Chapter 4

Functions differentiable on an interval

Two applications of differential calculus are particularly pervasive in the natural sciences. In one, we use the derivative of a function to deduce its maximum (or minimum) value and the input which produces that value. In the second, we are given information, not about the function itself, but about its derivative (which is often interpreted as a "rate of change" with respect to time). We then seek to reconstruct the function itself from this information. These two applications rely fundamentally on the Interior Extremum Theorem, and the Mean Value Theorem, respectively, and the main purpose of this section is to state and prove these. Both are statements concerning functions which are differentiable (at least) on an open interval (a, b). The functions in question may be differentiable elsewhere too (perhaps on the whole of $\mathbb{R}$, in fact), so it's important when you read "f is differentiable on (a, b)" not to add the unwarranted assumption "and only on (a, b)".

4.1 The Interior Extremum Theorem

Recall that $f : D \to \mathbb{R}$ attains a **maximum** at $a \in D$ if $f(x) \leq f(a)$ for all $x \in D$. Similarly, f attains a **minimum** at $a \in D$ if $f(x) \geq f(a)$ for all $x \in D$. Note that these definitions have *absolutely nothing to do with calculus!* In fact, they make perfectly good sense whatever set D, the domain of f, is: it need not be a subset of $\mathbb{R}$!

We say that f attains an **extremum** at $a \in D$ if it attains either a maximum or minimum at a. (The word "extremum" means "maximum or minimum", in much the same way that "monotonic" means "increasing or decreasing". Its plural is "extrema".)

Theorem 4.1 (The Interior Extremum Theorem) *Let $f : (a, b) \to \mathbb{R}$ be differentiable and f attain an extremum at $c \in (a, b)$. Then $f'(c) = 0$.*

Proof: Consider first the case that f attains a maximum at $c \in (a, b)$. By assumption, $f'(c)$ exists. Hence, by Theorem 2.14, given any sequence (x_n) in $(a, b)\backslash\{c\}$

which converges to c,

$$s(x_n) := \frac{f(x_n) - f(c)}{x_n - c} \to f'(c). \tag{4.1}$$

This holds, in particular, for the sequence $x_n = c + (b-c)/(n+1) \in (c,b)$. Note that $x_n > c$ for all c, so $x_n - c > 0$, and $f(c)$ is the maximum value of f, so $f(x_n) - f(c) \leq 0$. Hence $s(x_n) \leq 0$ for all n so, by Proposition 1.7, its limit is non-positive, i.e. $f'(c) \leq 0$.

But (4.1) also holds in the case where $x_n = c-(c-a)/(n+1) \in (a,c)$. Now $x_n - c < 0$ for all n, and $f(x_n) - f(c) \leq 0$ (since, as before, f attains its maximum at c), so $s(x_n) \geq 0$ for all n. Hence, by Proposition 1.7, its limit is non-negative, i.e. $f'(c) \geq 0$.

Hence, $f'(c) = 0$, as was to be shown.

Consider now the case that f attains a minimum at $c \in (a,b)$. Let $g = -f$. Then g is differentiable on (a,b) and attains a maximum at c, so, as we just proved, $g'(c) = 0$. Hence $f'(c) = -g'(c) = 0$. $\square$

Definition 4.2 Let $f : D \to \mathbb{R}$ be differentiable. Then $c \in D$ is a **critical point** of f if $f'(c) = 0$.

Given a differentiable function $f : [a,b] \to \mathbb{R}$, we know it attains both a maximum and a minimum by the Extreme Value Theorem. Either or both of these extrema might occur at an endpoint of the interval $[a,b]$. But if they don't, they occur at some interior point $c \in (a,b)$, and this point must be a *criticial point* of f by the Interior Extremum Theorem. So to find the maximum and minimum values attained by f, we just need to evaluate it at the endpoints a, b and any interior critical points, and extract the extreme values.

Example 4.3 Find the maximum and minimum values attained by the function

$$f : [0,3] \to \mathbb{R}, \qquad f(x) = x^2 - 2x + 7.$$

Solution: First note that, since f is continuous, it certainly attains both a maximum and a minimum by the Extreme Value Theorem (Theorem 2.26). Since f is differentiable on $(0,3)$, it follows from the Interior Extremum Theorem that each extremum occurs either at an endpoint, that is, 0 or 3, or at an interior critical point of f. Now

$$f'(x) = 2x - 2 = 0$$

if and only if $x = 1$, so 1 is the only critical point of f. Since $f(1) = 6$, $f(0) = 7$, and $f(3) = 10$, we deduce that f attains a maximum value of 10 at the right endpoint 3, and a minimum value of 6 at the interior critical point 1.

