Appendix A

Differentiation in one dimension

A.1 Introduction

We'll start by heuristically motivating the idea of a tangent and derivative of a function. Suppose $f : (a, b) \to \mathbb{R}$ is some function which maps an interval (a, b) with $a < b$ to the real line $\mathbb{R}$. For the time being, let's assume that f is continuous. It's often useful to represent f graphically by sketching its graph (see Figure A.1)

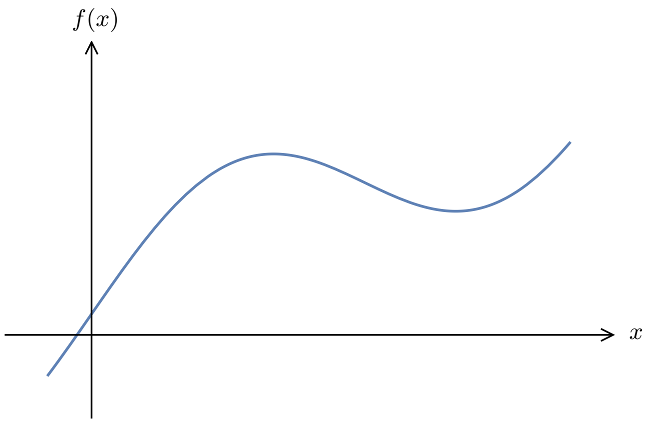

Figure A.1 The graph of a function f.

A very useful concept is the notion of the *tangent* to the graph at a point $p \in (a, b)$. We approximate this by constructing a straight line passing through $(p, f(p))$ and $(p + h, f(p + h))$. This line is the graph of the function:

$$x \mapsto \left(\frac{f(p + h) - f(p)}{h} \right) (x - p) + f(p).$$

Such a straight line is called a chord (see Figure A.2). As h becomes smaller, we might hope that the chords should get closer and closer together and n the limit where $h \to 0$ approach some line passing through p. We call this line the tangent to f at p (see Figure

A.3). The gradient of the tangent to f at p is an important quantity, called the *derivative* of f at p and denoted $f'(p)$.

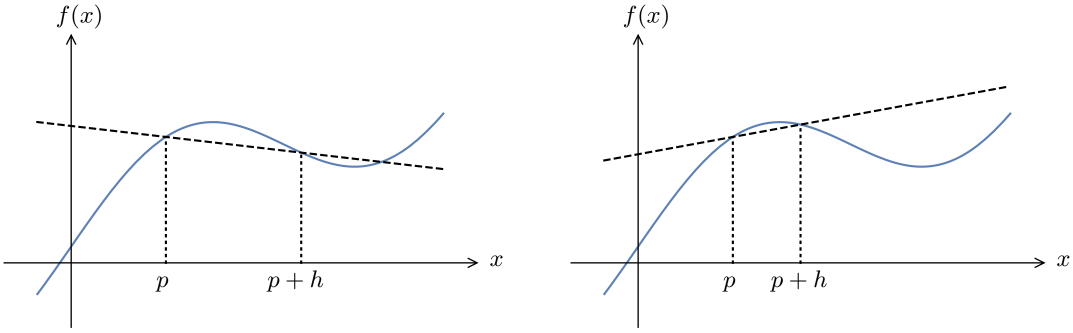

Figure A.2 Chords approximating the tangent to f at p for two values of h.

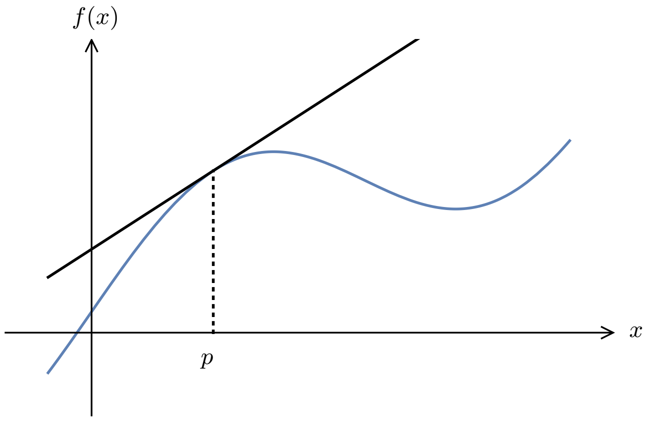

Figure A.3 The tangent to f at p.

