Chapter 4

The Derivative

4.1 The derivative

The idea of a derivative is the following. If the graph of a function looks locally like a straight line, then we can then talk about the slope of this line. The slope tells us the rate at which the value of the function is changing at that particular point. Of course, we are leaving out any function that has corners or discontinuities. Let us be precise.

4.1.1 Definition and basic properties

Definition 4.1. Let I be an interval, let $f\colon I \to \mathbb{R}$ be a function, and let $c \in I$. If the limit

$$L := \lim_{x \to c} \frac{f(x) - f(c)}{x - c}$$

exists, then we say f is *differentiable* at c, that L is the *derivative* of f at c, and write $f'(c) := L$.

If f is differentiable at all $c \in I$, then we simply say that f is *differentiable*, and then we obtain a function $f'\colon I \to \mathbb{R}$. The derivative is sometimes written as $\frac{df}{dx}$ or $\frac{d}{dx}\big(f(x)\big)$.

The expression $\frac{f(x)-f(c)}{x-c}$ is called the *difference quotient*.

The graphical interpretation of the derivative is depicted in Figure 4.1. The left-hand plot gives the line through $\big(c, f(c)\big)$ and $\big(x, f(x)\big)$ with slope $\frac{f(x)-f(c)}{x-c}$, that is, the so-called *secant line*. When we take the limit as x goes to c, we get the right-hand plot, where we see that the derivative of the function at the point c is the slope of the line tangent to the graph of f at the point $\big(c, f(c)\big)$.

We allow I to be a closed interval and we allow c to be an endpoint of I. Some calculus books do not allow c to be an endpoint of an interval, but all the theory still works by allowing it, and it makes our work easier.

Example 4.2: Let $f(x) := x^2$ defined on the whole real line. Let $c \in \mathbb{R}$ be arbitrary. We find that if $x \neq c$,

$$\frac{x^2 - c^2}{x - c} = \frac{(x + c)(x - c)}{x - c} = (x + c).$$

Therefore,

$$f'(c) = \lim_{x \to c} \frac{x^2 - c^2}{x - c} = \lim_{x \to c} (x + c) = 2c.$$

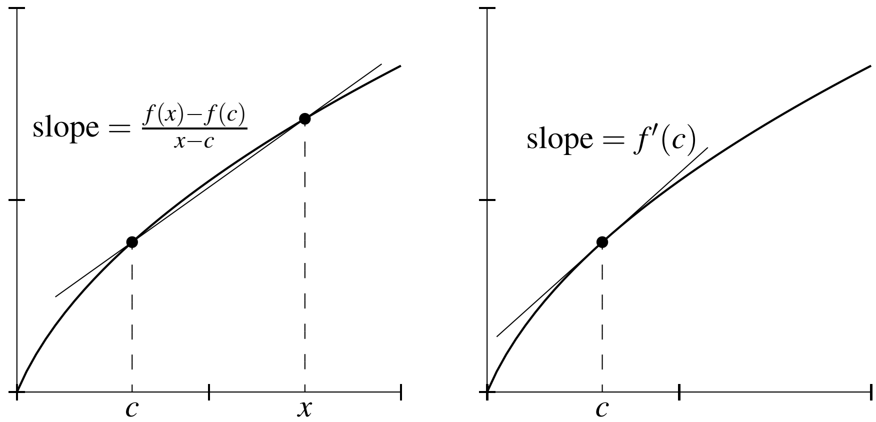

Figure 4.1: Graphical interpretation of the derivative.

Example 4.3: Let $f(x) := ax+b$ for numbers $a, b \in \mathbb{R}$. Let $c \in \mathbb{R}$ be arbitrary. For $x \neq c$,

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

Therefore,

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

In fact, every differentiable function "infinitesimally" behaves like the affine function $ax+b$. You can guess many results and formulas for derivatives, if you work them out for affine functions first.

Example 4.4: The function $f(x) := \sqrt{x}$ is differentiable for $x > 0$. To see this fact, fix $c > 0$, and take $x \neq c$, $x > 0$. Compute

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

Therefore,

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

Example 4.5: The function $f(x) := \lvert x \rvert$ is not differentiable at the origin. When $x > 0$,

$$\frac{\lvert x \rvert - \lvert 0 \rvert}{x-0} = \frac{x-0}{x-0} = 1.$$

When $x < 0$,

$$\frac{\lvert x \rvert - \lvert 0 \rvert}{x-0} = \frac{-x-0}{x-0} = -1.$$

A famous example of Weierstrass shows that there exists a continuous function that is not differentiable at *any* point. The construction of this function is beyond the scope of this chapter. On the other hand, a differentiable function is always continuous.

Proposition 4.6. *Let $f \colon I \to \mathbb{R}$ be differentiable at $c \in I$, then it is continuous at c.*

Proof. We know the limits

$$\lim_{x \to c} \frac{f(x) - f(c)}{x - c} = f'(c) \qquad \text{and} \qquad \lim_{x \to c} (x - c) = 0$$

exist. Furthermore,

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

Therefore, the limit of $f(x) - f(c)$ exists and

$$\lim_{x \to c} \bigl(f(x) - f(c) \bigr) = \left(\lim_{x \to c} \frac{f(x) - f(c)}{x - c} \right) \left(\lim_{x \to c} (x - c) \right) = f'(c) \cdot 0 = 0.$$

Hence $\lim\limits_{x \to c} f(x) = f(c)$, and f is continuous at c. □

An important property of the derivative is linearity. The derivative is the approximation of a function by a straight line. The slope of a line through two points changes linearly when the y-coordinates are changed linearly. By taking the limit, it makes sense that the derivative is linear.

Proposition 4.7. *Let I be an interval, let $f \colon I \to \mathbb{R}$ and $g \colon I \to \mathbb{R}$ be differentiable at $c \in I$, and let $\alpha \in \mathbb{R}$.*

(i) Define $h \colon I \to \mathbb{R}$ by $h(x) := \alpha f(x)$. Then h is differentiable at c and $h'(c) = \alpha f'(c)$.

(ii) Define $h \colon I \to \mathbb{R}$ by $h(x) := f(x) + g(x)$. Then h is differentiable at c and $h'(c) = f'(c) + g'(c)$.

Proof. First, let $h(x) := \alpha f(x)$. For $x \in I$, $x \neq c$,

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

The limit as x goes to c exists on the right-hand side by Corollary 3.12. We get

$$\lim_{x \to c} \frac{h(x) - h(c)}{x - c} = \alpha \lim_{x \to c} \frac{f(x) - f(c)}{x - c}.$$

Therefore, h is differentiable at c, and the derivative is computed as given.

Next, define $h(x) := f(x) + g(x)$. For $x \in I$, $x \neq c$, we have

$$\frac{h(x) - h(c)}{x - c} = \frac{\bigl(f(x) + g(x) \bigr) - \bigl(f(c) + g(c) \bigr)}{x - c} = \frac{f(x) - f(c)}{x - c} + \frac{g(x) - g(c)}{x - c}.$$

The limit as x goes to c exists on the right-hand side by Corollary 3.12. We get

$$\lim_{x \to c} \frac{h(x) - h(c)}{x - c} = \lim_{x \to c} \frac{f(x) - f(c)}{x - c} + \lim_{x \to c} \frac{g(x) - g(c)}{x - c}.$$

Therefore, h is differentiable at c, and the derivative is computed as given. □

It is not true that the derivative of a multiple of two functions is the multiple of the derivatives. Instead we get the so-called *product rule* or the *Leibniz rule*[1].

Proposition 4.8 (Product rule)**.** *Let I be an interval, let $f\colon I \to \mathbb{R}$ and $g\colon I \to \mathbb{R}$ be functions differentiable at c. If $h\colon I \to \mathbb{R}$ is defined by*

$$h(x) := f(x)g(x),$$

then h is differentiable at c and

$$h'(c) = f(c)g'(c) + f'(c)g(c).$$

The proof of the product rule is left as an exercise. The key to the proof is the identity $f(x)g(x) - f(c)g(c) = f(x)\bigl(g(x) - g(c)\bigr) + \bigl(f(x) - f(c)\bigr)g(c)$, which is illustrated in Figure 4.2.