Remark In this case, using calculus was really overkill, since we could have deduced the same information by simply completing the square:

$$f(x) = (x-1)^2 + 6.$$

We can also use information about the derivative to show, in some circumstances, that the extrema of a differentiable function on $[a,b]$ do *not* occur at its endpoints.

Proposition 4.4 *Let $f : [a,b] \to \mathbb{R}$ be differentiable.*

(i) If $f'(a) > 0$, then $f(a)$ is not the maximum value attained by f.

(ii) If $f'(a) < 0$, then $f(a)$ is not the minimum value attained by f.

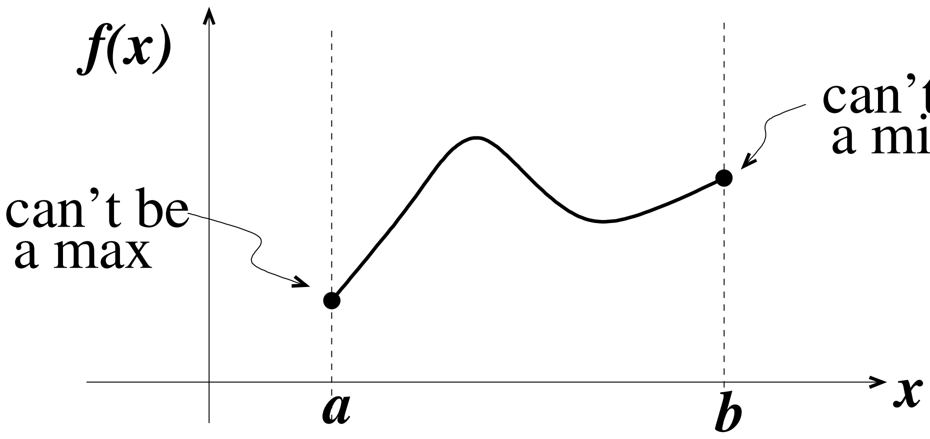
(iii) If $f'(b) > 0$, then $f(b)$ is not the minimum value attained by f.

(iv) If $f'(b) < 0$, then $f(b)$ is not the maximum value attained by f.

Proof: The proofs of each part are very similar, so we will just prove (i).
Assume $f'(a) > 0$. Assume further, towards a contradiction, that f does attain its maximum at a. Since f is differentiable at a,

$$s_n = \frac{f(x_n) - f(a)}{x_n - a} \to f'(a) > 0,$$

for any sequence (x_n) in $(a,b]$ converging to a (by Theorem 2.14). But $f(x_n) \leq f(a)$ by assumption, and $x_n > a$, so $s_n \leq 0$ for all n. Hence $f'(a) \leq 0$ by Proposition 1.7, a contradiction. $\square$



Exercise 4.5 Write out the proof of (at least) one of the other parts of Proposition 4.4.

4.2 The Mean Value Theorem

Our next theorem has hypotheses (assumptions on f) which look very restrictive – so restrictive that one might wonder whether the theorem is of any practical use at all. In fact, as we shall see, it has very powerful and useful consequences.

Theorem 4.6 (Rolle's Theorem) *Let f be continuous on $[a,b]$ and differentiable on (a,b). Assume $f(a) = f(b)$. Then there exists $c \in (a,b)$ such that $f'(c) = 0$.*

Proof: By the Extreme Value Theorem, f attains both a maximum and a minimum value on $[a,b]$. If both of these occur at the endpoints of $[a,b]$ then, since $f(a) = f(b)$, the maximum value coincides with the minimum value, whence it follows that

$f : [a, b] \to \mathbb{R}$ is constant. But then $f'(x) = 0$ for all $x \in (a, b)$, so $c = (b + a)/2$, for example, has $f'(c) = 0$.

