The notion of a function $f : X \to \mathbf{R}$ attaining a maximum or minimum at a point $x_0 \in X$ was defined in Definition 9.6.5. We now localize this definition:

Definition 10.2.1 (Local maxima and minima). Let $f : X \to \mathbf{R}$ be a function, and let $x \in X$. We say that f attains a *local maximum* at x_0 iff there exists a $\delta > 0$ such that the restriction $f|_{X\cap(x_0-\delta,x_0+\delta)}$ of f to $X \cap (x_0 - \delta, x_0 + \delta)$ attains a maximum at x_0. We say that f attains a *local minimum* at x_0 iff there exists a $\delta > 0$ such that the restriction $f|_{X\cap(x_0-\delta,x_0+\delta)}$ of f to $X \cap (x_0 - \delta, x_0 + \delta)$ attains a minimum at x_0.

Remark 10.2.2. If f attains a maximum at x_0, we sometimes say that f attains a *global* maximum at x_0, in order to distinguish it from the local maxima defined here. Note that if f attains a global maximum at x_0, then it certainly also attains a local maximum at this x_0, and similarly for minima.

Example 10.2.3. Let $f : \mathbf{R} \to \mathbf{R}$ denote the function $f(x) := x^2 - x^4$. This function does not attain a global minimum at 0, since for example $f(2) = -12 < 0 = f(0)$, however it does attain a local minimum, for if we choose $\delta := 1$ and restrict f to the interval $(-1, 1)$, then for all $x \in (-1, 1)$ we have $x^4 \leq x^2$ and thus $f(x) = x^2 - x^4 \geq 0 = f(0)$, and so $f|_{(-1,1)}$ has a local minimum at 0.

Example 10.2.4. Let $f : \mathbf{Z} \to \mathbf{R}$ be the function $f(x) = x$, defined on the integers only. Then f has no global maximum or global minimum (why?), but attains both a local maximum and local minimum at every integer n (why?).

Remark 10.2.5. If $f : X \to \mathbf{R}$ attains a local maximum at a point x_0 in X, and $Y \subset X$ is a subset of X which contains x_0, then the restriction $f|_Y : Y \to \mathbf{R}$ also attains a local maximum at x_0 (why?). Similarly for minima.

The connection between local maxima, minima and derivatives is the following.

Proposition 10.2.6 (Local extrema are stationary). *Let $a < b$ be real numbers, and let $f : (a, b) \to \mathbf{R}$ be a function. If $x_0 \in (a, b)$, f is differentiable at x_0, and f attains either a local maximum or local minimum at x_0, then $f'(x_0) = 0$.*

Proof. See Exercise 10.2.1. □

Note that f must be differentiable for this proposition to work; see Exercise 10.2.2. Also, this proposition does not work if the open interval (a, b) is replaced by a closed interval $[a, b]$. For instance, the function $f : [1, 2] \to \mathbf{R}$ defined by $f(x) := x$ has a local maximum at $x_0 = 2$ and a local minimum $x_0 = 1$ (in fact, these local extrema are global extrema), but at both points the derivative is $f'(x_0) = 1$, not $f'(x_0) = 0$. Thus the endpoints of an interval can be local maxima or minima even if the derivative is not zero there. Finally, the converse of this proposition is false (Exercise 10.2.3).

By combining Proposition 10.2.6 with the maximum principle, one can obtain

Theorem 10.2.7 (Rolle's theorem). *Let $a < b$ be real numbers, and let $g : [a, b] \to \mathbf{R}$ be a continuous function which is differentiable on (a, b). Suppose also that $g(a) = g(b)$. Then there exists an $x \in (a, b)$ such that $g'(x) = 0$.*

Proof. See Exercise 10.2.4. □

Remark 10.2.8. Note that we only assume f is differentiable on the open interval (a, b), though of course the theorem also holds if we assume f is differentiable on the closed interval $[a, b]$, since this is larger than (a, b).

Rolle's theorem has an important corollary.