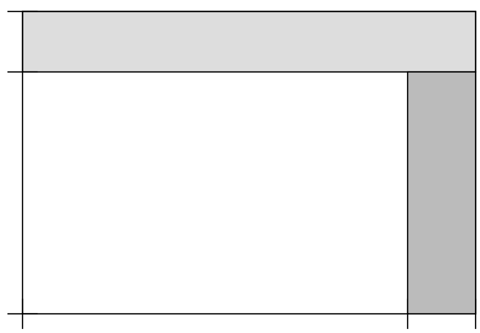

Figure 4.2: The idea of product rule. The area of the entire rectangle $f(x)g(x)$ differs from the area of the white rectangle $f(c)g(c)$ by the area of the lightly shaded rectangle $f(x)\bigl(g(x) - g(c)\bigr)$ plus the darker rectangle $\bigl(f(x) - f(c)\bigr)g(c)$. In other words, $\Delta(f \cdot g) = f \cdot \Delta g + \Delta f \cdot g$.

Proposition 4.9 (Quotient Rule)**.** *Let I be an interval, let $f\colon I \to \mathbb{R}$ and $g\colon I \to \mathbb{R}$ be differentiable at c and $g(x) \neq 0$ for all $x \in I$. If $h\colon I \to \mathbb{R}$ is defined by*

$$h(x) := \frac{f(x)}{g(x)},$$

then h is differentiable at c and

$$h'(c) = \frac{f'(c)g(c) - f(c)g'(c)}{\bigl(g(c)\bigr)^2}.$$

Again, the proof is left as an exercise.

[1]Named for the German mathematician Gottfried Wilhelm Leibniz (1646–1716).

4.1.2 Chain rule

More complicated functions are often obtained by composition, which is differentiated via the chain rule. The rule also tells us how a derivative changes if we change variables.

Proposition 4.10 (Chain Rule)**.** *Let I_1, I_2 be intervals, let $g\colon I_1 \to I_2$ be differentiable at $c \in I_1$, and $f\colon I_2 \to \mathbb{R}$ be differentiable at $g(c)$. If $h\colon I_1 \to \mathbb{R}$ is defined by*

$$h(x) := (f \circ g)(x) = f\bigl(g(x)\bigr),$$

then h is differentiable at c and

$$h'(c) = f'\bigl(g(c)\bigr)g'(c).$$

Proof. Let $d := g(c)$. Define $u\colon I_2 \to \mathbb{R}$ and $v\colon I_1 \to \mathbb{R}$ by

$$u(y) := \begin{cases} \frac{f(y)-f(d)}{y-d} & \text{if } y \neq d, \\ f'(d) & \text{if } y = d, \end{cases} \qquad v(x) := \begin{cases} \frac{g(x)-g(c)}{x-c} & \text{if } x \neq c, \\ g'(c) & \text{if } x = c. \end{cases}$$

Because f is differentiable at $d = g(c)$, we find that u is continuous at d. Similarly, v is continuous at c. For any x and y,

$$f(y) - f(d) = u(y)(y-d) \qquad \text{and} \qquad g(x) - g(c) = v(x)(x-c).$$

Plug in to obtain

$$h(x) - h(c) = f\bigl(g(x)\bigr) - f\bigl(g(c)\bigr) = u\bigl(g(x)\bigr)\bigl(g(x) - g(c)\bigr) = u\bigl(g(x)\bigr)\bigl(v(x)(x-c)\bigr).$$

Therefore, if $x \neq c$,

$$\frac{h(x) - h(c)}{x - c} = u\bigl(g(x)\bigr)v(x). \tag{4.1}$$

By continuity of u and v at d and c respectively, we find $\lim_{y\to d} u(y) = f'(d) = f'\bigl(g(c)\bigr)$ and $\lim_{x\to c} v(x) = g'(c)$. The function g is continuous at c, and so $\lim_{x\to c} g(x) = g(c)$. Hence the limit of the right-hand side of (4.1) as x goes to c exists and is equal to $f'\bigl(g(c)\bigr)g'(c)$. Thus h is differentiable at c and the limit is $f'\bigl(g(c)\bigr)g'(c)$. □

4.1.3 Exercises

Exercise* 4.1*: *Prove the product rule. Hint: Prove and use $f(x)g(x) - f(c)g(c) = f(x)\bigl(g(x) - g(c)\bigr) + \bigl(f(x) - f(c)\bigr)g(c)$.*

Exercise* 4.2*: *Prove the quotient rule. Hint: You can do this directly, but it may be easier to find the derivative of $\frac{1}{x}$ and then use the chain rule and the product rule.*

Exercise* 4.3*: *For $n \in \mathbb{Z}$, prove that x^n is differentiable and find the derivative, unless, of course, $n < 0$ and $x = 0$. Hint: Use the product rule.*

Exercise* 4.4*: *Prove that a polynomial is differentiable and find the derivative. Hint: Use the previous exercise.*

***Exercise* 4.5:** *Define* $f: \mathbb{R} \to \mathbb{R}$ *by*

$$f(x) := \begin{cases} x^2 & \text{if } x \in \mathbb{Q}, \\ 0 & \text{otherwise.} \end{cases}$$

Prove that f *is differentiable at* 0*, but discontinuous at all points except* 0*.*

***Exercise* 4.6:** *Assume the inequality* $|x - \sin(x)| \leq x^2$*. Prove that* $\sin$ *is differentiable at* 0*, and find the derivative at* 0*.*

***Exercise* 4.7:** *Using the previous exercise, prove that* $\sin$ *is differentiable at all* x *and that the derivative is* $\cos(x)$*. Hint: Use the sum-to-product trigonometric identity as we did before.*

***Exercise* 4.8:** *Let* $f: I \to \mathbb{R}$ *be differentiable. For* $n \in \mathbb{Z}$*, let* f^n *be the function defined by* $f^n(x) := \bigl(f(x)\bigr)^n$*. If* $n < 0$*, assume* $f(x) \neq 0$ *for all* $x \in I$*. Prove that* $(f^n)'(x) = n\bigl(f(x)\bigr)^{n-1} f'(x)$*.*

***Exercise* 4.9:** *Suppose* $f: \mathbb{R} \to \mathbb{R}$ *is a differentiable Lipschitz continuous function. Prove that* f' *is a bounded function.*

***Exercise* 4.10:** *Let* I_1, I_2 *be intervals. Let* $f: I_1 \to I_2$ *be a bijective function and* $g: I_2 \to I_1$ *be the inverse. Suppose that both* f *is differentiable at* $c \in I_1$ *and* $f'(c) \neq 0$ *and* g *is differentiable at* $f(c)$*. Use the chain rule to find a formula for* $g'\bigl(f(c)\bigr)$ *(in terms of* $f'(c)$*).*

***Exercise* 4.11:** *Suppose* $f: I \to \mathbb{R}$ *is bounded,* $g: I \to \mathbb{R}$ *is differentiable at* $c \in I$*, and* $g(c) = g'(c) = 0$*. Show that* $h(x) := f(x)g(x)$ *is differentiable at* c*. Hint: You cannot apply the product rule.*

***Exercise* 4.12:** *Suppose* $f: I \to \mathbb{R}$*,* $g: I \to \mathbb{R}$*, and* $h: I \to \mathbb{R}$*, are functions. Suppose* $c \in I$ *is such that* $f(c) = g(c) = h(c)$*,* g *and* h *are differentiable at* c*, and* $g'(c) = h'(c)$*. Furthermore suppose* $h(x) \leq f(x) \leq g(x)$ *for all* $x \in I$*. Prove* f *is differentiable at* c *and* $f'(c) = g'(c) = h'(c)$*.*

***Exercise* 4.13:** *Suppose* $f: (-1,1) \to \mathbb{R}$ *is a function such that* $f(x) = xh(x)$ *for a bounded function* h*.*

a) Show that $g(x) := \bigl(f(x)\bigr)^2$ *is differentiable at the origin and* $g'(0) = 0$*.*

b) Find an example of a continuous function $f: (-1,1) \to \mathbb{R}$ *with* $f(0) = 0$*, but such that* $g(x) := \bigl(f(x)\bigr)^2$ *is not differentiable at the origin.*

***Exercise* 4.14:** *Suppose* $f: I \to \mathbb{R}$ *is differentiable at* $c \in I$*. Prove there exist numbers* a *and* b *with the property that for every* $\epsilon > 0$*, there is a* $\delta > 0$*, such that* $|a + b(x-c) - f(x)| \leq \epsilon\,|x-c|$*, whenever* $x \in I$ *and* $|x - c| < \delta$*. In other words, show that there exists a function* $g: I \to \mathbb{R}$ *such that* $\lim_{x\to c} g(x) = 0$ *and* $|a + b(x-c) - f(x)| = g(x)\,|x-c|$*.*

***Exercise* 4.15:** *Prove the following simple version of L'Hôpital's rule. Suppose* $f: (a,b) \to \mathbb{R}$ *and* $g: (a,b) \to \mathbb{R}$ *are differentiable functions whose derivatives* f' *and* g' *are continuous functions. Suppose that at* $c \in (a,b)$*,* $f(c) = 0$*,* $g(c) = 0$*,* $g'(x) \neq 0$ *for all* $x \in (a,b)$*, and* $g(x) \neq 0$ *whenever* $x \neq c$*. Note that the limit of* $\frac{f'(x)}{g'(x)}$ *as* x *goes to* c *exists. Show that*

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

***Exercise* 4.16:** *Suppose* $f: (a,b) \to \mathbb{R}$ *is differentiable at* $c \in (a,b)$*,* $f(c) = 0$*, and* $f'(c) > 0$*. Prove that there is a* $\delta > 0$ *such that* $f(x) < 0$ *whenever* $c - \delta < x < c$ *and* $f(x) > 0$ *whenever* $c < x < c + \delta$*.*

4.2 Mean value theorem

4.2.1 Relative minima and maxima

We talked about absolute maxima and minima. These are the tallest peaks and lowest valleys in the whole mountain range. What about peaks of individual mountains and bottoms of individual valleys? The derivative, being a local concept, is like walking around in a fog; it can't tell you if you're on the highest peak, but it can help you find all the individual peaks.