Hence, we may assume that either the maximum or the minimum value does not occur at an endpoint. But then f attains an extremum at some interior point $c \in (a, b)$, so $f'(c) = 0$ by the Interior Extremum Theorem. □

Theorem 4.7 (The Mean Value Theorem) *Let f be 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: Let $\alpha = (f(b) - f(a))/(b - a)$ and define $g : [a, b] \to \mathbb{R}$ such that

$$g(x) = f(x) - \alpha(x - a).$$

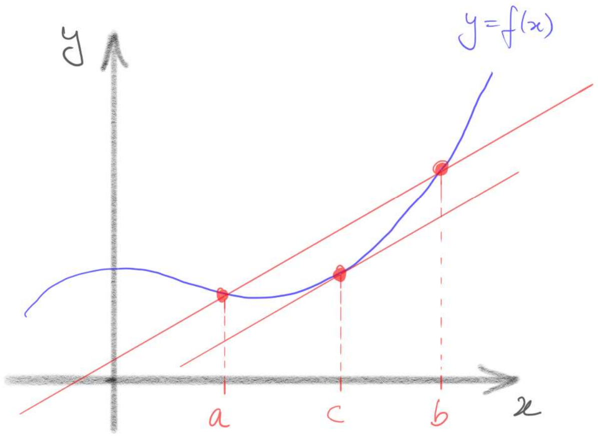
Then g is continuous, and is differentiable on (a, b). Furthermore $g(a) = f(a)$ and $g(b) = f(a)$. So g satisfies the hypotheses of Rolle's Theorem, and we deduce that there exists $c \in (a, b)$ such that $g'(c) = 0$. But

$$g'(c) = f'(c) - \alpha$$

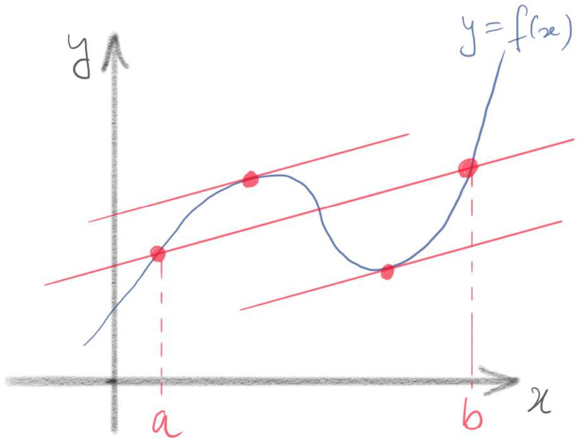
which completes the proof. □

Remarks

- The Mean Value Theorem has a nice geometric interpretation. Consider the graph $y = f(x)$ of a differentiable function $f : \mathbb{R} \to \mathbb{R}$. For each pair of distinct numbers $a < b$, we can construct the chord (straight line) passing through the points $(a, f(a))$ and $(b, f(b))$ on the graph. Its slope is $(f(b) - f(a))/(b - a)$. The Mean Value Theorem asserts that, at some point $(c, f(c))$ on the graph between these two points, the tangent line to the graph is parallel to the chord.



- The Mean Value Theorem guarantees the existence of a point c with the required $f'(c)$. It says nothing about uniqueness: there could be more than one such point.



- Clearly Rolle's Theorem is a special case of the Mean Value Theorem (geometrically, the case where the chord is horizontal). It's slightly surprising at first sight, therefore, that Rolle's Theorem actually implies the Mean Value Theorem.

You perhaps don't realize it, but you have probably been using the Mean Value Theorem rather a lot. We have seen (Example 3.2) that if a function is constant then its derivative is zero everywhere. When one solves a differential equation, one often uses the converse fact: if a function has zero derivative everywhere, then it must be constant. But how do we *know* this? It's a consequence of the Mean Value Theorem!

Example 4.8 Suppose we know that $f : \mathbb{R} \to \mathbb{R}$ is differentiable and satisfies $f'(x) = x^2$ for all x, and $f(0) = 2$. Then surely

$$f(x) = \frac{x^3}{3} + 2,$$

right? But how do we *know* this? Certainly, $F(x) := \frac{x^3}{3} + 2$ is an example of a function whose derivative at each x is x^2, and whose value at 0 is 2, so it's *possible* that $f = F$. But, perhaps there's more than one function satisfying those properties?

