline{x}, \overline{x} \in [a,b]$ such that*
$$f(\underline{x}) \le f(x) \le f(\overline{x}), \qquad a \le x \le b.$$
So
$$f(\underline{x}) = \inf\{f(x) | x \in [a,b]\}, \qquad f(\overline{x}) = \sup\{f(x) | x \in [a,b]\}.$$

Proof

1. We show that f is bounded above. Suppose not. Then for any $M > 0$ there is a point $x \in [a,b]$ such that $f(x) > M$. In particular, for every n there is a point $x_n \in [a,b]$ such that $f(x_n) \ge n$. The sequence $(x_n)_{n\in\mathbf{N}}$ is bounded, so there is a convergent subsequence x_{n_k}; let us denote its limit by x_0. Note that $x_0 \in [a,b]$. By sequential continuity of f at x_0,
$$\lim_{k\to\infty} f(x_{n_k}) = f(x_0)$$
But $f(x_{n_k}) \ge n_k \ge k$ and so
$$\lim_{k\to\infty} f(x_{n_k}) = \infty.$$

This gives a contradiction. The contradiction originated in the supposition that f is not bounded above. We conclude that f *must be* bounded from above.

2. Next we show that f attains its maximum value. Let

$$A = \{f(x) | x \in [a, b]\}.$$

Since A is not empty and bounded above, it has a least upper bound. Let

$$M_0 = \sup A.$$

By definition of supremum, if $\varepsilon > 0$ then $M - \varepsilon$ cannot be an upper bound of A, so there is a point $f(x) \in A$ such that $M_0 - \varepsilon \le f(x) \le M_0$. Taking $\varepsilon = \frac{1}{n}$ we obtain a sequence $x_n \in [a, b]$ such that

$$M_0 - \frac{1}{n} \le f(x_n) \le M_0.$$

By the Sandwich theorem,

$$\lim_{n\to\infty} f(x_n) = M_0.$$

Since $a \le x_n \le b$, it has a convergent subsequence x_{n_k} with limit $\bar{x} \in [a, b]$. By sequential continuity of f at x_0,

$$f(\bar{x}) = \lim_{k\to\infty} f(x_{n_k}) = M_0.$$

That is, f attains its maximum at $\bar{x}$.

3. Let $g(x) = -f(x)$. By step 1, g is bounded above and so f is bounded below. Note that

$$\sup\{g(x) | x \in [a, b]\} = -\inf\{f(x) | x \in [a, b]\}.$$

By the conclusion of step 2, there is a point $\underline{x}$ such that $g(\underline{x}) = \sup\{g(x) | x \in [a, b]\}$. This means that $f(\underline{x}) = \inf\{f(x) | x \in [a, b]\}$.

□

Remark 4.3.2 *By the extreme value theorem, if* $f : [a, b] \to \mathbf{R}$ *is a continuous function then the range of* f *is the closed interval* $[f(\underline{x}), f(\bar{x})]$. *By the Intermediate Value Theorem (IVT),* $f : [a, b] \to [f(\underline{x}), f(\bar{x})]$ *is surjective.*

Discussion on the assumptions:

Example 4.3.3 • *Let* $f : (0,1] \to \mathbf{R}$*,* $f(x) = \frac{1}{x}$*. Note that* f *is not bounded. The condition that the domain of* f *be a closed interval is violated.*

- *Let* $g : [1,2) \to \mathbf{R}$*,* $g(x) = x$*. It is bounded above and below. But the value*

$$\sup\{g(x)|x \in [1,2)\} = 2$$

is not attained on $[1,2)$*. Again, the condition "closed interval" is violated.*

Example 4.3.4 *Let* $f : [0,\pi] \cap \mathbf{Q}$*,* $f(x) = x$*. Let* $A = \{f(x)|x \in [0,\pi] \cap \mathbf{Q}\}$*. Then* $\sup A = \pi$*. But* π *is not attained on* $[0,\pi] \cap \mathbf{Q}$*. The condition which is violated here is that the domain of* f *must be an interval.*

Example 4.3.5 *Let* $f : \mathbf{R} \to \mathbf{R}$*,* $f(x) = x$*. Then* f *is not bounded. The condition "bounded interval" is violated.*

Example 4.3.6 *Let* $f : [-1,1] \to \mathbf{R}$*,*

$$f(x) = \begin{cases} \frac{1}{x}, & x \neq 0 \\ 2, & x = 0 \end{cases}$$

Then f *is not bounded. The condition "*f *is continuous" is violated.*

Chapter 5

The Inverse Function Theorem for continuous functions

We discuss the following question: if a continuous function f has an inverse, f^{-1}, is f^{-1} continuous?

5.1 The Inverse of a Function

Definition 5.1.1 *Let E and B be sets and let $f : E \to B$ be a function. We say*

1. *f is injective if $x \neq y \implies f(x) \neq f(y)$, for $x, y \in E$.*
2. *f is surjective, if for every $y \in B$ there is a point $x \in E$ such that $f(x) = y$.*
3. *f is bijective if it is surjective and injective.*

Definition 5.1.2 *If $f : E \to B$ is a bijection, the inverse of f is the function $f^{-1} : B \to E$ defined by*

$$f^{-1}(y) = x \quad \text{if} \quad f(x) = y.$$

N.B.

1) If $f : E \to B$ is a bijection, then so is its inverse $f^{-1} : B \to E$.

2) If $f : E \to \mathbf{R}$ is injective, then it is a bijection from its domain E to its range $f(E)$. So it has an inverse: $f^{-1} : f(E) \to E$.

3) If f^{-1} is the inverse of f, then f is the inverse of f^{-1}.

5.2 Monotone Functions

The simplest injective function is an increasing function, or a decreasing function. They are called monotone functions.

Definition 5.2.1 *Consider the function $f : E \to \mathbf{R}$.*

1. *We say f is increasing, if $f(x) < f(y)$ whenever $x < y$ and $x, y \in E$.*

2. *We say f is decreasing, if $f(x) > f(y)$ whenever $x < y$ and $x, y \in E$.*

3. *It is monotone or 'strictly monotone' if it is either increasing or decreasing.*

Compare this with the following definition:

Definition 5.2.2 *Consider the function $f : E \to \mathbf{R}$.*

1. *We say f is non-decreasing if for any pair of points x, y with $x < y$, we have $f(x) \leq f(y)$.*

2. *We say f is non-increasing if for any pair of points x, y with $x < y$, we have $f(x) \geq f(y)$.*

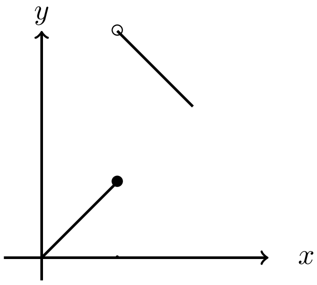
Some authors use the term 'strictly increasing' for 'increasing', and use the term 'increasing' where we use 'non-decreasing'. *We will always use 'increasing' and 'decreasing' in the strict sense defined in 5.2.1.*

If $f : [a, b] \to Range(f)$ is an increasing function, its inverse f^{-1} is also increasing. If f is decreasing, f^{-1} is also decreasing (Prove this!).

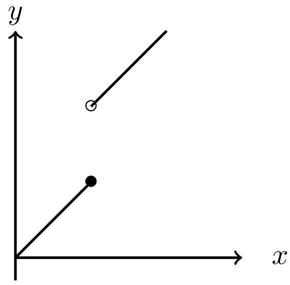
5.3 Continuous Injective and Monotone Surjective Functions

Increasing functions and decreasing functions are injective. Are there any other injective functions?

Yes: the function indicated in Graph A below is injective but not monotone. Note that f is not continuous. Surprisingly, if f is continuous and injective then it must be monotone.



Graph A



Graph B

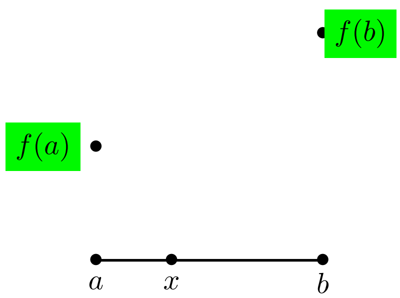
If $f : [a, b]$ is increasing, is $f : [a, b] \to [f(a), f(b)]$ surjective? Is it necessarily continuous? The answer to both questions is No, see Graph B. Again, continuity plays a role here. We show below that for an increasing function, 'being surjective' is equivalent to 'being continuous'!

Theorem 5.3.1 *Let $a < b$ and let $f : [a, b] \to \mathbf{R}$ be a continuous injective function. Then f is either strictly increasing or strictly decreasing.*

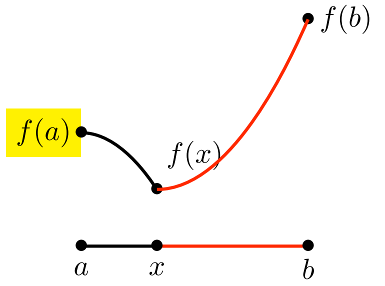
Proof First note that $f(a) \neq f(b)$ by injectivity.

- Case 1: Assume that $f(a) < f(b)$.
  Step 1: We first show that if $a < x < b$, then $f(a) < f(x) < f(b)$.

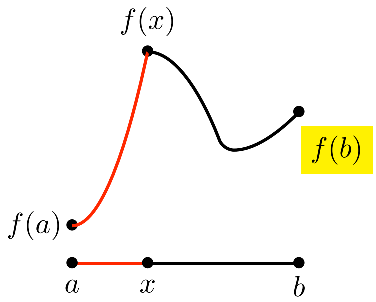
  Note that $f(x) \neq f(a)$ and $f(x) \neq f(b)$ by injectivity. If it is not true that $f(a) < f(x) < f(b)$, then either $f(x) < f(a)$ or $f(b) < f(x)$.



  1. In the case where $f(x) < f(a)$, we have $f(x) < f(a) < f(b)$. Take $v = f(a)$. By IVT for f on $[x, b]$, there exists $c \in (x, b)$ with $f(c) = v = f(a)$. Since $c \neq a$, this violates injectivity.



2. In the case where $f(b) < f(x)$, we have $f(a) < f(b) < f(x)$. Take $v = f(b)$. By the IVT for f on $[a, x]$, there exists $c \in (a, x)$ with $f(c) = v = f(b)$. This again violates injectivity.



We have therefore shown

$$\text{if } f(a) < f(b), \text{ then } a < x < b \implies f(a) < f(x) < f(b). \tag{5.3.1}$$

To show that f is strictly increasing on $[a, b]$, suppose that $a \leq x_1 < x_2 \leq b$. We must show that $f(x_1) < f(x_2)$. To do this, we simply apply the argument of Step 1, but now replacing a by x_1 and x by x_2. Making these replacements in (5.3.1) we obtain

$$\text{if } f(x_1) < f(b), \text{ then } x_1 < x_2 < b \implies f(x_1) < f(x_2) < f(b). \tag{5.3.2}$$

Since we certainly have $f(x_1) < f(b)$, by Step 1, and we are assuming $x_1 < x_2 < b$, the conclusion $f(x_1) < f(x_2) < f(b)$ holds, and in particular $f(x_1) < f(x_2)$. This completes the proof that if $f(a) < f(b)$ then f is strictly increasing.

- Case 2: If $f(a) > f(b)$ we show f is decreasing. Let $g = -f$. Then $g(a) < g(b)$. By Case 1 g is increasing and so f is decreasing.

□

Theorem 5.3.2 *If $f : [a, b] \to [f(a), f(b)]$ is increasing and surjective, it is continuous.*

Proof Fix $c \in (a, b)$. Take $\varepsilon > 0$. We wish to find the set of x such that $|f(x) - f(c)| < \varepsilon$, or

$$f(c) - \varepsilon < f(x) < f(c) + \varepsilon. \tag{5.3.3}$$

Let $\varepsilon' = \min\{\varepsilon, f(b) - f(c), f(c) - f(a)\}$. Then

$$f(a) < f(c) - \varepsilon' < f(c), \qquad f(c) < f(c) + \varepsilon' < f(b).$$

Since $f : [a, b] \to [f(a), f(b)]$ is surjective we may define

$$a_1 = f^{-1}(f(c) - \varepsilon') \quad b_1 = f^{-1}(f(c) + \varepsilon').$$

Since f is increasing,

$$a_1 < x < b_1 \implies f(c) - \varepsilon' < f(x) < f(c) + \varepsilon'.$$

As $\varepsilon' \leq \varepsilon$, we conclude

$$a_1 < x < b_1 \implies f(c) - \varepsilon < f(x) < f(c) + \varepsilon.$$

We have proved that f is continuous at $c \in (a, b)$. The continuity of f at a and b can be proved similarly. □

5.4 The Inverse Function Theorem (Continuous Version)

Suppose that $f : [a, b] \to \mathbf{R}$ is continuous and injective. Then $f : [a, b] \to$ *range*(f) is a continuous bijection and has inverse $f^{-1} :$ *range*$(f) \to [a, b]$.

By Theorem 5.3.1 f is either increasing or decreasing. If it is increasing,

$$f : [a, b] \to [f(a), f(b)]$$

is surjective by the IVT. It has inverse

$$f^{-1} : [f(a), f(b)] \to [a, b]$$

which is also increasing and surjective, and therefore continuous, by Theorem 5.3.2. If f is decreasing,

$$f : [a, b] \to [f(b), f(a)]$$

is surjective, again by the IVT. It has inverse

$$f^{-1} : [f(b), f(a)] \to [a, b]$$

which is also decreasing and surjective, and therefore continuous, again by Theorem 5.3.2.

We have proved

Theorem 5.4.1 *[The Inverse Function Theorem: continuous version]*

If $f : [a, b] \to \mathbf{R}$ is continuous and injective, its inverse $f^{-1} : range(f) \to [a, b]$ is continuous.

Example 5.4.2 *For $n \in \mathbf{N}$, the function $f : [0, \infty) \to [0, \infty)$ defined by $f(x) = x^{\frac{1}{n}} : [0, 1] \to \mathbf{R}$ is continuous.*

Proof Let $g(x) = x^n$. Then $g : [0, a] \to [0, a^n]$ is increasing and continuous, and therefore has continuous inverse. Every point $c \in (0, \infty)$ lies in the interior of some interval $[0, a^2]$. Thus f is continuous at c. □

Example 5.4.3 *Consider* $\sin : [-\frac{\pi}{2}, \frac{\pi}{2}] \to [-1, 1]$. *It is increasing and continuous. Define its inverse* $\arcsin : [-1, 1] \to [-\frac{\pi}{2}, \frac{\pi}{2}]$. *Then* $\arcsin$ *is continuous.*

Chapter 6

Differentiation

Definition 6.0.4 *Let $f : (a, b) \to \mathbf{R}$ be a function and let $x \in (a, b)$. We say f is* **differentiable** *at x if*

$$\lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

exists and is a real number (not $\pm\infty$). If this is the case, the value of the limit is the **derivative** *of f at x, which is usually denoted $f'(x)$.*

Example 6.0.5 *(i) Let $f(x) = x^2$. Then*

$$\frac{f(x+h) - f(x)}{h} = \frac{(x+h)^2 - x^2}{h} = \frac{2xh + h^2}{h} = 2x + h. \tag{6.0.1}$$

As h gets smaller, this tends to $2x$:

$$\lim_{h \to 0} \frac{f(x+h) - f(x)}{h} = 2x. \tag{6.0.2}$$

(ii) Let $f(x) = |x|$ and let $x = 0$. Then

$$\frac{f(0+h) - f(0)}{(0+h) - 0} = \frac{|h|}{h} = \begin{cases} 1 & \text{if } h > 0 \\ -1 & \text{if } h < 0 \end{cases}$$

It follows that

$$\lim_{\Delta x \to 0-} \frac{f(0+h) - f(0)}{h} = -1, \qquad \lim_{h \to 0+} \frac{f(0+h) - f(0)}{h} = 1,$$

and, since the limit from the left and the limit from the right do not agree, $\lim_{h \to 0} \frac{f(0+h) - f(0)}{h}$ does not exist.

Thus, the function $f(x) = x^2$ is differentiable at every point x, with derivative $f'(x) = 2x$. The function $f(x) = |x|$ is not differentiable at $x = 0$. In fact it is differentiable everywhere else: at any point $x \neq 0$, there is a neighbourhood in which f is equal to the function $g(x) = x$, or the function $h(x) = -x$. Both of these are differentiable, with $g'(x) = 1$ for all x,and $h'(x) = -1$ for all x. The differentiability of f at x is determined by its behaviour in the neighbourhood of the point x, so if f coincides with a differentiable function near x then it too is differentiable there.

We list some of the motivations for studying derivatives:

- Calculate the angle between two curves where they intersect. (Descartes)
- Find local minima and maxima (Fermat 1638)
- Describe velocity and acceleration of movement (Galileo 1638, Newton 1686).
- Express some physical laws (Kepler, Newton).
- Determine the shape of the curve given by $y = f(x)$.

Example 6.0.6 1. $f : \mathbf{R}/\{0\} \to \mathbf{R}$, $f(x) = \frac{1}{x}$ is differentiable at $x_0 \neq 0$. For

$$\begin{aligned}\lim_{x\to x_0} \frac{f(x) - f(x_0)}{x - x_0} &= \lim_{x\to x_0} \frac{\frac{1}{x} - \frac{1}{x_0}}{x - x_0}\\ &= \lim_{x\to x_0} \frac{\frac{x_0 - x}{x_0 x}}{x - x_0}\\ &= \lim_{x\to x_0} -\frac{1}{x_0 x} = -\frac{1}{x_0^2}.\end{aligned}$$

In the last step we have used the fact that $\frac{1}{x}$ is continuous at $x_0 \neq 0$.

2. $\sin x : \mathbf{R} \to \mathbf{R}$ is differentiable everywhere and $(\sin x)' = \cos x$. For let $x_0 \in \mathbf{R}$.

$$\begin{aligned}\lim_{h\to 0} \frac{\sin(x_0 + h) - \sin x_0}{h} &= \lim_{h\to 0} \frac{\sin(x_0)\cos h + \cos(x_0)\sin h - \sin x_0}{h}\\ &= \lim_{h\to 0} \frac{\sin(x_0)(\cos h - 1) + \cos(x_0)\sin h}{h}\\ &= \left(\sin(x_0) \lim_{h\to 0} \frac{\cos h - 1}{h} + \cos(x_0) \lim_{h\to 0} \frac{\sin h}{h}\right)\\ &= \sin(x_0) \cdot 0 + \cos(x_0) \cdot 1 = \cos(x_0).\end{aligned}$$

We have used the previously calculated limits:

$$\lim_{h\to 0}\frac{(\cos h-1)}{h}=0, \qquad \lim_{h\to 0}\frac{\sin h}{h}=1.$$

3. The function

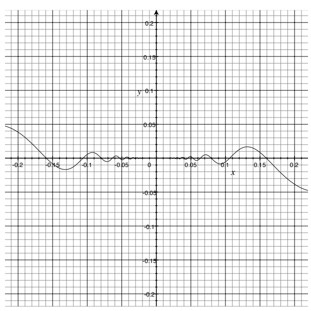
$$f(x)=\begin{cases} x\sin(1/x), & x\neq 0\\ 0, & x=0\end{cases}$$

is continuous everywhere. But it is not differentiable at $x_0=0$.
It is continuous at $x\neq 0$ by an easy application of the algebra of continuity. That $f(x)$ is continuous at $x=0$ is discussed in Example 3.1.15.

It is not differentiable at 0 because $(f(h)-f(0))/h = sin(1/h)$, and $\sin(1/h)$ does not have a limit as h tends to 0. (Take $x_n=\frac{1}{2\pi n}\to 0$, $y_n=\frac{1}{2\pi n+\frac{\pi}{2}}\to 0$. Then $\sin(1/x_n)=0$, $\sin(1/y_n)=1$.)

4. The graph below is of

$$f(x)=\begin{cases} x^2\sin(1/x) & \text{if} \quad x\neq 0\\ 0 & \text{if} \quad x=0\end{cases} \tag{6.0.3}$$



.

We will refer to this function as $x^2\sin(1/x)$ even though its value at 0 requires a separate line of definition. It is differentiable at 0, with derivative $f'(0)=0$ (**Exercise**). In fact f differentiable at every point $x\in\mathbf{R}$, but we will not be able to show this until we have proved the chain rule for differentiation.

6.1 The Weierstrass-Carathéodory Formulation for Differentiability

Below we give the formulation for the differentiability of f at x_0 given by Carathéodory in 1950.

Theorem 6.1.1 *[Weierstrass-Carathéodory Formulation] Consider* $f : (a, b) \to \mathbf{R}$ *and* $x_0 \in (a, b)$*. The following statements are equivalent:*

1. *f is differentiable at x_0*
2. *There is a function ϕ continuous at x_0 such that*

$$f(x) = f(x_0) + \phi(x)(x - x_0). \tag{6.1.1}$$

Furthermore $f'(x_0) = \phi(x_0)$*.*

Proof

- 1) $\Rightarrow$ 2) Suppose that f is differentiable at x_0. Set

  $$\phi(x) = \begin{cases} \frac{f(x)-f(x_0)}{x-x_0}, & x \neq x_0 \\ f'(x_0), & x = x_0 \end{cases}.$$

  Then

  $$f(x) = f(x_0) + \phi(x)(x - x_0).$$

  Since f is differentiable at x_0, ϕ is continuous at x_0.

- 2) $\Rightarrow$ 1). Assume that there is a function ϕ continuous at x_0 such that

  $$f(x) = f(x_0) + \phi(x)(x - x_0).$$

  Then

  $$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{x \to x_0} \phi(x) = \phi(x_0).$$

  The last step follows from the continuity of ϕ at x_0. Thus $f'(x_0)$ exists and is equal to $\phi(x_0)$.

□

Remark: Compare the above formulation with the geometric interpretation of the derivative. The tangent line at x_0 is

$$y = f(x_0) + f'(x_0)(x - x_0).$$

If f is differentiable at x_0,

$$\begin{aligned} f(x) &= f(x_0) + \phi(x)(x - x_0) \\ &= f(x_0) + \phi(x_0)(x - x_0) + [\phi(x) - \phi(x_0)](x - x_0) \\ &= f(x_0) + f'(x_0)(x - x_0) + [\phi(x) - \phi(x_0)](x - x_0). \end{aligned}$$

The last step follows from $\phi(x_0) = f'(x_0)$. Observe that

$$\lim_{x \to x_0} [\phi(x) - \phi(x_0)] = 0$$

and so $[\phi(x) - \phi(x_0)](x - x_0)$ is insignificant compared to the the first two terms. We may conclude that the tangent line $y = f(x_0) + f'(x_0)(x - x_0)$ is indeed a linear approximation of $f(x)$.

Corollary 6.1.2 *If f is differentiable at x_0 then it is continuous at x_0.*

Proof If f is differentiable at x_0, $f(x) = f(x_0) + \phi(x)(x - x_0)$ where ϕ is a function continuous at x_0. By algebra of continuity f is continuous at x_0.

□

Example 6.1.3 *The converse to the Corollary does not hold.*

Let $f(x) = |x|$. It is continuous at 0, but fails to be differentiable at 0.

Example 6.1.4 *Consider $f : \mathbf{R} \to \mathbf{R}$ given by*

$$f(x) = \begin{cases} x^2, & x \in \mathbf{Q} \\ 0, & x \notin \mathbf{Q} \end{cases}.$$

Claim: f is differentiable only at the point 0.

Proof *Take $x_0 \neq 0$. Then f is not continuous at x_0, as we learnt earlier, and so not differentiable at x_0.*

Take $x_0 = 0$, let

$$g(x) = \frac{f(x) - f(0)}{x} = \begin{cases} x, & x \in \mathbf{Q} \\ 0, & x \notin \mathbf{Q} \end{cases}.$$

We learnt earlier that g is continuous at 0 and hence has limit $g(0) = 0$. Thus $f'(0)$ exists and equals 0.