Definition 4.11. Let $S \subset \mathbb{R}$ be a set and let $f : S \to \mathbb{R}$ be a function. The function f is said to have a *relative maximum* at $c \in S$ if there exists a $\delta > 0$ such that for all $x \in S$ where $|x - c| < \delta$, we have $f(x) \leq f(c)$. The definition of *relative minimum* is analogous.

Lemma 4.12. *Suppose $f : [a, b] \to \mathbb{R}$ is differentiable at $c \in (a, b)$, and f has a relative minimum or a relative maximum at c. Then $f'(c) = 0$.*

Proof. We prove the statement for a maximum. For a minimum the statement follows by considering the function $-f$.

Let c be a relative maximum of f. That is, there is a $\delta > 0$ such that as long as $|x - c| < \delta$, we have $f(x) - f(c) \leq 0$. We look at the difference quotient. If $c < x < c + \delta$, then

$$\frac{f(x) - f(c)}{x - c} \leq 0,$$

and if $c - \delta < y < c$, then

$$\frac{f(y) - f(c)}{y - c} \geq 0.$$

See Figure 4.3 for an illustration.

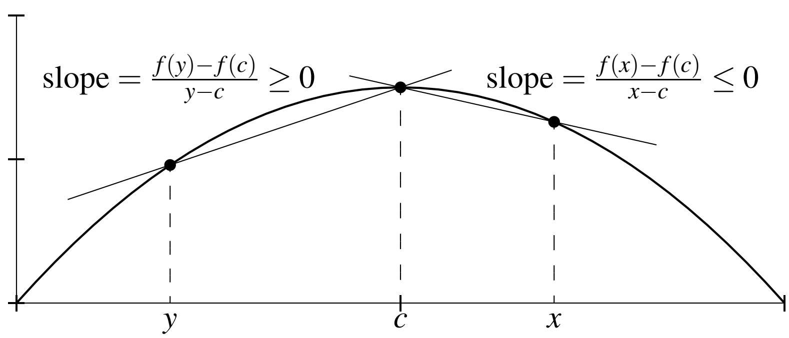

Figure 4.3: Slopes of secants at a relative maximum.

As $a < c < b$, there exist sequences $\{x_n\}$ and $\{y_n\}$ in $[a, b]$ and within δ of c, such that $x_n > c$, and $y_n < c$ for all $n \in \mathbb{N}$, and such that $\lim x_n = \lim y_n = c$. Since f is differentiable at c,

$$0 \geq \lim_{n\to\infty} \frac{f(x_n) - f(c)}{x_n - c} = f'(c) = \lim_{n\to\infty} \frac{f(y_n) - f(c)}{y_n - c} \geq 0. \qquad \square$$

For a differentiable function, a point where $f'(c) = 0$ is called a *critical point*. When f is not differentiable at some points, it is common to also say that c is a critical point if $f'(c)$ does not exist. The theorem says that a relative minimum or maximum at an interior point of an interval must be a critical point. As you remember from calculus, finding minima and maxima of a function can be done by finding all the critical points together with the endpoints of the interval and simply checking at which of these points is the function biggest or smallest.

4.2.2 Rolle's theorem

Suppose a function has the same value at both endpoints of an interval. Intuitively it ought to attain a minimum or a maximum in the interior of the interval, then at such a minimum or a maximum, the derivative should be zero. See Figure 4.4 for the geometric idea. This is the content of the so-called Rolle's theorem[2].

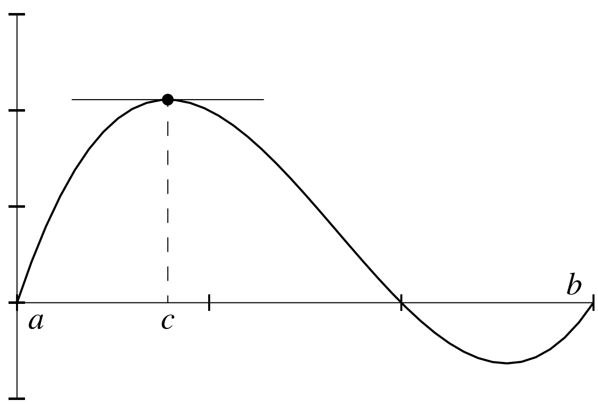

Figure 4.4: Point where the tangent line is horizontal, that is $f'(c) = 0$.

Theorem 4.13 (Rolle)**.** *Let $f \colon [a,b] \to \mathbb{R}$ be continuous function differentiable on (a,b) such that $f(a) = f(b)$. Then there exists a $c \in (a,b)$ such that $f'(c) = 0$.*

Proof. As f is continuous on $[a,b]$, it attains an absolute minimum and an absolute maximum in $[a,b]$. We wish to apply Lemma 4.12, and so we need to find some $c \in (a,b)$ where f attains a minimum or a maximum. Write $K := f(a) = f(b)$. If there exists an x such that $f(x) > K$, then the absolute maximum is bigger than K and hence occurs at some $c \in (a,b)$, and therefore $f'(c) = 0$. On the other hand, if there exists an x such that $f(x) < K$, then the absolute minimum occurs at some $c \in (a,b)$, and so $f'(c) = 0$. If there is no x such that $f(x) > K$ or $f(x) < K$, then $f(x) = K$ for all x and then $f'(x) = 0$ for all $x \in [a,b]$, so any $c \in (a,b)$ works. $\square$

It is absolutely necessary for the derivative to exist for all $x \in (a,b)$. Consider the function $f(x) := |x|$ on $[-1,1]$. Clearly $f(-1) = f(1)$, but there is no point c where $f'(c) = 0$.

[2]Named after the French mathematician Michel Rolle (1652–1719).

4.2.3 Mean value theorem

We extend Rolle's theorem to functions that attain different values at the endpoints.

Theorem 4.14 (Mean value theorem)**.** *Let $f: [a,b] \to \mathbb{R}$ be a continuous function differentiable on (a,b). Then there exists a point $c \in (a,b)$ such that*

$$f(b) - f(a) = f'(c)(b-a).$$

For a geometric interpretation of the mean value theorem, see Figure 4.5. The idea is that the value $\frac{f(b)-f(a)}{b-a}$ is the slope of the line between the points $\bigl(a, f(a)\bigr)$ and $\bigl(b, f(b)\bigr)$. Then c is the point such that $f'(c) = \frac{f(b)-f(a)}{b-a}$, that is, the tangent line at the point $\bigl(c, f(c)\bigr)$ has the same slope as the line between $\bigl(a, f(a)\bigr)$ and $\bigl(b, f(b)\bigr)$. The theorem follows from Rolle's theorem, by subtracting from f the affine linear function with the derivative $\frac{f(b)-f(a)}{b-a}$ with the same values at a and b as f. That is, we subtract the function whose graph is the straight line $\bigl(a, f(a)\bigr)$ and $\bigl(b, f(b)\bigr)$. Then we are looking for a point where this new function has derivative zero.

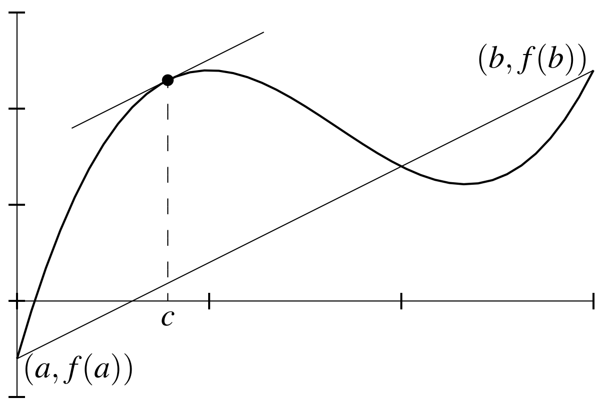

Figure 4.5: Graphical interpretation of the mean value theorem.