In fact, we can *prove* that $f = F$, as follows. Let $g : \mathbb{R} \to \mathbb{R}$ be

$$g(x) = f(x) - F(x) = f(x) - \frac{x^3}{3} - 2.$$

Then, for all x,

$$g'(x) = f'(x) - x^2 = 0.$$

Now assume, towards a contradiction, that g is not constant. Then there exist $a, b \in \mathbb{R}$ with $a < b$ such that $g(a) \neq g(b)$. But then, by the Mean Value Theorem applied to g on $[a, b]$, there exists $c \in (a, b)$ such that

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

a contradiction since $g'(x) = 0$ everywhere. Hence, g *is* constant, so for all $x \in \mathbb{R}$ $g(x) = g(0) = 0$, that is, $f = F$. $\square$

When we first learn calculus, we are taught to think of $f'(x)$ as the "rate of change of f with respect to the variable x". So if, at a particular point $a \in D$, $f'(a) > 0$, we expect that f should be strictly increasing, at least sufficiently close to a. You've already seen (Example 3.9) that this expectation can be false: the relationship between f' and the increasing/decreasing behaviour of f can be subtle. Before proceeding further, let's make absolutely precise what we mean by *increasing* and *decreasing* functions.

Definition 4.9 $f : D \to \mathbb{R}$ is **increasing** if for all $x, y \in D$ with $x < y$, $f(x) \leq f(y)$. It is **strictly increasing** if for all $x, y \in D$ with $x < y$, $f(x) < f(y)$. Similarly, f is **decreasing** if for all $x, y \in D$ with $x < y$, $f(x) \geq f(y)$, and **strictly decreasing** if for all $x, y \in D$ with $x < y$, $f(x) > f(y)$.

Example 4.10 $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^3$. This is strictly increasing: if $x < y$ then

$$\begin{aligned} f(x) - f(y) &= (x - y)(x^2 + xy + y^2) \\ &= (x - y)((x + y/2)^2 + 3y^2/4) \\ &< 0 \end{aligned}$$

since $((x + y/2)^2 + 3y^2/4) > 0$ unless $x = y = 0$.

Note that Definition 4.9 makes no mention of the derivative of f and may, in principle, apply when f is not differentiable, or even continuous. For example, the step function defined in Example 2.17, is an increasing function. Our next result uses the Mean Value Theorem to establish some useful links between f' and the increasing/decreasing behaviour of f.

Proposition 4.11 *Let $I \subseteq \mathbb{R}$ be an interval and $f : I \to \mathbb{R}$ be differentiable. Then*

(i) f is increasing if and only if $f'(x) \geq 0$ for all $x \in I$.

(ii) f is decreasing if and only if $f'(x) \leq 0$ for all $x \in I$.

(iii) f is constant if and only if $f'(x) = 0$ for all $x \in I$.

(iv) If $f'(x) > 0$ for all $x \in I$ then f is strictly increasing.

(v) If $f'(x) < 0$ for all $x \in I$ then f is strictly decreasing.

Proof:

(i) Assume $f'(x) \geq 0$ for all x in I, but that f is *not* increasing. Then there exist $a, b \in I$ with $a < b$ such that $f(a) > f(b)$. But then, by the MVT, at some $c \in (a, b)$,

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

contradicting the assumption that $f' \geq 0$. Hence f is increasing.

Conversely, assume that f is increasing. If I has a left endpoint $a \in I$, then f attains a minimum at a, so $f'(a) \geq 0$ by Proposition 4.4. Similarly, if I has a right endpoint $b \in I$ then f attains a maximum at b so $f'(b) \geq 0$ by Proposition 4.4. Let x be an interior point of I, so $x \in (a, b) \subseteq I$. Consider the sequence

$$x_n = x + \frac{b-x}{n+1} \in (x, b) \subset I.$$

This converges to x, and f is differentiable at x, so

$$s_n := \frac{f(x_n) - f(x)}{x_n - x} \to f'(x).$$

But $s_n \geq 0$ for all n, since f is increasing and $x_n > x$. Hence it's limit $f'(x) \geq 0$ by Proposition 1.7.

(ii) Exercise (just modify the proof of part (i) in the obvious way).

(iii) If $f(x) = c$ for all x then $f'(x) = 0$ (Example 3.2). Conversely, if $f'(x) = 0$ for all x, then f is increasing by part (i) and decreasing by part (ii). So for all $x, y \in I$ with $x < y$, $f(x) \leq f(y)$ and $f(x) \geq f(y)$, whence $f(x) = f(y)$. Hence f is constant.