□

Example 6.1.5

$$f_1(x) = \begin{cases} x^2 \sin(1/x) & x \neq 0 \\ 0, & x = 0 \end{cases}$$

is differentiable at $x = 0$*.*

Proof Let

$$\phi(x) = \begin{cases} x \sin(1/x) & x \neq 0 \\ 0, & x = 0 \end{cases}$$

Then

$$f_1(x) = \phi(x)x = f_1(0) + \phi(x)(x - 0).$$

Since ϕ is continuous at $x = 0$, f_1 is differentiable at $x = 0$.
[We proved that ϕ is continuous in Example 3.1.15.] □

6.2 Properties of Differentiatiable Functions

Rules of differentiation can be easily deduced from properties of continuous limits. They can be proved by the sequential limit method or directly from the definition in terms of ε and δ.

Theorem 6.2.1 (Algebra of Differentiability) *Suppose that* $x_0 \in (a, b)$ *and* $f, g : (a, b) \to \mathbf{R}$ *are differentiable at* x_0*.*

1. *Then* $f + g$ *is differentiable at* x_0 *and*

$$(f + g)'(x_0) = f'(x_0) + g'(x_0);$$

2. *(product rule)* fg *is differentiable at* x_0 *and*

$$(fg)'(x_0) = f'(x_0)g(x_0) + f(x_0)g'(x_0) = (f'g + fg')(x_0).$$

Proof If f and g are differentiable at x_0 there are two functions ϕ and ψ continuous at x_0 such that

$$f(x) = f(x_0) + \phi(x)(x - x_0) \qquad (6.2.1)$$
$$g(x) = g(x_0) + \psi(x)(x - x_0). \qquad (6.2.2)$$

Furthermore

$$f'(x_0) = \phi(x_0) \qquad g'(x_0) = \psi(x_0).$$

1. Adding the two equations (6.2.1) and (6.2.2), we obtain

$$\begin{aligned}(f+g)(x) &= f(x_0)+\phi(x)(x-x_0)+g(x_0)+\psi(x)(x-x_0)\\ &= (f+g)(x_0)+[\phi(x)+\psi(x)](x-x_0).\end{aligned}$$

Since $\phi+\psi$ is continuous at x_0, $f+g$ is differentiable at x_0. And

$$(f+g)'(x_0)=(\phi+\psi)(x_0)=f'(x_0)+g'(x_0).$$

2. Multiplying the two equations (6.2.1) and (6.2.2), we obtain

$$\begin{aligned}(fg)(x) &= \Big(f(x_0)+\phi(x)(x-x_0)\Big)\Big(g(x_0)+\psi(x)(x-x_0)\Big)\\ &= (fg)(x_0)\\ &\quad+\Big(g(x_0)\phi(x)+f(x_0)\psi(x)+\phi(x_0)\psi(x_0)(x-x_0)\Big)(x-x_0).\end{aligned}$$

Let

$$\theta(x)=g(x_0)\phi(x)+f(x_0)\psi(x)+\phi(x_0)\psi(x_0)(x-x_0).$$

Since ϕ,ψ are continuous at x_0, θ is continuous at x_0 by the algebra of continuity. It follows that fg is differentiable at x_0. Furthermore

$$(fg)'(x_0)=\theta(x_0)=g(x_0)\phi(x_0)+f(x_0)\psi(x_0)=g(x_0)f'(x_0)+f(x_0)g'(x_0).$$

□

In Lemma 1.3.6, we showed that if g is continuous at a point x_0 and $g(x_0)\neq 0$, then there is a neighbourhood of x_0,

$$U_{x_0}^r=(x_0-r,x_0+r)$$

on which $f(x)\neq 0$. Here r is a positive number. This means f/g is well defined on $U_{x_0}^r$ and below in the theorem we only need to consider f restricted to $U_{x_0}^r$.

Theorem 6.2.2 (Quotient Rule) *Suppose that $x_0\in(a,b)$ and $f,g:(a,b)\to\mathbf{R}$ are differentiable at x_0. Suppose that $g(x_0)\neq 0$, then $\frac{f}{g}$ is differentiable at x_0 and*

$$\left(\frac{f}{g}\right)'(x_0)=\frac{f'(x_0)g(x_0)-f(x_0)g'(x_0)}{g^2(x_0)}=\frac{f'g-g'f}{g^2}(x_0).$$

Proof There are two functions ϕ and ψ continuous at x_0 such that

$$\begin{aligned} f(x) &= f(x_0) + \phi(x)(x - x_0) \\ f'(x_0) &= \phi(x_0) \\ g(x) &= g(x_0) + \psi(x)(x - x_0) \\ g'(x_0) &= \psi(x_0). \end{aligned}$$

Since g is differentiable at x_0, it is continuous at x_0. By Lemma 1.3.6,

$$g(x) \neq 0, \qquad x \in (x_0 - r, x_0 + r)$$

for some $r > 0$ such that $(x_0 - r, x_0 + r) \subset (a, b)$. We may divide f by g to see that

$$\begin{aligned} (\frac{f}{g})(x) &= \frac{f(x_0) + \phi(x)(x - x_0)}{g(x_0) + \psi(x)(x - x_0)} \\ &= \frac{f(x_0)}{g(x_0)} - \frac{f(x_0)[g(x_0) + \psi(x)(x - x_0)]}{g(x_0)[g(x_0) + \psi(x)(x - x_0)]} + \frac{g(x_0)[f(x_0) + \phi(x)(x - x_0)]}{g(x_0)[g(x_0) + \psi(x)(x - x_0)]} \\ &= \frac{f(x_0)}{g(x_0)} + \frac{g(x_0)\phi(x) - f(x_0)\psi(x)}{g(x_0)[g(x_0) + \psi(x)(x - x_0)]}(x - x_0). \end{aligned}$$

Define

$$\theta(x) = \frac{g(x_0)\phi(x) - f(x_0)\psi(x)}{g(x_0)[g(x_0) + \psi(x)(x - x_0)]}.$$

Since $g(x_0) \neq 0$, θ is continuous at x_0 and (f/g) is differentiable at x_0. And

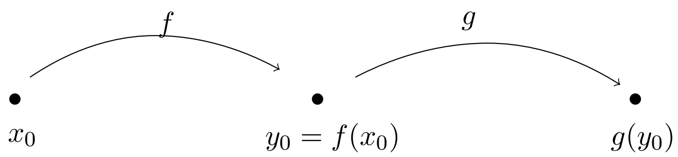
$$(\frac{f}{g})'(x_0) = \theta(x_0) = \frac{g(x_0)f'(x_0) + f(x_0)g'(x_0)}{[g(x_0)]^2}.$$

□

Exercise 6.2.3 *Prove the sum, product rule and quotient rules for the derivative directly from the definition of the derivative as a limit, without using the Carathéodory formulation.*

Theorem 6.2.4 (chain rule) *Let $x_0 \in (a, b)$. Suppose that $f : (a, b) \to (c, d)$ is differentiable at x_0 and $g : (c, d) \to \mathbf{R}$ is differentiable at $f(x_0)$. Then $g \circ f$ is differentiable at x_0 and*

$$(g \circ f)'(x_0) = g'(f(x_0)) \cdot f'(x_0).$$



Proof There is a function ϕ continuous at x_0 such that

$$f(x) = f(x_0) + \phi(x)(x - x_0).$$

Let $y_0 = f(x_0)$. There is a function ψ continuous at y_0 such that

$$g(y) = g(y_0) + \psi(y)(y - y_0). \tag{6.2.3}$$

Substituting y and y_0 in (6.2.3) by $f(x)$ and $f(x_0)$, we have

$$\begin{aligned} g\Big(f(x)\Big) &= g(f(x_0)) + \psi\Big(f(x)\Big)\Big(f(x) - f(x_0)\Big) \\ &= g(f(x_0)) + \psi\Big(f(x)\Big)\phi(x)(x - x_0). \end{aligned}$$

Let $\theta(x) = \big(\psi \circ f(x)\big)\phi(x)$. Then (6.2.4) gives

$$g \circ f(x) = g \circ f(x_0) + \theta(x)(x - x_0). \tag{6.2.4}$$

Since f is continuous at x_0 and ψ is continuous at $f(x_0)$, the composition $\psi \circ f$ is continuous at x_0. Then θ, as a product of continuous functions, is continuous at x_0. Using the Carathéodory formulation of differentiability, 6.1.1, it therefore follows from (6.2.4) that $g \circ f$ is differentiable at x_0, with

$$(g \circ f)'(x_0) = \theta(x_0) = \psi\Big(f(x_0)\Big) \cdot \phi(x_0) = g'\Big(f(x_0)\Big) \cdot f'(x_0).$$

□

Exercise 6.2.5 The following proof of the Chain Rule works under one additional assumption. What is it?

Assume f differentiable at x_0 and g differentiable at $f(x_0)$. We have

$$\frac{g(f(x_0 + h)) - g(f(x_0))}{h} = \left(\frac{g(f(x_0 + h)) - g(f(x_0))}{f(x_0 + h) - f(x_0)}\right)\left(\frac{f(x_0 + h) - f(x_0)}{h}\right) \tag{6.2.5}$$

Write $y_0 = f(x_0)$ *and* $t = f(x_0+h) - f(x_0)$. *Then* $f(x_0+h) = y_0 + t$ *and*

$$\frac{g(f(x_0+h)) - g(f(x_0))}{f(x_0+h) - f(x_0)} = \frac{g(y_0+t) - g(y_0)}{t}.$$

If f *is differentiable at* x_0, *then it is continuous there. Hence when* $h \to 0$, $t \to 0$ *also. It follows that as* $h \to 0$, $(g(y_0+t) - g(y_0))/t$ *tends to* $g'(f(x_0))$. *Therefore by* (6.2.5),

$$\begin{aligned}
&\lim_{h\to 0} \frac{g(f(x_0+h)) - g(f(x_0))}{h} \\
&\qquad = \lim_{h\to 0} \frac{g(f(x_0+h)) - g(f(x_0)}{f(x_0+h) - f(x_0)} \lim_{h\to 0} \frac{f(x_0+h) - f(x_0)}{h} \\
&\qquad = g'(f(x_0))f'(x_0).
\end{aligned}$$

Example 6.2.6 *We can use the product and chain rules to prove the quotient rule. Since* $g(x_0) \neq 0$, *it does not vanish on some interval* $(x_0 - \delta, x_0 + \delta)$. *The function* f/g *is defined everywhere on this interval. Let* $h(x) = \frac{1}{x}$. *Then*

$$\frac{f}{g} = f \times h \circ g.$$

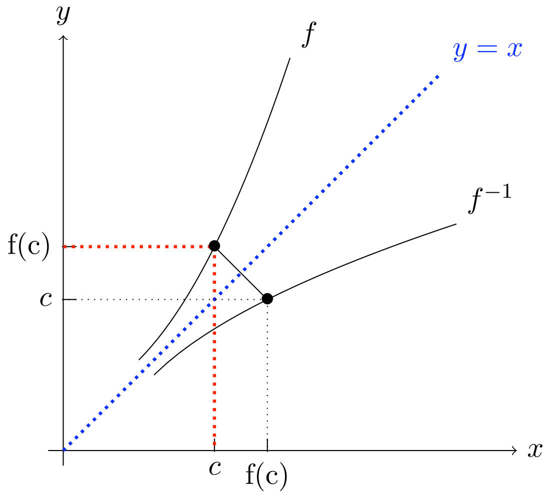
Since g *does not vanish anywhere on* $(x_0 - \delta, x_0 + \delta)$, $h \circ g$ *is differentiable at* x_0 *and therefore so is* $f \times h \circ g$.

For the value of the derivative, note that $h'(y) = -1/y^2$, *so*

$$\begin{aligned}
\left(\frac{f}{g}\right)'(x_0) &= f'(x_0)h(g(x_0)) + f(x_0)h'(g(x_0))g'(x_0) \\
&= \frac{f'(x_0)}{g(x_0)} + f(x_0)\frac{-1}{g^2(x_0)}g'(x_0) \\
&= \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g^2(x_0)}.
\end{aligned}$$

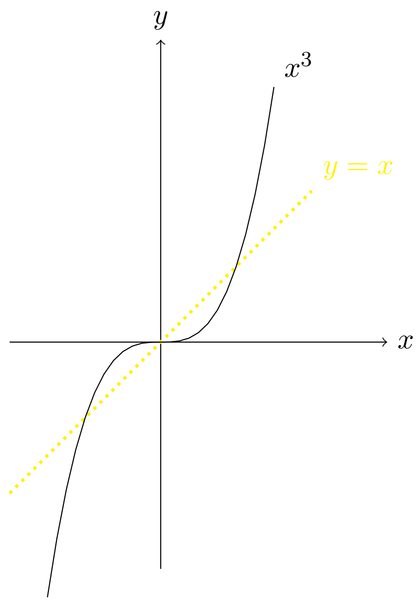
6.3 The Inverse Function Theorem

Is the inverse of a differentiable bijective function differentiable?



The graph of f is the graph of f^{-1}, reflected by the line $y = x$. We might therefore guess that f^{-1} is as smooth as f. This is almost right, but there is one important exception: where $f'(x_0) = 0$.

Conisder the following graph, of $f(x) = x^3$. Its tangent line at $(0, 0)$ is the line $\{y = 0\}$.



If we reflect the graph in the line $y = x$, the tangent line at $(0, 0)$ becomes the vertical line $\{x = 0\}$, which has infinite slope. Indeed, the inverse of f

is the function $f^{-1}(x) = x^{1/3}$, which is not differentiable at 0:

$$\lim_{h\to 0} \frac{(x+h)^{1/3} - x^{1/3}}{h} = \infty.$$

Exercise 6.3.1 *Prove this formula for the limit.*

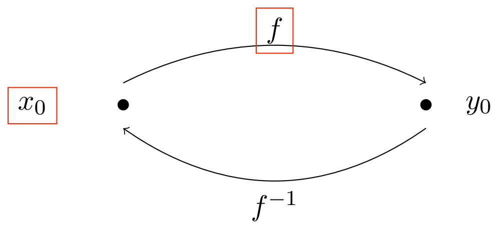
Recall that if f is continuous and injective, the inverse function theorem for continuous functions, 5.4.1, states that f is either increasing or decreasing, and that

- if f is increasing then $f : [a, b] \to [f(a), f(b)]$ is a bijection,
- if f is decreasing then $f : [a, b] \to [f(b), f(a)]$ is a bijection,

and in both cases f^{-1} is continuous.

Theorem 6.3.2 (The inverse Function Theorem, II) *Let $f : [a, b] \to [c, d]$ be a continuous bijection. Let $x_0 \in (a, b)$ and suppose that f is differentiable at x_0 and $f'(x_0) \neq 0$. Then f^{-1} is differentiable at $y_0 = f(x_0)$. Furthermore,*

$$(f^{-1})'(y_0) = \frac{1}{f'(x_0)}.$$



Proof Since f is differentiable at x_0,

$$f(x) = f(x_0) + \phi(x)(x - x_0),$$

where ϕ is continuous at x_0. Letting $x = f^{-1}(y)$,

$$f(f^{-1}(y)) = f(f^{-1}(y_0)) + \phi(f^{-1}(y))(f^{-1}(y) - f^{-1}(y_0)).$$

So

$$y - y_0 = \phi(f^{-1}(y))(f^{-1}(y) - f^{-1}(y_0)).$$

Since f is a continuous injective map its inverse f^{-1} is continuous. The composition $\phi \circ f^{-1}$ is continuous at y_0. By the assumption $\phi \circ f^{-1}(y_0) = \phi(x_0) = f'(x_0) \neq 0$, $\phi(x) \neq 0$ for x close to x_0. Define

$$\theta(y) = \frac{1}{\phi(f^{-1}(y))}.$$

It follows that θ is continuous at y_0 and

$$f^{-1}(y) = f^{-1}(y_0) + \theta(y)(y - y_0).$$

Consequently, f^{-1} is differentiable at y_0 and

$$(f^{-1})'(y_0) = \theta(y_0) = \frac{1}{f'(f^{-1}(y_0))} = \frac{1}{f'(x_0)}.$$

□

Recall the following lemma (Lemma 3.1.21).

Lemma 6.3.3 *Let $g : (a, b)/\{c\} \to \mathbf{R}$, and $c \in (a, b)$. Suppose that*

$$\lim_{x \to c} g(x) = \ell.$$

1. *If $\ell > 0$, then $g(x) > 0$ for x close to c.*
2. *If $\ell < 0$, then $g(x) < 0$ for x close to c.*

If $f'(x_0) \neq 0$, by the non-vanishing Lemma above, applied to $g(x) = \frac{f(x)-f(x_0)}{x-x_0}$, we may assume that on $(x_0 - r, x_0 + r)$

$$\frac{f(x) - f(x_0)}{x - x_0} > 0.$$

Remark 6.3.4

1. *If $f' > 0$ on an interval $(x_0 - r, x_0 + r)$, by Corollary 7.5.1 to the Mean Value Theorem in the next chapter, f is an increasing function. Thus*

$$f : [x_0 - \frac{r}{2}, x_0 + \frac{r}{2}] \to: [f(x_0 - \frac{r}{2}), f(x_0 + \frac{r}{2})]$$

is a bijection. In summary if $f'(x) > 0$ for all $x \in (a, b)$, then f is invertible.

2. *Suppose that f is differentiable on (a, b) and f' is continuous on (a, b). If $x_0 \in (a, b)$ and $f'(x_0) > 0$ then f' is positive on an interval near x_0 and on which the function is invertible.*

6.4 One Sided Derivatives

Definition 6.4.1 *1. A function $f : (a, x_0] \to \mathbf{R}$ is said to have left derivative at x_0 if*

$$\lim_{x \to x_0-} \frac{f(x) - f(x_0)}{x - x_0}$$

exists. The limit will be denoted by $f'(x_0-)$, and called the **left derivative** *of f at x_0.*

2. A function $f : (x_0, b) \to \mathbf{R}$ is said to have right derivative at x_0 if

$$\lim_{x \to x_0+} \frac{f(x) - f(x_0)}{x - x_0}$$

exists. The limit will be denoted by $f'(x_0+)$, and called the **right derivative** *of f at x_0.*

Theorem 6.4.2 *A function f has a derivative at x_0 if and only if both $f'(x_0+)$ and $f'(x_0-)$ exist and are equal.*

Example 6.4.3

$$f(x) = \begin{cases} x \sin(1/x) & x > 0 \\ 0, & x \le 0 \end{cases}$$

Claim: $f'(0-) = 0$ but $f'(0+)$ does not exist.
Proof

$$\lim_{x \to 0-} \frac{f(x) - f(0)}{x - 0} = \lim_{x \to 0-} \frac{0 - 0}{x} = 0.$$

But

$$\lim_{x \to 0+} \frac{f(x) - 0}{x - 0} = \lim_{x \to 0+} \sin(1/x)$$

does not exist, as can be seen below. Take

$$x_n = \frac{1}{2n\pi} \ge 0, \qquad y_n = \frac{1}{\frac{\pi}{2} + 2n\pi} \ge 0.$$

Both sequences converge to 0 from the right. But

$$\sin(1/x_n) = 0, \qquad \sin(1/y_n) = 1.$$

They have different limits so $\lim_{x \to 0+} \sin(1/x)$ does not exist, and so $f'(0+)$ does not exist.

□

Example 6.4.4

$$g(x) = \begin{cases} \sqrt{x}\sin(1/x) & x > 0 \\ 0, & x \le 0 \end{cases}$$

Claim: $g'(0+)$ does not exist.
Proof *Note that*

$$\lim_{x\to 0+} \frac{g(x) - g(0)}{x - 0} = \lim_{x\to 0+} \frac{1}{\sqrt{x}} \sin(1/x).$$

Take

$$y_n = \frac{1}{\frac{\pi}{2} + 2n\pi} \to 0.$$

But

$$\frac{1}{\sqrt{y_n}} \sin(1/y_n) = \sqrt{\frac{\pi}{2} + 2n\pi} \to +\infty.$$

Hence $\frac{1}{\sqrt{x}}\sin(1/x)$ cannot converge to a finite number as x approaches 0 from the right. We conclude that $g'(0+)$ does not exist. □

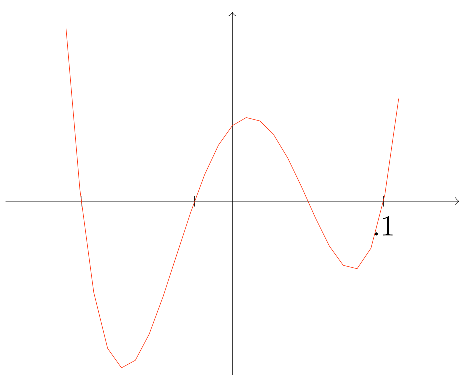
Chapter 7

The mean value theorem

If we have some information on the derivative of a function what can we say about the function itself?

What do we think when our minister of the economy tells us that "The rate of inflation is slowing down"? This means that if $f(t)$ is the price of a commodity at time t, its derivative is decreasing, but not necessarily the price itself (in fact most surely not, otherwise we would be told so directly)! I heard the following on the radio: "the current trend of increasing unemployment is slowing down". What does it really mean?

Exercise 7.0.5 *Consider the following graph.*



To do: *(i) Suppose that this is the graph of some function f, and make a sketch of the graph of f'.*
(ii) Suppose instead that this is the graph of the derivative f', and make a sketch of the graph of f.

Hint for (ii): First find the critical points (where $f' = 0$). For each *one, decide whether it is a local maximum or a local minimum: if x is a critical point and f' is increasing at x, then x it is a local minimum, and if f' is decreasing, then x is a local maximum. Identify inflection points (where f' reaches local minimum or local maximum). At these points the graph changes from convex to concave or vice versa. Try to decide whether the graph between two consecutive critical points is increasing or decreasing, convex or concave.*

7.1 Local Extrema

Definition 7.1.1 *Consider $f : [a, b] \to \mathbf{R}$ and $x_0 \in [a, b]$.*

1. *We say that f has a local maximum at x_0, if for all x in some neighbourhood $(x_0 - \delta, x_0 + \delta)$ (where $\delta > 0$) of x_0, we have $f(x_0) \geq f(x)$. If $f(x_0) > f(x)$ for $x \neq x_0$ in this neighbourhood, we say that f has a strict local maximum at x_0.*

2. *We say that f has a local minimum at x_0, if for all x in some neighbourhood $(x_0 - \delta, x_0 + \delta)$ (where delta > 0) of x_0, we have $f(x_0) \leq f(x)$. If $f(x_0) < f(x)$ for $x \neq x_0$ in this neighbourhood, we say that f has a strict local minimum at x_0.*

3. *We say that f has a local extremum at x_0 if it either has a local minimum or a local maximum at x_0.*

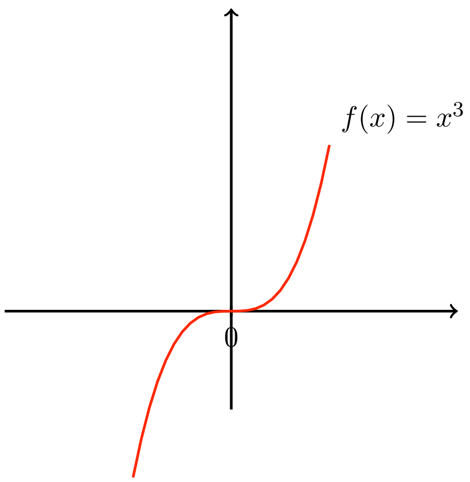
Example 7.1.2 *Let $f(x) = \frac{1}{4}x^4 - \frac{1}{2}x^2$. Then*

$$f(x) = \frac{1}{4}x^2(x^2 - 2).$$

It has a local maximum at $x = 0$, since $f(0) = 0$ and near $x = 0$, $f(x) < 0$.