Proof. Define the function $g: [a,b] \to \mathbb{R}$ by

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

The function g is differentiable on (a,b), continuous on $[a,b]$, such that $g(a) = 0$ and $g(b) = 0$. Thus there exists a $c \in (a,b)$ such that $g'(c) = 0$, that is,

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

In other words, $f'(c)(b-a) = f(b) - f(a)$. $\square$

The proof generalizes. By considering $g(x) := f(x) - f(b) - \frac{f(b)-f(a)}{\varphi(b)-\varphi(a)}\bigl(\varphi(x) - \varphi(b)\bigr)$, one can prove the following version. We leave the proof as an exercise.

Theorem 4.15 (Cauchy's mean value theorem)**.** *Let $f\colon [a,b] \to \mathbb{R}$ and $\varphi\colon [a,b] \to \mathbb{R}$ be continuous functions differentiable on (a,b). Then there exists a point $c \in (a,b)$ such that*

$$\big(f(b)-f(a)\big)\varphi'(c) = f'(c)\big(\varphi(b)-\varphi(a)\big).$$

The mean value theorem has the distinction of being one of the few theorems commonly cited in court. That is, when police measure the speed of cars by aircraft, or via cameras reading license plates, they measure the time the car takes to go between two points. The mean value theorem then says that the car must have somewhere attained the speed you get by dividing the difference in distance by the difference in time.

4.2.4 Applications

We now solve our very first differential equation.

Proposition 4.16. *Let I be an interval and let $f\colon I \to \mathbb{R}$ be a differentiable function such that $f'(x) = 0$ for all $x \in I$. Then f is constant.*

Proof. Take arbitrary $x, y \in I$ with $x < y$. As I is an interval, $[x,y] \subset I$. Then f restricted to $[x,y]$ satisfies the hypotheses of the mean value theorem. Therefore, there is a $c \in (x,y)$ such that

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

As $f'(c) = 0$, we have $f(y) = f(x)$. Hence, the function is constant. □

Now that we know what it means for the function to stay constant, let us look at increasing and decreasing functions. We say $f\colon I \to \mathbb{R}$ is *increasing* (resp. *strictly increasing*) if $x < y$ implies $f(x) \leq f(y)$ (resp. $f(x) < f(y)$). We define *decreasing* and *strictly decreasing* in the same way by switching the inequalities for f.

Proposition 4.17. *Let I be an interval and let $f\colon I \to \mathbb{R}$ be a differentiable function.*

(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 $x \in I$.

Proof. Let us prove the first item. Suppose f is increasing, then for all $x, c \in I$ with $x \neq c$, we have

$$\frac{f(x)-f(c)}{x-c} \geq 0.$$

Taking a limit as x goes to c we see that $f'(c) \geq 0$.

For the other direction, suppose $f'(x) \geq 0$ for all $x \in I$. Take any $x, y \in I$ where $x < y$, and note that $[x,y] \subset I$. By the mean value theorem, there is some $c \in (x,y)$ such that

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

As $f'(c) \geq 0$, and $y - x > 0$, then $f(y) - f(x) \geq 0$ or $f(x) \leq f(y)$ and so f is increasing.

We leave the decreasing part to the reader as exercise. □

A similar but weaker statement is true for strictly increasing and decreasing functions.

Proposition 4.18. *Let I be an interval and let $f\colon I \to \mathbb{R}$ be a differentiable function.*

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

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

The proof of i is left as an exercise. Then ii follows from i by considering $-f$ instead. The converse of this proposition is not true. The function $f(x) := x^3$ is strictly increasing, but $f'(0) = 0$.

Another application of the mean value theorem is the following result about location of extrema, sometimes called the *first derivative test*. The theorem is stated for an absolute minimum and maximum. To apply it to find relative minima and maxima, restrict f to an interval $(c - \delta, c + \delta)$.

Proposition 4.19. *Let $f : (a, b) \to \mathbb{R}$ be continuous. Let $c \in (a, b)$ and suppose f is differentiable on (a, c) and (c, b).*

(i) If $f'(x) \leq 0$ for $x \in (a, c)$ and $f'(x) \geq 0$ for $x \in (c, b)$, then f has an absolute minimum at c.

(ii) If $f'(x) \geq 0$ for $x \in (a, c)$ and $f'(x) \leq 0$ for $x \in (c, b)$, then f has an absolute maximum at c.

Proof. We prove the first item and leave the second to the reader. Take $x \in (a, c)$ and a sequence $\{y_n\}$ such that $x < y_n < c$ for all n and $\lim y_n = c$. By the preceding proposition, f is decreasing on (a, c) so $f(x) \geq f(y_n)$. As f is continuous at c, we take the limit to get $f(x) \geq f(c)$ for all $x \in (a, c)$.

Similarly, take $x \in (c, b)$ and $\{y_n\}$ a sequence such that $c < y_n < x$ and $\lim y_n = c$. The function is increasing on (c, b) so $f(x) \geq f(y_n)$. By continuity of f we get $f(x) \geq f(c)$ for all $x \in (c, b)$. Thus $f(x) \geq f(c)$ for all $x \in (a, b)$. □

The converse of the proposition does not hold. See Example 4.22 below.

Another often used application of the mean value theorem you have possibly seen in calculus is the following result on differentiability at the end points of an interval. The proof is Exercise 4.29.

Proposition 4.20.

(i) Suppose $f : [a, b) \to \mathbb{R}$ is continuous, differentiable in (a, b), and $\lim_{x \to a} f'(x) = L$. Then f is differentiable at a and $f'(a) = L$.

(ii) Suppose $f : (a, b] \to \mathbb{R}$ is continuous, differentiable in (a, b), and $\lim_{x \to b} f'(x) = L$. Then f is differentiable at b and $f'(b) = L$.

In fact, using the extension result Proposition 3.48, you do not need to assume that f is defined at the end point. See Exercise 4.30.

4.2.5 Continuity of derivatives and the intermediate value theorem

Derivatives of functions satisfy an intermediate value property.

Theorem 4.21 (Darboux)**.** *Let $f\colon [a,b] \to \mathbb{R}$ be differentiable. Suppose $y \in \mathbb{R}$ is such that $f'(a) < y < f'(b)$ or $f'(a) > y > f'(b)$. Then there exists a $c \in (a,b)$ such that $f'(c) = y$.*

The proof follows by subtracting f and a linear function with derivative y. The new function g reduces the problem to the case $y = 0$, where $g'(a) > 0 > g'(b)$. That is, g is increasing at a and decreasing at b, so it must attain a maximum inside (a,b), where the derivative is zero. See Figure 4.6.

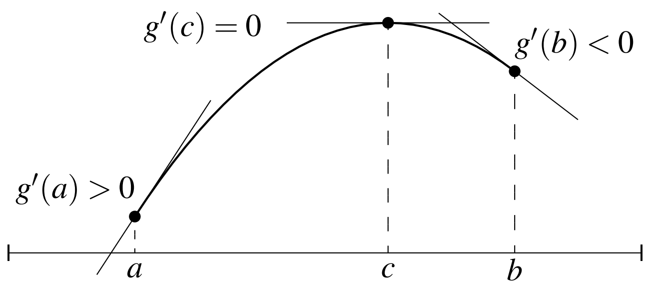

Figure 4.6: Idea of the proof of Darboux theorem.

Proof. Suppose $f'(a) < y < f'(b)$. Define

$$g(x) := yx - f(x).$$

The function g is continuous on $[a,b]$, and so g attains a maximum at some $c \in [a,b]$.

The function g is also differentiable on $[a,b]$. Compute $g'(x) = y - f'(x)$. Thus $g'(a) > 0$. As the derivative is the limit of difference quotients and is positive, there must be some difference quotient that is positive. That is, there must exist an $x > a$ such that

$$\frac{g(x) - g(a)}{x - a} > 0,$$

or $g(x) > g(a)$. Thus g cannot possibly have a maximum at a. Similarly, as $g'(b) < 0$, we find an $x < b$ (a different x) such that $\frac{g(x)-g(b)}{x-b} < 0$ or that $g(x) > g(b)$, thus g cannot possibly have a maximum at b. Therefore, $c \in (a,b)$, and Lemma 4.12 applies: As g attains a maximum at c we find $g'(c) = 0$ and so $f'(c) = y$.

