Chapter 5

Differentiation

5.1 Introduction

The processes of *differentiation* and *integration* constitute the two corner–stones of the *calculus* which revolutionised mathematics (and its applications), starting from the groundbreaking work of Newton and Leibniz in the seventeenth century. In this chapter we will focus on understanding differentiation from a rigorous analytic viewpoint, using the knowledge that we have gained about limits in previous chapters. Integration will be dealt with next semester. Before we start this process let us remind ourselves what differentiation is for.

The geometric motivation for differentiation is to find the *slope* or *gradient* of the tangent to a curve at a point lying on it. If the curve is given by a formula $y = f(x)$, then the gradient of the tangent at the point (x, y) appears to be well–approximated by the slope of a chord connecting the very nearby points (x, y) and $(x + \Delta x, y + \Delta y)$. This slope is given by the ratio:

$$\frac{\Delta y}{\Delta x} = \frac{f(x + \Delta x) - f(x)}{\Delta x}.$$

But to get from the slope of the chord to the slope of the tangent, it seems that we must put $\Delta x = 0$ and this gives us the meaningless ratio $0/0$.

The dynamic motivation for differentiation is to find the *instantaneous rate of change* of a one quantity with respect to another. Let us assume that we are dealing with a physical quantity $F(t)$ that changes as a function of time t, for example $F(t)$ could be position of a moving particle at time t, in which case the required rate of change is the velocity, or $F(t)$ could be the charge on a conductor at time t, in which case the rate of change is the current. Then over a very small time interval Δt, the average rate of change is:

$$\frac{\Delta F(t)}{\Delta t} = \frac{F(t + \Delta t) - F(t)}{\Delta t},$$

and we again want to know what happens when $\Delta t = 0$.

Of course, we now know that we must solve both of these problems by taking a limit. When Newton first discovered the calculus, there was no notion of "function". He called a dynamical quantity like $F(t)$ a "fluent", and its instantaneous rate of change was called a "fluxion". Without the modern ideas of either functions, or limits, he struggled to give a precise meaning to the process of differentiation. The following is taken from his essay "The Quadrature of Curves", written in 1693[1]:

"Fluxions are very nearly the Arguments of the Fluents, generated in equal, but infinitely small parts of Time; and to speak exactly, are in the Prime Ratio of the nascent Augments:...Tis the same thing if the Fluxions be taken in the Ultimate Ratio of the Evanescent Parts."

5.2 Differentiation as a Limit

Definition. Let $f : \mathbb{R} \to \mathbb{R}$ be a function with domain D_f. We say that f is *differentiable* at $a \in D_f$ if $\lim_{x\to a} \dfrac{f(x) - f(a)}{x - a}$ exists (and is finite). In this case we write

$$f'(a) = \lim_{x\to a} \frac{f(x) - f(a)}{x - a} = \lim_{h\to 0} \frac{f(a + h) - f(a)}{h}, \qquad (5.2.1)$$

and we call $f'(a) \in \mathbb{R}$, the *derivative of f at a*. We say that f is *differentiable* on $S \subseteq D_f$ if it is differentiable at every point $a \in S$. Then the mapping $a \to f'(a)$ defines a function from $\mathbb{R}$ to $\mathbb{R}$ with domain S which is called the *derivative* of f, and denoted by f'. We may then write $S = D_{f'}$, so

$$D_{f'} = \{x \in D_f; f'(x) \text{ exists}\}.$$

In applied mathematics, we may often write $y = f(x)$, and write[2] the function f' as dy/dx. Then $f'(a) = \frac{dy}{dx}\big|_{x=a}$. When we do analysis, we do not find the dy/dx notation so helpful; it is much more natural to work with f'.

We can of course, iterate the notion of differentiability in the usual way. Suppose that $a \in D_{f'}$ and that f' is differentiable at a, then we define the

[1] Quoted in P.E.Kopp, "Analysis", p102.