This somewhat hand-waving motivation justifies the following definition of differentiability, which we shall unpack shortly:

Definition A.1. A function $f : (a, b) \to \mathbb{R}$ is *differentiable at* $p \in (a, b)$ if the limit

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

exists. If this is the case, we call the value of this limit the *derivative* of f at p and write:

$$f'(p) := \frac{f(x) - f(p)}{x - p}.$$

We can expand this definition by making use of the usual $\epsilon - \delta$ formulation for limits:

Definition A.2 (Alternative definition of differentiability)**.** A function $f : (a, b) \to \mathbb{R}$ is *differentiable at* $p \in (a, b)$ if there exists $\lambda \in \mathbb{R}$ such that for any $\epsilon > 0$ we can find $\delta > 0$ with:

$$\left| \frac{f(x) - f(p)}{x - p} - \lambda \right| < \epsilon$$

for all $x \in (a, b)$ with $0 < |x - p| < \delta$.

The notation:

$$\frac{df}{dx}(p) := f'(p)$$

is sometimes used for the derivative at p.

Exercise(∗). Show that Definitions A.1, A.2 are equivalent.

We will use either formulation of differentiability interchangeably. Notice that at the heart of the definition of differentiability is the assertion that a certain limit exists. In general this may or may not be the case as we shall see now in some examples.

Example A.1. Fix $\alpha, \beta \in \mathbb{R}$. Consider $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = \alpha x + \beta$, and pick $p \in \mathbb{R}$. We have:

$$f(x) - f(p) = \alpha(x - p)$$

so that

$$\frac{f(x) - f(p)}{x - p} = \alpha.$$

We conclude that f is differentiable at p for any $p \in \mathbb{R}$, and the derivative of f is given by $f'(p) = \alpha$.

Example A.2. Consider $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = |x|$, and take $p = 0$. We have:

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

Thus for $\epsilon < 1$, there can exist no λ and $\delta > 0$ such that:

$$\left| \frac{f(x) - f(0)}{x} - \lambda \right| < \epsilon$$

for all x with $0 < |x| < \delta$. We conclude that $f(x) = |x|$ is not differentiable at $x = 0$.

This example shows that there exist continuous functions which are not differentiable. In the other direction, we have the following result:

Lemma A.1. *Suppose $f : (a, b) \to \mathbb{R}$ is differentiable at $p \in (a, b)$. Then f is continuous at p.*

Proof. Fix $\epsilon > 0$. Note that:

$$|f(x) - f(p)| = |x - p| \left| \frac{f(x) - f(p)}{x - p} - f'(p) + f'(p) \right|$$

Applying the triangle inequality, we have:

$$|f(x) - f(p)| \leq |x - p| \left| \frac{f(x) - f(p)}{x - p} - f'(p) \right| + |x - p| \left| f'(p) \right|$$

Now, since f is differentiable at p, we know that there exists δ' such that if $x \neq p$ satisfies $|x - p| < \delta'$ we have:

$$\left| \frac{f(x) - f(p)}{x - p} - f'(p) \right| < 1,$$

so that

$$|f(x) - f(p)| \leq |x - p| \left(1 + \left|f'(p)\right|\right).$$

Thus, if we take

$$|x - p| < \min \left\{ \frac{\epsilon}{1 + |f'(p)|}, \delta' \right\} =: \delta,$$

we have:

$$|f(x) - f(p)| < \epsilon.$$

This is precisely the statement that f is continuous at p. □

This Lemma tells us that differentiability is a strictly stronger condition than continuity. Not all functions that are continuous at p are differentiable at p, but all functions that are differentiable at p are continuous.

We can build differentiable functions by combining functions that we already know to be differentiable in various ways.

Theorem A.2. *i) Suppose that $f, g : (a, b) \to \mathbb{R}$ are both differentiable at $p \in (a, b)$. Then*

$$\begin{array}{rcrcl} u & : & (a,b) & \to & \mathbb{R}, \\ & & x & \mapsto & f(x) + g(x) \end{array}$$

is differentiable at p with

$$u'(p) = f'(p) + g'(p),$$

ii) [Product rule] *Suppose that $f, g : (a, b) \to \mathbb{R}$ are both differentiable at $p \in (a, b)$. Then*

$$\begin{array}{rcrcl} v & : & (a,b) & \to & \mathbb{R}, \\ & & x & \mapsto & f(x)g(x) \end{array}$$

is differentiable at p with

$$v'(p) = f'(p)g(p) + f(p)g'(p).$$

iii) [Quotient rule] *Suppose that $f, g : (a, b) \to \mathbb{R}$ are both differentiable at $p \in (a, b)$ and that $g(p) \neq 0$. Then taking δ sufficiently small,*

$$\begin{array}{rcccl} w & : & (p-\delta, p+\delta) & \to & \mathbb{R}, \\ & & x & \mapsto & \frac{f(x)}{g(x)} \end{array}$$

is well defined. Moreover, w is differentiable at p and

$$w'(p) = \frac{f'(p)g(p) - f(p)g'(p)}{g(p)^2}$$

Proof. We will use quite heavily the results of Theorem I.7.

i) We write:

$$\frac{u(x) - u(p)}{x - p} = \frac{f(x) - f(p)}{x - p} + \frac{g(x) - g(p)}{x - p}.$$