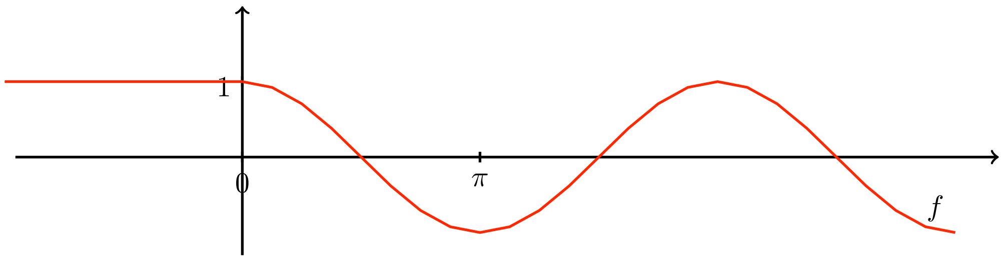
It has critical points at $x = \pm 1$. We will prove that they are local minima later.

Example 7.1.3 *$f(x) = x^3$. This function is strictly increasing, and there is no local minimum or local maximum at 0, even though $f'(0) = 0$.*



Example 7.1.4

$$f(x) = \begin{cases} 1, & x \in [-\pi, 0], \\ \cos(x), & x \in [0, 3\pi]. \end{cases} .$$



Then $x = 2\pi$ is a strict local maximum, and $x = -\pi$ is a strict local minimum; the function is constant on $[-\pi, 0]$, so each point in the interior of this interval is both a local minimum and a local maximum! The point $x = 0$ is a local maximum but not a local minimum. The point $-\pi$ is both a local minimum and a local maximum.

Lemma 7.1.5 *Consider $f : (a, b) \to \mathbf{R}$. Suppose that $x_0 \in (a, b)$ is either a local minimum or a local maximum of f. If f is differentiable at x_0 then $f'(x_0) = 0$.*

Proof Case 1. We first assume that f has a local maximum at x_0:

$$f(x_0) \geq f(x), \qquad x \in (x_0 - \delta_1, x_0 + \delta_1),$$

for some $\delta_1 > 0$. Then

$$f'(x_0+) = \lim_{h \to 0+} \frac{f(x_0 + h) - f(x_0)}{h} \leq 0$$

since $h \in (0, \delta_1)$ eventually. Similarly

$$f'(x_0-) = \lim_{h \to 0-} \frac{f(x_0 + h) - f(x_0)}{h} \geq 0.$$

Since the derivative exists we must have

$$f'(x_0) = f'_+(x_0) = f'_-(x_0).$$

Finally we deduce that $f'(x_0) = 0$.

Case 2. If x_0 is a local minimum of f, take $g = -f$. Then x_0 is a local maximum of g and $g'(x_0) = -f'(x_0)$ exists. It follows that $g'(x_0) = 0$ and consequently $f'(x_0) = 0$. □

7.2 Global Maximum and Minimum

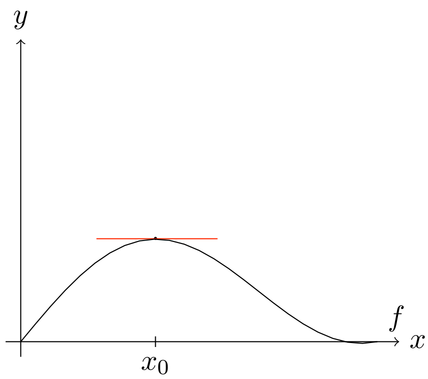
If $f : [a, b] \to \mathbf{R}$ is continuous by the extreme value theorem, it has a minimum and a maximum. How do we find them?

What we learnt in the last section suggested that we find all critical points.

Definition 7.2.1 *A point c is a critical point of f if either $f'(c) = 0$ or $f'(c)$ does not exist.*

To find the (global) maximum and minimum of f, evaluate the values of f at a, b and at all the critical points. Select the largest value.

7.3 Rolle's Theorem



Theorem 7.3.1 (Rolle's Theorem) *Suppose that*

1. *f is continuous on $[a, b]$*
2. *f is differentiable on (a, b)*
3. *$f(a) = f(b)$.*

Then there is a point $x_0 \in (a, b)$ such that

$$f'(x_0) = 0.$$

Proof If f is constant, Rolle's Theorem holds.

Otherwise by the Extreme Value Theorem, there are points $\underline{x}, \bar{x} \in [a, b]$ with

$$f(\bar{x}) \le f(x) \le f(\bar{x}), \quad \forall x \in [a, b].$$

Since f is not constant $f(\underline{x}) \neq f(\bar{x})$.

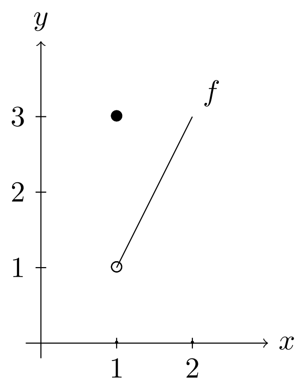
Since $f(a) = f(b)$, one of the point $\bar{x}$ or $\underline{x}$ is in the open interval (a, b). Denote this point by x_0. By the previous lemma $f'(x_0) = 0$. □

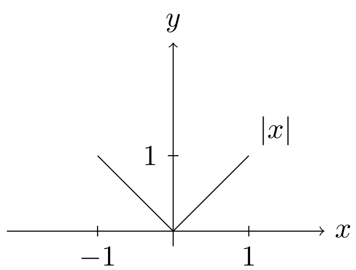
Example 7.3.2 *Discussions on Assumptions:*

1. *Continuity on the closed interval $[a, b]$ is necessary. For example consider $f : [1, 2] \to \mathbf{R}$.*

$$f(x) = \begin{cases} f(x) = 2x - 1, & x \in (1, 2] \\ f(1) = 3, & x = 1. \end{cases}$$

The function f is continuous on $(1, 2]$ and is differentiable on (a, b). Also $f(2) = 3 = f(1)$. But there is no point in $x_0 \in (1, 2)$ such that $f'(x_0) = 0$.

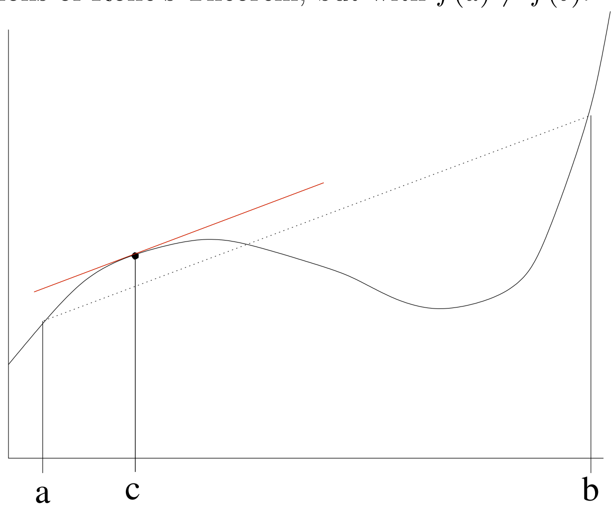




2. *Differentiability is necessary. e.g. take* $f : [-1,1] \to \mathbf{R}$, $f(x) = |x|$. *Then* f *is continuous on* $[-1,1]$, $f(-1) = f(1)$. *But there is no point on* $x_0 \in (-1,1)$ *with* $f'(x_0) = 0$. *This point ought to be* $x = 0$ *but the function fails to be differentiable at* x_0.

7.4 The Mean Value Theorem

Let us now consider the graph of a function f satisfying the regularity conditions of Rolle's Theorem, but with $f(a) \neq f(b)$.



Theorem 7.4.1 (The Mean Value Theorem) *Suppose that* f *is continuous on* $[a,b]$ *and is differentiable on* (a,b), *then there is a point* $c \in (a,b)$ *such that*

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

Note: $\frac{f(b)-f(a)}{b-a}$ is the slope of the chord joining the points $(a, f(a))$ and $(b, f(b))$ on the graph of f. So the theorem says that there is a point $c \in (a,b)$ such that the tangent to the graph at $(c, f(c))$ is parallel to this chord.

Proof The equation of the chord joining $(a, f(a))$ and $(b, f(b))$ is

$$y = \frac{f(b) - f(a)}{b - a}(x - a).$$

Let

$$g(x) = f(x) - \frac{f(b) - f(a)}{b - a}(x - a)$$

Then $g(b) = g(a)$, and g is continuous on $[a, b]$ and differentiable on (a, b). By Rolle's Theorem applied to g on $[a, b]$, there exists $c \in (a, b)$ with $g'(c) = 0$. But

$$g'(c) = f'(c) - \frac{f(b) - f(a)}{b - a}.$$

□

Corollary 7.4.2 *Suppose that f is continuous on $[a, b]$ and is differentiable on (a, b). Suppose that $f'(c) = 0$, for all $c \in (a, b)$. Then f is constant on $[a, b]$.*

Proof Let $x \in (a, b]$. By the Mean Value Theorem on $[a, x]$, there exists $c \in (a, x)$ with

$$\frac{f(x) - f(a)}{x - a} = f'(c).$$

Since $f'(c) = 0$, this means that $f(x) = f(a)$. □

Example 7.4.3 *If $f'(x) = 1$ for all x then $f(x) = x + c$, where c is some constant. To see this, let $g(x) = f(x) - x$. Then $g'(x) = 0$ for all x. Hence $g(x)$ is a constant: $g(x) = g(a) = f(a) - a$. Consequently $f(x) = x + g(x) = f(a) + x - a$.*

Exercise 7.4.4 *If $f'(x) = x$ for all x, what does f look like?*

Exercise 7.4.5 *Show that if f, g are continuous on $[a, b]$ and differentiable on (a, b), and $f' = g'$ on (a, b), then $f = g + C$ where $C = f(a) - g(a)$.*

Remark 7.4.6 When f' is continuous, we can deduce the Mean Value Theorem from the Intermediate Value Theorem, together with some facts about integration which will be proved in Analysis III.

If f' is continuous on $[a, b]$, it is integrable. Since f' is continuous, by the Extreme Value Theorem, there exist $x_1, x_2 \in [a, b]$ such that $f'(x_1)$ and $f'(x_2)$ are respectively the minimum and the maximum values of f' on the interval $[a, b]$. Hence we have

$$f'(x_1)(b - a) \quad \le \quad \int_a^b f'(x)dx \quad \le \quad f'(x_2)(b - a).$$

Divide by $b - a$ to see that

$$f'(x_1) \quad \le \quad \frac{1}{b - a}\int_a^b f'(x)dx \quad \le \quad f'(x_2).$$

Now the Intermediate Value Theorem, applied to f' on the interval between x_1 and x_2, says there exists c between x_1 and x_2 such that

$$f'(c) = \frac{1}{b-a}\int_a^b f'(x)dx.$$

Finally, the fundamental theorem of calculus, again borrowed from Analysis III, tells us that

$$\int_a^b f'(x) = f(b) - f(a).$$

7.5 Monotonicity and Derivatives

If $f : [a, b] \to \mathbf{R}$ is increasing and differentiable at x_0, then $f'(x_0) \geq 0$. This is because

$$f'(x_0) = \lim_{x \to x_0+} \frac{f(x) - f(x_0)}{x - x_0}.$$

Both numerator and denominator in this expression are non-negative, so the quotient is non-negative also, and hence so is its limit.

Corollary 7.5.1 *Suppose that $f : \mathbf{R} \to \mathbf{R}$ is continuous on $[a, b]$ and differentiable on (a, b).*

1. *If $f'(x) \geq 0$ for all $x \in (a, b)$, then f is non-decreasing on $[a, b]$.*
2. *If $f'(x) > 0$ for all $x \in (a, b)$, then f is increasing on $[a, b]$.*
3. *If $f'(x) \leq 0$ for all $x \in (a, b)$, then f is non-increasing on $[a, b]$.*
4. *If $f'(x) < 0$ for all $x \in (a, b)$, then f is decreasing on $[a, b]$.*

Proof Part 1. Suppose $f'(x) \geq 0$ on (a, b). Take $x, y \in [a, b]$ with $x < y$. By the mean value theorem applied on $[x, y]$, there exists $c \in (x, y)$ such that

$$\frac{f(y) - f(x)}{y - x} = f'(c)$$

and

$$f(y) - f(x) = f'(c)(y - x) \geq 0.$$

Thus $f(y) \geq f(x)$.

Part 2. If $f'(x) > 0$ everywhere then as in part 1, for $x < y$,

$$f(y) - f(x) = f'(c)(y - x) > 0.$$

□

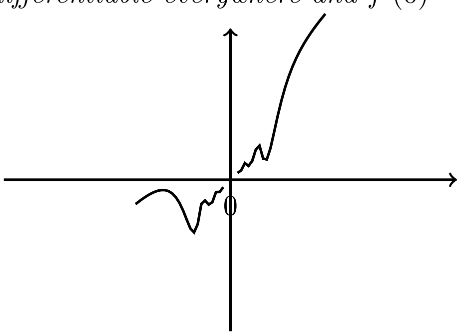
Suppose that $f:(a,b)\to\mathbf{R}$ is differentiable at x_0 and that $f'(x_0)>0$. Does this imply that f is increasing on some neighbourhood of x_0? If you attempt to answer this question by drawing pictures and exploring the possibilities, it is quite likely that you will come to the conclusion that the answer is Yes. But it is not true! We will see this shortly, by means of an example. First, notice that if we assume not only that $f'(x_0)>0$, but also that f is differentiable at all points x in some neighbourhood of x_0, and that f' is *continuous* at x_0, then in this case it *is* true that f is increasing on a neighbourhood of x_0. This is simply because the continuity of f' at x_0 implies that $f'(x)>0$ for all x in some neighbourhood of x_0, and then Part 2 of Corollary 7.5.1 implies that f is increasing on this neighbourhood.

To see that it is NOT true without this extra assumption, consider the following example.

Example 7.5.2 *The function*

$$f(x)=\begin{cases} x+x^2\sin(\frac{1}{x^2}), & x\neq 0, x\in[-1,1] \\ 0, & x=0 \end{cases}.$$

is differentiable everywhere and $f'(0)=1$.



Proof *Since the sum, product, composition and quotient of differentiable functions are differentiable, provided the denominator is not zero, f is differentiable at $x\neq 0$, and $f'(x)=1+2x\sin(\frac{1}{x^2})-\frac{2}{x}\cos(\frac{1}{x^2})$. At $x=0$,*

$$\begin{aligned} f'(0) &= \lim_{x\to 0}\frac{f(x)-f(0)}{x-0}=\lim_{x\to 0}\frac{f(x)}{x} \\ &= \lim_{x\to 0}\frac{x+x^2\sin(\frac{1}{x^2})}{x} \\ &= \lim_{x\to 0}(1+x\sin(\frac{1}{x^2}))=1. \end{aligned}$$

□

Notice, however, that although f is everywhere differentiable, the function $f'(x)$ is not continuous at $x = 0$. This is clear: $f'(x)$ does not tend to a limit as x tends to 0. And in fact, Claim: Even though $f'(0) = 1 > 0$, there is no interval $(-\delta, \delta)$ on which f is increasing.

Proof

$$f'(x) = \begin{cases} 1 + 2x\sin(\frac{1}{x^2}) - \frac{2}{x}\cos(\frac{1}{x^2}), & x \neq 0 \\ 1, & x = 0. \end{cases}$$

Consider the intervals:

$$I_n = \left[\frac{1}{\sqrt{2\pi n + \frac{\pi}{4}}}, \frac{1}{\sqrt{2\pi n}}\right].$$

If $x \in I_n$, $\sqrt{2\pi n} \le \frac{1}{x} \le \sqrt{2\pi n + \frac{\pi}{4}}$ and

$$\cos\frac{1}{x^2} \ge \cos\frac{\pi}{4} = \frac{1}{\sqrt{2}}.$$

Thus if $x \in I_n$

$$f'(x) \le 1 + 2\frac{1}{\sqrt{2\pi n}} - 2\sqrt{2\pi n}\frac{1}{\sqrt{2}} = \frac{\sqrt{\pi n} + \sqrt{2} - 2\pi n}{\sqrt{\pi n}} < 0.$$

Consequently f is decreasing on I_n. As every open neighbourhood of 0 contains an interval I_n, the claim is proved.

□

Even though, as this example shows, if $f'(x_0) > 0$ then we cannot conclude that f is increasing on some neighbourhood of x_0, a weaker conclusion does hold:

Proposition 7.5.3 *If $f : (a, b) \to \mathbf{R}$ and $f'(x_0) > 0$ then there exists $\delta > 0$ such that*

$$\text{if } x_0 - \delta < x < x_0 \quad \text{then} \quad f(x) < f(x_0)$$

and

$$\text{if } x_0 < x < x_0 + \delta \quad \text{then} \quad f(x_0) < f(x).$$

Proof Because

$$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h} > 0,$$

there exists $\delta > 0$ such that if $0 < |h| < \delta$ then

$$\frac{f(x_0 + h) - f(x_0)}{h} > 0.$$

Writing x in place of $x_0 + h$, so that h is replaced by $x - x_0$, this says if $0 < |x - x_0| < \delta$ then

$$\frac{f(x) - f(x_0)}{x - x_0} > 0.$$

When $x_0 - \delta < x < x_0$, the denominator in the last quotient is negative. Since the quotient is positive, the numerator must be negative too.

When $x_0 < x < x_0 + \delta$, the denominator in the last quotient is positive. Since the quotient is positive, the numerator must be positive too. □

Notice that this proposition compares the value of f at the variable point x with its value at the fixed point x_0. A statement that f is increasing would have to compare the values of f at variable points x_1, x_2.

Example 7.5.2 shows that f' need not be continuous. However, it turns out that the kind of discontinuities it may have are rather different from the discontinuities of a function like $g(x) = [x]$. To begin with, even though it may not be continuous, rather surprisingly f' has the "intermediate value property":

Exercise 7.5.4 *Let f be differentiable on (a, b).*

1. *Suppose that $a < c < d < b$, and that $f'(c) < v < f'(d)$. Show that there exists $x_0 \in (c, d)$ such that $f'(x_0) = v$.*

   Hint: let $g(x) = f(x) - vx$. Then $g'(c) < 0 < g'(d)$. By Proposition 7.5.3, $g(x) < g(c)$ for x slightly greater than c, and $g(x) < g(d)$ for x slightly less than d. It follows that g achieves its minimum value on $[c, d]$ at some point in the interior (c, d).

2. *Suppose that $\lim_{x \to x_0 -} f'(x_0)$ exists. Use the intermediate value property to show that it is equal to $f'(x_0)$. Prove also the analogous statement for $\lim_{x \to x_0 +} f'(x)$. Deduce that if both limits exist then f' is continuous at x_0.*

7.6 Inequalities from the MVT

We wish to use information on f' to deduce information on f. Suppose we have a bound for f' on the interval $[a, x]$. Applying the mean value theorem to the interval $[a, x]$, we have

$$\frac{f(x) - f(a)}{x - a} = f'(c)$$

for some $c \in (a, x)$. From the bound on f' we can deduce something about $f(x)$.

Example 7.6.1 *For any $x > 1$,*

$$1 - \frac{1}{x} < \ln(x) < x - 1.$$

Proof Consider the function ln on $[1, x]$. We will prove later that $\ln : (0, \infty) \to \mathbf{R}$ is differentiable everywhere and $\ln'(x) = \frac{1}{x}$. We already know for any $b > 0$, $\ln : [1, b] \to \mathbf{R}$ is continuous, as it is the inverse of the continuous function $e^x : [0, \ln b] \to [1, b]$.

Fix $x > 1$. By the mean value theorem applied to the function $f(y) = \ln y$ on the interval $[1, x]$, there exists $c \in (1, x)$ such that

$$\frac{\ln(x) - \ln 1}{x - 1} = f'(c) = \frac{1}{c}.$$

Since $\frac{1}{x} < \frac{1}{c} < 1$,

$$\frac{1}{x} < \frac{\ln x}{x - 1} < 1.$$

Multiplying through by $x - 1 > 0$, we see

$$\frac{x - 1}{x} < \ln x < x - 1.$$

□

Example 7.6.2 *Let $f(x) = 1 + 2x + \frac{2}{x}$. Show that $f(x) \le 23.1 + 2x$ on $[0.1, \infty)$.*

Proof First $f'(x) = 2 - \frac{2}{x^2}$. Since $f'(x) < 0$ on $[0.1, 1)$, by Corollary 7.5.1 of the MVT, f decreases on $[0.1, 1]$. So the maximum value of f on $[0.1, 1]$ is $f(0.1) = 23.1$. On $[1, \infty)$, $f(x) \le 1 + 2x + 2 = 3 + 2x$. □

7.7 Asymptotics of f at Infinity

Example 7.7.1 *Let* $f : [1, \infty) \to \mathbf{R}$ *be continuous. Suppose that it is differentiable on* $(1, \infty)$. *If* $\lim_{x \to +\infty} f'(x) = 1$ *there is a constant* K *such that*

$$f(x) \le 2x + K.$$

Proof

1. Since $\lim_{x \to +\infty} f'(x) = 1$, taking $\varepsilon = 1$, there is a real number $A > 1$ such that if $x > A$, $|f'(x) - 1| < 1$. Hence

   $$f'(x) < 2 \qquad \text{if } x > A.$$

   We now split the interval $[1, \infty)$ into $[1, A] \cup (A, \infty)$ and consider f on each interval separately.

2. Case 1: $x \in [1, A]$. By the Extreme Value Theorem, f has an upper bound K on $[1, A]$. If $x \in [1, A]$, $f(x) \le K$. Since $x > 0$,

   $$f(x) \le K + 2x, \qquad x \in [1, A].$$

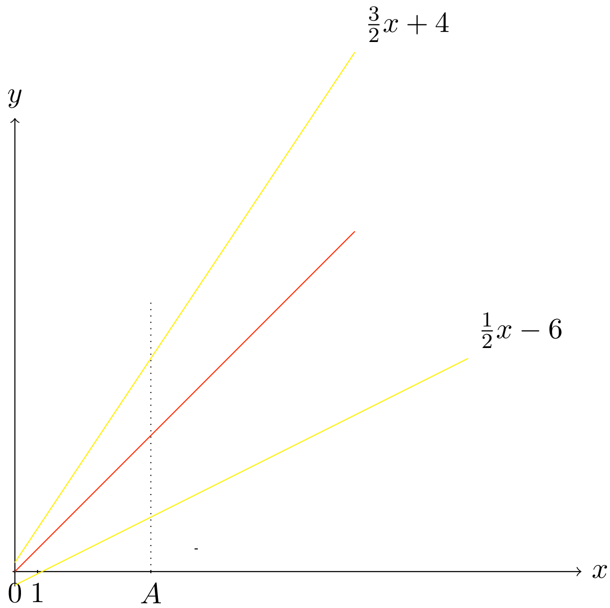
3. Case 2. $x > A$. By the Mean Value Theorem, applied to f on $[A, x]$, there exists $c \in (A, x)$ such that

   $$\frac{f(x) - f(A)}{x - A} = f'(c) < 2,$$

   by part 1). So if $x > A$,

   $$f(x) \le f(A) + 2(x - A) \le f(A) + 2x \le K + 2x.$$

The required conclusion follows. □



The graph of f lies in the wedge-shaped region with the yellow boundaries.