[2] This notation was introduced by Leibniz.

second derivative $f''(a)$ of f at a by

$$f''(a) = (f')'(a).$$

More generally if $n \in \mathbb{N}$ with $n > 2$ we define the nth derivative of f at a by

$$f^{(n)}(a) = (f^{(n-1)})'(a),$$

whenever the limit on the right hand side exists. We say that f is *infinitely differentiable* or *smooth* at a if $f^{(n)}(a)$ exists (and is finite) for all $n \in \mathbb{N}$. It is also sometimes useful (especially when considering Taylor series, see section 5.4), to employ the notation $f(a) = f^{(0)}(a)$.

Example 5.1 If $f(x) = c$ where c is $\mathbb{R}$ is constant, it is easy to check directly from the definition (5.2.1) that $D_{f'} = D_f = \mathbb{R}$, and $f'(a) = 0$ for all $a \in \mathbb{R}$.

Example 5.2 Let $f(x) = x^n$ for $x \in \mathbb{R}$, where $n \in \mathbb{N}$ is fixed. Then for all $a \in \mathbb{R}$, using the binomial theorem

$$\begin{aligned}\frac{f(a+h) - f(a)}{h} &= \frac{(a+h)^n - a^n}{h} \\ &= \frac{a^n + na^{n-1}h + \frac{1}{2}n(n-1)a^{n-2}h^2 + \cdots + nah^{n-1} + h^n - a^n}{h} \\ &= na^{n-1} + \frac{1}{2}n(n-1)a^{n-2}h + \cdots + nah^{n-2} + h^{n-1} \\ &\to na^{n-1} \text{ as } h \to 0,\end{aligned}$$

so $f'(x) = nx^{n-1}$ for all $x \in \mathbb{R}$, and $D_{f'} = D_f = \mathbb{R}$.

Next we turn our attention to the relationship between differentiability and continuity.

Theorem 5.2.1. *If $f : \mathbb{R} \to \mathbb{R}$ is differentiable at $a \in D_f$ then f is continuous at a.*

Proof. We need to show that $\lim_{x\to a} f(x) = f(a)$. For $x \neq a$, write

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

Since f is differentiable at a, $\lim_{x\to a} \dfrac{f(x) - f(a)}{x - a} = f'(a)$, and of course $\lim_{x\to a}(x - a) = 0$. Hence by algebra of limits (Theorem 3.3.1),

$\lim_{x\to a} \dfrac{f(x) - f(a)}{x - a}.(x-a) = 0$. So $\lim_{x\to a}(f(x) - f(a))$ exists and equals zero, and the result follows. □

The converse to Theorem 5.2.1, that every function that is continuous at a is differentiable at a, is false. A counter-example is:

Example 5.3 Consider the function $f(x) = |x|$ with $D_f = \mathbb{R}$. It is continuous at every point in $\mathbb{R}$. It is also easy to see that it is differentiable at every $x \neq 0$. We'll establish that it is not differentiable at zero, by showing that left and right limits are different there. In fact

$$\lim_{h \uparrow 0} \frac{f(0+h) - f(0)}{h} = \lim_{h \uparrow 0} \frac{|h|}{h} = \lim_{h \uparrow 0} \frac{-h}{h} = -1.$$

$$\lim_{h \downarrow 0} \frac{f(0+h) - f(0)}{h} = \lim_{h \downarrow 0} \frac{|h|}{h} = \lim_{h \downarrow 0} \frac{h}{h} = 1.$$

Generalising the last example, we say that the mapping $f : \mathbb{R} \to \mathbb{R}$ has a *left derivative* at $a \in D_f$ if $f'_-(a) = \lim_{h \uparrow 0} \dfrac{f(a+h) - f(a)}{h}$ exists (and is finite), and that it has a *right derivative* at $a \in D_f$ if $f'_+(a) = \lim_{h \downarrow 0} \dfrac{f(a+h) - f(a)}{h}$ exists (and is finite)

Theorem 5.2.2. *The mapping $f : \mathbb{R} \to \mathbb{R}$ is differentiable at $a \in D_f$ if and only if both the left and right derivatives exist and are equal. In this case*

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

The proof is left for you to do as Problem 87.

5.3 Rules For Differentiation

The results in this section should all be familiar from MAS110, but now we can make the proofs rigorous.