Now, since f is differentiable, we know that the first term on the RHS has a limit, $f'(p)$ as $x \to p$. Similarly, the second term has a limit, $g'(p)$. Applying Theorem I.7 part i), we deduce the result.

ii) We write:

$$\begin{aligned} \frac{u(x) - u(p)}{x - p} &= \frac{f(x)g(x) - f(p)g(p)}{x - p} \\ &= \frac{[f(x) - f(p)]\,[g(x) - g(p)] + [f(x) - f(p)]\,g(p) + [g(x) - g(p)]\,f(p)}{x - p} \\ &= \frac{f(x) - f(p)}{x - p}\,[g(x) - g(p)] + \frac{f(x) - f(p)}{x - p} g(p) + \frac{g(x) - g(p)}{x - p} f(p) \end{aligned} \tag{A.1}$$

Now, as $x \to p$, we have that

$$\lim_{x \to p} \frac{f(x) - f(p)}{x - p} = f'(p), \qquad \lim_{x \to p} [g(x) - g(p)] = 0,$$

so by Theorem I.7 part ii), the first term of (A.1) vanishes as $x \to p$. The other two terms are also easily dealt with in a similar fashion, and we find:

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

iii) First note that since g is differentiable at p, it is continuous at p, which implies that if $g(p) \neq 0$, then there exists $\delta > 0$ such that $g(x) \neq 0$ for $x \in (p - \delta, p + \delta)$. Moving on to differentiability, by making use of part ii), it's enough to consider the case where $f(x) \equiv 1$. Then we have, for $x \in (p - \delta, p + \delta)$:

$$\frac{w(x) - w(p)}{x - p} = \frac{\frac{1}{g(x)} - \frac{1}{g(p)}}{x - p} = -\frac{1}{g(x)g(p)} \frac{g(x) - g(p)}{x - p}$$

Now as $x \to p$ we have, making use of Theorem I.7 part iii):

$$\lim_{x\to p} -\frac{1}{g(x)g(p)} = \frac{-1}{g(p)^2}, \qquad \lim_{x\to p} \frac{g(x)-g(p)}{x-p} = g'(p).$$

By Theorem I.7 part ii) we conclude:

$$\lim_{x\to p} \frac{w(x)-w(p)}{x-p} = -\frac{g'(p)}{g(p)^2}.$$

□

Exercise(∗). Give a proof of Theorem A.2 parts $i), ii)$ directly using the $\epsilon-\delta$ formulation of differentiability (Definition A.2).

The next result tells us how to differentiate functions which are formed as the composition of two functions. The proof is slightly subtle, but in this form will be useful later when we come to the chain rule for functions of many variables.

Theorem A.3 (Chain Rule). *Suppose that $g : (a, b) \to (c, d)$ and $f : (c, d) \to \mathbb{R}$. Suppose further that g is differentiable at p and f is differentiable at $g(p)$. Then*

$$\begin{array}{cccc} z & : & (a,b) & \to & \mathbb{R}, \\ & & x & \mapsto & f(g(x)) \end{array}$$

is differentiable at p and:

$$z'(p) = g'(p)f'(g(p))$$

Proof. Since f is differentiable at p, we know that:

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

We can package this piece of information up in a slightly different form by defining a new function $F : (c, d) \to \mathbb{R}$ by:

$$F(x) := \begin{cases} \frac{f(x)-f(p)}{x-p} - f'(p) & x \neq p \\ 0 & x = p. \end{cases}$$

The statement that f is differentiable at p is equivalent to the statement that F is continuous at p (check that you understand this). Notice that we can write:

$$f(x) = f(p) + \left[f'(p) + F(x)\right](x-p). \tag{A.2}$$

Similarly, letting $q = f(p)$, we define $G : (a, b) \to \mathbb{R}$ by:

$$G(y) := \begin{cases} \frac{g(y)-g(q)}{y-q} - f'(p) & y \neq q \\ 0 & y = q. \end{cases}$$

The statement that g is differentiable at $f(p)$ is equivalent to the statement that G is continuous at 0. We also have:

$$g(y) = g(q) + \left[g'(q) + G(y)\right](y - q). \tag{A.3}$$

Now, let's rewrite:

$$\begin{aligned} g(f(x)) - g(f(p)) &= g(q) + \left[g'(q) + G(f(x))\right](f(x) - q) - g(f(p)) \\ &= \left[g'(q) + G(f(x))\right](f(x) - q) \end{aligned}$$