Exercise 7.7.2 *Let* $f(x) = x + \sin(\sqrt{x})$*. Find* $\lim_{x\to+\infty} f'(x)$*.*

Exercise 7.7.3 *In Example 7.7.1, can you find a lower bound for* f*?*

Example 7.7.4 *Let* $f : [1,\infty) \to \mathbf{R}$ *be continuous. Suppose that it is differentiable on* $(1,\infty)$*. If* $\lim_{x\to+\infty} f'(x) = 1$ *then there are real numbers* m *and* M *such that for all* $x \in [1,\infty)$*,*

$$m + \frac{1}{2}x \le f(x) \le M + \frac{3}{2}x.$$

Proof

1. By assumption $\lim_{x\to+\infty} f'(x) = 1$. For $\varepsilon = \frac{1}{2}$, there exists $A > 1$ such that if $x > A$,
$$\frac{1}{2} = 1 - \varepsilon < f'(x) < 1 + \varepsilon = \frac{3}{2}.$$
2. Let $x > A$. By the Mean Value Theorem, applied to f on $[A, x]$, there exists $c \in (A, x)$ such that
$$f'(c) = \frac{f(x) - f(A)}{x - A}.$$

Hence

$$\frac{1}{2} < \frac{f(x) - f(A)}{x - A} < \frac{3}{2}.$$

So if $x > A$,

$$\begin{aligned} f(A) + \frac{1}{2}(x - A) &\leq f(x) \leq & f(A) + \frac{3}{2}(x - A) \\ f(A) - \frac{1}{2}A + \frac{1}{2}x &\leq f(x) \leq & f(A) + \frac{3}{2}x. \end{aligned}$$

3. On the finite interval $[1, A]$, we may apply the Extreme Value Theorem to $f(x) - \frac{1}{2}x$ and to $f(x)$ so there are m, M such that

$$m \leq f(x) - \frac{1}{2}x, \qquad f(x) \leq M.$$

Then if $x \in [1, A]$,

$$\begin{aligned} f(x) &\leq M \leq M + \frac{3}{2}x \\ f(x) &= f(x) - \frac{1}{2}x + \frac{1}{2}x \geq m + \frac{1}{2}x. \end{aligned}$$

4. Since $m \leq f(A) - \frac{1}{2}A$, the required identity also holds if $x > A$.

□

Exercise 7.7.5 *In Example 7.7.4 can you show that there exist real numbers m and M such that for all $x \in [0, \infty)$,*

$$m + 0.99x \leq f(x) \leq 1.01x + M?$$

Exercise 7.7.6 *Let $f : [1, \infty) \to \mathbf{R}$ be continuous. Suppose that it is differentiable on $(1, \infty)$. Suppose that $\lim_{x\to+\infty} f'(x) = k$. Show that for any $\varepsilon > 0$ there are numbers m, M such that*

$$m + (k - \varepsilon)x \leq f(x) \leq M + (k + \varepsilon)x.$$

Comment: If $\lim_{x\to+\infty} f'(x) = k$, the graph of f lies in the wedge formed by the lines $y = M + (k + \varepsilon)x$ and $y = m + (k - \varepsilon)x$. This wedge contains and closely follows the line with slope k. But we have to allow for some oscillation and hence the ε, m and M. In the same way if $|f'(x)| \leq Cx^\alpha$ when x is sufficiently large then $f(x)$ is controlled by $x^{1+\alpha}$ with allowance for the oscillation.

What can we say if $\lim_{x\to\infty} f'(x) = 0$?

Exercise 7.7.7 *(i) Draw the graph of a function f such that*

1. $\lim_{x\to\infty} f'(x) = 0$, *but* $\lim_{x\to\infty} f(x) = \infty$
2. $\lim_{x\to\infty} f(x) = 0$ *but* $\lim_{x\to\infty} f'(x)$ *does not exist.*

(ii) Show that if $\lim_{x\to\infty} f(x)$ *and* $\lim_{x\to\infty} f'(x)$ *both exist and are finite, then* $\lim_{x\to\infty} f'(x) = 0$.

7.8 Continuously Differentiable Functions

When f is differentiable at every point of an interval, we investigate the properties of its derivative $f'(x)$.

Definition 7.8.1 *We say that* $f : [a, b] \to \mathbf{R}$ *is continuously differentiable if it is differentiable on* $[a, b]$ *and* f' *is continuous on* $[a, b]$. *Instead of 'continuously differentiable' we also say that* f *is* C^1 *on* $[a, b]$.

Note that by $f'(a)$ we mean $f'(a+)$ and by $f'(b)$ we mean $f'(b-)$. A similar notion exists for f defined on (a, b), or on (a, ∞), on $(-\infty, b)$, or on $(-\infty, \infty)$.

Example 7.8.2 *The function g defined by* $g(x) = x^2 \sin(1/x^2)$ *for* $x \neq 0$, $g(0) = 0$, *is differentiable on all of* $\mathbf{R}$, *but not* C^1 *— its derivative is not continuous at* 0. *See Example 7.5.2 (which deals with the function* $f(x) = x + g(x)$*) for details.*

Example 7.8.3 *Let* $0 < \alpha$ *and*

$$h(x) = \begin{cases} x^{2+\alpha} \sin(1/x) & x > 0 \\ 0, & x = 0 \end{cases}$$

Claim: The function h is C^1 on $[0, \infty)$.
Proof *We have*

$$h'(0) = \lim_{x\to 0+} \frac{x^{2+\alpha} \sin(1/x)}{x - 0} = \lim_{x\to 0+} x^{1+\alpha} \sin(1/x) = 0,$$

and

$$h'(x) = (2 + \alpha)x^{1+\alpha} \sin(1/x) - x^{\alpha} \cos(1/x)$$

for $x > 0$. *Since*

$$|(2 + \alpha)x^{1+\alpha} \sin(1/x) - x^{\alpha} \cos(1/x)| \leq (2 + \alpha)|x|^{1+\alpha} + |x|^{\alpha},$$

by the sandwich theorem, $\lim_{x\to 0+} h'(x) = 0 = h'(0)$. *Hence h is* C^1. □

Show that h' is not differentiable at $x = 0$, if $\alpha \in (0, 1)$.

Definition 7.8.4 *The set of all continuously differentiable functions on* $[a, b]$ *is denoted by* $C^1([a, b]; \mathbf{R})$ *or by* $C^1([a, b])$ *for short.*

Similarly we may define $C^1((a, b); \mathbf{R})$, $C^1((a, b]; \mathbf{R})$, $C^1([a, b); \mathbf{R})$, $C^1([a, \infty); \mathbf{R})$, $C^1((-\infty, b]; \mathbf{R})$, $C^1((-\infty, \infty); \mathbf{R})$ etc.

7.9 Higher Order Derivatives

Definition 7.9.1 *Let* $f : (a, b) \to \mathbf{R}$ *be differentiable. We say* f *is twice differentiable at* x_0 *if* $f' : (a, b) \to \mathbf{R}$ *is differentiable at* x_0. *We write:*

$$f''(x_0) = (f')'(x_0).$$

It is called the second derivative of f *at* x_0. *It is also denoted by* $f^{(2)}(x_0)$.

Example 7.9.2 *Let*

$$g(x) = \begin{cases} x^3 \sin(1/x) & x \neq 0 \\ 0, & x = 0 \end{cases}$$

Claim: The function g *is twice differentiable on all of* $\mathbf{R}$.
Proof *It is clear that* g *is differentiable on* $\mathbf{R} \setminus \{0\}$. *It follows directly from the defintion of derivative that* g *is differentiable at* 0 *and* $g'(0) = 0$. *So* g *is differentiable everywhere and its derivative is*

$$g'(x) = \begin{cases} 3x^2 \sin(1/x) - x\cos(1/x) & x \neq 0 \\ 0, & x = 0 \end{cases}$$

By the sandwich theorem, $\lim_{x\to 0} 3x^2 \sin(1/x) - x\cos(1/x) = 0 = g'(0)$. *Hence* g *is* C^1.

□

If the derivative of f is also differentiable, we consider the derivative of f'.

Definition 7.9.3 *We say that* f *is* n *times differentiable at* x_0 *if* $f^{(n-1)}$ *is differentiable at* x_0. *The the nth derivative of* f *at* x_0 *is:*

$$f^{(n)}(x_0) = (f^{(n-1)})'(x_0).$$

Definition 7.9.4 *We say that f is C^∞, or smooth on (a,b) if it is n times differentiable for every $n \in \mathbf{N}$.*

Definition 7.9.5 *We say that f is n times continously differentiable on $[a,b]$ if $f^{(n)}$ exists and is continuous on $[a,b]$.*

This means all the lower order derivatives and f itself are continuous, of course - otherwise $f^{(n)}$ would not be defined..

Definition 7.9.6 *The set of all functions which are n times differentiable on (a,b) and such that the derivative $f^{(0)}, f^{(1)}, \dots f^{(n)}$ are continuous functions on (a,b) is denoted by $C^n(a,b)$.*

There is a similar definition with $[a,b]$, or other types of intervals, in place of (a,b).

7.10 Distinguishing Local Minima from Local Maxima

Theorem 7.10.1 *Let $f : (a,b) \to \mathbf{R}$ be differentiable in a neighbourhood of the point x_0. Suppose that $f'(x_0) = 0$, and that f' also is differentiable at x_0. Then if $f''(x_0) > 0$, f has a local minimum at x_0, and if $f''(x_0) < 0$, f has a local maximum at x_0.*

Proof Suppose that $f''(x_0) > 0$. That is, since $f'(x_0) = 0$,

$$\lim_{h\to 0} \frac{f'(x_0+h)}{h} > 0.$$

Hence there exists $\delta > 0$ such that for $0 < |h| < \delta$,

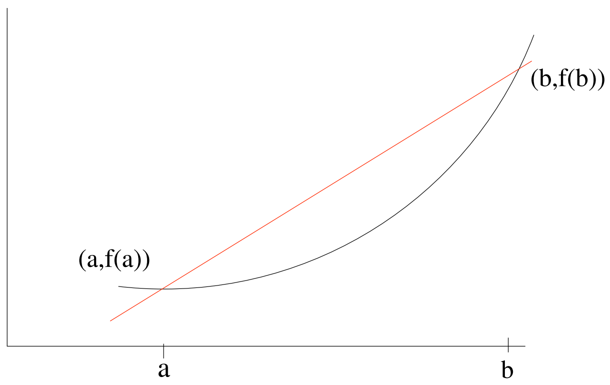
$$\frac{f'(x_0+h)}{h} > 0. \tag{7.10.1}$$

This means that if $0 < h < \delta$, then $f'(x_0+h) > 0$, while if $-\delta < h < 0$, then $f'(x_0+h) < 0$. (We are repeating the argument from the proof of Proposition 7.5.3.) It follows by the MVT that f is *decreasing* in $(x_0-\delta, x_0]$ and *increasing* in $[x_0, x_0+\delta)$. Hence x_0 is a strict local minimum. The case where $f''(x_0) < 0$ follows from this by taking $g = -f$. □

Corollary 7.10.2 *If $f : [a, b] \to \mathbf{R}$ is such that $f''(x) > 0$ for all $x \in (a, b)$ and f is continuous on $[a, b]$ then for all $x \in (a, b)$,*

$$f(x) \le f(a) + \frac{f(b) - f(a)}{b - a}(x - a).$$

In other words, the graph of f between $(a, f(a))$ and $(b, f(b))$ lies below the line segment joining these two points.



Proof Let

$$g(x) = f(a) + \frac{f(b) - f(a)}{b - a}(x - a).$$

The graph of g is the straight line joining the points $(a, f(a))$ and $(b, f(b))$. Let $h(x) = f(x) - g(x)$. Then

1. $h(a) = h(b) = 0$;
2. h is continuous on $[a, b]$ and twice differentiable on (a, b);
3. $h''(x) = f''(x) - g''(x) = f''(x) > 0$ for all $x \in (a, b)$.

By the extreme value theorem there are points $x_1, x_2 \in [a, b]$ such that

$$h(x_1) \le h(x) \le h(x_2), \qquad \forall x \in [a, b].$$

If $x_2 \in (a, b)$, it is a local maximum and hence $h'(x_2) = 0$ (see Lemma 7.1.5). But then by 7.10.1, since $h''(x_2) > 0$, x_2 would also be a strict local minimum! This is absurd. So x_2 cannot be in the interior of the interval $[a, b]$ – it must be an end-point, a or b. In either case $h(x_2) = 0$. We have proved that 0 is the maximum value of h and so $h(x) \le 0$ and $f(x) \le g(x)$. □

Still higher derivatives

What if $f'(x_0) = f''(x_0) = 0$ and $f'''(x_0) \neq 0$? What can we conclude about the behaviour of f in the neighbourhood of x_0? Let us reason as we did in the proof of 7.10.1. Suppose $f'''(x_0) > 0$. Then there exists $\delta > 0$ such that

$$\text{if } \; x_0 - \delta < x < x_0 \quad \text{then} \quad f''(x) < f''(x_0) \text{ and therefore } f''(x) < 0$$

and

$$\text{if } \; x_0 < x < x_0 + \delta \quad \text{then} \quad f''(x_0) < f''(x) \text{ and therefore } 0 < f''(x).$$

It follows that f' is *decreasing* in $[x_0 - \delta, x_0]$ and *increasing* in $[x_0, x_0 + \delta]$. Since $f'(x_0) = 0$, we conclude that f' is positive in $[x_0 - \delta, x_0)$ and also positive in $(0, x_0 + \delta]$. We have proved

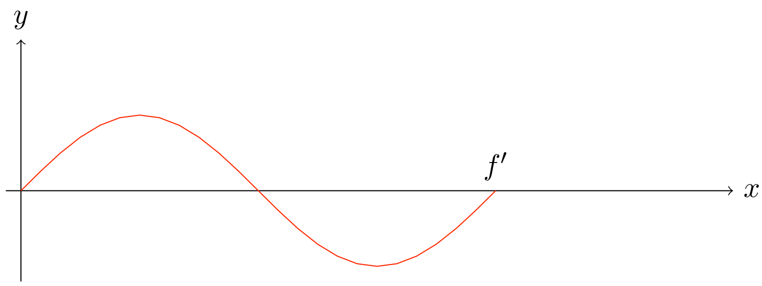
Proposition 7.10.3 *If $f'(x_0) = f''(x_0) = 0$ and $f'''(x_0) > 0$ then there exists $\delta > 0$ such that f is increasing on $[x_0 - \delta, x_0 + \delta]$.* □

Observe that this is exactly the behaviour we see in $f(x) = x^3$ in the neighbourhood of 0.

Exercise 7.10.4

1. Prove that if $f'(x_0) = f''(x_0) = 0$ and $f'''(x_0) < 0$ then there exists $\delta > 0$ such that f is decreasing on $[x_0 - \delta, x_0 + \delta]$.
2. What happens if $f'(x_0) = f''(x_0) = f'''(x_0) = 0$ and $f^{(4)}(x_0) > 0$?
3. What happens if $f'(x_0) = \cdots = f^{(k)}(x_0) = 0$ and $f^{(k+1)}(x_0) \neq 0$?

Exercise 7.10.5 *Sketch the graph of f if f' is the function whose graph is*



shown and $f(0) = 0$.

Chapter 8

Power Series

Definition 8.0.6 *By a power series we mean an expression of the type*

$$\sum_{n=n_0}^{\infty} a_n x^n \quad \textit{or} \quad \sum_{n=n_0}^{\infty} a_n (x - x_0)^n$$

where each a_n is in $\mathbf{R}$ (or, later on in the subject, in the complex domain $\mathbf{C}$). Usually the initial index n_0 is 0 or 1.

By convention, $x^0 = 1$ for all x (including 0) and $0^n = 0$ for all $n \in \mathbf{N}$ with $n > 0$

Example 8.0.7 *1.* $\sum_{n=0}^{\infty} (\frac{x}{3})^n$, $\quad a_n = \frac{1}{3^n} \quad n_0 = 0$

2. $\sum_{n=1}^{\infty} (-1)^n \frac{x^n}{n}$, $\quad n_0 = 0$, $a_n = (-1)^n \frac{1}{n}$, $n_0 = 1$.

3. $\sum_{n=0}^{\infty} \frac{x^n}{n!}$, $\quad a_n = \frac{1}{n!}$, $n_0 = 0$. *By convention* $0! = 1$.

4. $\sum_{n=0}^{\infty} n! x^n$, $\quad a_n = n!$, $\quad n_0 = 0$.

If we substitute a real number in place of x, we obtain an infinite series. For which values of x does the infinite series converge?

Lemma 8.0.8 (The Ratio test) *The series* $\sum_{n=n_0}^{\infty} b_n$

1. *converges absolutely if there is a number* $r < 1$ *such that* $\left|\frac{b_{n+1}}{b_n}\right| < r$ *eventually.*

2. *diverges if there is a number* $r > 1$ *such that* $\left|\frac{b_{n+1}}{b_n}\right| > r$ *eventually.*

Recall that by a statement holding 'eventually' we mean that there exists N_0 such that the statement holds for all $n > N_0$.

Lemma 8.0.9 (The Limiting Ratio test) *The series* $\sum_{n=0}^{\infty} b_n$

1. *converges absolutely if* $\lim_{n\to\infty} \left|\frac{b_{n+1}}{b_n}\right| < 1$;
2. *diverges if* $\lim_{n\to\infty} \left|\frac{b_{n+1}}{b_n}\right| > 1$.

Example 8.0.10 *Consider* $\sum_{n=0}^{\infty} \left(\frac{x}{3}\right)^n$. *Let* $b_n = \left(\frac{x}{3}\right)^n$. *Then*

$$\left|\frac{b_{n+1}}{b_n}\right| = \frac{\left|(\frac{x}{3})^{n+1}\right|}{\left|(\frac{x}{3})^n\right|} = \frac{|x|}{3}, \qquad \begin{cases} < 1, & \text{if } |x| < 3 \\ > 1, & \text{if } |x| > 3. \end{cases}$$

By the ratio test, the power series is absolutely convergent if $|x| < 3$, *and divergent if* $|x| > 3$.
If $x = 3$, *the power series is equal to* $1 + 1 + \dots$, *which is divergent.*
If $x = -3$ *the power series is equal to* $(-1) + 1 + (-1) + 1 + \dots$, *and again is divergent.*

Example 8.0.11 *Consider* $\sum_{n=0}^{\infty} \frac{(-1)^n}{n} x^n$. *Let* $b_n = \frac{(-1)^n}{n} x^n$. *Then*

$$\left|\frac{b_{n+1}}{b_n}\right| = \frac{\left|\frac{(-1)^{n+1}}{n+1} x^{n+1}\right|}{\left|\frac{(-1)^n}{n} x^n\right|} = |x| \left(\frac{n}{n+1}\right) \xrightarrow{n\to\infty} |x|, \qquad \begin{cases} < 1, & \text{if } |x| < 1 \\ 1, & \text{if } |x| > 1. \end{cases}$$

The series is convergent for $|x| < 1$, *and divergent for* $|x| > 1$.
For $x = 1$, *we have* $\sum \frac{(-1)^n}{n}$ *which is convergent.*
For $x = -1$ *we have* $\sum \frac{1}{n}$ *which is divergent.*

Example 8.0.12 *Consider the power series* $\sum_{n=0}^{\infty} \frac{x^n}{n!}$. *We have*

$$\left|\frac{\frac{x^{(n+1)}}{(n+1)!}}{\frac{x^n}{n!}}\right| = \frac{|x|}{n+1} \to 0 < 1$$

The power series converges for all $x \in (-\infty, \infty)$.
$\sum_{n=0}^{\infty} n! x^n$ *converges only at* $x = 0$. *Indeed if* $x \neq 0$,

$$\frac{|(n+1)x^{(n+1)|}}{|n! x^n|} = (n+1)|x| \to \infty$$

as $n \to \infty$.

If for some $x \in \mathbf{R}$, the infinite series $\sum_{n=0}^{\infty} a_n x^n$ converges, we define

$$f(x) = \sum_{n=0}^{\infty} a_n x^n.$$

The domain of f consists of all x such that $\sum_{n=1}^{\infty} a_n x^n$ is convergent. If $x = 0$,

$$\sum_{n=0}^{\infty} a_n 0^n = a_0$$

is convergent, and so $f(0) = a_0$. We wish to identify the set E of points for which the power series is convergent. And is the function so defined on E continuous and even differentiable?

If the power series $\sum_{n=0}^{\infty} a_n x^n$ converges for $x \in E$, then the power series $\sum_{n=0}^{\infty} a_n (x - x_0)^n$ converges for $x \in x_0 + E$. So in what follows we focus on series of the form $\sum_{n=0}^{\infty} a_n x^n$.

8.1 Radius of Convergence

Lemma 8.1.1 *If for some number x_0, $\sum_{n=0}^{\infty} a_n x_0^n$ converges, then $\sum_{n=0}^{\infty} a_n C^n$ converges absolutely for any C with $|C| < |x_0|$.*

Proof As $\sum a_n x_0^n$ is convergent,

$$\lim_{n\to\infty} |a_n||x_0|^n = 0.$$

Convergent sequences are bounded. So there is a number M such that for all n, $|a_n x_0^n| \le M$.

Suppose $|C| < |x_0|$. Let $r = \frac{|C|}{|x_0|}$. Then $r < 1$, and

$$|a_n C^n| = \left| a_n x_0^n \left(\frac{C^n}{x_0^n} \right) \right| \le M r^n.$$

By the comparison theorem, $\sum_{n=0}^{\infty} a_n C^n$ converges absolutely. □

Theorem 8.1.2 *Consider a power series $\sum_{n=0}^{\infty} a_n x^n$. Then one of the following holds:*

(1) The series only converges at $x = 0$.

(2) The series converges for all $x \in (-\infty, \infty)$.

(3) *There is a positive number* $0 < R < \infty$*, such that* $\sum_{n=0}^{\infty} a_n x^n$ *converges for all* x *with* $|x| < R$ *and diverges for all* x *with* $|x| > R$*.*

Proof

- Let
$$S = \left\{ x \;\middle|\; \sum_{n=0}^{\infty} a_n x^n \text{ is convergent} \right\}.$$
Since $0 \in S$, S is not empty. If S is not bounded above, then by Lemma 8.1.1, $\sum_{n=0}^{\infty} a_n x^n$ converges for all x. This is case (2).

- Either there is a point $y_0 \neq 0$ such that $\sum_{n=0}^{\infty} a_n y_0^n$ is convergent, or $R = 0$ and we are in case (1).

- If S is strictly bigger than just $\{0\}$, and is bounded, let
$$R = \sup\{|x| : x \in S\}.$$
  - If x is such that $|x| < R$, then $|x|$ is *not* an upper bound for S (after all, R is the *least* upper bound). Thus there is a number $b \in S$ with $|x| < b$. Because $\sum_{n=0}^{\infty} a_n b^n$ converges, by Lemma 8.1.1 it follows that $\sum_{n=0}^{\infty} a_n x^n$ converges.
  - If $|x| > R$, then $x \notin S$ hence $\sum_{n=0}^{\infty} a_n x^n$ diverges.

□

This number R is called the *radius of convergence* of the power series. In cases (1) and (2) we *define* the radius of convergence to be 0 and ∞ respectively. With this definition, we have

Corollary 8.1.3 *Let* R *be the radius of convergence of the power series* $\sum_{n=n_0}^{\infty} a_n x^n$*. Then*

$$R = \sup\{x \in \mathbf{R} : \sum_{n=n_0}^{\infty} a_n x^n \ \textit{converges}\}.$$

□

When $x = -R$ and $x = R$, we cannot say anything about the convergence of the series $\sum_{n=0}^{\infty} a_n x^n$ without further examination.

In the next section we will develop a method for finding the value of R.