Similarly, if $f'(a) > y > f'(b)$, consider $g(x) := f(x) - yx$. □

We have seen already that there exist discontinuous functions that have the intermediate value property. While it is hard to imagine at first, there also exist functions that are differentiable everywhere and the derivative is not continuous.

Example 4.22: Let $f\colon \mathbb{R} \to \mathbb{R}$ be the function defined by

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

We claim that f is differentiable everywhere, but $f' \colon \mathbb{R} \to \mathbb{R}$ is not continuous at the origin. Furthermore, f has a minimum at 0, but the derivative changes sign infinitely often near the origin. See Figure 4.7.

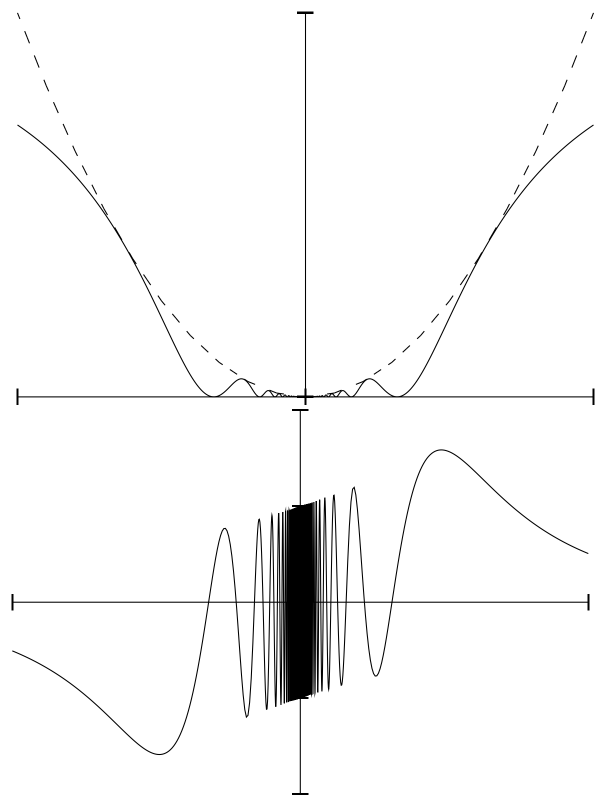

Figure 4.7: A function with a discontinuous derivative. The function f is on the left and f' is on the right. Notice that $f(x) \leq x^2$ on the left graph.

Proof: It is immediate from the definition that f has an absolute minimum at 0; we know $f(x) \geq 0$ for all x and $f(0) = 0$.

The function f is differentiable for $x \neq 0$, and the derivative is $2\sin(\frac{1}{x})\bigl(x\sin(\frac{1}{x}) - \cos(\frac{1}{x})\bigr)$. As an exercise, show that for $x_n = \frac{4}{(8n+1)\pi}$, we have $\lim f'(x_n) = -1$, and for $y_n = \frac{4}{(8n+3)\pi}$, we have $\lim f'(y_n) = 1$. Hence if f' exists at 0, then it cannot be continuous.

Let us show that f' exists at 0. We claim that the derivative is zero. In other words, $\left\lvert \frac{f(x)-f(0)}{x-0} - 0 \right\rvert$ goes to zero as x goes to zero. For $x \neq 0$,

$$\left\lvert \frac{f(x)-f(0)}{x-0} - 0 \right\rvert = \left\lvert \frac{x^2\sin^2(\frac{1}{x})}{x} \right\rvert = \left\lvert x\sin^2(\frac{1}{x}) \right\rvert \leq \lvert x \rvert .$$

And, of course, as x tends to zero, $\lvert x \rvert$ tends to zero, and hence $\left\lvert \frac{f(x)-f(0)}{x-0} - 0 \right\rvert$ goes to zero. Therefore, f is differentiable at 0 and the derivative at 0 is 0. A key point in the calculation above is that $\lvert f(x) \rvert \leq x^2$, see also Exercises 4.11 and 4.12.

It is sometimes useful to assume the derivative of a differentiable function is continuous. If $f \colon I \to \mathbb{R}$ is differentiable and the derivative f' is continuous on I, then we say f

is *continuously differentiable*. It is common to write $C^1(I)$ for the set of continuously differentiable functions on I.

4.2.6 Exercises

***Exercise* 4.17:** *Finish the proof of Proposition 4.17.*

***Exercise* 4.18:** *Finish the proof of Proposition 4.19.*

***Exercise* 4.19:** *Suppose $f\colon \mathbb{R} \to \mathbb{R}$ is a differentiable function such that f' is a bounded function. Prove that f is a Lipschitz continuous function.*

***Exercise* 4.20:** *Suppose $f\colon [a,b] \to \mathbb{R}$ is differentiable and $c \in [a,b]$. Show there exists a sequence $\{x_n\}$ converging to c, $x_n \neq c$ for all n, such that*

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

Do note this does not *imply that f' is continuous (why?).*

***Exercise* 4.21:** *Suppose $f\colon \mathbb{R} \to \mathbb{R}$ is a function such that $\lvert f(x) - f(y) \rvert \leq \lvert x-y \rvert^2$ for all x and y. Show that $f(x) = C$ for some constant C. Hint: Show that f is differentiable at all points and compute the derivative.*

***Exercise* 4.22:** *Finish the proof of Proposition 4.18. That is, suppose I is an interval and $f\colon I \to \mathbb{R}$ is a differentiable function such that $f'(x) > 0$ for all $x \in I$. Show that f is strictly increasing.*

***Exercise* 4.23:** *Suppose $f\colon (a,b) \to \mathbb{R}$ is a differentiable function such that $f'(x) \neq 0$ for all $x \in (a,b)$. Suppose there exists a point $c \in (a,b)$ such that $f'(c) > 0$. Prove $f'(x) > 0$ for all $x \in (a,b)$.*

***Exercise* 4.24:** *Suppose $f\colon (a,b) \to \mathbb{R}$ and $g\colon (a,b) \to \mathbb{R}$ are differentiable functions such that $f'(x) = g'(x)$ for all $x \in (a,b)$, then show that there exists a constant C such that $f(x) = g(x)+C$.*

***Exercise* 4.25:** *Prove the following version of L'Hôpital's rule. Suppose $f\colon (a,b) \to \mathbb{R}$ and $g\colon (a,b) \to \mathbb{R}$ are differentiable functions and $c \in (a,b)$. Suppose that $f(c) = 0$, $g(c) = 0$, $g'(x) \neq 0$ when $x \neq c$, and that the limit of $\frac{f'(x)}{g'(x)}$ as x goes to c exists. Show that*

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

Compare to Exercise 4.15. Note: Before you do anything else, prove that $g(x) \neq 0$ when $x \neq c$.

***Exercise* 4.26:** *Let $f\colon (a,b) \to \mathbb{R}$ be an unbounded differentiable function. Show $f'\colon (a,b) \to \mathbb{R}$ is unbounded.*

***Exercise* 4.27:** *Prove the theorem Rolle actually proved in 1691:* If f is a polynomial, $f'(a) = f'(b) = 0$ for some $a < b$, and there is no $c \in (a,b)$ such that $f'(c) = 0$, then there is at most one root of f in (a,b), that is at most one $x \in (a,b)$ such that $f(x) = 0$. *In other words, between any two consecutive roots of f' is at most one root of f. Hint: Suppose there are two roots and see what happens.*

Exercise* 4.28*: *Suppose $a, b \in \mathbb{R}$ and $f \colon \mathbb{R} \to \mathbb{R}$ is differentiable, $f'(x) = a$ for all x, and $f(0) = b$. Find f and prove that it is the unique differentiable function with this property.*

Exercise* 4.29*:

a) *Prove Proposition 4.20.*

b) *Suppose $f \colon (a, b) \to \mathbb{R}$ is continuous, and suppose f is differentiable everywhere except at $c \in (a, b)$ and $\lim_{x \to c} f'(x) = L$. Prove that f is differentiable at c and $f'(c) = L$.*

Exercise* 4.30*: *Suppose $f \colon (0, 1) \to \mathbb{R}$ is differentiable and f' is bounded.*

a) *Show that there exists a continuous function $g \colon [0, 1) \to \mathbb{R}$ such that $f(x) = g(x)$ for all $x \neq 0$.*
Hint: Proposition 3.48 and Exercise 4.19.

b) *Find an example where the g is not differentiable at $x = 0$.*
Hint: Consider something based on $\sin(\ln x)$, and assume you know basic properties of $\sin$ and $\ln$ from calculus.

c) *Instead of assuming that f' is bounded, assume that $\lim_{x \to 0} f'(x) = L$. Prove that not only does g exist but it is differentiable at 0 and $g'(0) = L$.*

Exercise* 4.31*: *Prove Theorem 4.15.*

4.3 Taylor's theorem

4.3.1 Derivatives of higher orders

When $f\colon I \to \mathbb{R}$ is differentiable, we obtain a function $f'\colon I \to \mathbb{R}$. The function f' is called the *first derivative* of f. If f' is differentiable, we denote by $f''\colon I \to \mathbb{R}$ the derivative of f'. The function f'' is called the *second derivative* of f. We similarly obtain f''', f'''', and so on. With a larger number of derivatives the notation would get out of hand; we denote by $f^{(n)}$ the *nth derivative* of f.