Here we have used (A.3) to rewrite $g(y)$ with $y = f(x)$, and used $q = f(p)$. Next, we expand the $f(x)$ appearing in the second bracket using (A.2):

$$\begin{aligned} g(f(x)) - g(f(p)) &= \left[g'(q) + G(f(x))\right](f(x) - q) \\ &= \left[g'(q) + G(f(x))\right]\left[f'(p) + F(x)\right](x - p), \end{aligned}$$

again using $q = f(p)$ to get a cancellation. Thus we have:

$$\frac{g(f(x)) - g(f(p))}{x - p} = \left[g'(q) + G(f(x))\right]\left[f'(p) + F(x)\right].$$

Now, since F is continuous at p, we have $\lim_{x\to p} F(x) = 0$. Moreover, since f is continuous at p and G is continuous at $q = f(p)$, we have:

$$\lim_{x\to p} G(f(x)) = G\left(\lim_{x\to p} f(x)\right) = G(q) = 0.$$

Thus we have:

$$\lim_{x\to p} \frac{g(f(x)) - g(f(p))}{x - p} = f'(x)g'(f(x)),$$

the result we require. □

The results above allow us to establish that quite a few functions are differentiable. An important further result on differentiability concerns power series. We shan't prove this result at this stage, since the proof will follow as a corollary of some results we establish later in the course (see Corollary 2.18).

Theorem A.4. *Suppose f is given by the power series:*

$$f(x) = \sum_{k=0}^{\infty} a_k x^k,$$

with radius of convergence $R > 0$. Then f is differentiable at all points $x \in (-R, R)$ and moreover:

$$f'(x) = \sum_{k=1}^{\infty} k a_k x^{k-1}.$$

In other words, provided we are inside the radius of convergence, we can differentiate a power series term by term.

Exercise A.1. An alternative proposed proof of the chain rule runs as follows. We rewrite:

$$\frac{g(f(x)) - g(f(p))}{x-p} = \frac{g(f(x)) - g(f(p))}{f(x) - f(p)} \frac{f(x) - f(p)}{x-p},$$

and we claim:

$$\lim_{x \to p} \frac{g(f(x)) - g(f(p))}{f(x) - f(p)} = \lim_{y \to f(x)} \frac{g(y) - g(f(p))}{y - f(p)} = g'(f(p))$$

which together with:

$$\lim_{x \to p} \frac{f(x) - f(p)}{x-p} = f'(p)$$

gives the result. Why is this 'proof' not valid?

Exercise A.2. a) Suppose that $n \in \mathbb{N}$ and $f : \mathbb{R} \to \mathbb{R}$ is given by:

$$f(x) = x^n.$$

Show that f is differentiable at all points of $\mathbb{R}$ and $f'(x) = nx^{n-1}$.

b) Show that if $f : \mathbb{R} \to \mathbb{R}$ is any polynomial function of x, then f is differentiable at all points of $\mathbb{R}$.

Exercise A.3. Prove[1] that $f : \mathbb{R} \to \mathbb{R}$ is differentiable everywhere in $\mathbb{R}$, and find the derivative of:

a) $f(x) = \sin x$

b) $f(x) = \cos x$

c) $f(x) = \exp x$

d) $f(x) = e^{-x^2}$

Exercise A.4. Consider the function $f : \mathbb{R} \to \mathbb{R}$ given by:

$$f(x) := \begin{cases} x^2 \sin \frac{1}{x} & x \neq 0 \\ 0 & x = 0. \end{cases}$$

a) Show that for $x \neq 0$, f is differentiable at x with derivative:

$$f'(x) = -\cos \frac{1}{x} + 2x \sin \frac{1}{x}.$$

b) Show that f is differentiable at $x = 0$ with derivative:

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

c) Show that the function $x \mapsto f'(x)$ is not continuous at $x = 0$.

[1] You may assume

$$\sin x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1}, \quad \cos x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n}, \quad \exp x = \sum_{n=0}^{\infty} \frac{1}{n!} x^n.$$

A.1.1 Higher derivatives

Suppose $f : (a, b) \to \mathbb{R}$ is differentiable at each point $x \in (a, b)$. Then we can obviously consider the function:

$$\begin{array}{rcccl} f' & : & (a,b) & \to & \mathbb{R}, \\ & & x & \mapsto & f'(x). \end{array}$$

A priori, this function can be very badly behaved. From the example in Exercise A.4, we know that f' need not be continuous. If f' is continuous, we say that f is continuously differentiable. If f' is itself differentiable, we say that f is twice differentiable and we denote by f'' or $f^{(2)}$ the derivative of f' (the second derivative of f). Obviously we can inductively define a $k-$times differentiable function to be a function for which f' is $(k-1)-$times differentiable, and we denote the k^{th} derivative by $f^{(k)}$. If a function is $k-$times differentiable for all $k \in \mathbb{N}$, we say that f is *smooth* (or infinitely differentiable).