Corollary 10.2.9 (Mean value theorem). *Let $a < b$ be real numbers, and let $f : [a, b] \to \mathbf{R}$ be a function which is continuous on $[a, b]$ and differentiable on (a, b). Then there exists an $x \in (a, b)$ such that $f'(x) = \frac{f(b)-f(a)}{b-a}$.*

Proof. See Exercise 10.2.5. □

Exercise 10.2.1. Prove Proposition 10.2.6.

Exercise 10.2.2. Give an example of a function $f : (-1, 1) \to \mathbf{R}$ which is continuous and attains a global maximum at 0, but which is not differentiable at 0. Explain why this does not contradict Proposition 10.2.6.

Exercise 10.2.3. Give an example of a function $f : (-1, 1) \to \mathbf{R}$ which is differentiable, and whose derivative equals 0 at 0, but such that 0 is neither a local minimum nor a local maximum. Explain why this does not contradict Proposition 10.2.6.

Exercise 10.2.4. Prove Theorem 10.2.7. (Hint: use Corollary 10.1.12 and the maximum principle, Proposition 9.6.7, followed by Proposition 10.2.6. Note that the maximum principle does not tell you whether the maximum or minimum is in the open interval (a, b) or is one of the boundary points a, b, so you have to divide into cases and use the hypothesis $g(a) = g(b)$ somehow.)

Exercise 10.2.5. Use Theorem 10.2.7 to prove Corollary 10.2.9. (Hint: consider a function of the form $f(x) - cx$ for some carefully chosen real number c.)

Exercise 10.2.6. Let $M > 0$, and let $f : [a, b] \to \mathbf{R}$ be a function which is continuous on $[a, b]$ and differentiable on (a, b), and such that $|f'(x)| \leq M$ for all $x \in (a, b)$ (i.e., the derivative of f is bounded). Show that for any $x, y \in [a, b]$ we have the inequality $|f(x) - f(y)| \leq M|x - y|$. (Hint: apply the mean value theorem (Corollary 10.2.9) to a suitable restriction of f.) Functions which obey the bound $|f(x) - f(y)| \leq M|x - y|$ are known as *Lipschitz continuous functions* with *Lipschitz constant* M; thus this exercise shows that functions with bounded derivative are Lipschitz continuous.

Exercise 10.2.7. Let $f : \mathbf{R} \to \mathbf{R}$ be a differentiable function such that f' is bounded. Show that f is uniformly continuous. (Hint: use the preceding exercise.)

10.3 Monotone functions and derivatives

In your elementary calculus courses, you may have come across the assertion that a positive derivative meant an increasing function, and a negative derivative meant a decreasing function. This statement is not completely accurate, but it is pretty close; we now give the precise version of these statements below.

Proposition 10.3.1. *Let X be a subset of $\mathbf{R}$, let x_0 be a limit point of X, and let $f : X \to \mathbf{R}$ be a function. If f is monotone increasing and f is differentiable at x_0, then $f'(x_0) \geq 0$. If f is monotone decreasing and f is differentiable at x_0, then $f'(x_0) \leq 0$.*

Proof. See Exercise 10.3.1. □

Remark 10.3.2. We have to assume that f is differentiable at x_0; There exist monotone functions which are not always differentiable (see Exercise 10.3.2), and of course if f is not differentiable at x_0 we cannot possibly conclude that $f'(x_0) \geq 0$ or $f'(x_0) \leq 0$.

One might naively guess that if f were *strictly* monotone increasing, and f was differentiable at x_0, then the derivative $f'(x_0)$ would be strictly positive instead of merely non-negative. Unfortunately, this is not always the case (Exercise 10.3.3).

On the other hand, we do have a converse result: if function has strictly positive derivative, then it must be strictly monotone increasing:

Proposition 10.3.3. *Let $a < b$, and let $f : [a, b] \to \mathbf{R}$ be a differentiable function. If $f'(x) > 0$ for all $x \in [a, b]$, then f is strictly monotone increasing. If $f'(x) < 0$ for all $x \in [a, b]$, then f is strictly monotone decreasing. If $f'(x) = 0$ for all $x \in [a, b]$, then f is a constant function.*

Proof. See Exercise 10.3.4. □

Exercise 10.3.1. Prove Proposition 10.3.1.