When f possesses n derivatives, we say f is *n times differentiable*.

4.3.2 Taylor's theorem

Taylor's theorem[3] is a generalization of the mean value theorem. Mean value theorem says that up to a small error $f(x)$ for x near x_0 can be approximated by $f(x_0)$, that is

$$f(x) = f(x_0) + f'(c)(x - x_0),$$

where the "error" is measured in terms of the first derivative at some point c between x and x_0. Taylor's theorem generalizes this result to higher derivatives. It tells us that up to a small error, any n times differentiable function can be approximated at a point x_0 by a polynomial. The error of this approximation behaves like $(x - x_0)^n$ near the point x_0. To see why this is a good approximation notice that for a big n, $(x - x_0)^n$ is very small in a small interval around x_0.

Definition 4.23. For an n times differentiable function f defined near a point $x_0 \in \mathbb{R}$, define the nth order *Taylor polynomial* for f at x_0 as

$$\begin{aligned} P_n^{x_0}(x) &:= \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k \\ &= f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2 + \frac{f^{(3)}(x_0)}{6}(x - x_0)^3 + \cdots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n. \end{aligned}$$

See Figure 4.8 for the odd degree Taylor polynomials for the sine function at $x_0 = 0$. The even degree terms are all zero, as even derivatives of sine are again a sine, which are zero at the origin.

Taylor's theorem says a function behaves like its nth Taylor polynomial. The mean value theorem is really Taylor's theorem for the first derivative.

Theorem 4.24 (Taylor)**.** *Suppose $f\colon [a,b] \to \mathbb{R}$ is a function with n continuous derivatives on $[a,b]$ and such that $f^{(n+1)}$ exists on (a,b). Given distinct points x_0 and x in $[a,b]$, we can find a point c between x_0 and x such that*

$$f(x) = P_n^{x_0}(x) + \frac{f^{(n+1)}(c)}{(n+1)!}(x - x_0)^{n+1}.$$

[3] Named for the English mathematician Brook Taylor (1685–1731). It was first found by the Scottish mathematician James Gregory (1638–1675). The statement we give was proved by Joseph-Louis Lagrange (1736–1813)

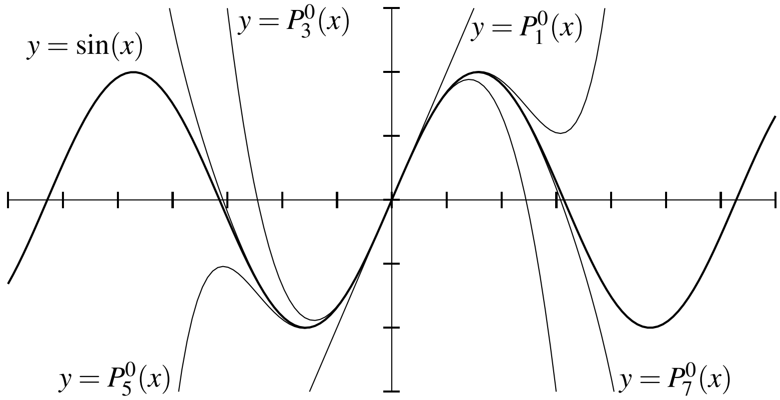

Figure 4.8: The odd degree Taylor polynomials for the sine function.

The term $R_n^{x_0}(x) := \frac{f^{(n+1)}(c)}{(n+1)!}(x-x_0)^{n+1}$ is called the *remainder term*. This form of the remainder term is called the *Lagrange form* of the remainder. There are other ways to write the remainder term, but we skip those. Note that c depends on both x and x_0.

Proof. Find a number M_{x,x_0} (depending on x and x_0) solving the equation

$$f(x) = P_n^{x_0}(x) + M_{x,x_0}(x-x_0)^{n+1}.$$

Define a function $g(s)$ by

$$g(s) := f(s) - P_n^{x_0}(s) - M_{x,x_0}(s-x_0)^{n+1}.$$

We compute the kth derivative at x_0 of the Taylor polynomial $(P_n^{x_0})^{(k)}(x_0) = f^{(k)}(x_0)$ for $k = 0,1,2,\ldots,n$ (the zeroth derivative of a function is the function itself). Therefore,

$$g(x_0) = g'(x_0) = g''(x_0) = \cdots = g^{(n)}(x_0) = 0.$$

In particular, $g(x_0) = 0$. On the other hand $g(x) = 0$. By the mean value theorem there exists an x_1 between x_0 and x such that $g'(x_1) = 0$. Applying the mean value theorem to g' we obtain that there exists x_2 between x_0 and x_1 (and therefore between x_0 and x) such that $g''(x_2) = 0$. We repeat the argument $n+1$ times to obtain a number x_{n+1} between x_0 and x_n (and therefore between x_0 and x) such that $g^{(n+1)}(x_{n+1}) = 0$.

Let $c := x_{n+1}$. We compute the $(n+1)$th derivative of g to find

$$g^{(n+1)}(s) = f^{(n+1)}(s) - (n+1)!\, M_{x,x_0}.$$

Plugging in c for s we obtain $M_{x,x_0} = \frac{f^{(n+1)}(c)}{(n+1)!}$, and we are done. □

In the proof, we have computed $(P_n^{x_0})^{(k)}(x_0) = f^{(k)}(x_0)$ for $k = 0,1,2,\ldots,n$. Therefore, the Taylor polynomial has the same derivatives as f at x_0 up to the nth derivative. That is why the Taylor polynomial is a good approximation to f. Notice how in Figure 4.8 the Taylor polynomials are reasonably good approximations to the sine near $x = 0$.

We do not necessarily get good approximations by the Taylor polynomial everywhere. Consider expanding the function $f(x) := \frac{x}{1-x}$ around 0, for $x < 1$, we get the graphs in Figure 4.9. The dotted lines are the first, second, and third degree approximations. The dashed line is the 20th degree polynomial, and yet the approximation only seems to get better with the degree for $x > -1$, and for smaller x, it in fact gets worse. The polynomials are the partial sums of the geometric series $\sum_{n=1}^{\infty} x^n$, and the series only converges on $(-1, 1)$. See the discussion of power series §2.6.

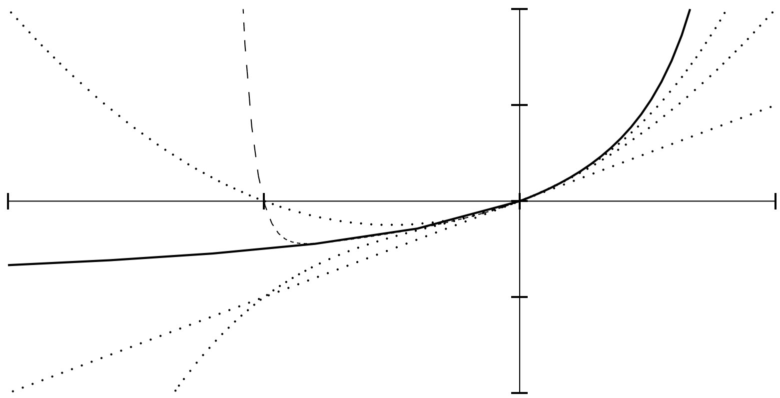

Figure 4.9: The function $\frac{x}{1-x}$, and the Taylor polynomials P_1^0, P_2^0, P_3^0 (all dotted), and the polynomial P_{20}^0 (dashed).

If f is *infinitely differentiable*, that is, if f can be differentiated any number of times, then we define the *Taylor series*:

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k.$$

There is no guarantee that this series converges for any $x \neq x_0$. And even where it does converge, there is no guarantee that it converges to the function f. Functions f whose Taylor series at every point x_0 converges to f in some open interval containing x_0 are called *analytic functions*. Most functions one tends to see in practice are analytic. See Exercise 5.54, for an example of a non-analytic function.