Lemma 8.1.4 *Suppose that $\sum_{n=n_0}^{\infty} a_n x^n$ has radius of convergence R and $\sum_{n=n_0}^{\infty} b_n x^n$ has radius of convergence S, let $c \in \mathbf{R} \setminus \{0\}$, and let T and U be the radii of convergence of $\sum_{n=n_0}^{\infty} (a_n + b_n) x^n$ and $\sum_{n=n_0}^{\infty} c a_n x^n$ respectively. Then $T \geq \min\{R, S\}$ and $U = R$.* □

Proof If $\sum_{n=n_0}^{\infty} a_n x^n$ and $\sum_{n=n_0}^{\infty} b_n x^n$ both converge, then so does $\sum_{n=n_0} a_n x^n + b_n x^n$. For

$$\sum_{n=n_0}^{N} (a_n x^n + b_n x^n) = \sum_{n=n_0}^{N} a_n x^n + \sum_{n=n_0}^{N} b_n x^n$$

and therefore as $N \to \infty$, the sequence on the left of this equation tends to the sum of the limits of the sequences on the right.

It follows by Corollary 8.1.3 that $T \geq \min\{S, T\}$. The proof for $\sum c a_n x^n$ is similar. □

Exercise 8.1.5 *Give an example where $T > \min\{R, S\}$.*

Remark 8.1.6 Complex Analysis studies power series in which both the variable and the coefficients may be complex numbers. We will refer to them as *complex power series*. The proof of Lemma 8.1.1 works unchanged for complex power series, and in consequence there is a version of Theorem 8.1.2 for complex power series: every complex power series $\sum_{n=0}^{\infty} a_n (z - z_0)^n$ has a radius of convergence, a number $R \in \mathbf{R} \cup \{\infty\}$ such that if $|z - z_0| < R$ then the series converges absolutely. The set of z for which the series converges is now a disc centred on z_0, rather than an interval as in the real case. Also different from the real case is the fact (which we do not prove here) that a complex power series with radius of convergence $0 < R < \infty$ necessarily diverges for some z with $|z| = R$.

8.2 The Limit Superior of a Sequence

In this section we show how to find the radius of convergence of a power series. The method is based on the *root test* for the convergence of a series. This says that the series $\sum_{n=0}^{\infty} a_n$ converges if $\lim_{n\to\infty} |a_n|^{1/n} < 1$ and diverges if $\lim_{n\to\infty} |a_n|^{1/n} > 1$. The problem with this test is that in many cases, such as where $a_n = (1 + (-1)^n)$, the sequence $|a_n|^{1/n}$ does not have a limit as $n \to \infty$.

The way round this problem is to replace (a_n) by a related sequence which always does have a limit. Recall that if A is a set then

$$\sup A = \begin{cases} \text{the least upper bound of } A, & \text{if } A \text{ is bounded above} \\ +\infty, & \text{if } A \text{ is not bounded above} \end{cases}$$

$$\inf A = \begin{cases} \text{the greatest lower bound of } A, & \text{if } A \text{ is bounded below} \\ -\infty, & \text{if } A \text{ is not bounded below} \end{cases}$$

Given a sequence (a_n), we define a new sequence (s_n) by

$$\begin{array}{lcl} s_1 & = & \sup\{a_1, a_2, a_3, \ldots\} \\ s_2 & = & \sup\{a_2, a_3, a_4, \ldots\} \\ \cdots & = & \cdots \\ s_n & = & \sup\{a_n, a_{n+1}, \ldots\} \\ \cdots & = & \cdots \end{array}$$

Remarkably, as we will see, the sequence (s_n) always has a limit as $n \to \infty$, provided we allow this limit to be either a real number or $\pm\infty$. The reason for this is that (s_n) is a monotone sequence: $s_{n+1} \leq s_n$ for all n. This inequality is more or less obvious from the definition of s_n: we have

$$\{a_{n+1}, a_{n+2}, \ldots\} \subseteq \{a_n, a_{n+1}, \ldots\}$$

and so

$$\sup\{a_{n+1}, a_{n+2}, \ldots\} \leq \sup\{a_n, a_{n+1}, \ldots\}$$

If this inequality is not obvious to you, draw a picture of two subsets of $\mathbf{R}$, with one contained in the other, and look at where the supremum of each must lie.

Lemma 8.2.1 *Let (a_n) be a sequence of real numbers, and let s_n be defined as above. Then one of the following occurs:*

1. *(a_n) is not bounded above. In this case $s_n = \infty$ for all n, and so*

$$\lim_{n\to\infty} s_n = \infty.$$

2. *(a_n) is bounded above and (s_n) is bounded below. In this case*

$$\lim_{n\to\infty} s_n \in \mathbf{R}.$$

3. (a_n) is bounded above, but (s_n) is not bounded below. In this case

$$\lim_{n\to\infty} s_n = -\infty.$$

Proof If $\{a_1, a_2, \ldots\}$ is not bounded above, then nor is the set $\{a_n, a_{n+1}, \ldots\}$ for any n. Thus $s_n = \infty$ for all n.

If (a_n) is bounded above, then $s_n \in \mathbf{R}$ for all n. Moreover, the sequence (s_n) is non-increasing, as observed above. If it is bounded below, then as it is non-increasing, it must converge to a real limit – in fact it converges to its infimum:

$$\lim_{n\to\infty} s_n = \inf\{s_n : n \in \mathbf{N}\}.$$

If the sequence (s_n) is not bounded below then since it is non-increasing, it converges to $-\infty$. □

Note that if (a_n) is bounded above and below, then so is (s_n), for

$$L \leq a_n \leq U \text{ for all } n$$

implies

$$L \leq s_n \leq U \text{ for all } n.$$

Definition 8.2.2 *Let (a_n) be a sequence. With (s_n) as above, we define*

$$\lim_{n\to\infty} \sup a_n = \lim_{n\to\infty} s_n.$$

Note that the limit on the right exists, by Lemma 8.2.1.

For a *convergent* sequence, this definition is nothing new:

Proposition 8.2.3 *If $\lim_{n\to\infty} a_n = \ell$ then $\lim_{n\to\infty} \sup a_n = \ell$ also.*

Proof Let $\varepsilon > 0$. By the convergence of (a_n), there exists N such that if $n \geq N$ then

$$\ell - \varepsilon < a_n < \ell + \varepsilon.$$

Since $\ell + \varepsilon$ is an upper bound for $\{a_N, a_{N+1}, \ldots\}$, the supremum s_N of this set is less than or equal to $a + \varepsilon$. As $s_n \leq s_N$ for $n \geq N$, we have

$$s_n \leq \ell + \varepsilon$$

for $n \geq N$. And since $s_n \geq a_n$ for all n, if $n \geq N$ we have

$$\ell - \varepsilon < a_n \leq s_n.$$

Thus, if $n \geq N$, then

$$\ell - \varepsilon < s_n \leq \ell + \varepsilon.$$

This shows that $s_n \to \ell$ as $n \to \infty$. □

Example 8.2.4 *1. Let* $b_n = 1 - \frac{1}{n}$. *Then*

$$\sup\left\{1 - \frac{1}{n}, 1 - \frac{1}{n+1}, \dots\right\} = 1.$$

Hence

$$\limsup_{n\to\infty}(1 - \frac{1}{n}) = \lim_{n\to\infty} 1 = 1.$$

2. Let $b_n = 1 + \frac{1}{n}$. *Then* $\sup\{1 + \frac{1}{n}, 1 + \frac{1}{n+1}, \dots\} = 1 + \frac{1}{n}$. *Hence*

$$\limsup_{n\to\infty}(1+\frac{1}{n}) = \lim_{n\to\infty} \sup\left\{1 + \frac{1}{n}, 1 + \frac{1}{n+1}, \dots\right\} = \lim_{n\to\infty}(1+\frac{1}{n}) = 1.$$

Exercises 8.2.5 1. Show that if $b \in \mathbf{R}_{\geq 0}$ then $\limsup ba_n = b \limsup a_n$. What can you say if $b < 0$?

2. Suppose $\ell = \limsup_{n\to\infty} b_n$ is finite, and let $\varepsilon > 0$. Show

   (a) There is a natural number N_ε such that if $n > N$ then $b_n < \ell + \varepsilon$. That is, b_n is bounded above by $\ell + \varepsilon$, eventually.

   (b) For any natural number N, there exists $n > N$ s.t. $b_n > \ell - \varepsilon$. That is, there is a sub-sequence b_{n_k} with $b_{n_k} > \ell - \varepsilon$.

   Deduce that if $\ell = \limsup_{n\to\infty} b_n$ there is a sub-sequence b_{n_k} such that $\lim_{k\to\infty} b_{n_k} = \ell$.

3. We say that ℓ is a *limit point* of the sequence (x_n) if there is a sub-sequence (x_{n_k}) tending to ℓ. Let $\mathcal{L}((x_n))$ be the set of limit points of (x_n). Show that

$$\limsup_{n\to\infty} b_n = \max \mathcal{L}((x_n))$$

$$\liminf_{n\to\infty} b_n = \min \mathcal{L}((x_n)).$$

4. Suppose that (b_n) is a sequence tending to $b \in \mathbf{R}$, and (x_n) is another sequence, not necessarily convergent. Show that

$$\limsup(b_n x_n) = b \limsup x_n.$$

8.3 Hadamard's Test

Theorem 8.3.1 *(Cauchy's root test) The infinite sum* $\sum_{n=1}^{\infty} z_n$ *is*

1. *convergent if* $\limsup_{n\to\infty} |z_n|^{\frac{1}{n}} < 1$.
2. *divergent if* $\limsup_{n\to\infty} |z_n|^{\frac{1}{n}} > 1$.

Proof Let

$$a = \limsup_{n\to\infty} |z_n|^{\frac{1}{n}}.$$

1. Suppose that $a < 1$. Choose $\varepsilon > 0$ so that $a+\varepsilon < 1$. Then there exists N_0 such that whenever $n > N_0$, $|z_n| < (a+\varepsilon)^n$. As $\sum_{n=1}^{\infty}(a+\epsilon)^n$ is convergent, $\sum_{n=N_0}^{\infty} |z_n|$ is convergent by comparison. And

$$\sum_{n=1}^{\infty} |z_n| = \sum_{n=1}^{N_0-1} |z_n| + \sum_{n=N_0}^{\infty} |z_n|$$

is convergent.

2. Suppose that $a > 1$. Choose ε so that $a - \varepsilon > 1$. Then for infinitely many n, $|z_n| > (a-\varepsilon)^n > 1$. Consequently $|z_n| \not\to 0$ and hence $\sum_{n=1}^{\infty} z_n$ is divergent.

□

Theorem 8.3.2 *(Hadamard's Test) Let* $\sum_{n=0}^{\infty} a_n x^n$ *be a power series.*

1. *If* $\limsup_{n\to\infty} |a_n|^{\frac{1}{n}} = \infty$, *the power series converges only at* $x = 0$.
2. *If* $\limsup_{n\to\infty} |a_n|^{\frac{1}{n}} = 0$, *the power series converges for all* $x \in (-\infty, \infty)$.
3. *Let* $r = \limsup_{n\to\infty} |a_n|^{\frac{1}{n}}$. *If* $0 < r < \infty$, *the radius of convergence* R *of the power series is equal to* $1/r$.

Proof We only prove case 3. If $|x| < R$, $\limsup_{n\to\infty} |a_n x^n|^{\frac{1}{n}} = r|x| < 1$ and by Cauchy's root test the series $\sum_{n=0}^{\infty} a_n x^n$ converges.

Suppose that $|x| > R$, then $\limsup_{n\to\infty} |a_n x^n|^{\frac{1}{n}} = r|x| > 1$, and by Cauchy's root test, the series $\sum_{n=0}^{\infty} a_n x^n$ diverges.

□

Example 8.3.3 *Consider* $\sum_{n=0}^{\infty} \frac{(-3)^n x^n}{\sqrt{n+1}}$. *We have*

$$\lim_{n\to\infty} \left(\frac{|(-3)^n|}{\sqrt{n+1}} \right)^{\frac{1}{n}} = 3 \lim_{n\to\infty} \frac{1}{(n+1)^{\frac{1}{2n}}} = 3.$$

So $R = \frac{1}{3}$. *If* $x = -1/3$, *we have a divergent sequence. The interval of convergence is* $x = (-1/3, 1/3]$.

Example 8.3.4 *Consider* $\sum_{n=1}^{\infty} a_n x^n$ *where* $a_{2k} = 2^{2k}$ *and* $a_{2k+1} = 5(3^{2k+1})$. *This means*

$$a_n = \begin{cases} 2^n, & n \text{ is even} \\ 5(3^n), & n \text{ is odd} \end{cases}$$

Thus

$$|a_n|^{\frac{1}{n}} = \begin{cases} 2, & n \text{ is even} \\ 3(5^{\frac{1}{n}}), & n \text{ is odd} \end{cases},$$

which has two limit points: 2 *and* 3.

$$\limsup |a_n|^{\frac{1}{n}} = 3.$$

and $R = 1/3$. *For* $x = \frac{1}{3}$, $x = -\frac{1}{3}$, $|a_n x^n| \neq 0$ *and does not converge to* 0 *(check). So the interval of convergence is* $(-1/3, 1/3)$.

Example 8.3.5 *Consider* $\sum_{n=1}^{\infty} a_n x^n$ *where* $a_{2k} = (\frac{1}{5})^{2k}$ *and* $a_{2k+1} = \left(-\frac{1}{3}\right)^{2k+1}$.

$$\limsup |a_n|^{\frac{1}{n}} = 1/3.$$

So $R = 3$. *For* $x = 3$, $x = -3$, $|a_n x^n| \geq 1$ *and does not tend to* 0. *So the interval of convergence is* $(-3, 3)$.

Example 8.3.6 *Consider the series* $\sum_{k=1}^{\infty} k^2 x^{k^2}$. *We have*

$$a_n = \begin{cases} n, & \text{if } n \text{ is the square of a natural number} \\ 0, & \text{otherwise.} \end{cases}$$

Thus $\limsup |a_n|^{\frac{1}{n}} = 1$ *and* $R = 1$. *Does the series converge when* $x = \pm 1$*?*

8.4 Term by Term Differentiation

If $f_N(x) = \sum_{n=0}^N a_n x^n = a_0 + a_1 x + \cdots + a_N x^N$ then

$$f_N'(x) = \sum_{n=0}^N n a_n x^{n-1}, \qquad f_N''(x) = \sum_{n=0}^N n(n-1) a_n x^{n-2}.$$

Does it hold, if $f(x) = \sum_{n=0}^\infty a_n x^n$, that

$$\begin{aligned} f'(x) &\equiv \frac{d}{dx}\left(\sum_{n=0}^\infty a_n x^n\right) = \sum_{n=0}^\infty n a_n x^{n-1} \\ f''(x) &\equiv \frac{d^2}{dx^2}\left(\sum_{n=0}^\infty a_n x^n\right) = \sum_{n=0}^\infty n(n-1) a_n x^{n-2} \quad ? \end{aligned}$$

These equalities hold if it is true that the derivative of a power series is the sum of the derivatives of its terms; that is, if

$$\frac{d}{dx}\sum_{n=0}^\infty a_n x^n = \sum_{n=0}^\infty \frac{d}{dx} a_n x^n.$$

In other words, we are asking whether it is correct to differentiate a power series *term by term*. As preliminary question, we ask: if R is the radius of convergence of $\sum_{n=0}^\infty a_n x^n$, what are the radii of convergence of the power series $\sum_{n=0}^\infty n a_n x^{n-1}$ and $\sum_{n=0}^\infty n(n-1) a_n x^{n-2}$?

Lemma 8.4.1 *The power series $\sum_{n=0}^\infty a_n x^n$, $\sum_{n=0}^\infty n a_n x^{n-1}$ and $\sum_{n=0}^\infty n(n-1) a_n x^{n-2}$ have the same radius of convergence.*

Proof Let R_1, R_2, R_3 be respectively the radii of convergence of the first, the second and the third series. Recall Hadamard's Theorem: if $\ell = \limsup_{n\to\infty} |a_n|^{\frac{1}{n}}$, then $R_1 = \frac{1}{\ell}$. We interpret $\frac{1}{0}$ as ∞ and $\frac{1}{\infty}$ as 0 to cover all three cases for the radius of convergence.

Observe that

$$x\left(\sum_{n=0}^\infty n a_n x^{n-1}\right) = \sum_{n=0}^\infty n a_n x^n$$

so $\sum_{n=0}^\infty n a_n x^{n-1}$ and $\sum_{n=0}^\infty n a_n x^n$ have the same radius of convergence R_2. By Hadamard's Theorem

$$R_2 = \frac{1}{\limsup_{n\to\infty} |n a_n|^{\frac{1}{n}}}$$

and note that

$$\limsup_{n\to\infty} |na_n|^{\frac{1}{n}} = \lim_{n\to\infty} n^{\frac{1}{n}} \limsup_{n\to\infty} |a_n|^{\frac{1}{n}} = \ell.$$

The last step follows from the fact that $\lim_{n\to\infty} n^{\frac{1}{n}} = 1$ and from Exercise 8.2.5(4). So $R_1 = R_2$.

$\square$

Exercise: show that $R_3 = R_1$. You may use the fact that

$$\lim_{n\to\infty} (n-1)^{\frac{1}{n}} = 1.$$

Below we prove that a function defined by a power series can be differentiated term by term, and in the process show that f is differentiable on $(-R, R)$. From this theorem it follows that f is continuous.

Theorem 8.4.2 *(Term by Term Differentiation) Let R be the radius of convergence of the power series $\sum_{n=0}^{\infty} a_n x^n$. Let $f : (-R, R) \to \mathbf{R}$ be the function defined by this power series,*

$$f(x) = \sum_{n=0}^{\infty} a_n x^n.$$

Then f is differentiable at every point of $(-R, R)$ and

$$f'(x) = \sum_{n=0}^{\infty} n a_n x^{n-1}.$$

Proof Let $x_0 \in (-R, R)$. We show that f is differentiable at x_0 and $f'(x_0) = \sum_{n=0}^{\infty} n a_n x_0^{n-1}$. We wish to show that for any $\varepsilon > 0$ there exists $\delta > 0$ such that if $|x - x_0| < \delta$,

$$\left| \frac{|f(x) - f(x_0)}{x - x_0} - f'(x_0) \right| < \varepsilon. \tag{8.4.1}$$

We simplify the left hand side:

$$\begin{aligned}
&\left|\frac{|f(x)-f(x_0)}{x-x_0}-f'(x_0)\right| \\
=&\left|\frac{\sum_{n=0}^\infty a_n x^n-\sum_{n=0}^\infty a_n x_0^n}{x-x_0}-\sum_{n=0}^\infty n a_n x_0^{n-1}\right| \\
=&\left|\sum_{n=1}^\infty \frac{a_n(x^n-x_0^n)}{x-x_0}-\sum_{n=1}^\infty n a_n x_0^{n-1}\right| \\
=&\left|\sum_{n=2}^\infty a_n\left[\frac{x^n-x_0^n}{x-x_0}-n x_0^{n-1}\right]\right|
\end{aligned}$$

$$\leq \sum_{n=2}^\infty |a_n|\left|\frac{x^n-x_0^n}{x-x_0}-n x_0^{n-1}\right|. \tag{8.4.2}$$

To prove (8.4.1), we thus need to control each term $\left|\frac{x^n-x_0^n}{x-x_0}-nx_0^{n-1}\right|$. We do this by means of the following lemma.

Note that if $x\in(a,b)$ then $|x|\leq \max(a,b)$; similarly

$$\max\{|x| : x\in[x_0-\delta, x_0+\delta]\} = \max\{|x_0-\delta|, |x_0+\delta|\}.$$

Lemma 8.4.3 *Let $x_0\in\mathbf{R}$ and let $n\in\mathbf{N}$. If $\delta>0$ and $0<|x-x_0|<\delta$, then*

$$\left|\frac{x^n-x_0^n}{x-x_0}-nx_0^{n-1}\right|\leq n(n-1)\rho^{n-2}|x-x_0|,$$

where $\rho=\max(|x_0-\delta|, |x_0+\delta|)$.

Proof We assume that $x>x_0$. The case that $x<x_0$ can be proved analogously. Apply the Mean Value Theorem to $f_n(x)=x^n$, on $[x_0,x]$: there exists $c_n\in(x_0,x)$ such that

$$\frac{f_n(x)-f_n(x_0)}{x-x_0}=f_n'(c_n).$$

Since $f_n'(x)=nx^{n-1}$ this becomes

$$\frac{x^n-x_0^n}{x-x_0}=nc_n^{n-1}.$$

It follows that

$$\begin{aligned}\left|\frac{x^n - x_0^n}{x - x_0} - nx_0^{n-1}\right| &= |nc_n^{n-1} - nx_0^{n-1}| \\ &= n|c_n^{n-1} - x_0^{n-1}|.\end{aligned}$$

Apply the mean value theorem again, this time to $f_{n-1}(x) = x^{n-1}$ on $[x_0, c_n]$: there exists $\xi_n \in (x_0, c_n)$ such that

$$\frac{f_{n-1}(x) - f_{n-1}(x_0)}{x - x_0} = f'_{n-1}(\xi_n).$$

Since $f'_{n-1}(x) = \frac{d}{dx}(x^{n-1}) = (n-1)x^{n-2}$,

$$\frac{c_n^{n-1} - x_0^{n-1}}{c_n - x_0} = (n-1)(\xi_n)^{n-2}$$

and

$$|c_n^{n-1} - x_0^{n-1}| = (n-1)|\xi_n|^{n-2}|c_n - x_0|.$$

We see that

$$\begin{aligned}\left|\frac{x^n - x_0^n}{x - x_0} - nx_0^{n-1}\right| &= n|c_n^{n-1} - x_0^{n-1}| \\ &= n(n-1)|\xi_n|^{n-2}|c_n - x_0|.\end{aligned}$$

Since $\xi_n \in (x_0, c_n) \subset (x_0, x)$, $|\xi_n| \leq \rho$ and $|c_n - x_0| \leq |x - x_0|$ we finally obtain that

$$\left|\frac{x^n - x_0^n}{x - x_0} - nx_0^{n-1}\right| \leq n(n-1)\rho^{n-2}|x - x_0|.$$

□

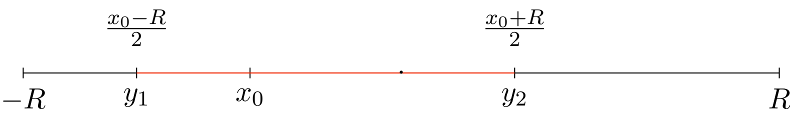
Proof of 8.4.2 (continued): Applying the Lemma to each of the terms in (8.4.2), we get

$$\left|\frac{f(x) - f(x_0)}{x - x_0} - f'(x_0)\right| \leq \left[\sum_{n=2}^{\infty} n(n-1)|a_n|\rho^{n-2}\right]|x - x_0|.$$

We need to chose δ so that if $\rho = \max\{|x| : x \in [x_0 - \delta, x_0 + \delta]\}$, then the power series $\sum_{n=2}^{\infty} n(n-1)|a_n|\rho^{n-2}$ converges. Since, by Lemma 8.4.1, this power series has radius of convergence R, it converges if $\rho < R$.

We must now choose δ. We use the midpoint y_1 of $[-R, x_0]$ and the midpoint y_2 of $[x_0, R]$. We then choose a sub-interval

$$(x_0 - \delta_0, x_0 + \delta_0) \subset (y_1, y_2).$$



Let

$$\delta_0 := \min\{\frac{1}{2}|R - x_0|, \frac{1}{2}|x_0 + R|\}$$

and

$$\rho := \max\{\frac{1}{2}|R - x_0|, \frac{1}{2}|x_0 + R|\}.$$

Then $\rho < R$, as required.