Theorem 5.3.1. *Let f and g be functions from $\mathbb{R}$ to $\mathbb{R}$ that are differentiable at $a \in D_f \cap D_g$. Then*

1. *For each $\alpha, \beta \in \mathbb{R}$, the function $\alpha f + \beta g$ is differentiable at a and*

$$(\alpha f + \beta g)'(a) = \alpha f'(a) + \beta g'(a),$$

2. *(The Product Rule) The function fg is differentiable at a and*

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

3. *(The Quotient Rule) If $g(a) \neq 0$ then f/g is differentiable at a and*

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

Proof. 1. This is an easy application of the algebra of limits.

2. For all $h \neq 0$,

$$\begin{aligned} & \frac{(fg)(a+h) - (fg)(a)}{h} \\ = & \frac{f(a+h)g(a+h) - f(a)g(a+h)}{h} + \frac{f(a)g(a+h) - f(a)g(a)}{h} \\ = & \left(\frac{f(a+h) - f(a)}{h}\right) g(a+h) + f(a)\left(\frac{g(a+h) - g(a)}{h}\right). \end{aligned}$$

The result follows by taking limits as $h \to 0$ and using (5.2.1) and the algebra of limits, together with the fact that at a, g is differentiable, hence continuous by Theorem 5.2.1, and so $\lim_{h\to 0} g(a+h) = g(a)$.

3. First observe that by Problem 61, there exists $\delta > 0$ so that $g(x) \neq 0$ for all $x \in (a - \delta, a + \delta)$. In the following, we will only consider $h \in \mathbb{R}$ such that $|h| < \delta$. Then

$$\begin{aligned} & \frac{1}{h}\left\{\left(\frac{f}{g}\right)(a+h) - \left(\frac{f}{g}\right)(a)\right\} \\ = & \frac{1}{h}\left\{\frac{f(a+h)g(a) - f(a)g(a+h)}{g(a)g(a+h)}\right\} \\ = & \frac{1}{g(a)g(a+h)}\left\{\frac{f(a+h) - f(a)}{h}g(a) - f(a)\frac{g(a+h) - g(a)}{h}\right\}, \end{aligned}$$

and the result follows by algebra of limits, using the fact that (as in (2)), $\lim_{h\to 0} g(a+h) = g(a)$.

□

Theorem 5.3.2 (The Chain Rule)**.** *Let f, g be functions from $\mathbb{R}$ to $\mathbb{R}$ for which Ran$(g) \subseteq$ Dom(f). Suppose that g is differentiable at a and that f is differentiable at $g(a)$. Then $f \circ g$ is differentiable at a and*

$$(f \circ g)'(a) = f'(g(a))g'(a).$$

Proof. We begin by arguing formally. For $h \in \mathbb{R}$,

$$\begin{aligned} & \frac{(f \circ g)(a+h) - (f \circ g)(a)}{h} \\ = & \frac{f(g(a+h)) - f(g(a))}{h} \\ = & \frac{f(g(a+h)) - f(g(a))}{g(a+h) - g(a)} \frac{g(a+h) - g(a)}{h}, \end{aligned}$$

but we cannot take limits at this stage, as the first factor appears to be tending to 0/0. To overcome this problem, for each $y \in D_{f'}$ define $\Phi_y : \mathbb{R} \to \mathbb{R}$ by

$$\Phi_y(k) = \begin{cases} \frac{f(y+k)-f(y)}{k} & \text{if } k \neq 0 \text{ and } k + y \in D_f \\ f'(y) & \text{if } k = 0, \\ 0 & \text{otherwise} \end{cases}$$

then $\lim_{k \to 0} \Phi_y(k) = f'(y) = \Phi_y(0)$, so Φ_y is continuous at 0. Now let $k(h) = g(a+h) - g(a)$. Since g is differentiable at a, it is continuous there by Theorem 5.2.1 and so we can rewrite the above display as

$$\begin{aligned} \frac{(f \circ g)(a+h) - (f \circ g)(a)}{h} &= \Phi_{g(a)}(k(h)) \frac{g(a+h) - g(a)}{h} \\ &\to f'(g(a))g'(a), \end{aligned}$$

as $h \to 0$, where we have used Theorem 4.1.3 to deduce that $\lim_{h \to 0} \Phi_{g(a)}(k(h)) = f'(g(a))$. □