Exercise A.5. Show that if the power series:

$$f(x) = \sum_{k=0}^{\infty} a_k x^k,$$

has radius of convergence $R > 0$, then the power series:

$$g(x) = \sum_{k=1}^{\infty} k a_k x^{k-1}$$

also has radius of convergence R. Deduce that a power series is smooth inside its radius of convergence.

A.2 Rolle's theorem

Suppose we have a function which is continuous on the closed interval $[a, b]$ and differentiable on the open interval (a, b), and moreover, $f(a) = f(b)$. Intuitively, we expect that (unless f is constant), the graph of f will have to 'turn around' at some point between a and b (see Figure A.4). We'd also expect that this might mean the derivative of f should vanish somewhere in the interval (a,b). Rolle's theorem asserts that this is indeed the case.

Theorem A.5 (Rolle's theorem)**.** *Suppose $f : [a, b] \to \mathbb{R}$ is continuous on $[a, b]$ and differentiable at each $x \in (a, b)$. Suppose further that $f(a) = f(b)$. Then there exists $c \in (a, b)$ such that $f'(c) = 0$.*

Proof. By the extreme value theorem (Theorem I.9) we know that f achieves a maximum and a minimum in $[a, b]$. Now, either f is constant (in which case the result holds trivially), or else one of the maxima or minima occurs at some $c \in (a, b)$. Suppose that $c \in (a, b)$ is a maximum, and consider for $x \neq c$:

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

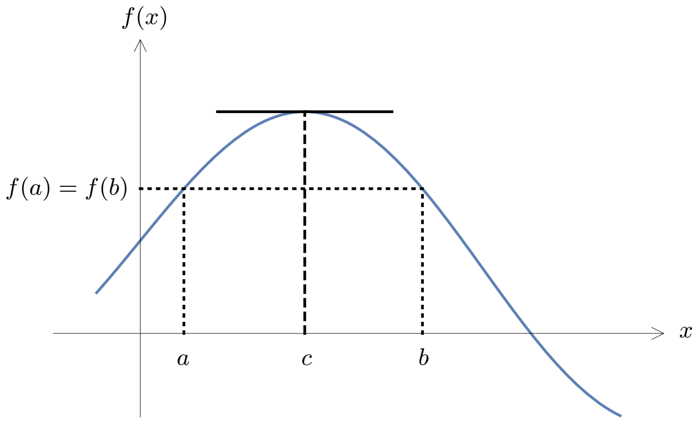

Figure A.4 The set-up for Rolle's theorem

Since f achieves its maximum value at c, we have $f(x) - f(c) \leq 0$. Thus we have $g(x) \geq 0$ for $x < c$ and $g(x) \leq 0$ for $x > c$. We also know that the limit $\lim_{x \to c} g(x)$ exists and is equal to $f'(c)$. We deduce that $f'(c) = 0$.

In the case where $c \in (a, b)$ is a minimum, we can apply the same argument to $-f(x)$, which has a maximum at c. □

A.3 Mean value theorem

The mean value theorem can be thought of as a 'tilted' version of Rolle's theorem. It asserts that a function which is continuous on $[a, b]$ and differentiable on (a, b) must somewhere have a slope equal to the 'average slope' over the interval, i.e. $\frac{f(b)-f(a)}{b-a}$ (see Figure A.5). Put in more human terms, this says:

> If you drive 50 miles, starting from London at 1pm and arriving in Brighton at 2pm, then at some time between 1pm and 2pm your car must have been travelling at 50mph.

This is true regardless of what style of driving one adopts.

Theorem A.6 (Mean value theorem)**.** *Suppose $f : [a, b] \to \mathbb{R}$ is continuous on $[a, b]$ and differentiable at each $x \in (a, b)$. Then there exists $c \in (a, b)$ such that:*

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

Proof. We will use a clever trick to reduce this proof to an application of Rolle's theorem. Consider the function:

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

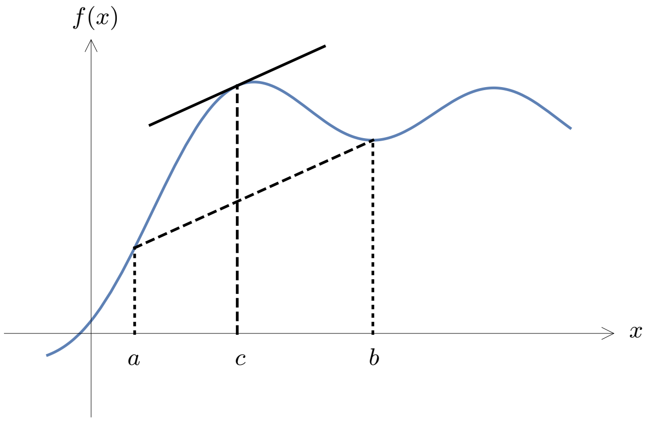

Figure A.5 The set-up for the mean value theorem

Clearly g is continuous on $[a,b]$ and differentiable on (a,b). Moreover, $g(a) = f(a) = g(b)$. Thus we can apply Rolle's theorem to g to deduce that there exists $c \in (a,b)$ such that $g'(c) = 0$. However,

$$0 = g'(c) = f'(c) - \frac{f(b) - f(a)}{b - a},$$

so we deduce that

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

Corollary A.7. *Suppose that $f : (a,b) \to \mathbb{R}$ satisfies $f'(x) = 0$ for $x \in (a,b)$. Then $f(x) = C$ for some $C \in \mathbb{R}$.*