Exercise 10.3.2. Give an example of a function $f : (-1, 1) \to \mathbf{R}$ which is continuous and monotone increasing, but which is not differentiable at 0. Explain why this does not contradict Proposition 10.3.1.

Exercise 10.3.3. Give an example of a function $f : \mathbf{R} \to \mathbf{R}$ which is strictly monotone increasing and differentiable, but whose derivative at 0 is zero. Explain why this does not contradict Proposition 10.3.1 or Proposition 10.3.3. (Hint: look at Exercise 10.2.3.)

Exercise 10.3.4. Prove Proposition 10.3.3. (Hint: you do not have integrals or the fundamental theorem of calculus yet, so these tools cannot be used. However, one can proceed via the mean-value theorem, Corollary 10.2.9.)

Exercise 10.3.5. Give an example of a subset $X \subset \mathbf{R}$ and a function $f : X \to \mathbf{R}$ which is differentiable on X, is such that $f'(x) > 0$ for all $x \in X$, but f is not strictly monotone increasing. (Hint: the conditions here are subtly different from those in Proposition 10.3.3. What is the difference, and how can one exploit that difference to obtain the example?)

10.4 Inverse functions and derivatives

We now ask the following question: if we know that a function $f : X \to Y$ is differentiable, and it has an inverse $f^{-1} : Y \to X$, what can we say about the differentiability of f^{-1}? This will be useful for many applications, for instance if we want to differentiate the function $f(x) := x^{1/n}$.

We begin with a preliminary result.

Lemma 10.4.1. *Let $f : X \to Y$ be an invertible function, with inverse $f^{-1} : Y \to X$. Suppose that $x_0 \in X$ and $y_0 \in Y$ are such that $y_0 = f(x_0)$ (which also implies that $x_0 = f^{-1}(y_0)$). If f is differentiable at x_0, and f^{-1} is differentiable at y_0, then*

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

Proof. From the chain rule (Theorem 10.1.15) we have

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

But $f^{-1} \circ f$ is the identity function on X, and hence by Theorem 10.1.13(b) $(f^{-1} \circ f)'(x_0) = 1$. The claim follows. $\square$

As a particular corollary of Lemma 10.4.1, we see that if f is differentiable at x_0 with $f'(x_0) = 0$, then f^{-1} cannot be differentiable at $y_0 = f(x_0)$, since $1/f'(x_0)$ is undefined in that case. Thus for instance, the function $g : [0, \infty) \to [0, \infty)$ defined by $g(y) := y^{1/3}$ cannot be differentiable at 0, since this function is the inverse $g = f^{-1}$ of the function $f : [0, \infty) \to [0, \infty)$ defined by $f(x) := x^3$, and this function has a derivative of 0 at $f^{-1}(0) = 0$.

If one writes $y = f(x)$, so that $x = f^{-1}(y)$, then one can write the conclusion of Lemma 10.4.1 in the more appealing form $dx/dy = 1/(dy/dx)$. However, as mentioned before, this way of writing things, while very convenient and easy to remember, can be misleading and cause errors if applied too carelessly (especially when one begins to work in the calculus of several variables).

Lemma 10.4.1 seems to answer the question of how to differentiate the inverse of a function, however it has one significant drawback: the lemma only works if one assumes *a priori* that f^{-1} is differentiable. Thus, if one does not already know that f^{-1} is differentiable, one cannot use Lemma 10.4.1 to compute the derivative of f^{-1}.

However, the following improved version of Lemma 10.4.1 will compensate for this fact, by relaxing the requirement on f^{-1} from differentiability to continuity.

Theorem 10.4.2 (Inverse function theorem). *Let $f : X \to Y$ be an invertible function, with inverse $f^{-1} : Y \to X$. Suppose that $x_0 \in X$ and $y_0 \in Y$ are such that $f(x_0) = y_0$. If f is differentiable at x_0, f^{-1} is continuous at y_0, and $f'(x_0) \neq 0$, then f^{-1} is differentiable at y_0 and*

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

Proof. We have to show that

$$\lim_{y \to y_0; y \in Y - \{y_0\}} \frac{f^{-1}(y) - f^{-1}(y_0)}{y - y_0} = \frac{1}{f'(x_0)}.$$

By Proposition 9.3.9, it suffices to show that

$$\lim_{n \to \infty} \frac{f^{-1}(y_n) - f^{-1}(y_0)}{y_n - y_0} = \frac{1}{f'(x_0)}$$

for any sequence $(y_n)_{n=1}^{\infty}$ of elements in $Y - \{y_0\}$ which converge to y_0.