The definition of derivative says that a function is differentiable if it is locally approximated by a line. We mention in passing that there exists a converse to Taylor's theorem, which we will neither state nor prove, saying that if a function is locally approximated in a certain way by a polynomial of degree d, then it has d derivatives.

Taylor's theorem gives us a quick proof of a version of the second derivative test. By a *strict relative minimum* of f at c, we mean that there exists a $\delta > 0$ such that $f(x) > f(c)$ for all $x \in (c - \delta, c + \delta)$ where $x \neq c$. A *strict relative maximum* is defined similarly. Continuity of the second derivative is not needed, but the proof is more difficult and is left as an exercise. The proof also generalizes immediately into the nth derivative test, which is also left as an exercise.

Proposition 4.25 (Second derivative test)**.** *Suppose $f: (a, b) \to \mathbb{R}$ is twice continuously differentiable, $x_0 \in (a, b)$, $f'(x_0) = 0$ and $f''(x_0) > 0$. Then f has a strict relative minimum at x_0.*

Proof. As f'' is continuous, there exists a $\delta > 0$ such that $f''(c) > 0$ for all $c \in (x_0-\delta, x_0+\delta)$, see Exercise 3.27. Take $x \in (x_0 - \delta, x_0 + \delta)$, $x \neq x_0$. Taylor's theorem says that for some c between x_0 and x,

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(c)}{2}(x - x_0)^2 = f(x_0) + \frac{f''(c)}{2}(x - x_0)^2.$$

As $f''(c) > 0$, and $(x - x_0)^2 > 0$, then $f(x) > f(x_0)$. □

4.3.3 Exercises

***Exercise* 4.32:** *Compute the nth Taylor polynomial at 0 for the exponential function.*

***Exercise* 4.33:** *Suppose p is a polynomial of degree d. Given $x_0 \in \mathbb{R}$, show that the dth Taylor polynomial for p at x_0 is equal to p.*

***Exercise* 4.34:** *Let $f(x) := |x|^3$. Compute $f'(x)$ and $f''(x)$ for all x, but show that $f^{(3)}(0)$ does not exist.*

***Exercise* 4.35:** *Suppose $f\colon \mathbb{R} \to \mathbb{R}$ has n continuous derivatives. Show that for every $x_0 \in \mathbb{R}$, there exist polynomials P and Q of degree n and an $\epsilon > 0$ such that $P(x) \leq f(x) \leq Q(x)$ for all $x \in [x_0, x_0 + \epsilon]$ and $Q(x) - P(x) = \lambda(x - x_0)^n$ for some $\lambda \geq 0$.*

***Exercise* 4.36:** *If $f\colon [a, b] \to \mathbb{R}$ has $n+1$ continuous derivatives and $x_0 \in [a, b]$, prove $\lim\limits_{x\to x_0} \frac{R_n^{x_0}(x)}{(x-x_0)^n} = 0$.*

***Exercise* 4.37:** *Suppose $f\colon [a, b] \to \mathbb{R}$ has $n + 1$ continuous derivatives and $x_0 \in (a, b)$. Prove: $f^{(k)}(x_0) = 0$ for all $k = 0, 1, 2, \ldots, n$ if and only if $g(x) := \frac{f(x)}{(x-x_0)^{n+1}}$ is continuous at x_0.*

***Exercise* 4.38:** *Suppose $a, b, c \in \mathbb{R}$ and $f\colon \mathbb{R} \to \mathbb{R}$ is differentiable, $f''(x) = a$ for all x, $f'(0) = b$, and $f(0) = c$. Find f and prove that it is the unique differentiable function with this property.*

***Exercise* 4.39** (Challenging)**:** *Show that a simple converse to Taylor's theorem does not hold. Find a function $f\colon \mathbb{R} \to \mathbb{R}$ with no second derivative at $x = 0$ such that $|f(x)| \leq |x^3|$, that is, f goes to zero at 0 faster than x^3, and while $f'(0)$ exists, $f''(0)$ does not.*

***Exercise* 4.40:** *Suppose $f\colon (0, 1) \to \mathbb{R}$ is differentiable and f'' is bounded.*

a) Show that there exists a once differentiable function $g\colon [0, 1) \to \mathbb{R}$ such that $f(x) = g(x)$ for all $x \neq 0$. Hint: See Exercise 4.30.

b) Find an example where the g is not twice differentiable at $x = 0$.

***Exercise* 4.41:** *Prove the* nth derivative test. *Suppose $n \in \mathbb{N}$, $x_0 \in (a, b)$, and $f\colon (a, b) \to \mathbb{R}$ is n times continuously differentiable, with $f^{(k)}(x_0) = 0$ for $k = 1, 2, \ldots, n - 1$, and $f^{(n)}(x_0) \neq 0$. Prove:*

a) If n is odd, then f has neither a relative minimum, nor a maximum at x_0.

b) If n is even, then f has a strict relative minimum at x_0 if $f^{(n)}(x_0) > 0$ and a strict relative maximum at x_0 if $f^{(n)}(x_0) < 0$.

***Exercise* 4.42:** *Prove the more general version of the second derivative test. Suppose $f\colon (a, b) \to \mathbb{R}$ is differentiable and $x_0 \in (a, b)$ is such that, $f'(x_0) = 0$, $f''(x_0)$ exists, and $f''(x_0) > 0$. Prove that f has a strict relative minimum at x_0. Hint: Consider the limit definition of $f''(x_0)$.*

4.4 Inverse function theorem

4.4.1 Inverse function theorem

We start with a simple example. Consider the function $f(x) := ax$ for a number $a \neq 0$. Then $f\colon \mathbb{R} \to \mathbb{R}$ is bijective, and the inverse is $f^{-1}(y) = \frac{1}{a}y$. In particular, $f'(x) = a$ and $(f^{-1})'(y) = \frac{1}{a}$. As differentiable functions are "infinitesimally like" linear functions, we expect the same behavior from the inverse function. The main idea of differentiating inverse functions is the following lemma.

Lemma 4.26. *Let $I, J \subset \mathbb{R}$ be intervals. If $f\colon I \to J$ is strictly monotone (hence one-to-one), onto ($f(I) = J$), differentiable at $x_0 \in I$, and $f'(x_0) \neq 0$, then the inverse f^{-1} is differentiable at $y_0 = f(x_0)$ and*

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

If f is continuously differentiable and f' is never zero, then f^{-1} is continuously differentiable.

Proof. By Proposition 3.67, f has a continuous inverse. For convenience call the inverse $g\colon J \to I$. Let x_0, y_0 be as in the statement. For $x \in I$ write $y := f(x)$. If $x \neq x_0$ and so $y \neq y_0$, we find

$$\frac{g(y) - g(y_0)}{y - y_0} = \frac{g\bigl(f(x)\bigr) - g\bigl(f(x_0)\bigr)}{f(x) - f(x_0)} = \frac{x - x_0}{f(x) - f(x_0)}.$$

See Figure 4.10 for the geometric idea.

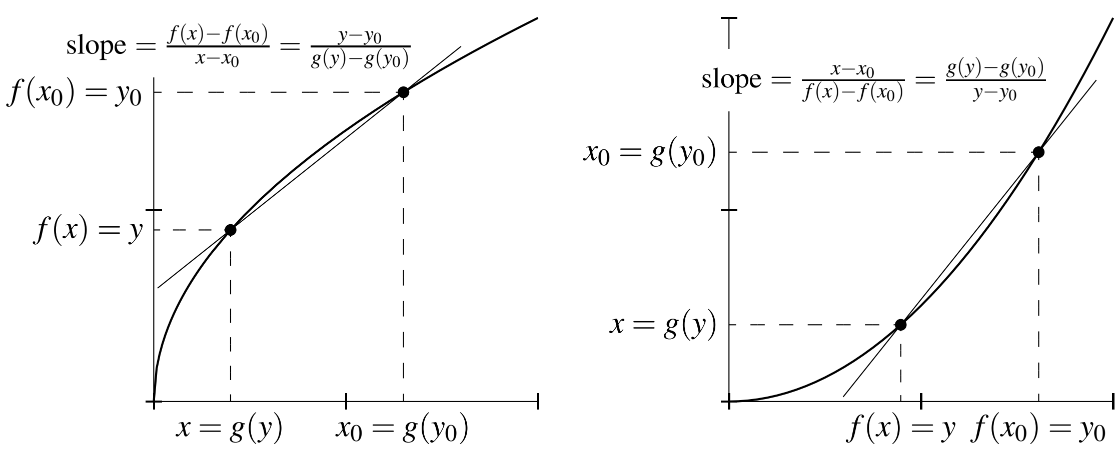

Figure 4.10: Interpretation of the derivative of the inverse function.