5.4 Turning Points and Rollé's Theorem

A function $f : \mathbb{R} \to \mathbb{R}$ has a *local minimum* at $a \in D_f$ if there exists $\delta > 0$ such that $f(x) \geq f(a)$ for all $a - \delta < x < a + \delta$, and a *local maximum* at $a \in D_f$ if there exists $\delta > 0$ such that $f(x) \leq f(a)$ for all $a - \delta < x < a + \delta$. An *extreme point* (or *turning point*) for f is a point in its domain that is either a local maximum or a local minimum. Its important to distinguish carefully between local maxima and minima and global maxima and minima, when the latter exist. For example if $f : [a, b] \to \mathbb{R}$ is continuous, then we know by Theorem 4.3.4 that it attains both its supremum and infimum on $[a, b]$. So these are the global maximum, and minimum respectively. But they are not necessarily extreme points.

Example 5.4 Consider the function $f : [-3, 2] \to \mathbb{R}$ defined by

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

Here the global maximum is attained at $x = 2$, the global minimum is attained at $x = -3$. Neither of these are extreme points. There is also a local minimum at $x = 0$.

Theorem 5.4.1. *If f is differentiable at $a \in D_f$ and a is an extreme point for f, then $f'(a) = 0$.*

Proof. Suppose that f has a local minimum at a (the argument in the case of a local maximum is very similar), then there exists $\delta > 0$ so that whenever $a - \delta < x < a + \delta$, then $f(a) \leq f(x)$. Hence

$$\frac{f(x) - f(a)}{x - a} \leq 0 \text{ for all } a - \delta < x < a, \text{ and}$$
$$\frac{f(x) - f(a)}{x - a} \geq 0 \text{ for all } a < x < a + \delta.$$

Taking one-sided limits as $x \to a$ and using Problem 31, we deduce that $f'_-(a) \leq 0$ and $f'_+(a) \geq 0$. But f is differentiable at a, so by Theorem 5.2.2, $f'_-(a) = f'_+(a) = f'(a)$, and so we conclude that $f'(a) = 0$, as required. □

Of course the converse to Theorem 5.4.1 is false, consider for example $f(x) = x^3$. Then $f'(0) = 0$ but 0 is neither a local maximum nor a local minimum. We will not pursue the story of classifying extreme points further here. You have seen this before in MAS110, and it is revisited in Problem 96. Instead we will use Theorem 5.4.1 to explore some new territory.

Theorem 5.4.2 (Rollé's Theorem)**.** *Let f be continuous on $[a, b]$ and differentiable on (a, b) with $f(a) = f(b)$. Then there exists $c \in (a, b)$ such that $f'(c) = 0$.*

Proof. If f is constant, the result is obvious, so assume that f takes at least two distinct values. By Theorem 4.3.4, f is bounded on $[a, b]$ and attains both its supremum and infimum. It cannot attain both of these at the end-points, as then f would be constant. So there must be a $c \in (a, b)$ where either the supremum or infimum is attained. But then c is an extreme point, and so $f'(c) = 0$ by Theorem 5.4.1. □

5.5 Mean Value Theorems

The next result can be seen as a precursor to Taylor's theorem. But it is also an important result in its own right.

Theorem 5.5.1 (The Mean Value Theorem)**.** *If f is continuous on $[a, b]$ and differentiable on (a, b), then there exists $c \in (a, b)$ such that*

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

Proof. For all $x \in [a, b]$, define $g(x) = f(x) - \alpha(x - a)$, where $\alpha = \frac{f(b)-f(a)}{b-a}$. Then g is continuous on $[a, b]$, and differentiable on (a, b). You can check easily that $g(a) = g(b) = f(a)$, and so we may apply Rollé's theorem (Theorem 5.4.2) to deduce that there exists $c \in (a, b)$ such that $g'(c) = 0$. Hence $f'(c) = \alpha$, as required. □

The mean value theorem has many interesting consequences, for example:

Corollary 5.5.2. *[Monotonicity Revisited] Suppose that $f : \mathbb{R} \to \mathbb{R}$ is continuous on $[a, b]$ and differentiable on (a, b). If for all $x \in (a, b)$ we have*

$f'(x) \geq 0$, then f is monotonic increasing on $[a, b]$,
$f'(x) > 0$, then f is strictly monotonic increasing on $[a, b]$,
$f'(x) \leq 0$, then f is monotonic decreasing on $[a, b]$,
$f'(x) < 0$, then f is strictly monotonic decreasing on $[a, b]$.