(iv) Assume $f'(x) > 0$ for all x in I, but that f is *not* strictly increasing. Then there exist $a, b \in I$ with $a < b$ such that $f(a) \geq f(b)$. But then, by the MVT, at some $c \in (a, b)$,

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

contradicting the assumption that $f' > 0$. Hence f is strictly increasing. (Note this is almost exactly the same as the proof of the "if" direction of part (i).)

(v) Exercise (just modify the proof of part (iv) in the obvious way).

□

Remarks

- Note that the converses of parts (iv) and (v) of Proposition 4.11 are **false**! Example 4.10 is a counterexample: $f(x) = x^3$ is strictly increasing, but $f'(0) = 0$, so its derivative is not positive everywhere.

- The condition that I is an **interval** is important for Proposition 4.11. For example,

$$f : \mathbb{R}\backslash\{0\} \to \mathbb{R}, \qquad f(x) = \frac{1}{x}$$

has $f'(x) = -1/x^2 < 0$ everywhere, but $f(-1) = -1 < 1 = f(1)$, so f is not a decreasing function. This doesn't contradict Proposition 4.11 since $\mathbb{R}\backslash\{0\}$ is not an interval.

- Since strictly increasing and strictly decreasing functions are automatically injective, Proposition 4.11 gives us a sneaky way of showing that some functions are injective.

Example 4.12 Let $f : \mathbb{R} \to \mathbb{R}$ such that $f(x) = 2x + \sin x$. Then f is strictly increasing, and hence injective.

Proof: The function f is differentiable, with derivative[1]

$$f'(x) = 2 + \cos x \geq 1 > 0.$$

Hence, by Proposition 4.11, f is strictly increasing, and hence injective. □

Looking back at Example 4.8, we can see that we really used part (iii) of Proposition 4.11: since $g(x) = f(x) - F(x)$ has derivative 0 everywhere on the interval $\mathbb{R}$, it must be constant, and $g(0) = 2 - 2 = 0$, so $f(x) = F(x)$.

We finish this section by studying a function $f : \mathbb{R} \to \mathbb{R}$ with some very counter-intuitive properties.

Example 4.13 Consider the function $f : \mathbb{R} \to \mathbb{R}$ defined so that

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

Claim: f is differentiable everywhere, has $f'(0) = 1 > 0$, but is not increasing on any neighbourhood of 0.

Proof: f coincides with the function $g : U \to \mathbb{R}$, $g(x) = x + 2x^2 \sin(1/x)$ on the open set $U = \mathbb{R}\backslash\{0\}$. g is differentiable by the Product and Chain Rules, so f is differentiable on U by the Localization Lemma, and for all $a \neq 0$,

$$f'(a) = g'(a) = 1 + 4a \sin(1/a) - 2\cos(1/a). \tag{4.2}$$

Consider now the case $a = 0$. Let (x_n) be any sequence in $\mathbb{R}\backslash\{0\}$ converging to 0. Then

$$s_n := \frac{f(x_n) - f(0)}{x_n - 0} = 1 + 2x_n \sin(1/x_n).$$

Now $-|x_n| \leq x_n \sin(1/x_n) \leq |x_n|$, so $x_n \sin(1/x_n) \to 0$ by the Squeeze Rule, and hence $s_n \to 1$ (by the Algebra of Limits). Hence, f is differentiable at 0 and $f'(0) = 1$.

For the final part, assume, to the contrary, that there exists $\varepsilon > 0$ such that $f : (-\varepsilon, \varepsilon) \to \mathbb{R}$ is increasing. Then, by Proposition 4.11, $f'(x) \geq 0$ for all $x \in (-\varepsilon, \varepsilon)$. But there exists $m \in \mathbb{Z}^+$ such that $m > 1/(2\pi\varepsilon)$. Then $x = 1/(2\pi m) \in (0, \varepsilon)$ and, by equation (4.2),

$$f'(x) = 1 + 4x \times 0 - 2 \times 1 = -1 < 0,$$

a contradiction. □

Note that we proved f can't be increasing on any interval containing 0 by showing that its derivative takes negative values on every such interval. So we are again using Proposition 4.11.

[1] Strictly speaking, you haven't yet proved that sin is differentiable, with derivative cos, so this argument requires you to suspend disbelief for the time being.

4.3 The Extended Mean Value Theorem and L'Hospital's Rule

Sometimes a more general version of the Mean Value Theorem is useful.