Since the radius of convergence of $\sum_{n=2}^{\infty} n(n-1)a_n x^{n-2}$ is also R, and $\rho < R$,

$$\sum_{n=2}^{\infty} n(n-1)a_n \rho^{n-2}$$

is absolutely convergent. Let

$$A := \sum_{n=2}^{\infty} n(n-1)|a_n|\rho^{n-2} < \infty.$$

Finally, let $\delta = \min\{\frac{\varepsilon}{A}, \delta_0\}$. Then if $0 < |x - x_0| < \delta$,

$$\left|\frac{f(x) - f(x_0)}{x - x_0} - f'(x_0)\right| < A|x - x_0| \le \varepsilon.$$

□

Since the derivative of f is a power series with the same radius of convergence, we apply Theorem 8.4.2 to f' to see that f is twice differentiable with the second derivative again a power series with the same radius of convergence. Iterating this procedure we obtain the following corollary.

Corollary 8.4.4 *Let* $f : (-R, R) \to \mathbf{R}$,

$$f(x) = \sum_{n=0}^{\infty} a_n x^n,$$

where R is the radius of convergence of the power series. Then f is in $C^\infty(-R, R)$.

Example 8.4.5 *If $x \in (-1, 1)$, evaluate $\sum_{n=0}^{\infty} nx^n$.*

Solution. The power series $\sum_{n=0}^{\infty} x^n$ has $R = 1$. But if $|x| < 1$ the geometric series has a limit and

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}.$$

By term by term differentiation, for $x \in (-1, 1)$,

$$\frac{d}{dx}\left(\sum_{n=0}^{\infty} x^n\right) = \sum_{n=1}^{\infty} nx^{n-1}.$$

Hence

$$\begin{aligned}\sum_{n=0}^{\infty} nx^n &= x\frac{d}{dx}\left(\sum_{n=0}^{\infty} x^n\right)\\ &= x\frac{d}{dx}\left(\frac{1}{1-x}\right) = \frac{x}{(1-x)^2}.\end{aligned}$$

Chapter 9

Classical Functions of Analysis

The following power series have radius of convergence R equal to ∞.

$$\sum_{n=0}^{\infty} \frac{x^n}{n!}, \qquad \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}, \qquad \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}.$$

You can check this yourself using Theorem 8.3.2.

Definition 9.0.6 *1. We define the exponential function* $\exp : \mathbf{R} \to \mathbf{R}$ *by*

$$\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots, \qquad \forall x \in \mathbf{R}.$$

2. We define the sine function $\sin : \mathbf{R} \to \mathbf{R}$ *by*

$$\begin{aligned} \sin x &= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \ldots \\ &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1}. \end{aligned}$$

3. We define the cosine function $\cos : \mathbf{R} \to \mathbf{R}$ *by*

$$\begin{aligned} \cos x &= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \ldots \\ &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n} \end{aligned}$$

9.1 The Exponential and the Natural Logarithm Function

Consider the exponential function $\exp : \mathbf{R} \to \mathbf{R}$,

$$\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots, \qquad \forall x \in \mathbf{R}.$$

Note that $\exp(0) = 1$. By the term-by-term differentiation theorem 8.4.2, $\frac{d}{dx}\exp(x) = \exp(x)$, and so exp is infinitely differentiable.

Proposition 9.1.1 *For all $x, y \in \mathbf{R}$,*

$$\begin{aligned} \exp(x+y) &= \exp(x)\exp(y), \\ \exp(-x) &= \frac{1}{\exp(x)}. \end{aligned}$$

Proof For any $y \in \mathbf{R}$ considered as a fixed number, let

$$f(x) = \exp(x+y)\exp(-x).$$

Then

$$\begin{aligned} f'(x) &= \frac{d}{dx}\exp(x+y)\exp(-x) + \exp(x+y)\frac{d}{dx}\exp(-x) \\ &= \exp(x+y)\exp(-x) + \exp(x+y)[-\exp(-x)] \\ &= 0. \end{aligned}$$

By the corollary to the Mean Value Theorem, $f(x)$ is a constant. Since $f(0) = \exp(y)$, $f(x) = \exp(y)$, i.e.

$$\exp(x+y)\exp(-x) = \exp(y).$$

Take $y = 0$, we have

$$\exp(x+0)\exp(-x) = \exp(0) = 1.$$

So

$$\exp(-x) = \frac{1}{\exp(x)}$$

and

$$\exp(x+y) = \exp(y)\frac{1}{\exp(-x)} = \exp(y)\exp(x).$$

□

A neat argument, no?

Proposition 9.1.2 exp *is the only solution to the ordinary differential equation (ODE)*

$$\begin{cases} f'(x) & = f(x) \\ f(0) & = 1. \end{cases}$$

Proof Since $\frac{d}{dx}\exp(x) = \exp(x)$ and $\exp(0) = 1$, the exponential function is *one* solution of the ODE. Let $f(x)$ be *any* solution. Define $g(x) = f(x)\exp(-x)$.Then

$$\begin{aligned} g'(x) &= f'(x)\exp(-x) + f(x)\frac{d}{dx}[\exp(-x)] \\ &= f(x)\exp(-x) - f(x)\exp(-x) \\ &= 0. \end{aligned}$$

Hence for all x, $g(x) = g(0) = f(0)\exp(0) = 1 \cdot 1 = 1$. Thus

$$f(x)\exp(-x) = 1$$

and any solution $f(x)$ must be equal to $\exp(x)$. □

It is obvious from the power series that $\exp(x) > 0$ for all $x \geq 0$. Since $\exp(-x) = 1/\exp(x)$, it follows that $\exp(x) > 0$ for all $x \in \mathbf{R}$.

Exercise 9.1.3
1. *Prove that the range of* exp *is all of* $\mathbf{R}_{>0}$. *Hint: If* $\exp(x) > 1$ *for some* x *then* $\exp(x^n)$ *can be made as large, or as small, as you wish, by suitable choice of* $n \in \mathbf{Z}$.
2. *Show that* $\exp : \mathbf{R} \to \mathbf{R}_{>0}$ *is a bijection.*

We would like to say next:

Proposition 9.1.4 *For all* $x \in \mathbf{R}$,

$$\exp(x) = e^x$$

where $e = \sum_{n=0}^{\infty} \frac{1}{n!}$.

But what does " e^x " mean? We all know that " e^2 " means "$e \times e$", but what about " e^π "? Operations of this kind need a definition!

It is easy to extend the simplest definition, that of raising a number to an *integer* power, to define what it means to raise a number to a rational power: first, for each $n \in \mathbf{N} \setminus \{0\}$ we define an "n'th root" function

$$\sqrt[n]{} : \mathbf{R}_{\geq 0} \to \mathbf{R}$$

as the inverse function of the (strictly increasing, infinitely differentiable) function

$$x \mapsto x^n.$$

Then for any $m/n \in \mathbf{Q}$, we set

$$x^{m/n} = \left(\sqrt[n]{x}\right)^m,$$

where we assume, as we may, that $n > 0$. But this does not tell us what e^π is. We could try approximation: choose a sequence of rational numbers x_n tending to π, and define

$$e^\pi = \lim_{n\to\infty} e^{x_n}.$$

We would have to show that the result is independent of the choice of sequence x_n (i.e. depends only on its limit). This can be done. But then proving that the function $f(x) = a^x$ is differentiable, and finding its derivative, are rather hard. There is a much more elegant approach. First, we define $\ln : \mathbf{R}_{>0} \to \mathbf{R}$ as the inverse to the function $\exp : \mathbf{R} \to \mathbf{R}_{>0}$, which we know to be injective since its derivative is everywhere strictly positive, and surjective by Exercise 9.1.3. Then we make the following definition:

Definition 9.1.5 *For any $x \in \mathbf{R}$ and $a \in \mathbf{R}_{>0}$,*

$$a^x := \exp(x \ln a).$$

Before anything else we should check that this agrees with the original definition of a^x where it applies, i.e. where $x \in \mathbf{Q}$. This is easy: because ln is the inverse of exp, and (by Proposition 9.1.1) exp turns addition into multiplication, it follows that ln turns multiplication into addition:

$$\ln(a \times b) = \ln\, a + \ln\, b,$$

from which we easily deduce that $\ln(a^m) = m \ln a$ (for $m \in \mathbf{N}$) and then that $\ln(a^{m/n}) = m/n \ln\, a$ for $m, n \in \mathbf{Z}$. Thus

$$\exp(\frac{m}{n} \ln a) = \exp(\ln(a^{m/n}) = a^{m/n},$$

the last equality because exp and ln are mutually inverse.

We have given a meaning to " e^x ", and we have shown that when $x \in \mathbf{Q}$ this new meaning coincides with the old meaning. Now that Proposition 9.1.4 is meaningful, we will prove it.

Proof In the light of Definition 9.1.5, Proposition 9.1.4 reduces to

$$\exp(x) = \exp(x \ln e). \tag{9.1.1}$$

But $\ln\, e = 1$, since $\exp(1) = e$; thus (9.1.1) is obvious! □

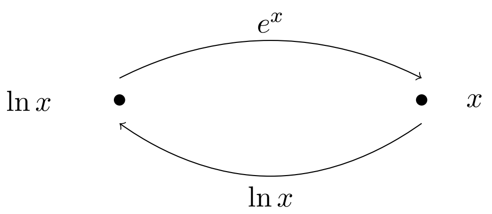
Exercise 9.1.6 *Show that Definition 9.1.5, and the definition of " a^x " sketched in the paragraph preceding Definition 9.1.5 agree with one another.*

Proposition 9.1.7 *The natural logarithm function* $\ln : (0, \infty) \to \mathbf{R}$ *is differentiable and*

$$\ln'(x) = \frac{1}{x}.$$

Proof Since the differentiable bijective map $\exp(x)$ has $\exp'(x) \neq 0$ for all x, the differentiability of its inverse follows from the Inverse Function Theorem. And

$$\ln'(x) = \frac{1}{\exp'(y)}\Big|_{y=\ln(x)} = \frac{1}{\exp(\ln(x))} = \frac{1}{x}.$$



□

Remark: In some computer programs, eg. gnuplot, $x^{\frac{1}{n}}$ is defined as following, $x^{\frac{1}{n}} = \exp(\ln(x^{\frac{1}{n}})) = \exp(\frac{1}{n}\ln x)$. Note that $\ln(x)$ is defined only for $x > 0$. This is the reason that typing in $(-2)^{\frac{1}{3}}$ returns an error message: it is not defined!

9.2 The Sine and Cosine Functions

We have defined

$$\begin{aligned}
\sin x &= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots \\
&= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} \\
\cos x &= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots \\
&= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n}
\end{aligned}$$

Term by term differentiation shows that $\sin'(x) = \cos(x)$ and $\cos'(x) = -\sin(x)$.

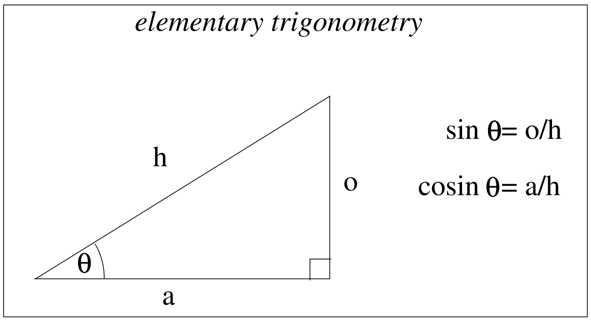
Exercise 9.2.1 *For a fixed $y \in \mathbf{R}$, put*

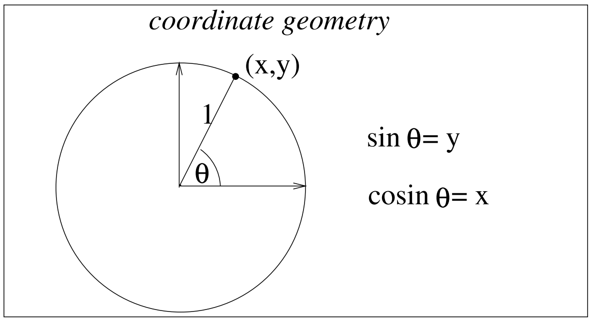
$$f(x) = \sin(x+y) - \sin x \cos y - \cos x \sin y.$$

Compute $f'(x)$ and $f''(x)$. Then let $E(x) = (f(x))^2 + (f'(x))^2$. What is E'? Apply the Mean Value Theorem to E, and hence prove the addition formulae

$$\begin{aligned} \sin(x+y) &= \sin x \cos y + \cos x \sin y \\ \cos(x+y) &= \cos x \cos y - \sin x \sin y. \end{aligned}$$

The sine and cosine functions defined by means of power series behave very well, and it is gratifying to be able to prove things about them so easily. But what relation do they bear to the sine and cosine functions defined in elementary geometry?





We cannot answer this question now, but will be able to answer it after discussing Taylor's Theorem in the next chapter.

Chapter 10

Taylor's Theorem and Taylor Series

We have seen that some important functions can be defined by means of power series, and that any such function is infinitely differentiable.

If we are given an infinitely differentiable function, does there always exist a power series which converges to it? And if such a power series exists, how can we determine its coefficients?

We approach an answer to the question by first looking at polynomials, i.e. at *finite* power series, and proving *Taylor's Theorem.* This is concerned with the following problem. Suppose that the function f is n times differentiable in the interval (a, b), and that $x_0 \in (a, b)$. If we know the values of f and its first n derivatives at x_0, how much can we say about the behaviour of f in the rest of (a, b)? It turns out (and is easy to prove) that there is a unique polynomial P of degree n such that $f(x_0) = P(x_0)$, $f'(x_0) = P'(x_0), \ldots, f^{(n)}(x_0) = P^{(n)}(x_0)$. Taylor's theorem gives a way of estimating the difference between $f(x)$ and $P(x)$ at other points $x \in (a, b)$.

Proposition 10.0.2 *Let $f : (a, b) \to \mathbf{R}$ be n times differentiable and let $x_0 \in (a, b)$. Then there is a unique polynomial of degree n, $P_n(x)$, such that*

$$f(x_0) = P_n(x_0), f'(x_0) = P_n'(x_0), \ldots, f^{(n)}(x_0) = P_n^{(n)}(x_0), \qquad (10.0.1)$$

namely

$$\begin{aligned} P_n(x) &= f(x_0) + f'(x_0)(x-x_0) + \frac{f^{(2)}(x_0)}{2!}(x-x_0)^2 + \cdots + \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n \\ &= \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k. \qquad (10.0.2) \end{aligned}$$

□

Proof Evaluate the polynomial P_n at x_0; it is clear that it satisfies $P_n(x_0) = f(x_0)$. Now differentiate and again evaluate at x_0. It is clear that $P_n'(x_0) = f'(x_0)$. By repeatedly differentiating and evaluating at x_0, we see that P_n satisfies (10.0.1). □

The polynomial

$$P_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)(x-x_0)^k}{k!}$$

is called the *degree n Taylor polynomial of f about x_0*.

Warning *$P_n(x)$ depends on the choice of point x_0.* Consider the function $f(x) = \sin x$. Let us determine the polynomial $P_2(x)$ for two different values of x_0.

1. Taking $x_0 = 0$ we get

$$P_2(x) = \sin(0) + \sin'(0)\,x + \frac{\sin''(0)}{2}x^2 = x.$$

2. Taking $x_0 = \pi/2$ we get

$$P_2(x) = \sin(\pi/2) + \sin'(\pi/2)(x-\pi/2) + \frac{\sin''(\pi/2)}{2}(x-\pi/2)x^2 = 1 - \frac{1}{2}(x-\pi/2)^2.$$

Questions:

1. Let $R_n(x) = f(x) - P_n(x)$ be the "error term". It measures how well the polynomial P_n approximates the value of f. How large is the error term? Taylor's Theorem is concerned with estimating the value of the error $R_n(x)$.

2. Is it true that $\lim_{n\to\infty} R_n(x) = 0$? If so, then

$$f(x) = \lim_{n\to\infty} P_n(x) = \lim_{n\to\infty} \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k.$$

and we have a "power series expansion" for f.

3. Does every infinitely differentiable function have a power series expansion?

Definition 10.0.3 *If f is C^∞ on its domain, the infinite series*

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)(x-x_0)^k}{k!}$$

is called the Taylor series for f about x_0, or around x_0.

Does the Taylor series of f about x_0 converge on some neighbourhood of x_0? If so, it defines an infinitely differentiable function on this neighbourhood. Is this function equal to f?

The answers to both questions are yes for some functions and no for some others.

Definition 10.0.4 *If f is C^∞ on its domain (a,b) and $x_0 \in (a,b)$, we say Taylor's formula holds for f near x_0 if for all x in some neighbourhood of x_0,*

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)(x-x_0)^k}{k!}.$$

The following is an example of a C^∞ function whose Taylor series converges, but does not converge to f:

Example 10.0.5 *(Cauchy's Example (1823))*

$$f(x) = \begin{cases} \exp(-\frac{1}{x^2}), & x \neq 0 \\ 0 & x = 0 \end{cases}.$$

It is easy to see that if $x \neq 0$ then $f(x) \neq 0$. Below we sketch a proof that $f^{(k)}(0) = 0$ for all k. The Taylor series for f about 0 is therefore:

$$\sum_{k=0}^{\infty} \frac{0}{k!} x^k = 0 + 0 + 0 + \dots.$$

This Taylor series converges everywhere, obviously, and the function it defines is the zero function. So it does not converge to $f(x)$ unless $x = 0$.

How do we show that $f^{(k)}(0) = 0$ for all k? Let $y = \frac{1}{x}$,

$$f'_+(0) = \lim_{x \to 0+} \frac{\exp(-\frac{1}{x^2})}{x} = \lim_{y \to +\infty} \frac{y}{\exp(y^2)}.$$

Since $\exp(y^2) \geq y^2$, $\lim_{y \to +\infty} \frac{y}{\exp(y^2)} = 0$ and $\lim_{y \to -\infty} \frac{y}{\exp(y^2)} = 0$. It follows that

$$f'_+(0) = f'_-(0) = 0.$$

A similar argument gives the conclusion for $k > 1$. An induction is needed. Details are given in the last exercise in Section C of Assignment 8.

Remark 10.0.6 *By Taylor's Theorem below, if*

$$R_n(x) = \frac{x^{n+1}}{(n+1)!} f^{(n+1)}(\xi_n) \to 0$$

where ξ_n is between 0 and x, Taylor's formula holds. It follows that, for the function above, either $\lim_{n\to\infty} R_n(x)$ does not exist, or, if it exists, it cannot be 0. Convince yourself of this (without rigorous proof) by observing that $Q_{n+1}(y)$ contains a term of the form $(-1)^{n+1}(n+2)!y^{-(n+2)}$. And indeed

$$\frac{x^{n+1}}{(n+1)!}(-1)^{n+1}(n+2)!\xi_n^{-n-2}$$

may not converge to 0 as $n \to \infty$.

Definition 10.0.7 ** A C^∞ function $f : (a, b) \to \mathbf{R}$ is said to be* real analytic *if for each $x_0 \in (a, b)$, its Taylor series about x_0 converges, and converges to f, in some neighbourhood of x_0.*

The preceding example shows that the function e^{-1/x^2} is not real analytic in any interval containing 0, even though it is infinitely differentiable. Complex analytic functions are studied in the third year Complex Analysis course. Surprisingly, *every* complex differentiable function is complex analytic.

Example 10.0.8 *Find the Taylor series for* $\exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!}$ *about $x_0 = 1$. For which x does the Taylor series converge? For which x does Taylor's formula hold?*

Solution. *For all k,* $\exp^{(k)}(1) = e^1 = e$. *So the Taylor series for* exp *about 1 is*

$$\sum_{k=0}^{\infty} \frac{\exp^{(k)}(1)(x-1)^k}{k!} = \sum_{k=0}^{\infty} \frac{e(x-1)^k}{k!} = e\sum_{k=0}^{\infty} \frac{(x-1)^k}{k!}.$$

The radius of convergence of this series is $+\infty$. Hence the Taylor series converges everywhere.

Furthermore, since $\exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!}$*, it follows that*

$$\sum_{k=0}^{\infty} \frac{(x-1)^k}{k!} = \exp(x-1)$$

and thus

$$e\sum_{k=0}^{\infty} \frac{(x-1)^k}{k!} = e\exp(x-1) = \exp(1)\exp(x-1) = \exp(x)$$

and Taylor's formula holds.

Exercise 10.0.9 *Show that* exp *is real analytic on all of* $\mathbf{R}$.

10.1 More on Power Series

Definition 10.1.1 *If $x_0 \in \mathbf{R}$, a formal power series centred at x_0 is an expression of the form:*

$$\sum_{n=0}^{\infty} a_n(x-x_0)^n = a_0 + a_1(x-x_0) + a_2(x-x_0)^2 + a_3(x-x_0)^3 + \dots.$$

Example 10.1.2 *Take $\sum_{n=0}^{\infty} \frac{1}{3^n}(x-2)^n$. If $x=1$,*

$$\sum_{n=0}^{\infty} \frac{1}{3^n}(x-2)^n = \sum_{n=0}^{\infty} \frac{1}{3^n}(1-2)^n = \sum_{n=0}^{\infty}(-1)^n \frac{1}{3^n}$$

is a convergent series.

What we learnt about power series $\sum_{n=0}^{\infty} a_n x^n$ can be applied here.
For example,

1. There is a radius of convergence R such that the series $\sum_{n=0}^{\infty} a_n(x-x_0)^n$ converges when $|x-x_0| < R$, and diverges when $|x-x_0| > R$.

2. R can be determined by the Hadamard Theorem:

$$R = \frac{1}{r}, \qquad r = \limsup_{n\to\infty} |a_n|^{\frac{1}{n}}.$$

3. The open interval of convergence for $\sum_{n=0}^{\infty} a_n(x-x_0)^n$ is

$$(x_0 - R, x_0 + R).$$

The functions

$$f(x) = \sum_{n=0}^{\infty} a_n(x-x_0)^n \qquad \text{and} \qquad g(x) = \sum_{n=0}^{\infty} a_n x^n$$

are related by

$$f(x) = g(x - x_0).$$

4. The term by term differentiation theorem holds for x in (x_0-R, x_0+R) and

$$f'(x) = \sum_{n=0}^{\infty} n a_n (x-x_0)^{n-1} = a_1 + 2a_2(x-x_0) + 3a_3(x-x_0)^2 + \ldots.$$

Example 10.1.3 *Consider* $\sum_{n=0}^{\infty} \frac{1}{3^n}(x-2)^n$.

Then

$$\limsup (\frac{1}{3^n})^{\frac{1}{n}} = \frac{1}{3}, \qquad so \qquad R = 3.$$

The power series converges for all x *with* $|x-2| < 3$.
For $x \in (-1,5)$, $f(x) = \sum_{n=0}^{\infty} \frac{1}{3^n}(x-2)^n$ *is differentiable and*

$$f'(x) = \sum_{n=0}^{\infty} n \frac{1}{3^n}(x-2)^{n-1}.$$

10.2 Uniqueness of Power Series Expansion

Example 10.2.1 *What is the Taylor's series for* $\sin(x) = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!}$ *about* $x = 0$*?*
Solution. $\sin(0) = 0$, $\sin'(0) = \cos 0 = 1$, $\sin^{(2)}(0) = -\sin 0 = 0$, $\cos^{(3)}(0) = -\cos 0 = -1$, *By induction,* $\sin^{(2n)}(0) = 0$, $\sin^{(2n+1)} = (-1)^n$. *Hence the Taylor series of* $\sin(x)$ *at* 0 *is*

$$\sum_{k=0}^{\infty} (-1)^k \frac{1}{(2k+1)!} x^{2k+1}.$$

This series converges everywhere and equals $\sin(x)$.