Proof. We'll just do the first of these, as the others are so similar. Choose arbitrary $a \leq \alpha < \beta \leq b$. By the mean value theorem (Theorem 5.5.1), there exists $c \in (\alpha, \beta)$ so that

$$\frac{f(\beta) - f(\alpha)}{\beta - \alpha} = f'(c) \geq 0.$$

Hence $f(\beta) \geq f(\alpha)$ and so f is monotonic increasing, as required. □

Corollary 5.5.2 becomes a powerful tool to study inverses, when used in conjunction with Theorem 4.3.6. You can see this for yourself in Problem 93. We also have the following result:

Theorem 5.5.3. *[Inverses Revisited] Suppose that $f : \mathbb{R} \to \mathbb{R}$ is continuous on $[a, b]$ and differentiable on (a, b), and that f' is continuous at $c \in (a, b)$. If $f'(c) \neq 0$ then*

1. *There exists $\delta > 0$ so that f is invertible on $[c - \delta, c + \delta]$, and f^{-1} is continuous on $(f(c-\delta), f(c+\delta))$ if $f^{-1}(c) > 0$, and on $(f(c+\delta), f(c-\delta))$ if $f^{-1}(c) < 0$.*

2. The mapping f^{-1} *is differentiable at* $f(c)$ *and*

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

Proof. Assume without loss of generality, that $f'(c) < 0$.

1. Since f' is continuous at c, by Problem 61 there exists $\delta > 0$ so that $f'(x) < 0$ for all $x \in (c-\delta, c+\delta)$, and we can certainly ensure (by choosing a smaller δ, if necessary), that $(c-\delta, c+\delta) \subseteq (a,b)$. By Corollary 5.5.2, f is strictly decreasing on $[c-\delta, c+\delta]$, and the result then follows from Theorem 4.3.7.

2. Let $y = f(c)$, then for arbitrary $d \in (c-\delta, c) \cup (c, c+\delta)$, we have $f(d) \neq f(c)$ since f is invertible, hence injective. Write $x = f(d)$. Then

$$\begin{aligned} \frac{f^{-1}(y) - f^{-1}(x)}{y-x} &= \frac{c-d}{y-x} \\ &= \frac{1}{\frac{y-x}{c-d}} = \frac{1}{\frac{f(c)-f(d)}{c-d}}. \end{aligned}$$

Now since f^{-1} is continuous on $(f(c+\delta), f(c-\delta))$, as $x \to y$, we have $d \to c$ and so

$$\lim_{x\to y} \frac{f^{-1}(y) - f^{-1}(x)}{y-x} = \lim_{d\to c} \frac{1}{\frac{f(c)-f(d)}{c-d}} = \frac{1}{f'(c)},$$

as was required.

□

Note. Theorem 5.5.3 should be familiar to you from calculus as the rule

$$\frac{dx}{dy} = \frac{1}{\frac{dy}{dx}}.$$

The next result is a useful variation on the “mean value theorem theme”.