Theorem 4.14 (Extended Mean Value Theorem) *Let f and g be real functions which are continuous on $[a, b]$ and differentiable on (a, b), and assume that, for all $x \in (a, b)$, $g'(x) \neq 0$. Then there exists $c \in (a, b)$ such that*

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

Proof: Exercise. Here are some hints:

(i) Define the constant $\alpha = \dfrac{f(b) - f(a)}{g(b) - g(a)}$. (Question: how do we know $g(b) \neq g(a)$?)

(ii) Apply Rolle's Theorem to $h(x) = f(x) - \alpha g(x)$.

If we apply Theorem 4.14 in the case where $g(x) = x$, we immediately deduce the usual Mean Value Theorem[2]. We will see that other choices of g can be useful. For

[2] Or, even better, apply it in the case $g(x) = x - a$. Then the *proof* of Theorem 4.14 reduces to the proof of Theorem 4.7.

example, we can use Theorem 4.14 to rigorously justify a popular trick for computing limits called "L'Hospital's Rule".

Theorem 4.15 (L'Hospital's Rule) *Let $I \subseteq \mathbb{R}$ be an open interval, $f, g : I \to \mathbb{R}$ be differentiable functions, $a \in I$, and $f(a) = g(a) = 0$. Assume that, for all $x \in I \backslash \{a\}$, $g(x) \neq 0$ and $g'(x) \neq 0$. Then*

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \lim_{x \to a} \frac{f'(x)}{g'(x)}$$

provided the limit on the right exists.

Proof: Let (x_n) be any sequence in $I \backslash \{a\}$ converging to a, and for each n, consider f, g restricted to the closed interval with endpoints a and x_n. These functions satisfy the hypotheses of Theorem 4.14, so there exists c_n, between a and x_n, such that

$$\frac{f'(c_n)}{g'(c_n)} = \frac{f(x_n) - f(a)}{g(x_n) - g(a)} = \frac{f(x_n)}{g(x_n)}.$$

By the Squeeze Rule, $c_n \to a$, and so, assuming the limit

$$\lim_{x \to a} \frac{f'(x)}{g'(x)}$$

exists, call it $L \in \mathbb{R}$, say, $f'(c_n)/g'(c_n) \to L$ (by the definition of limit). Hence $f(x_n)/g(x_n) \to L$, as was to be shown. □

Example 4.16 Claim: $\displaystyle\lim_{x \to 0} \frac{1 - \cos x}{\sin x} = 0.$

Proof: Let $I = (-\pi/2, \pi/2)$, $f(x) = 1 - \cos x$, $g(x) = \sin x$, and $a = 0 \in I$. Then f, g satisfy the hypotheses of L'Hospital's Rule on I.

$$\lim_{x \to 0} \frac{f'(x)}{g'(x)} = \lim_{x \to 0} \frac{\sin x}{\cos x} = \frac{0}{1} = 0.$$

Hence, by L'Hospital's Rule,

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

also. □

Example 4.17 Claim: $\displaystyle\lim_{x \to \pi} \frac{\sin x}{e^x - e^\pi} = -e^{-\pi}.$

Proof: To fit this problem into the template of Theorem 4.15, we define

$$\begin{aligned} f(x) &= \\ g(x) &= \\ I &= \\ a &= \end{aligned}$$

These satisfy the hypotheses of the theorem. Then

$$\begin{aligned} f'(x) &= \\ g'(x) &= \\ \Rightarrow \lim_{x\to\pi} \frac{f'(x)}{g'(x)} &= \\ &= \\ &= \end{aligned}$$

Hence, by Theorem 4.15,

$$\lim_{x\to\pi} \frac{\sin x}{e^x - e^\pi} = -e^{-\pi}$$

□

4.4 Higher derivatives and Taylor's Theorem

If a function $f : D \to \mathbb{R}$ is differentiable (everywhere on D) then its derivative defines another function $f' : D \to \mathbb{R}$, and we can ask whether this function, in turn, is differentiable (recall that, in general, it might not even be continuous). If f' is differentiable, then *its* derivative is denoted $f'' : D \to \mathbb{R}$ and called the *second derivative* of f. Similarly, if $f'' : D \to \mathbb{R}$ is differentiable, its derivative $f''' : D \to \mathbb{R}$ is the *third derivative* of f. Proceeding inductively, we can define the n^{th} derivative of f, denoted $f^{(n)}$ to be the derivative (if it exists), of $f^{(n-1)}$, where $f^{(0)}$ is, by definition, just the function f itself (so $f^{(1)} = f'$, $f^{(2)} = f''$ etc.). If $f^{(n)} : D \to \mathbb{R}$ exists, we say that f is n times differentiable, and if f is n times differentiable for all $n \in \mathbb{Z}^+$, we say that f is *smooth*. It is convenient to develop some terminology and notation for describing the differentiability of functions:

Definition 4.18 A function $f : D \to \mathbb{R}$ is **continuously differentiable** if it is differentiable and $f' : D \to \mathbb{R}$ is continuous. The function is n **times continuously differentiable** if all its derivatives up to the n^{th} exist everywhere on D, and are continuous. We denote the set of such functions by $C^n(D)$ and say that "f is C^n" if $f \in C^n(D)$. The function f is **smooth** if it is C^n for all $n \in \mathbb{Z}^+$. The set of smooth functions on D is denoted $C^\infty(D)$. For completeness, we denote the set of **continuous** functions $D \to \mathbb{R}$ by $C^0(D)$.

Example 4.19 (i) Every polynomial function $p : \mathbb{R} \to \mathbb{R}$, $p(x) = a_0 + a_1 x + \cdots + a_m x^m$, is smooth. This follows immediately from Proposition 3.17.

(ii) The function $f : \mathbb{R} \to \mathbb{R}$, $f(x) = |x|^3$ is C^2 but not C^3. (Exercise: prove it.)

(iii) The function $f : \mathbb{R} \to \mathbb{R}$ defined in Example 4.13 is differentiable, but not C^1 (because f' is discontinuous).

An interesting interpretation of the Mean Value Theorem can be given as follows. Let $I \subseteq \mathbb{R}$ be an open interval, $f : I \to \mathbb{R}$ be differentiable, and choose and fix some $a \in I$. Then for any other $x \in I$, we may apply the Mean Value Theorem to $f : [a, x] \to \mathbb{R}$ (if $x > a$) or $f : [x, a] \to \mathbb{R}$ (if $x < a$) to deduce that there exists c between a and x such that

$$f(x) = f(a) + f'(c)(x - a).$$

Note that the same equation holds trivially in the case where $x = a$ (with $c = x = a$). So, for x close to a (meaning that $|x - a|$ is small) we can approximate f by the constant $f(a)$, and the error in this approximation is of size $|f'(c)||x - a|$, where c is somewhere between a and x (and, of course, depends on x).

The moral is that differentiability (once) of f on an open interval allows one to approximate f by a degree 0 polynomial function (i.e. a constant), with an error controlled by f'. It turns out that if f is n times differentiable one can do better: one can approximate f by a degree $(n-1)$ polynomial function, with an error controlled by $f^{(n)}$. To prove this, we will again use the Extended Mean Value Theorem (Theorem 4.14).

Theorem 4.20 (Taylor's Theorem) *Let $I \subseteq \mathbb{R}$ be an interval, $f : I \to \mathbb{R}$ be $(n+1)$ times differentiable, and $a, x \in I$. Then there exists c between a and x such that*

$$\begin{aligned} f(x) &= f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n \\ &\quad + \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1}. \end{aligned}$$

Proof: First note that the claimed equation holds trivially (with $c = a$) in the case where $x = a$. Choose and fix $a, x \in I$, $x \neq a$, and consider the function $F : I \to \mathbb{R}$,

$$F(t) = f(x) - f(t) - f'(t)(x-t) - \frac{f''(t)}{2!}(x-t)^2 - \cdots - \frac{f^{(n)}(t)}{n!}(x-t)^n.$$

Then the claim to be proved is that there exists c between a and x such that

$$F(a) = \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1}.$$

Now, by the Product Rule, F is differentiable, and, as is easily verified (exercise!),

$$F'(t) = -\frac{f^{(n+1)}(t)}{n!}(x-t)^n.$$

Let $G(t) = (x-t)^{n+1}$. Then F, G satisfy the hypotheses of Theorem 4.14 on the interval with endpoints a, x (note that $G'(t) = 0$ only if $t = x$), so there exists c between a and x such that

$$\begin{aligned} \frac{F'(c)}{G'(c)} &= \frac{F(x)-F(a)}{G(x)-G(a)} = \frac{F(a)}{G(a)} \\ \text{i.e.} \qquad \frac{-f^{(n+1)}(c)(x-c)^n}{n![-(n+1)(x-c)^n]} &= \frac{F(a)}{(x-a)^{n+1}}, \end{aligned}$$

whence the claimed expression for $F(a)$ immediately follows. □

The polynomial

$$p_n(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n$$

is often called the n^{th} **Taylor approximant** of the function f about a, and

$$\frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1}$$

is called the **remainder**.