The following proposition shows that we would have needed no computation to compute the Taylor series of $\sin x$ at 0:

Proposition 10.2.2 (Uniqueness of Power Series expansion) *Let $x_0 \in (a,b)$. Suppose that $f : (a,b) \subset (x_0 - R, x_0 + R) \to \mathbf{R}$ is such that*

$$f(x) = \sum_{n=0}^{\infty} a_n (x - x_0)^n,$$

where R is the radius of convergence. Then

$$a_0 = f(x_0),\ a_1 = f'(x_0), \ldots, a_k = \frac{f^{(k)}(x_0)}{k!}, \ldots.$$

Proof Take $x = x_0$ in

$$f(x) = \sum_{n=0}^{\infty} a_n (x - x_0)^n = a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + \ldots,$$

it is clear that $f(x_0) = a_0$. By term by term differentiation,

$$f'(x) = a_1 + 2a_2(x - x_0) + 3a_3(x - x_0)^2 + \ldots.$$

Take $x = x_0$ to see that $f'(x_0) = a_1$. Iterating this argument, we get

$$f^{(k)}(x) = a_k k! + a_{k+1}(k+1)!(x - x_0) + a_{k+2}\frac{(k+2)!}{2!}(x - x_0)^2 + a_{k+3}\frac{(k+3)!}{3!}(x - x_0)^3 + \ldots.$$

Take $x = x_0$ to see that

$$a_k = \frac{f^{(k)}(x_0)}{k!}.$$

□

Moral of the proposition: if the function is already in the form of a power series centred at x_0, this is the Taylor series of f about x_0.

Example 10.2.3 *What is the Taylor series of $f(x) = x^2 + x$ about $x_0 = 2$?*

Solution 1. $f(2) = 6$, $f'(2) = 5$, $f''(2) = 2$. Since $f'''(x) = 0$ for all x, the Taylor series about $x_0 = 2$ is

$$6 + \frac{5}{1}(x - 2) + \frac{2}{2!}(x - 2) = 6 + 5(x - 2) + (x - 2)^2.$$

Solution 2. On the other hand, we may re-write

$$f(x) = x^2 + x = (x - 2 + 2)^2 + (x - 2 + 2) = (x - 2)^2 + 5(x - 2) + 6.$$

The Taylor series for f at $x_0 = 2$ is $(x - 2)^2 + 5(x - 2) + 6$.

The Taylor series is identical to f and so Taylor's formula holds.

Example 10.2.4 *Find the Taylor series for $f(x) = \ln(x)$ about $x_0 = 1$.*

Solution. If $x > 0$,

$$\begin{aligned}
f(1) &= 0 \\
f'(x) &= \frac{1}{x}, \qquad f'(1) = 1 \\
f''(x) &= -\frac{1}{x^2}, \qquad f''(1) = -1 \\
f^{(3)}(x) &= 2\frac{1}{x^3}, \qquad f'''(1) = -2 \\
&\dots \\
f^{(k)}(x) &= (-1)^{k+1}(k-1)!\frac{1}{x^k}, \qquad f^{(k)}(1) = (-1)^{(k+1)}(k-1)! \\
f^{(k+1)}(x) &= (-1)^{k+2}k!\frac{1}{x^{k+1}}.
\end{aligned}$$

Hence the Taylor series is:

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(1)}{k!}(x-1)^k = \sum_{k=1}^{\infty} \frac{(-1)^{(k+1)}}{k}(x-1)^k = (x-1) - \frac{1}{2}(x-1)^2 + \frac{1}{3}(x-1)^3 - \dots.$$

It has radius of convergence $R = 1$.

Question: Does Taylor's formula hold? i.e. is it true that

$$\ln(x) = \sum_{k=1}^{\infty} \frac{(-1)^{(k+1)}}{k}(x-1)^k?$$

If we write $x - 1 = y$, we are asking whether

$$\ln(1+y) = \sum_{k=1}^{\infty} \frac{(-1)^{(k+1)}}{k} y^k.$$

We will need Taylor's Theorem (which tells us how good is the approximation) to answer this question.

Remark: We could borrow a little from Analysis III, Since $\sum_{k=0}^{\infty}(-x)^k = \frac{1}{1+x}$, term by term integration (which we are not yet able to justify) would give:

$$\sum_{k=0}^{\infty}(-1)^k \frac{x^{k+1}}{k+1} = \ln(1+x).$$

10.3 Taylor's Theorem with remainder in Lagrange form

If f is continuous on $[a, b]$ and differentiable on (a, b), the Mean Value Theorem states that

$$f(b) = f(a) + f'(c)(b - a)$$

for some $c \in (a, b)$. We could treat $f(a) = P_0$, a polynomial of degree 0.

We have seen earlier in Lemma 8.4.3, iterated use of the Mean Value Theorem gives nice estimates. Let us try that.

Let $f : [a, b] \to \mathbf{R}$ be continuously differentiable and twice differentiable on (a, b). Consider

$$f(b) = f(x) + f'(x)(b - x) + error.$$

The quantity $f(x) + f'(x)(b - x)$ describes the value at b of the tangent line of f at x. We now vary x. We guess that the error is of the order $(b - x)^2$. Define

$$g(x) := f(b) - f(x) - f'(x)(b - x) - A(b - x)^2$$

with A a number to be determined so that we can apply Rolle's Theorem to g on the interval $[a, b]$ — in other words, so that $g(a) = g(b)$.

Note that $g(b) = 0$ and $g(a) = f(b) - f(a) - f'(a)(b - a) + A(b - a)^2$. So set

$$A = \frac{f(b) - f(a) - f'(a)(b - a)}{(b - a)^2}.$$

Then also $g(a) = 0$. Applying Rolle's theorem to g we see that there exists a $\xi \in (a, b)$ such that $g'(\xi) = 0$. Since,

$$\begin{aligned} g'(x) &= -f'(x) - [f''(x)(b - x) - f'(x)] + 2\frac{f(b) - f(a) - f'(a)(b - a)}{(b - a)^2}(b - x) \\ &= -f''(x)(b - x) + 2\frac{f(b) - f(a) - f'(a)(b - a)}{(b - a)^2}(b - x). \end{aligned}$$

we see that

$$\frac{1}{2}f''(\xi) = \frac{f(b) - f(a) - f'(a)(b - a)}{(b - a)^2}.$$

This gives

$$f(b) = f(a) + f'(a)(b - a) + \frac{1}{2}f''(\xi)(b - a)^2.$$

The following is a theorem of Lagrange (1797). To prove it we apply MVT to $f^{(n)}$ and hence we need to assume that $f^{(n)}$ satisfies the conditions of the Mean Value Theorem. By "f is n times continuously differentiable on $[x_0, x]$" we mean that its nth derivative is defined and continuous (so that all the earlier derivatives must also be defined, of course). We denote by $C^n([a,b])$ the set of all n times continuously differentiable functions on the interval $[a, b]$.

Theorem 10.3.1 *[Taylor's theorem with Lagrange Remainder Form]*

1. *Let $x > x_0$. Suppose that f is n times continuously differentiable on $[x_0, x]$ and $n+1$ times differentiable on (x_0, x). Then*

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k + R_n(x)$$

where

$$R_n(x) = \frac{(x - x_0)^{n+1}}{(n+1)!} f^{(n+1)}(\xi)$$

some $\xi \in (x_0, x)$.

2. *The conclusion holds also for $x < x_0$, if f is $(n+1)$ times continuously differentiable on $[x, x_0]$ and $n+1$ times differentiable on (x, x_0).*

Proof Let us vary the starting point of the interval and consider $[y, x]$ for any $y \in [x_0, x]$. We will regard x as fixed (it does not move!).

We are interested in the function with variable y:

$$g(y) = f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(y)(x - y)^k}{k!}.$$

In long form,

$$g(y) = f(x) - [f(y) + f'(y)(x-y) + f''(y)(x-y)^2/2 + \cdots + \frac{f^{(n)}(y)(x - y)^n}{n!}].$$

Then $g(x_0) = R_n(x)$ and $g(x) = 0$. Define

$$h(y) = g(y) - A(x - y)^{n+1}$$

where

$$A = \frac{g(x_0)}{(x - x_0)^{n+1}}.$$

We check that h satisfies the conditions of Rolle's Theorem on $[x_0, x]$:

- $h(x_0) = g(x_0) - A(x - x_0)^{n+1} = 0$
- $h(x) = g(x) = 0$.
- h is continuous on $[x_0, x]$ and differentiable on (x_0, x).

Hence, by Rolle's Theorem, there exists some $\xi \in (x_0, x)$ such that $h'(\xi) = 0$.

Now we calculate $h'(y)$. First, by the product rule,

$$\begin{aligned}
g'(y) &= -\frac{d}{dy}\sum_{k=0}^{n}\frac{f^{(k)}(y)(x-y)^k}{k!} \\
&= -\sum_{k=0}^{n}\frac{f^{(k+1)}(y)(x-y)^k}{k!} + \sum_{k=1}^{n}\frac{f^{(k)}(y)k(x-y)^{k-1}}{k!} \\
&= -\sum_{k=0}^{n}\frac{f^{(k+1)}(y)(x-y)^k}{k!} + \sum_{k=1}^{n}\frac{f^{(k)}(y)(x-y)^{k-1}}{(k-1)!} \\
&= -\frac{f^{(n+1)}(y)(x-y)^n}{n!}.
\end{aligned}$$

Next,

$$\frac{d}{dy}\left(A(x-y)^{n+1}\right) = (n+1)A(x-y)^n.$$

Thus,

$$h'(y) = -\frac{f^{(n+1)}(y)(x-y)^n}{n!} - (n+1)A(x-y)^n$$

and the vanishing of $h'(\xi)$ means that

$$\frac{f^{(n+1)}(\xi)(x-\xi)^n}{n!} = \left(\frac{g(x_0)}{(x-x_0)^{n+1}}\right)(n+1)(x-\xi)^n$$

(we have replaced A by $g(x_0)/(x-x_0)^{n+1}$). Since $R_n(x) = g(x_0)$, we deduce that

$$R_n(x) = \frac{(x-x_0)^{(n+1)}}{(n+1)!}f^{(n+1)}(\xi).$$

□

Corollary 10.3.2 *Let $x_0 \in (a, b)$. If f is C^∞ on $[a, b]$ and the derivatives of f are uniformly bounded on (a, b), i.e. there exists some K such that*

$$\text{for all } k \text{ and for all } x \in [a,b], \quad |f^{(n)}(x)| \le K$$

then Taylor's formula holds, i.e. for each $x, x_0 \in [a, b]$

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)(x - x_0)^k}{k!}, \qquad x \in [a, b].$$

In other words f *has a Taylor series expansion at every point of* (a, b)*. In yet other words,* f *is real analytic in* $[a, b]$*.*

Proof For each $x \in [a, b]$,

$$\begin{aligned} |R_n(x)| &= \left| \frac{(x - x_0)^{n+1}}{(n+1)!} f^{n+1}(\xi) \right| \\ &\leq K \frac{|x - x_0|^{n+1}}{(n+1)!} \end{aligned}$$

and this tends to 0 as $n \to \infty$. This means

$$\lim_{n\to\infty} \sum_{k=0}^{n} \frac{f^{(k)}(x_0)(x - x_0)^k}{k!} = f(x)$$

as required. $\square$

10.3.1 Taylor's Theorem in Integral Form

This section is not included in the lectures nor in the exam for this module. The integral form for the remainder term $R_n(x)$ is the best; the other forms (Lagrange's and Cauchy's) follow easily from it. But we need to know something of the Theory of Integration (Analysis III). With what we learnt in school and a little flexibility we should be able to digest the information. Please read on.

Theorem 10.3.3 *(Taylor's Theorem with remainder in integral form) If* f *is* C^{n+1} *on* $[a, x]$*, then*

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k}(x - a)^k + R_n(x)$$

where

$$R_n(x) = \int_a^x \frac{(x - t)^n}{n!} f^{(n+1)}(t)dt.$$

Proof See Spivack, *Calculus*, pages 415-417. $\square$

Lemma 10.3.4 *(Intermediate Value Theorem in Integral Form) If g is continuous on $[a, b]$ and f is positive everywhere and integrable, then for some $c \in (a, b)$,*

$$\int_a^b f(x)g(x)dx = g(c) \int_a^b f(x)dx.$$

□

By the lemma, the integral remainder term in Theorem 10.3.3 can take the folllowing form:

$$\begin{aligned} R_n(x) &= \int_a^x \frac{(x-t)^n}{n!} f^{(n+1)}(t)dt \\ &= f^{(n+1)}(c) \int_a^x \frac{(x-t)^n}{n!} dt \\ &= \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{(n+1)}. \end{aligned}$$

This is the **Lagrange form** of the remainder term $R_n(x)$ in Taylor's Theorem, as shown in 10.3.1 above.

The **Cauchy form** for the remainder $R_n(x)$ is

$$R_n = f^{(n+1)}(c)\frac{(x-c)^n}{n!}(x-a).$$

This can be obtained in the same way:

$$\begin{aligned} R_n(x) &= \int_a^x \frac{(x-t)^n}{n!} f^{(n+1)}(t)dt \\ &= f^{(n+1)}(c)(x-c)^n \int_a^x \frac{1}{n!} dt \\ &= f^{(n+1)}(c)(x-c)^n \frac{x-a}{n!}. \end{aligned}$$

Other integral techniques:

Example 10.3.5 *For $|x| < 1$, by term by term integration (proved in second* *year modules)*

$$\begin{aligned}
\arctan x &= \int_0^x \frac{dy}{1+y^2} = \int_0^x \sum_{n=0}^{\infty}(-y^2)^n dy \\
&= \sum_{n=0}^{\infty}\int_0^x (-y^2)^n dy \\
&= \sum_{n=0}^{\infty}(-1)^n \frac{y^{2n+1}}{2n+1}\Big|_{y=0}^{y=x} \\
&= \sum_{n=0}^{\infty}(-1)^n \frac{x^{2n+1}}{2n+1}
\end{aligned}$$

10.3.2 Trigonometry Again

Now we can answer the question we posed at the end of Section 10. Consider the sine and cosine functions defined using elementary geometry. We have seen in Example 6.0.6 that cosine is the derivative of sine and the derivative of cosine is minus sine. Since both |sine| and |cosine| are bounded by 1 on all of **R**, both sine and cosine satisfy the conditions of Corollary 10.3.2, with $K = 1$. It follows that both sine and cosine are analytic on all of **R**. Power series 2 and 3 of Definition 9.0.6 are the Taylor series of sine and cosine about $x_0 = 0$. Corollary 10.3.2 tells us that they converge to sine and cosine. So the definition by means of power series gives the same function as the definition in elementary geometry. Just in case you think that elementary geometry is for children and that grown-ups prefer power series, try proving, directly from the power series definition, that sine and cosine are periodic with period 2π.

Example 10.3.6 *Compute* $\sin(1)$ *to the precision* 0.001.

Solution: For some $\xi \in (0,1)$,

$$|R_n(1)| = \left|\frac{\sin^{(n+1)}(\xi)}{(n+1)!}1^{n+1}\right| \le \frac{1}{(n+1)!}.$$

We wish to have

$$\frac{1}{(n+1)!} \le 0.001 = \frac{1}{1000}.$$

Take n *such that* $n! \ge 1000$. *Take* $n = 7$, $7! = 5040$ *is sufficient. Then*

$$\sin(1) \sim [x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!}]|_{x=1} = 1 - \frac{1}{3!} + \frac{1}{5!} - \frac{1}{7!}.$$

Now use your calculator if needed.

Example 10.3.7 *Show that* $\lim_{x\to 0}\frac{\cos x-1}{x}=0$.

By Taylor's Theorem $\cos x - 1 = \frac{x^2}{2} + R_3(x)$. *For* $x \to 0$ *we may assume that* $x \in [-1,1]$. *Then for some* $c \in (-1,1)$,

$$|R_3(x)| = |\frac{\cos^{(3)}(c)}{3!}x^3| = |\frac{\sin(c)}{3!}x^3| \leq | x^3|.$$

It follows that

$$\left|\frac{\cos x - 1}{x}\right| \leq \frac{x^2/2 + |x|^3}{|x|} = |x/2| + |x^2| \to 0,$$

as $x \to 0$.

10.3.3 Taylor Series for Exponentials and Logarithms

How do we prove that $\lim_{n\to\infty}\frac{r^n}{n!}=0$ for any $r>0$? Note that

$$\sum_{n=0}^{\infty}\frac{r^n}{n!}$$

converges by the Ratio Test, and so its individual terms $\frac{r^n}{n!}$ must converge to 0.

Example 10.3.8 *We can determine the Taylor series for* e^x *about* $x_0 = 0$ *using only the facts that* $\frac{d}{dx}e^x = e^x$ *and* $e^0 = 1$. *For*

$$\frac{d^2}{dx^2}e^x = \frac{d}{dx}\left(\frac{d}{dx}e^x\right) = \frac{d}{dx}e^x = e^x,$$

and so, inductively, $\frac{d^n}{dx^n}e^x = e^x$ *for all* $n \in \mathbf{N}$. *Hence, evaluating these derivatives at* $x = 0$, *all take the value 1, and the Taylor series is*

$$\sum_{k=0}^{\infty}\frac{x^k}{k!}.$$

The Taylor series has radius of convergence ∞. *For any* $x \in \mathbf{R}$, *there is* $\xi \in (-x, x)$ *s.t.*

$$|R_n(x)| = \left|\frac{x^{n+1}}{(n+1)!}e^{\xi}\right| \leq \max(e^x, e^{-x})|\frac{x^{n+1}}{(n+1)!}| \to 0$$

as $n \to \infty$. *So*

$$e^x = \sum_{k=0^\infty} \frac{x^k}{k!} = \exp(x)$$

for all x.

Example 10.3.9 *The Taylor series for* $\ln(1+x)$ *about* 0 *is*

$$\sum_{k=1}^{\infty} \frac{(-1)^{(k+1)}}{k} x^k$$

with $R = 1$. *Do we have, for* $x \in (-1, 1)$,

$$\ln(1+x) = \sum_{k=1}^{\infty} \frac{(-1)^{(k+1)}}{k} x^k?$$

Solution. Let $f(x) = \ln(1+x)$, *then*

$$f^{(n)}(x) = (-1)^{n+1} \frac{(n-1)!}{(1+x)^n}.$$

and

$$\begin{aligned} |R_n(x)| &= \left| \frac{f^{(n+1)}(\xi)}{(n+1)!} x^{n+1} \right| \\ &= \left| \frac{(-1)^{n+2} n!}{(n+1)!(1+\xi)^{n+1}} x^{n+1} \right| \\ &= \frac{1}{(n+1)} \left| \frac{x}{1+\xi} \right|^{n+1} \end{aligned}$$

If $0 < x < 1$, *then* $\xi \in (0, x)$ *and* $\left| \frac{x}{1+\xi} \right| < 1$.

$$\lim_{n \to \infty} |R_n(x)| \le \lim_{n \to \infty} \frac{1}{(n+1)} = 0.$$

**For the case of* $-1 < x < 0$, *we use the integral form of the remainder term* $R_n(x)$. *Since*

$$f^{(n+1)}(x) = (-1)^n n! \frac{1}{(1+x)^{n+1}},$$

and $x < 0$,

$$\begin{aligned} R_n(x) &= \int_0^x \frac{(x-t)^n}{n!} f^{(n+1)}(t)dt \\ &= (-1)^n \int_0^x (x-t)^n \frac{1}{(1+t)^{n+1}} dt \\ &= -\int_x^0 (t-x)^n \frac{1}{(1+t)^{n+1}} dt. \end{aligned}$$

Since

$$\frac{d}{dt}\left(\frac{t-x}{1+t}\right) = \frac{(1+t)-(t-x)}{(1+t)^2} = \frac{(1+x)}{(1+t)^2} > 0,$$

$\frac{t-x}{1+t}$ *is an increasing function in* t *and*

$$\max_{x \le t \le 0} \frac{t-x}{1+t} = \frac{0-x}{1+0} = -x > 0.$$

$$\min_{x \le t \le 0} \frac{t-x}{1+t} = \frac{x-x}{1+x} = 0.$$

Note that $-x > 0$.

$$\begin{aligned} |R_n(x)| &= \int_x^0 (x-t)^n \frac{1}{(1+t)^{n+1}} dt \\ &\le \int_x^0 \frac{(-x)^n}{(1+t)} dt \\ &= (-x)^n[-\ln(1+x)]. \end{aligned}$$

Since $0 < x+1 < 1$, $\ln(1+x) < 0$ *and* $0 < -x < 1$. *It follows that*

$$\lim_{n\to\infty} |R_n(x)| \le \lim_{n\to\infty} (-x)^n[-\ln(1+x)] = 0.$$

Thus Taylor's formula hold for $x \in (-1, 1)$.

10.4 A Table

Table of standard Taylor expansions:

$$
\begin{aligned}
\frac{1}{1-x} &= \sum_{n=0}^{\infty} x^n = 1 + x + x^2 + x^3 + \dots, && |x| < 1 \\
e^x &= \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots, && \forall x \in \mathbf{R} \\
\log(1+x) &= \sum_{n=0}^{\infty} (-1)^n \frac{x^{n+1}}{n+1} = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \dots, && -1 < x \leq 1 \\
\sin x &= \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots, && \forall x \in \mathbf{R} \\
\cos x &= \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots, && \forall x \in \mathbf{R} \\
\arctan x &= \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)} = x - \frac{x^3}{3} + \frac{x^5}{5} - \dots, && \forall x \in \mathbf{R}, \qquad |x| \leq 1
\end{aligned}
$$

Stirling's Formula:

$$n! = \sqrt{2\pi n}\left(\frac{n}{e}\right)^n (1 + \frac{a_1}{n} + \frac{a_2}{n^2} + \dots).$$

Chapter 11

Techniques for Evaluating Limits

Let c be an extended real number (i.e. $c \in \mathbf{R}$ or $c = \pm\infty$). Suppose that $\lim_{x\to c} f(x)$ and $\lim_{x\to c} g(x)$ are both equal to 0, or both equal to ∞, or both equal to $-\infty$. What can we say about $\lim_{x\to c} f(x)/g(x)$? Which function tend to 0 (respectively $\pm\infty$) faster? We cover two techniques for answering this kind of question: Taylor's Theorem and L'Hôpital's Theorem.

11.1 Use of Taylor's Theorem

How do we compute limits using Taylor expansions?

Let $r < s$ be two real numbers and $a \in (r, s)$, and suppose that $f, g \in C^n([r, s])$. Suppose f and g are $n+1$ times differentiable on (r, s). Assume that

$$f(a) = f^{(1)}(a) = \dots f^{(n-1)}(a) = 0$$
$$g(a) = g^{(1)}(a) = \dots g^{(n-1)}(a) = 0.$$

Suppose that $f^{(n)}(a) = a_n$, $g^{(n)}(a) = b_n \neq 0$. By Taylor's Theorem,

$$\begin{aligned} f(x) &= \frac{a_n}{n!}(x-a)^n + R_n^f(x) \\ g(x) &= \frac{b_n}{n!}(x-a)^n + R_n^g(x). \end{aligned}$$

Here

$$R_n^f(x) = f^{(n+1)}(\xi_f)\frac{(x-a)^{n+1}}{(n+1)!}, \qquad R_n^g(x) = g^{(n+1)}(\xi_g)\frac{(x-a)^{n+1}}{(n+1)!}$$

for some ξ_f and ξ_g between a and x.