To prove this, we set $x_n := f^{-1}(y_n)$. Then $(x_n)_{n=1}^{\infty}$ is a sequence of elements in $X - \{x_0\}$. (Why? Note that f^{-1} is a bijection) Since f^{-1} is continuous by assumption, we know that $x_n = f^{-1}(y_n)$ converges to $f^{-1}(y_0) = x_0$ as $n \to \infty$. Thus, since f is differentiable at x_0, we have (by Proposition 9.3.9 again)

$$\lim_{n\to\infty} \frac{f(x_n) - f(x_0)}{x_n - x_0} = f'(x_0).$$

But since $x_n \neq x_0$ and f is a bijection, the fraction $\frac{f(x_n)-f(x_0)}{x_n-x_0}$ is non-zero. Also, by hypothesis $f'(x_0)$ is non-zero. So by limit laws

$$\lim_{n\to\infty} \frac{x_n - x_0}{f(x_n) - f(x_0)} = \frac{1}{f'(x_0)}.$$

But since $x_n = f^{-1}(y_0)$ and $x_0 = f^{-1}(y_0)$, we thus have

$$\lim_{n\to\infty} \frac{f^{-1}(y_n) - f^{-1}(y_0)}{y_n - y_0} = \frac{1}{f'(x_0)}$$

as desired. □

We give some applications of the inverse function theorem in the exercises below.

Exercise 10.4.1. Let $n \geq 1$ be a natural number, and let $g : (0, \infty) \to (0, \infty)$ be the function $g(x) := x^{1/n}$.

(a) Show that g is continuous on $(0, \infty)$. (Hint: use Proposition 9.8.3.)

(b) Show that g is differentiable on $(0, \infty)$, and that $g'(x) = \frac{1}{n}x^{\frac{1}{n}-1}$ for all $x \in (0, \infty)$. (Hint: use the inverse function theorem and (a).)

Exercise 10.4.2. Let q be a rational number, and let $f : (0, \infty) \to \mathbf{R}$ be the function $f(x) = x^q$.

(a) Show that f is differentiable on $(0, \infty)$ and that $f'(x) = qx^{q-1}$. (Hint: use Exercise 10.4.1 and the laws of differential calculus in Theorem 10.1.13 and Theorem 10.1.15.)

(b) Show that $\lim_{x\to 1; x\in(0,\infty)} \frac{x^q-1}{x-1} = q$ for every rational number q. (Hint: use part (a) and Definition 10.1.1. An alternate route is to apply L'Hôpital's rule from the next section.)

Exercise 10.4.3. Let α be a real number, and let $f : (0,\infty) \to \mathbf{R}$ be the function $f(x) = x^\alpha$.

(a) Show that $\lim_{x\to 1; x\in(0,\infty)} \frac{f(x)-f(1)}{x-1} = \alpha$. (Hint: use Exercise 10.4.2 and the comparison principle; you may need to consider right and left limits separately. Proposition 5.4.14 may also be helpful.)

(b) Show that f is differentiable on $(0,\infty)$ and that $f'(x) = \alpha x^{\alpha-1}$. (Hint: use (a), exponent laws (Proposition 6.7.3), and Definition 10.1.1.)

10.5 L'Hôpital's rule

Finally, we present a version of a rule you are all familiar with.

Proposition 10.5.1 (L'Hôpital's rule I). *Let X be a subset of $\mathbf{R}$, let $f : X \to \mathbf{R}$ and $g : X \to \mathbf{R}$ be functions, and let x_0 be a limit point of X. Suppose that $f(x_0) = g(x_0) = 0$, that f and g are both differentiable at x_0, but $g'(x_0) \neq 0$. Then there exists a $\delta > 0$ such that $g(x) \neq 0$ for all $x \in (X \cap (x_0 - \delta, x_0 + \delta)) - \{x_0\}$, and*

$$\lim_{x\to x_0; x\in(X\cap(x_0-\delta,x_0+\delta))-\{x_0\}} \frac{f(x)}{g(x)} = \frac{f'(x_0)}{g'(x_0)}.$$

Proof. See Exercise 10.5.1. □

The presence of the δ here may seem somewhat strange, but is needed because $g(x)$ might vanish at some points other than x_0, which would imply that quotient $\frac{f(x)}{g(x)}$ is not necessarily defined at all points in $X - \{x_0\}$.