Example 4.21 Let $f : [0, \infty) \to \mathbb{R}$ be the function $f(x) = \sqrt{x}$. Construct the third Taylor approximant for f about 4, and hence find an approximation to $f(5) = \sqrt{5}$. Use Taylor's Theorem to find upper and lower bounds on $\sqrt{5}$.

Solution The first four derivatives of f are

$$\begin{aligned} f'(x) &= \frac{x^{-1/2}}{2}, \\ f''(x) &= -\frac{x^{-3/2}}{4}, \\ f'''(x) &= \frac{3x^{-5/2}}{8}, \\ f^{(4)}(x) &= -\frac{15x^{-7/2}}{16}, \end{aligned}$$

so $f(4) = 2$, $f'(4) = \frac{1}{4}$, $f''(4) = -2^{-5}$, and $f'''(4) = 3/2^8$. It follows that the third Taylor approximant for f about 4 is

$$\begin{aligned} p_3(x) &= f(4) + f'(4)(x-4) + \frac{f''(4)}{2!}(x-4)^2 + \frac{f'''(4)}{3!}(x-4)^3 \\ &= 2 + \frac{1}{4}(x-4) - \frac{1}{64}(x-4)^2 + \frac{1}{512}(x-4)^3. \end{aligned}$$

Using the approximation $f(x) \approx p_3(x)$ gives

$$\sqrt{5} = f(5) \approx p_3(5) = 2 + \frac{1}{4} - \frac{1}{64} + \frac{1}{512} = 2.236328125.$$

By Taylor's Theorem, the true value of $f(5)$ is

$$f(5) = p_3(5) + \frac{f^{(4)}(c)}{4!}(5-4)^4 = p_3(5) - \frac{5}{2^7 c^{7/2}}$$

for some $c \in (4, 5)$. It follows that the true value of $\sqrt{5}$ is strictly less than $p_3(5)$, but strictly greater than $p_3(5) - 5/(2^7 4^{7/2})$. That is, we have established that

$$2.236022949 < \sqrt{5} < 2.236328125.$$

For comparison, my pocket calculator tells me that $\sqrt{5} \approx 2.236067977$. $\square$

Summary

- The **Interior Extremum Theorem**: If $f : (a, b) \to \mathbb{R}$ is differentiable, and has an extremum (maximum or minimum) at $c \in (a, b)$, then $f'(c) = 0$.
- The **Mean Value Theorem**: If $f : [a, b] \to \mathbb{R}$ is continuous, and differentiable on (a, b), then there exists $c \in (a, b)$ such that
$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$
- Let $f : I \to \mathbb{R}$ be differentiable, where $I \subseteq \mathbb{R}$ is an interval. We can use the Mean Value Theorem to prove that
  - If $f'(x) = 0$ for all $x \in I$ then f is constant.
  - If $f'(x) > 0$ for all $x \in I$ then f is strictly increasing.
  - If $f'(x) < 0$ for all $x \in I$ then f is strictly decreasing.

  We often use the first of these when we solve ordinary differential equations.
- The **Extended Mean Value Theorem**: If $f : [a, b] \to \mathbb{R}$ and $g : [a, b] \to \mathbb{R}$ are continuous, and differentiable on (a, b), and $g'(x) \neq 0$ for all $x \in (a, b)$, then there exists $c \in (a, b)$ such that
$$\frac{f'(c)}{g'(c)} = \frac{f(b) - f(a)}{g(b) - g(a)}.$$
- We can use the Extended Mean Value Theorem to prove Taylor's Theorem.
- **Taylor's Theorem**: If f is $(n + 1)$ times differentiable on some open interval I and $a, x \in I$, then there exists c between a and x such that
$$f(x) = f(a) + \frac{f'(a)}{1!}(x - a) + \cdots + \frac{f^{(n)}(a)}{n!}(x - a)^n + \frac{f^{(n+1)}(c)}{n!}(x - a)^{n+1}.$$
- Taylor's Theorem allows us to approximate $(n+1)$ times differentiable functions by a polynomial of degree n, with an error which is controlled by $f^{(n+1)}$.