Proof. Pick c, d with $a < c < d < b$. We can apply the mean value theorem on the interval (c,d) to deduce there exists $e \in (c,d)$ with

$$f'(e) = \frac{f(c) - f(d)}{c - d}.$$

However, we know that $f'(e) = 0$ and $c \neq d$, so we must have $f(c) = f(d)$. Since c, d were arbitrary points in (a,b) we deduce that f is constant, i.e. $f(x) = C$. $\square$

Exercise A.6. Suppose that $f, g : (a,b) \to \mathbb{R}$ are differentiable on (a,b) and satisfy:

$$f'(x) = g'(x), \qquad \text{for all } x \in (a,b).$$

Show that:

$$f(x) = g(x) + C.$$

Exercise A.7. Suppose that $f : \mathbb{R} \to \mathbb{R}$ is differentiable and satisfies:

$$f'(x) = f(x), \quad \text{for all } x \in \mathbb{R} \qquad f(0) = 1. \tag{A.4}$$

a) Show that the function:

$$f(x) = \exp x$$

is a solution of (A.4).

b) Let $g(x) = \frac{f(x)}{\exp(x)}$. Show that:

$$g'(x) = 0, \quad \text{for all } x \in \mathbb{R} \qquad g(0) = 1.$$

Conclude that $f(x) = \exp x$ is the *unique* solution of (A.4).

c) By considering $f(x) = \frac{1}{\exp(-x)}$ or otherwise, show that $\exp(-x) = \frac{1}{\exp x}$.

Exercise A.8. Let $p_k : \mathbb{R} \to \mathbb{R}$ be given by:

$$p_k(x) = a_k x^k + a_{k-1} x^{k-1} + \ldots + a_0$$

for some $a_i \in \mathbb{R}$. Show that there is a unique $f : \mathbb{R} \to \mathbb{R}$ which is differentiable everywhere in $\mathbb{R}$ and satisfies:

$$f'(x) = p_k, \qquad f(0) = 0,$$

and that f is a polynomial of degree $k + 1$.

Exercise A.9. Suppose that $f, g : (a, b) \to \mathbb{R}$ are $k-$times differentiable on (a, b) and satisfy:

$$f^{(k)}(x) = g^{(k)}(x), \qquad \text{for all } x \in (a, b).$$

Show that:

$$f(x) = g(x) + p_{k-1}(x).$$

for some polynomial p_{k-1} of order $k - 1$.

A.4 Taylor's theorem

An important feature of $k-$times differentiable functions is that they can be locally approximated by a polynomial of degree $k - 1$. This important fact is encapsulated in *Taylor's theorem*:

Theorem A.8 (Taylor's Theorem)**.** *Suppose that $f : (c, d) \to \mathbb{R}$ is $k-$times differentiable on (c, d) and suppose $[a, x] \subset (c, d)$. Then:*

$$f(x) = f(a) + f'(a)(x - a) + \frac{f''(a)}{2!}(x - a)^2 + \ldots + \frac{f^{(k-1)}(a)}{(k-1)!}(x - a)^{k-1} + R_k(x),$$

where:

$$R_k(x) = \frac{f^{(k)}(\xi)}{k!}(x - a)^k$$

for some $\xi \in (a, x)$. The same result holds if $[x, a] \subset (c, d)$, the only modification being that now $\xi \in (x, a)$.

We shall proceed to prove this by first proving two simpler Lemmas, such that Taylor's theorem follows as a Corollary. This version of the proof is adapted from one that I first saw on Tim Gowers blog[2].

Lemma A.9 (Higher order Rolle's Theorem)**.** *Suppose that f satisfies the conditions of Theorem A.8, and that moreover,*

$$f^{(l)}(a) = 0, \quad \textit{for all } \ 0 \le l \le k-1, \qquad \textit{and } f(x) = 0.$$

Then there exists $\xi \in (a, x)$ such that $f^{(k)}(\xi) = 0$.