Theorem 5.5.4 (Cauchy’s Mean Value Theorem)**.** *Let* f *and* g *each be continuous on* $[a,b]$ *and differentiable on* (a,b) *with* $g'(x) \neq 0$ *for all* $x \in (a,b)$. *Then there exists* $c \in (a,b)$ *so that*

$$\frac{f'(c)}{g'(c)} = \frac{f(b)-f(a)}{g(b)-g(a)}. \tag{5.5.2}$$

Proof. As $g'(x) \neq 0$ for all $x \in (a,b)$, we must have $g(a) - g(b) \neq 0$, by Rollé's theorem (Theorem 5.4.2). The rest of the proof follows along similar lines to that of the mean value theorem, and is left to you to do in Problem 98. □

Corollary 5.5.5 (l'Hôpital's Rule)**.** *Suppose that f and g are each differentiable on (a,b), with $g'(x) \neq 0$ for all $x \in (a,b)$.*

1. *If $c \in (a,b)$ with $f(c) = g(c) = 0$, then*
$$\lim_{x \to c} \frac{f(x)}{g(x)} = \lim_{x \to c} \frac{f'(x)}{g'(x)},$$
whenever the limit on the right hand side is finite.

2. *If $\lim_{x \downarrow a} f(x) = \lim_{x \downarrow a} g(x) = 0$, then*
$$\lim_{x \downarrow a} \frac{f(x)}{g(x)} = \lim_{x \downarrow a} \frac{f'(x)}{g'(x)},$$
whenever the limit on the right hand side is finite.

3. *If $\lim_{x \uparrow b} f(x) = \lim_{x \uparrow b} g(x) = 0$, then*
$$\lim_{x \uparrow b} \frac{f(x)}{g(x)} = \lim_{x \uparrow b} \frac{f'(x)}{g'(x)},$$
whenever the limit on the right hand side is finite.

Proof. We'll only prove (2) as the other proofs are so similar. Note that since f and g are both differentiable on (a,b), they are continuous there by Theorem 5.2.1. We extend f and g to the point a by defining $f(a) = g(a) = 0$. Then f and g are both right continuous at a. We now apply Cauchy's mean value theorem (Theorem 5.5.4) on the interval $[a,x]$ where $a < x < b$ to deduce that there exists $r(x) \in (a,x)$ such that
$$\frac{f(x)}{g(x)} = \frac{f(x) - f(a)}{g(x) - g(a)} = \frac{f'(r(x))}{g'(r(x))}.$$
Now $r(x) = a + \theta(x)(x-a)$ for some $0 < \theta(x) < 1$. Then
$$0 \leq |r(x) - a| = |\theta(x)(x-a)| < |x-a|,$$
so that (by the sandwich rule) $\lim_{x \downarrow a} r(x) = a$. Hence
$$\lim_{x \downarrow a} \frac{f(x)}{g(x)} = \lim_{x \downarrow a} \frac{f'(r(x))}{g'(r(x))} = \lim_{x \downarrow a} \frac{f'(x)}{g'(x)}.$$
□

In Problem 101 you can prove a variation on l'Hôpital's Rule, where instead of converging to zero, the functions diverge to infinity at the point of interest. We'll use this in the next example:

Example 5.5[3]

From e.g. MAS110, you know that for $a > 0$, the function $f(x) = a^x$ may be defined as $f(x) = e^{x \log_e(a)}$ for $x \in \mathbb{R}$. Similarly we may define $g(x) = x^x = e^{x \log_e(x)}$ for $x > 0$, But what happens as $x \downarrow 0$? There appear to be two competing tendencies, as for $x > 0, 0^x = 0$, but $x^0 = 1$. We use the version of l'Hôpital's Rule from Problem 101 and consider

$$\begin{aligned} \lim_{x \downarrow 0} x \log_e(x) &= -\lim_{x \downarrow 0} \frac{-\log_e(x)}{1/x} \\ &= -\lim_{x \downarrow 0} \frac{1/x}{1/x^2} = -\lim_{x \downarrow 0} x = 0, \end{aligned}$$

So by continuity of the exponential function,

$$\lim_{x \downarrow 0} x^x = \lim_{x \downarrow 0} e^{x \log_e(x)} = 1.$$

5.6 Taylor's Theorem

Let $[a, b]$ be a given interval in $\mathbb{R}$. For each $n \in \mathbb{N}$ we introduce the real vector space $C^n(a, b)$ (it is also a ring, and an algebra) of functions $f : [a, b] \to \mathbb{R}$ for which

- The nth derivative $f^{(n)}$ of f exists for all points in (a, b).
- $f^{(n)}$ is continuous on (a, b).