Let

$$Q(x) := \begin{cases} \frac{x-x_0}{f(x)-f(x_0)} & \text{if } x \neq x_0, \\ \frac{1}{f'(x_0)} & \text{if } x = x_0 \quad (\text{notice that } f'(x_0) \neq 0). \end{cases}$$

As f is differentiable at x_0,

$$\lim_{x \to x_0} Q(x) = \lim_{x \to x_0} \frac{x - x_0}{f(x) - f(x_0)} = \frac{1}{f'(x_0)} = Q(x_0),$$

that is, Q is continuous at x_0. As $g(y)$ is continuous at y_0, the composition $Q\bigl(g(y)\bigr) = \frac{g(y)-g(y_0)}{y-y_0}$ is continuous at y_0 by Proposition 3.24. Therefore,

$$\frac{1}{f'\bigl(g(y_0)\bigr)} = Q\bigl(g(y_0)\bigr) = \lim_{y \to y_0} Q\bigl(g(y)\bigr) = \lim_{y \to y_0} \frac{g(y) - g(y_0)}{y - y_0}.$$

So g is differentiable at y_0 and $g'(y_0) = \frac{1}{f'(g(y_0))}$.

If f' is continuous and nonzero at all $x \in I$, then the lemma applies at all $x \in I$. As g is also continuous (it is differentiable), the derivative $g'(y) = \frac{1}{f'(g(y))}$ must be continuous. $\square$

What is usually called the inverse function theorem is the following result.

Theorem 4.27 (Inverse function theorem)**.** *Let $f\colon (a,b) \to \mathbb{R}$ be a continuously differentiable function, $x_0 \in (a,b)$ a point where $f'(x_0) \neq 0$. Then there exists an open interval $I \subset (a,b)$ with $x_0 \in I$, the restriction $f|_I$ is injective with a continuously differentiable inverse $g\colon J \to I$ defined on an interval $J := f(I)$, and*

$$g'(y) = \frac{1}{f'\bigl(g(y)\bigr)} \qquad \textit{for all } y \in J.$$

Proof. Without loss of generality, suppose $f'(x_0) > 0$. As f' is continuous, there must exist an open interval $I = (x_0 - \delta, x_0 + \delta)$ such that $f'(x) > 0$ for all $x \in I$. See Exercise 3.27.

By Proposition 4.18, f is strictly increasing on I, and hence the restriction $f|_I$ is bijective onto $J := f(I)$. As f is continuous, then by the Corollary 3.64 (or directly via the intermediate value theorem) $f(I)$ is in interval. Now apply Lemma 4.26. $\square$

If you tried to prove the existence of roots directly as in Example 1.14, you saw how difficult that endeavor is. However, with the machinery we have built for inverse functions it becomes an almost trivial exercise, and with the lemma above we prove far more than mere existence.

Corollary 4.28. *Given $n \in \mathbb{N}$ and $x \geq 0$ there exists a unique number $y \geq 0$ (denoted $x^{1/n} := y$), such that $y^n = x$. Furthermore, the function $g\colon (0,\infty) \to (0,\infty)$ defined by $g(x) := x^{1/n}$ is continuously differentiable and*

$$g'(x) = \frac{1}{nx^{(n-1)/n}} = \frac{1}{n} x^{(1-n)/n},$$

using the convention $x^{m/n} := \bigl(x^{1/n}\bigr)^m$.

Proof. For $x = 0$ the existence of a unique root is trivial.

Let $f\colon (0,\infty) \to (0,\infty)$ be defined by $f(y) := y^n$. The function f is continuously differentiable and $f'(y) = ny^{n-1}$, see Exercise 4.3. For $y > 0$ the derivative f' is strictly positive and so again by Proposition 4.18, f is strictly increasing (this can also be proved directly). Given any $M > 1$, $f(M) = M^n \geq M$, and given any $1 > \epsilon > 0$, $f(\epsilon) =$

$\epsilon^n \leq \epsilon$. For every x with $\epsilon < x < M$, we have, by the intermediate value theorem, that $x \in f\big([\epsilon, M]\big) \subset f\big((0,\infty)\big)$. As M and ϵ were arbitrary, f is onto $(0,\infty)$, and hence f is bijective. Let g be the inverse of f, and we obtain the existence and uniqueness of positive nth roots. Lemma 4.26 says g has a continuous derivative and $g'(x) = \frac{1}{f'(g(x))} = \frac{1}{n(x^{1/n})^{n-1}}$. □

Example 4.29: The corollary provides a good example of where the inverse function theorem gives us an interval smaller than (a,b). Take $f\colon \mathbb{R} \to \mathbb{R}$ defined by $f(x) := x^2$. Then $f'(x) \neq 0$ as long as $x \neq 0$. If $x_0 > 0$, we can take $I = (0,\infty)$, but no larger.

Example 4.30: Another useful example is $f(x) := x^3$. The function $f\colon \mathbb{R} \to \mathbb{R}$ is one-to-one and onto, so $f^{-1}(y) = y^{1/3}$ exists on the entire real line including zero and negative y. The function f has a continuous derivative, but f^{-1} has no derivative at the origin. The point is that $f'(0) = 0$. See Figure 4.11 for a graph, notice the vertical tangent on the cube root at the origin. See also Exercise 4.46.

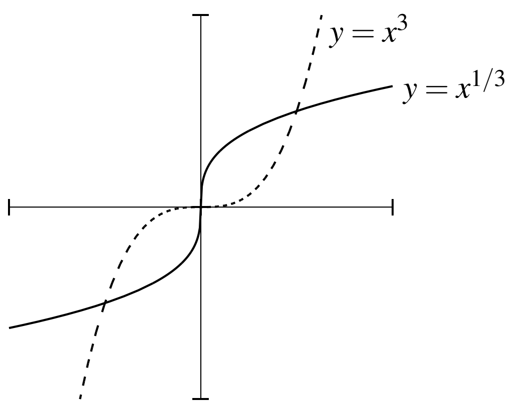

Figure 4.11: Graphs of x^3 and $x^{1/3}$.

4.4.2 Exercises

***Exercise* 4.43:** *Suppose $f\colon \mathbb{R} \to \mathbb{R}$ is continuously differentiable such that $f'(x) > 0$ for all x. Show that f is invertible on the interval $J = f(\mathbb{R})$, the inverse is continuously differentiable, and $(f^{-1})'(y) > 0$ for all $y \in f(\mathbb{R})$.*

***Exercise* 4.44:** *Suppose I, J are intervals and a monotone onto $f\colon I \to J$ has an inverse $g\colon J \to I$. Suppose you already know that both f and g are differentiable everywhere and f' is never zero. Using chain rule but not Lemma 4.26 prove the formula $g'(y) = \frac{1}{f'(g(y))}$.*

***Exercise* 4.45:** *Let $n \in \mathbb{N}$ be even. Prove that every $x > 0$ has a unique negative nth root. That is, there exists a negative number y such that $y^n = x$. Compute the derivative of the function $g(x) := y$.*

***Exercise* 4.46:** *Let $n \in \mathbb{N}$ be odd and $n \geq 3$. Prove that every x has a unique nth root. That is, there exists a number y such that $y^n = x$. Prove that the function defined by $g(x) := y$ is differentiable except at $x = 0$ and compute the derivative. Prove that g is not differentiable at $x = 0$.*

Exercise* 4.47** (requires §4.3):*** *Show that if in the inverse function theorem f has k continuous derivatives, then the inverse function g also has k continuous derivatives.*

Exercise* 4.48*: *Let $f(x) := x + 2x^2 \sin(\frac{1}{x})$ for $x \neq 0$ and $f(0) := 0$. Show that f is differentiable at all x, that $f'(0) > 0$, but that f is not invertible on any open interval containing the origin.*

Exercise* 4.49*:

a) Let $f \colon \mathbb{R} \to \mathbb{R}$ be a continuously differentiable function and $k > 0$ be a number such that $f'(x) \geq k$ for all $x \in \mathbb{R}$. Show f is one-to-one and onto, and has a continuously differentiable inverse $f^{-1} \colon \mathbb{R} \to \mathbb{R}$.

b) Find an example $f \colon \mathbb{R} \to \mathbb{R}$ where $f'(x) > 0$ for all x, but f is not onto.

Exercise* 4.50*: *Suppose I, J are intervals and a monotone onto $f \colon I \to J$ has an inverse $g \colon J \to I$. Suppose $x \in I$ and $y := f(x) \in J$, and that g is differentiable at y. Prove:*

a) If $g'(y) \neq 0$, then f is differentiable at x.

b) If $g'(y) = 0$, then f is not differentiable at x.