Proof. We proceed by induction on k. The case $k = 1$ is simply Rolle's theorem. Now suppose that the result holds for $k = K$, and that:

$$f^{(l)}(a) = 0, \quad \text{for all } \ 0 \le l \le K, \qquad \text{and } f(x) = 0.$$

Applying Rolle's theorem, we know that there exists $y \in (a, x)$ such that $f'(y) = 0$. Now let $g = f'$. The function g satisfies:

$$g^{(l)}(a) = 0, \quad \text{for all } \ 0 \le l \le K-1, \qquad \text{and } g(y) = 0.$$

By the induction hypothesis, there exists $\xi \in (a, y) \subset (a, x)$ such that $0 = g^{(K)}(\xi) = f^{(K+1)}(\xi)$, which establishes the result for $k = K + 1$. By induction we're done. □

Lemma A.10. *Suppose that f satisfies the conditions of Theorem A.8. There exists a polynomial of order k, P_k with the properties:*

i) $f^{(l)}(a) = P_k^{(l)}(a)$ *for all* $0 \le l \le k-1$.

ii) $f(x) = P_k(x)$.

Proof. First, note that for $m \in \mathbb{N}$, the function:

$$q_m : y \mapsto \frac{(y-a)^m}{m!}$$

satisfies:

$$q_m^{(l)}(a) = \begin{cases} 1 & m = l, \\ 0 & m \neq l. \end{cases}$$

Let us define Q_{k-1} to be the polynomial:

$$\begin{aligned} Q_{k-1}(y) &= \sum_{m=0}^{k-1} f^{(m)}(a) q_m(y) \\ &= f(a) + f'(a)(y-a) + \frac{f''(a)}{2!}(y-a)^2 + \ldots + \frac{f^{(k-1)}(a)}{(k-1)!}(y-a)^{k-1}. \end{aligned}$$

[2]https://gowers.wordpress.com/2014/02/11/taylors-theorem-with-the-lagrange-form-of-the-remainder/

This satisfies $f^{(l)}(a) = Q^{(l)}_{k-1}(a)$ for all $0 \leq l \leq k-1$. To construct P_k, we take:

$$P_k(y) = Q_{k-1}(y) + \lambda q_k(y) = Q_{k-1}(y) + \lambda \frac{(y-a)^k}{k!}.$$

This certainly satisfies $f^{(l)}(a) = P^{(l)}_k(a)$ for all $0 \leq l \leq k-1$. Setting $y = x$, we see that $f(x) = P_k(x)$ implies

$$\lambda = \frac{k!}{(x-a)^k}\left[f(x) - Q_{k-1}(x)\right].$$

□

Now, we are ready to prove Taylor's theorem, by combining these two results:

Proof of Taylor's Theorem. Consider the function $g = f - P_k$, where P_k is the polynomial constructed in Lemma A.10. This satisfies

$$g^{(l)}(a) = 0, \qquad \text{for all} \quad 0 \leq l \leq k-1, \qquad \text{and } g(x) = 0,$$

so by Lemma A.9, we know there exists $\xi \in (a, x)$ such that $g^{(k)}(\xi) = 0$. Now, we can calculate $g^{(k)}(\xi)$, and we find:

$$\begin{aligned} 0 = g^{(k)}(\xi) &= f^{(k)}(\xi) - P^{(k)}_k(\xi) = f^{(k)}(\xi) - \lambda \\ &= f^{(k)}(\xi) - \frac{k!}{(x-a)^k}\left[f(x) - Q_{k-1}(x)\right]. \end{aligned}$$

Rearranging this, we deduce:

$$\begin{aligned} f(x) &= Q_{k-1}(x) + f^{(k)}(\xi)\frac{(x-a)^k}{k!}, \\ &= f(a) + f'(a)(y-a) + \frac{f''(a)}{2!}(y-a)^2 + \ldots + \frac{f^{(k-1)}(a)}{(k-1)!}(y-a)^{k-1} + f^{(k)}(\xi)\frac{(x-a)^k}{k!}, \end{aligned}$$

which is precisely the result we seek. □

As an example of the use of Taylor's theorem, let us consider the function:

$$f : y \mapsto \frac{1}{1+y},$$

which is smooth on the interval $(-1, \infty)$. A short exercise shows that

$$f^{(l)}(y) = (-1)^l \frac{l!}{(1+y)^{l+1}}.$$

Taylor's theorem tells us that for $x \in (0, 1)$:

$$f(x) = \sum_{l=0}^{k-1} (-x)^l + R_k(x), \tag{A.5}$$

where

$$R_k(x) = \frac{(-x)^k}{(1+\xi)^{k+1}}$$

for some $\xi \in (0, x)$. We have no control over which value ξ takes. We can however estimate:

$$\left| \frac{(-1)^k}{(1+\xi)^{k+1}} \right| \leq 1$$

for $\xi \in (0, x)$. We deduce $|R_k(x)| \leq x^k$. Rearranging (A.5), we have:

$$\left| \frac{1}{1+x} - \sum_{l=0}^{k-1} (-x)^l \right| \leq x^k$$

for[3] $x \in [0, 1)$ and all k. As $k \to \infty$, the right hand side tends to zero for $x \in [0, 1)$. We conclude that for x in this range:

$$\frac{1}{1+x} = \sum_{l=0}^{\infty} (-x)^l.$$

Notice that Taylor's theorem doesn't fail to hold for $x \geq 1$. However, our control over the remainder term becomes too weak for us to sum the series.