We may also consider the vector space (which is again, also a ring, and an algebra) $C^\infty(a, b)$ of functions that are infinitely differentiable on (a, b). Clearly for all $n \in \mathbb{N}$, we have

$$C^\infty(a, b) \subseteq C^n(a, b) \subseteq C^{n-1}(a, b) \subseteq \cdots \subseteq C^1(a, b) \subseteq C(a, b),$$

where $C(a, b)$ is the space of continuous functions on (a, b).

Let $f \in C^n(a, b)$, for some $n \in \mathbb{N}$. Fix $x_0 \in (a, b)$ and consider the real numbers $f^{(k)}(x_0)/k!$, for $k = 0, 1, \ldots, n$. These are called the *Taylor*

[3] In this example, we will use properties of the exponential function which will be made rigorous in Semester 2 work.

coefficients of f at x_0. We define a function $T_f^{(n)} \in C^n(a,b)$ by

$$\begin{aligned} T_f^{(n)}(x) &= \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k \\ &= f(x_0) + f'(x_0)(x-x_0) + \frac{1}{2}f^{(2)}(x_0)(x-x_0)^2 + \cdots + \frac{1}{n!}f^{(n)}(x_0)(x-x_0)^n. \end{aligned}$$

The mapping $T_f^{(n)}$ is called *the Taylor polynomial of f of degree n around x_0.*

Theorem 5.6.1. *[Taylor's Theorem] Let $f \in C^{n+1}(a,b)$ and $x_0 \in (a,b)$. Then for all $x \in (a,b)$,*

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + R_f^{n+1}(x), \tag{5.6.3}$$

where $R_f^{n+1}(x) = \frac{f^{(n+1)}(c)}{(n+1)!}(x-x_0)^{n+1}$, with $c \in (x_0,x)$ if $x > x_0$ and $c \in (x,x_0)$ if $x < x_0$.[4]

Proof. Assume for convenience that $x > x_0$.

Define $M_f : (x_0,x) \to \mathbb{R}$ by

$$M_f(x) = \frac{(n+1)!}{(x-x_0)^{n+1}}[f(x) - T_f^{(n)}(x)], \tag{5.6.4}$$

and $g : [x_0,x] \to \mathbb{R}$ by

$$g(t) = -f(x) + f(t) + \sum_{k=1}^{n} \frac{f^{(k)}(t)}{k!}(x-t)^k + \frac{(x-t)^{n+1}}{(n+1)!}M_f(x).$$

Then g is clearly continuous on $[x_0,x]$ and differentiable on (x_0,x). You can easily check that $g(x_0) = g(x) = 0$. Then by Rollé's theorem (Theorem 5.4.2), there exists $c \in (x_0,x)$ with $g'(c) = 0$. Now for $t \in (x_0,x)$

$$\begin{aligned} g'(t) &= f'(t) - \sum_{k=1}^{n} \frac{f^{(k)}(t)}{(k-1)!}(x-t)^{k-1} + \sum_{k=1}^{n} \frac{f^{(k+1)}(t)}{k!}(x-t)^k - \frac{(x-t)^n}{n!}M_f(x) \\ &= \frac{(x-t)^n}{n!}(f^{(n+1)}(t) - M_f(x)). \end{aligned}$$

Then $g'(c) = 0$ tells us that c is such that $f^{(n+1)}(c) = M_f(x)$, and then (5.6.3) follows by straightforward algebra from (5.6.4). □

[4] Note that c depends on x.

Notes.

1. The term $R_f^{n+1}(x) = f(x) - T_f^{(n)}(x)$ measures the error in approximating f by its Taylor polynomial at f. It is called the *remainder term of degree* $n+1$.

2. If $0 \in (a,b)$, we can take $x_0 = 0$. In this special case, Theorem 5.6.1 is called *Maclaurin's theorem.*

Now suppose that $f \in C^\infty(a,b)$ and that the series $\sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k$ converges for all $x \in (a,b)$. If we may write

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k = \lim_{n\to\infty} \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k,$$

we say that f is *represented by its Taylor series* on (a,b). You will learn more about convergence of infinite series of both numbers and functions in Semester 2.