Theorem 11.1.1 *Suppose that $f, g \in C^{n+1}([r,s])$, $a \in (r,s)$. Suppose that*

$$f(a) = f^{(1)}(a) = \dots f^{(n-1)}(a) = 0,$$
$$g(a) = g^{(1)}(a) = \dots g^{(n-1)}(a) = 0.$$

Suppose that $g^{(n)}(a) \neq 0$. Then

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \frac{f^{(n)}(a)}{g^{(n)}(a)}.$$

Proof Write $g^{(n)}(a) = b_n$ and $f^{(n)}(a) = a_n$. By Taylor's Theorem,

$$\begin{aligned} f(x) &= \frac{a_n}{n!}(x-a)^n + R_n^f(x) \\ g(x) &= \frac{b_n}{n!}(x-a)^n + R_n^g(x). \end{aligned}$$

Here

$$R_n^f(x) = f^{(n+1)}(\xi_f)\frac{(x-a)^{n+1}}{(n+1)!}, \qquad R_n^g(x) = g^{(n+1)}(\xi_g)\frac{(x-a)^{n+1}}{(n+1)!},$$

where $\xi_f, \xi_g \in [r,s]$. If $f^{(n+1)}$ and $g^{(n+1)}$ are continuous, they are bounded on $[r,s]$ (by the Extreme Value Theorem). Then

$$\begin{aligned} \lim_{x\to a} f^{(n+1)}(\xi)\frac{(x-a)}{(n+1)} &= 0, \\ \lim_{x\to a} g^{(n+1)}(\xi)\frac{(x-a)}{(n+1)} &= 0. \end{aligned}$$

So

$$\begin{aligned} \lim_{x\to a} \frac{f(x)}{g(x)} &= \lim_{x\to a} \frac{\frac{a_n}{n!}(x-a)^n + R_n^f(x)}{\frac{b_n}{n!}(x-a)^n + R_n^g(x)} \\ &= \lim_{x\to a} \frac{a_n + f^{(n+1)}(\xi)\frac{(x-a)}{(n+1)}}{b_n + g^{(n+1)}(\xi)\frac{(x-a)}{(n+1)}} \\ &= \frac{a_n}{b_n}. \end{aligned}$$

□

Example 11.1.2

$$\lim_{x \to 0} \frac{\cos x - 1}{x} = 0$$

Since $\cos x - 1 = -\frac{x^2}{2} + \cdots = 0x - \frac{x^2}{2} + \ldots$.

$$\lim_{x \to 0} \frac{\cos x - 1}{x} = \frac{0}{2} = 0.$$

11.2 L'Hôpital's Rule

Question: Suppose that $\lim_{x \to c} \frac{f'(x)}{g'(x)} = \ell$ where $l \in \mathbf{R} \cup \{\pm\infty\}$. Suppose that $\lim_{x \to c} f(x)$ and $\lim_{x \to c} g(x)$ are both equal to 0, or to $\pm\infty$. Can we say something about $\lim_{x \to c} \frac{f(x)}{g(x)}$? We call limits where both f and g tend to 0, "$\frac{0}{0}$-type limits", and limits where both f and g tend to $\pm\infty$ " $\frac{\infty}{\infty}$-type limits". We will deal with both $\frac{0}{0}$- and $\frac{\infty}{\infty}$- type limits and would like to conclude in both cases that

$$\lim_{x \to c} \frac{f(x)}{g(x)} = \lim_{x \to c} \frac{f'(x)}{g'(x)} = \ell.$$

Theorem 11.2.1 (A simple L'Hôpital's rule) *Let* $x_0 \in (a, b)$. *Suppose that* f, g *are differentiable on* (a, b). *Suppose* $\lim_{x \to x_0} f(x) = \lim_{x \to x_0} g(x) = 0$ *and* $g'(x_0) \neq 0$. *Then*

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = \frac{f'(x_0)}{g'(x_0)}$$

Proof Since $g'(x_0) \neq 0$, there is an interval $(x_0 - r, x_0 + r) \subset (a, b)$ on which $g(x) - g(x_0) \neq 0$. By the algebra of limits, for $x \in (x_0 - r, x_0 + r)$,

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = \lim_{x \to x_0} \frac{\frac{f(x) - f(x_0)}{x - x_0}}{\frac{g(x) - g(x_0)}{x - x_0}} = \frac{\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}}{\lim_{x \to x_0} \frac{g(x) - g(x_0)}{x - x_0}} = \frac{f'(x_0)}{g'(x_0)}.$$

□

Example 11.2.2 *Evaluate* $\lim_{x \to 0} \frac{\sin x}{x}$. *This is a* $\frac{0}{0}$ *type limit. Moreover, taking* $g(x) = x$, *we have* $g'(x_0) \neq 0$. *So L'Hôpital's rule applies, and*

$$\lim_{x \to 0} \frac{\sin x}{x} = \frac{\cos x}{1}|_{x=0} = \frac{1}{1} = 1.$$

Example 11.2.3 *Evaluate* $\lim_{x\to 0} \frac{x^2}{x^2+\sin x}$.
Since $x^2|_{x=0} = 0$ *and* $(x^2 + \sin x)|_{x=0} = 0$ *we identify this as* $\frac{0}{0}$ *type. By L'Hôpital's rule,*

$$\lim_{x\to 0} \frac{x^2}{x^2+\sin x} = \frac{2x}{2x+\cos x}|_{x=0} = \frac{0}{0+1} = 0.$$

We will improve on this result by proving a version which needs fewer assumptions. We will need the following theorem of Cauchy.

Lemma 11.2.4 (Cauchy's Mean Value Theorem) *If* f *and* g *are continuous on* $[a, b]$ *and differentiable on* (a, b) *then there exists* $c \in (a, b)$ *with*

$$(f(b) - f(a))g'(c) = f'(c)(g(b) - g(a)). \tag{11.2.1}$$

If $g'(c) \neq 0 \neq g(b) - g(a)$, *then*

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(c)}{g'(c)}.$$

Proof Set

$$h(x) = g(x)(f(b) - f(a)) - f(x)(g(b) - g(a)).$$

Then $h(a) = h(b)$, h is continuous on $[a, b]$ and differentiable on (a, b). By Rolle's Theorem, there is a point $c \in (a, b)$ such that

$$0 = h'(c) = g'(c)(f(b) - f(a)) - f'(c)(g(b) - g(a)),$$

hence the conclusion. □

Now for the promised improved version of L'Hôpital's rule.

Theorem 11.2.5 *Let* $x_0 \in (a, b)$. *Consider* $f, g : (a, b) \setminus \{x_0\} \to \mathbf{R}$ *and assume that they are differentiable at every point of* $(a, b) \setminus \{x_0\}$. *Suppose that* $g'(x) \neq 0$ *for all* $x \in (a, b) \setminus \{x_0\}$.

1. *Suppose that*

$$\lim_{x\to x_0} f(x) = 0 = \lim_{x\to x_0} g(x).$$

If

$$\lim_{x\to x_0} \frac{f'(x)}{g'(x)} = \ell$$

(with $\ell \in \mathbf{R} \cup \{\pm\infty\}$*) then*

$$\lim_{x\to x_0} \frac{f(x)}{g(x)} = \ell.$$

2. Suppose that

$$\lim_{x\to x_0} f(x) = \pm\infty, \qquad \lim_{x\to x_0} g(x) = \pm\infty.$$

If

$$\lim_{x\to x_0} \frac{f'(x)}{g'(x)} = \ell$$

(with $\ell \in \mathbf{R} \cup \{\pm\infty\}$) then

$$\lim_{x\to x_0} \frac{f(x)}{g(x)} = \ell.$$

Proof For part 1, we prove only the case where $\ell \in \mathbf{R}$. The case where $\ell = \pm\infty$ is left as an exercise.

1. We may assume that f, g are defined and continuous at x_0 and that $f(x_0) = g(x_0) = 0$. For otherwise we simply define

$$f(x_0) = 0 = \lim_{x\to x_0} f(x), \qquad g(x_0) = 0 = \lim_{x\to x_0} g(x).$$

2. Let $x > x_0$. Our functions are continuous on $[x_0, x]$ and differentiable on (x_0, x). Apply Cauchy's Mean Value Theorem to see there exists $\xi \in (x_0, x)$ with

$$\frac{f(x) - f(x_0)}{g(x) - g(x_0)} = \frac{f'(\xi)}{g'(\xi)}.$$

   Hence

$$\lim_{x\to x_0+} \frac{f(x)}{g(x)} = \lim_{x\to x_0+} \frac{f(x) - f(x_0)}{g(x) - g(x_0)} = \lim_{x\to x_0+} \frac{f'(\xi)}{g'(\xi)} = \ell.$$

   This is because as $x \to x_0$, also $\xi \in (x_0, x) \to x_0$.

3. Let $x < x_0$. The same 'Cauchy's Mean Value Theorem' argument shows that

$$\lim_{x\to x_0-} \frac{f(x)}{g(x)} = \lim_{x\to x_0-} \frac{f(x) - f(x_0)}{g(x) - g(x_0)} = \lim_{x\to x_0-} \frac{f'(\xi)}{g'(\xi)} = \ell.$$

   In conclusion $\lim_{x\to x_0} \frac{f(x)}{g(x)} = \ell$.

For part 2. First the case $\ell = \lim_{x\to x_0} \frac{f'(x)}{g'(x)} \in \mathbf{R}$.

1. There exists $\delta > 0$ such that if $0 < |x - x_0| < \delta$,

$$\ell - \varepsilon < \frac{f'(x)}{g'(x)} < \ell + \varepsilon. \tag{11.2.2}$$

Since $g'(x) \neq 0$ on $(a, b)\backslash\{x_0\}$, the above expression makes sense.

2. Take x with $0 < |x - x_0| < \delta$, and $y \neq x$ such that $0 < |y - x_0| < \delta$, and such that $x - x_0$ and $y - x_0$ have the same sign (so that x_0 does not lie between x and y). Then $g(x) - g(y) \neq 0$ because if $g(x) = g(y)$ there would be some point ξ between x and y with $g'(\xi) = 0$, by the Mean Value Theorem, whereas by hypothesis g' is never zero on $(a, b) \setminus \{x_0\}$. Thus

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(y)}{g(x)} + \frac{f(y)}{g(x)} = \left(\frac{f(x) - f(y)}{g(x) - g(y)}\right) \frac{g(x) - g(y)}{g(x)} + \frac{f(y)}{g(x)} \tag{11.2.3}$$

3. Fix any such y and let x approach x_0. Since $\lim_{x \to x_0} f(x) = \lim_{x \to x_0} g(x) = +\infty$, we have

$$\lim_{x \to x_0} \frac{f(y)}{f(x)} = 0, \qquad \lim_{x \to x_0} \frac{g(x) - g(y)}{g(x)} = 1.$$

(exercise)

4. By Cauchy's Mean Value Theorem,(11.2.3) can be written

$$\frac{f(x)}{g(x)} = \left(\frac{f'(\xi)}{g'(\xi)}\right) \left(\frac{g(x) - g(y)}{g(x)}\right) + \frac{f(y)}{g(x)} \tag{11.2.4}$$

for some ξ between x and y. By choosing δ small enough, and x and y as in (b), we can make $\frac{f'(\xi)}{g'(\xi)}$ as close as we wish to ℓ. By taking x close enough to x_0 while keeping y fixed, we can make $\frac{g(x)-g(y)}{g(x)}$ as close as we like to 1, and $\frac{f(y)}{g(x)}$ as close as we wish to 0. It follows from (11.2.4) that

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = \ell.$$

Part 2), case $\ell = +\infty$. For any $A > 0$ there exists $\delta > 0$ such that

$$\frac{f'(x)}{g'(x)} > A$$

whenever $0 < |x - x_0| < \delta$. Consider the case $x > x_0$. Fix $y > x_0$ with $0 < |y - x_0| < \delta$. By Cauchy's mean value theorem,

$$\frac{f(x) - f(y)}{g(x) - g(y)} = \frac{f'(c)}{g'(c)}$$

for some c between x and y.

Clearly

$$\lim_{x \to x_0} \frac{f(y)}{f(x)} = 0, \qquad \lim_{x \to x_0} \frac{g(x) - g(y)}{g(x)} = 1.$$

It follows (taking $\varepsilon = 1/2$ in the definition of limit) that there exists δ_1 such that if $0 < |x - x_0| < \delta_1$,

$$\frac{1}{2} = 1 - \varepsilon < \frac{g(x) - g(y)}{g(x)} < 1 + \varepsilon = \frac{3}{2}, \qquad -\frac{1}{2} = -\varepsilon < \frac{f(y)}{f(x)} < \varepsilon = \frac{1}{2}.$$

Thus if $0 < |x - x_0| < \delta$,

$$\begin{aligned} \frac{f(x)}{g(x)} &= \left(\frac{f(x) - f(y)}{g(x) - g(y)}\right) \frac{g(x) - g(y)}{g(x)} + \frac{f(y)}{g(x)} \\ &\geq \frac{A}{2} - \frac{1}{2}. \end{aligned}$$

Since $\frac{A}{2} - \frac{1}{2}$ can be made arbitrarily large we proved that

$$\lim_{x \to x_0+} \frac{f(x)}{g(x)} = +\infty.$$

The proof for the case $\lim_{x \to x_0-} \frac{f(x)}{g(x)} = +\infty$ is similar.

□

Example 11.2.6 *1. Evaluate* $\lim_{x \to 0} \frac{\cos x - 1}{x}$.

This is of $\frac{0}{0}$ *type.*

$$\lim_{x \to 0} \frac{\cos x - 1}{x} = \lim_{x \to 0} \frac{-\sin x}{1} = \frac{0}{1} = 1.$$

2. Evaluate $\lim_{x \to 0} \frac{\arctan(e^x - 1)}{x}$.

The limit $\lim_{x \to 0} \frac{\arctan(e^x - 1)}{x}$ *is of* $\frac{0}{0}$ *type, since* $e^x - 1 \to 0$ *and* $\arctan(e^x - 1) \to \arctan(0) = 0$ *by the continuity of* $\arctan$.

$$\begin{aligned} \lim_{x \to 0} \frac{\arctan(e^x - 1)}{x} &= \lim_{x \to 0} \frac{\frac{1}{1 + (e^x - 1)^2} e^x}{1} \\ &= \frac{e^0}{1 + (e^0 - 1)^2} = 1. \end{aligned}$$

Example 11.2.7 *Evaluate* $\lim_{x\to 1} \frac{e^{2(x-1)}-1}{\ln x}$.
All functions concerned are continuous, $\frac{e^{2(x-1)}-1}{\ln x} = \frac{e^{2(1-1)}-1}{\ln 1} = \frac{0}{0}$.

$$\lim_{x\to 1} \frac{e^{2(x-1)}-1}{\ln x} = \lim_{x\to 1} \frac{2e^{2(x-1)}}{\frac{1}{x}} = 2.$$

Example 11.2.8 *Don't get carried away with l'Hôpital's rule. Is the following correct and why?*

$$\lim_{x\to 2} \frac{\sin x}{x^2} = \lim_{x\to 2} \frac{\cos x}{2x} = \lim_{x\to 2} \frac{-\sin x}{2} = \frac{-\sin 2}{2}?$$

Example 11.2.9 *Evaluate* $\lim_{x\to 0} \frac{\cos x - 1}{x^2}$. *This is of* $\frac{0}{0}$ *type.*

$$\begin{aligned} \lim_{x\to 0} \frac{\cos x - 1}{x^2} &= \lim_{x\to 0} \frac{-\sin x}{2x} \\ &= \lim_{x\to 0} \frac{-\cos x}{2} \\ &= -\frac{1}{2}. \end{aligned}$$

The limit $\lim_{x\to 0} \frac{-\sin x}{2x}$ *is again of* $\frac{0}{0}$ *type and again we applied L'Hôpital's rule.*

Example 11.2.10 *Take* $f(x) = \sin(x-1)$ *and*

$$g(x) = \begin{cases} x-1, & x \neq 1, \\ 0, & x = 1 \end{cases}$$

$$\lim_{x\to 1} f(x) = 0, \lim_{x\to 1} g(x) = 0.$$

$$\lim_{x\to 1} \frac{f'(x)}{g'(x)} = 1.$$

Hence

$$\lim_{x\to 1} \frac{f(x)}{g(x)} = 1.$$

Theorem 11.2.11 *Suppose that* f *and* g *are differentiable on* (a, x_0) *and* $g(x) \neq 0$, $g'(x) \neq 0$ *for all* $x \in (a, x_0)$. *Suppose*

$$\lim_{x\to x_0-} f(x) = \lim_{x\to x_0-} g(x)$$

is either 0 *or infinity. Then*

$$\lim_{x \to x_0-} \frac{f(x)}{g(x)} = \lim_{x \to x_0-} \frac{f'(x)}{g'(x)}.$$

provided that $\lim_{x \to x_0-} \frac{f'(x)}{g'(x)}$ *exists.*

Example 11.2.12 *Evaluate* $\lim_{x \to 0+} x \ln x$.
Since $\lim_{x \to 0+} \ln x = -\infty$ *we identify this as a "*$0 \cdot \infty$*- type" limit. We have not studied such limits, but can transform it into a* $\frac{\infty}{\infty}$*-type limit, which we know how to deal with:*

$$\lim_{x \to 0+} x \ln x = \lim_{x \to 0+} \frac{\ln x}{\frac{1}{x}} = \lim_{x \to 0+} \frac{\frac{1}{x}}{-\frac{1}{x^2}} = \lim_{x \to 0+} (-x) = 0.$$

Exercise 11.2.13 *Evaluate* $\lim_{y \to \infty} \frac{-\ln y}{y}$.

Theorem 11.2.14 *Suppose* f *and* g *are differentiable on* (a, ∞) *and* $g(x) \neq 0$, $g'(x) \neq 0$ *for all* x *sufficiently large.*

1. *Suppose*
$$\lim_{x \to +\infty} f(x) = \lim_{x \to +\infty} g(x) = 0$$
Then
$$\lim_{x \to +\infty} \frac{f(x)}{g(x)} = \lim_{x \to \infty} \frac{f'(x)}{g'(x)}.$$
provided that $\lim_{x \to +\infty} \frac{f'(x)}{g'(x)}$ *exists.*

2. *Suppose*
$$\lim_{x \to +\infty} f(x) = \lim_{x \to +\infty} g(x) = \infty$$
Then
$$\lim_{x \to +\infty} \frac{f(x)}{g(x)} = \lim_{x \to \infty} \frac{f'(x)}{g'(x)}.$$
provided that $\lim_{x \to +\infty} \frac{f'(x)}{g'(x)}$ *exists.*

Proof Idea of proof: Let $t = \frac{1}{x}$, $F(t) = f(\frac{1}{t})$ and $G(t) = g(\frac{1}{t})$

$$\lim_{x \to +\infty} f(x) = \lim_{t \to 0} f(\frac{1}{t}) = \lim_{t \to 0} F(t).$$

$$\lim_{x \to +\infty} g(x) = \lim_{t \to 0} g(\frac{1}{t}) = \lim_{t \to 0} G(t).$$

Also
$$\frac{F'(t)}{G'(t)} = \frac{f'(\frac{1}{t})(-\frac{1}{t^2})}{g'(\frac{1}{t})(-\frac{1}{t^2})} = \frac{f'(x)}{g'(x)}.$$
Note that the limit as $x \to \infty$ is into a limit as $t \to)^+$, which we do know how to deal with. □

Example 11.2.15 *Evaluate* $\lim_{x\to\infty} \frac{x^2}{e^x}$.
We identify this as $\frac{\infty}{\infty}$ *type.*
$$\lim_{x\to\infty} \frac{x^2}{e^x} = \lim_{x\to\infty} \frac{2x}{e^x} = \lim_{x\to\infty} \frac{2}{e^x} = 0.$$
We have to apply L'Hôpital's rule twice since the result after applying L'Hôpital's rule the first time is $\lim_{x\to\infty} \frac{2x}{e^x}$*, which is again* $\frac{\infty}{\infty}$ *type.*

Proposition 11.2.16 *Suppose that* $f, g : (a,b) \to \mathbf{R}$ *are differentiable, and* $f(c) = g(c) = 0$ *for some* $c \in (a,b)$*. Suppose that* $\lim_{x\to c} \frac{f'(x)}{g'(x)} = +\infty$*. Then* $\lim_{x\to c} \frac{f(x)}{g(x)} = +\infty$

Proof For any $M > 0$ there is $\delta > 0$ such that if $0 < |x - c| < \delta$,
$$\frac{f'(x)}{g'(x)} > M.$$

First take $x \in (c, c+\delta)$. By Cauchy's mean value theorem applied to the interval $[c,x]$, there exists ξ with $|\xi - c| < \delta$ such that
$$\begin{aligned} \frac{f(x)}{g(x)} &= \frac{f(x) - f(c)}{f(x) - g(c)} \\ &= \frac{f'(\xi)}{g'(\xi)} > M. \end{aligned}$$

hence $\lim_{x\to x_0+} \frac{f(x)}{g(x)} = +\infty$. A similar argument shows that $\lim_{x\to x_0-} \frac{f(x)}{g(x)} = +\infty$ □

Remark 11.2.17 *We summarise the cases: Let* $x_0 \in \mathbf{R} \cup \{\pm\infty\}$ *and* I *be an open interval which either contains* x_0 *or has* x_0 *as an end point. Suppose* f *and* g *are differentiable on* I *and* $g(x) \neq 0$*,* $g'(x) \neq 0$ *for all* $x \in I$*. Suppose*
$$\lim_{x\to x_0} f(x) = \lim_{x\to x_0} g(x) = 0$$

Then

$$\lim_{x\to x_0} \frac{f(x)}{g(x)} = \lim_{x\to x_0} \frac{f'(x_0)}{g'(x_0)}.$$

provided that $\lim_{x\to x_0} \frac{f'(x)}{g'(x)}$ *exists. Suitable one sided limits are used if* x_0 *is an end-point of the interval* I*.*

Recall that if α is a number, for $x > 0$ we define

$$x^\alpha = e^{\alpha \log x}.$$

So

$$\frac{d}{dx} x^\alpha = e^{\alpha \log x} \frac{\alpha}{x} = x^\alpha \frac{\alpha}{x} = \alpha x^{\alpha-1}.$$

Example 11.2.18 *For* $\alpha > 0$*,*

$$\lim_{x\to\infty} \frac{\log x}{x^\alpha} = \lim_{x\to\infty} \frac{\frac{1}{x}}{\alpha x^{\alpha-1}} = \lim_{x\to\infty} \frac{1}{\alpha x^\alpha} = 0.$$

Conclusion: As $x \to +\infty$*,* $\log x$ *goes to infinity slower than* x^α *any* $\alpha > 0$*.*

Example 11.2.19 *Evaluate* $\lim_{x\to 0}(\frac{1}{x} - \frac{1}{\sin x})$*.*
This is apparently of "$\infty - \infty$*-type". We have*

$$\begin{aligned} \lim_{x\to 0}(\frac{1}{x} - \frac{1}{\sin x}) &= \lim_{x\to 0} \frac{\sin x - x}{x \sin x} \\ &= \lim_{x\to 0} \frac{\cos x - 1}{\sin x + x\cos x} \\ &= \lim_{x\to 0} \frac{-\sin x}{2\cos x - x \sin x} \\ &= \frac{-0}{2-0} = 0. \end{aligned}$$

Example 11.2.20 *Evaluate* $\lim_{x\to 0}(\frac{1}{x} - \frac{1}{\sin x})$*.*

We try Taylor's method. For some ξ*,* $\sin x = x - \frac{x^3}{3!}\cos\xi$*:*

$$\begin{aligned} \lim_{x\to 0}(\frac{1}{x} - \frac{1}{\sin x}) &= \lim_{x\to 0} \frac{\sin x - x}{x \sin x} \\ &= \lim_{x\to 0} \frac{-\frac{x^3}{3!}\cos\xi}{x(x - \frac{x^3}{3!}\cos\xi)} \\ &= \lim_{x\to 0} \frac{-\frac{x}{3!}\cos\xi}{(1 - \frac{x}{3!}\cos\xi)} = 0. \end{aligned}$$