Taylor's theorem is very powerful, but is frequently misused, so lots of care must be taken in its application. Notice that Taylor's theorem makes a statement about a finite number of derivatives. One is often tempted to assume that a (smooth) function is equal to its Taylor series. The following example shows that this is not the case.

Consider the function:

$$f(x) = \begin{cases} 0 & x \leq 0, \\ e^{-\frac{1}{x}} & x > 0. \end{cases}$$

First of all, we note that on any interval which avoids $x = 0$, this function is certainly continuous, indeed smooth. In fact, I claim that this function is smooth on the whole real line. To show this, we'll prove a slightly more general result which will imply the statement we require.

Lemma A.11. *Suppose p_k is a polynomial of degree k. Consider the function:*

$$f(x) = \begin{cases} 0 & x \leq 0, \\ p_k\left(\frac{1}{x}\right) e^{-\frac{1}{x}} & x > 0. \end{cases}$$

This function is differentiable on $\mathbb{R}$, with derivative:

$$f'(x) = \begin{cases} 0 & x \leq 0, \\ p_{k+2}\left(\frac{1}{x}\right) e^{-\frac{1}{x}} & x > 0. \end{cases}$$

where p_{k+2} is a polynomial of order $k+2$ given by:

$$p_{k+2}(t) = t^2 \left(p_k(t) - p_k'(t)\right)$$

[3]Note the result holds trivially with $x = 0$

Proof. First we note that as $x \to 0$, we have:

$$\lim_{x \to 0} p_k \left(\frac{1}{x}\right) e^{-\frac{1}{x}} = 0,$$

by Exercise A.10, so f is continuous. Next, observe that for any $p < 0$, the function is constant in a neighbourhood of $x = p$, so $f'(p) = 0$ if $p < 0$. For $p > 0$, we have that if x is close enough to $x = p$, then:

$$f(x) = p_k \left(\frac{1}{x}\right) e^{-\frac{1}{x}}.$$

Invoking theorems A.2, A.3 and Exercise A.2, we have that:

$$f'(p) = \left(\frac{1}{p^2} p_k \left(\frac{1}{p}\right) - \frac{1}{p^2} p_k' \left(\frac{1}{p}\right)\right) e^{-\frac{1}{p}} = p_{k+2} \left(\frac{1}{p}\right) e^{-\frac{1}{p}}.$$

It remains to show that f is differentiable at $x = 0$. We need to consider the limit:

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

Now, we know that:

$$\frac{f(x)}{x} = \begin{cases} 0 & x \leq 0, \\ \frac{1}{x} p_k \left(\frac{1}{x}\right) e^{-\frac{1}{x}} & x > 0. \end{cases}$$

Again invoking Exercise A.10, we conclude that

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

This establishes the result we require. □

Corollary A.12. *The function:*

$$f(x) = \begin{cases} 0 & x \leq 0, \\ e^{-\frac{1}{x}} & x > 0. \end{cases}$$

is smooth on $\mathbb{R}$, *and* $f^{(k)}(0) = 0$, *for all* $k \in \mathbb{N}$.

Now consider the Taylor series for this function. We have that every term in the Taylor series is identically zero. Therefore the Taylor's series converges everywhere in $\mathbb{R}$ (as it is simply the zero series). On the other hand, $e^{-\frac{1}{y}} \neq 0$ for $y > 0$. This means that even though the Taylor series converges everywhere, the function only equals its Taylor series at the point $x = 0$!

Exercise A.10. a) By using the power series or otherwise, show that for any $N \in \mathbb{N}^*$ and $x \geq 0$ that:

$$1 + \frac{x^N}{N!} \leq \exp x.$$

b) Deduce that for $y > 0$:

$$e^{-\frac{1}{y}} \leq \frac{1}{1 + \frac{1}{N! y^N}} \leq N! y^N.$$

c) Suppose p_k is a polynomial of degree k, and consider the function:

$$f(x) = \begin{cases} 0 & x \leq 0, \\ p_k\left(\frac{1}{x}\right) e^{-\frac{1}{x}} & x > 0. \end{cases}$$

Show that:

$$\lim_{y \to 0} f(y) = 0.$$

Exercise A.11 (*)**.** Consider the function:

$$f(x) = \begin{cases} 0 & |x| \geq 1, \\ e^{-\frac{1}{1-x^2}} & |x| < 1. \end{cases}$$

Show that f is smooth and non-negative on $\mathbb{R}$.