A more sophisticated version of L'Hôpital's rule is the following.

Proposition 10.5.2 (L'Hôpital's rule II). *Let $a < b$ be real numbers, let $f : [a,b] \to \mathbf{R}$ and $g : [a,b] \to \mathbf{R}$ be functions which are differentiable on $[a,b]$. Suppose that $f(a) = g(a) = 0$, that*

g' is non-zero on $[a, b]$ (i.e., $g'(x) \neq 0$ for all $x \in [a, b]$), and $\lim_{x \to a; x \in (a,b]} \frac{f'(x)}{g'(x)}$ exists and equals L. Then $g(x) \neq 0$ for all $x \in (a, b]$, and $\lim_{x \to a; x \in (a,b]} \frac{f(x)}{g(x)}$ exists and equals L.

Remark 10.5.3. This proposition only considers limits to the right of a, but one can easily state and prove a similar proposition for limits to the left of a, or around both sides of a. Speaking very informally, the proposition states that

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

though one has to ensure all of the conditions of the proposition hold (in particular, that $f(a) = g(a) = 0$, and that the right-hand limit exists), before one can apply L'Hôpital's rule.

Proof. (Optional) We first show that $g(x) \neq 0$ for all $x \in (a, b]$. Suppose for sake of contradiction that $g(x) = 0$ for some $x \in (a, b]$. But since $g(a)$ is also zero, we can apply Rolle's theorem to obtain $g'(y) = 0$ for some $a < y < x$, but this contradicts the hypothesis that g' is non-zero on $[a, b]$.

Now we show that $\lim_{x \to a; x \in (a,b]} \frac{f(x)}{g(x)} = L$. By Proposition 9.3.9, it will suffice to show that

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

for any sequence $(x_n)_{n=1}^{\infty}$ taking values in $(a, b]$ which converges to x.

Consider a single x_n, and consider the function $h_n : [a, x_n] \to \mathbf{R}$ defined by

$$h_n(x) := f(x)g(x_n) - g(x)f(x_n).$$

Observe that h_n is continuous on $[a, x_n]$ and equals 0 at both a and x_n, and is differentiable on (a, x_n) with derivative $h_n'(x) = f'(x)g(x_n) - g'(x)f(x_n)$. (Note that $f(x_n)$ and $g(x_n)$ are constants with respect to x.) By Rolle's theorem (Theorem 10.2.7), we can thus find $y_n \in (a, x_n)$ such that $h_n'(y_n) = 0$, which implies that

$$\frac{f(x_n)}{g(x_n)} = \frac{f'(y_n)}{g'(y_n)}.$$

Since $y_n \in (a, x_n)$ for all n, and x_n converges to a as $n \to \infty$, we see from the squeeze test (Corollary 6.4.14) that y_n also converges to a as $n \to \infty$. Thus $\frac{f'(y_n)}{g'(y_n)}$ converges to L, and thus $\frac{f(x_n)}{g(x_n)}$ also converges to L, as desired. $\square$

Exercise 10.5.1. Prove Proposition 10.5.1. (Hint: to show that $g(x) \neq 0$ near x_0, you may wish to use Newton's approximation (Proposition 10.1.7). For the rest of the proposition, use limit laws, Proposition 9.3.14.)

Exercise 10.5.2. Explain why Exercise 1.2.12 does not contradict either of the propositions in this section.

Chapter 11

The Riemann integral

In the previous chapter we reviewed *differentiation* - one of the two pillars of single variable calculus. The other pillar is, of course, *integration*, which is the focus of the current chapter. More precisely, we will turn to the *definite integral*, the integral of a function on a fixed interval, as opposed to the *indefinite integral*, otherwise known as the antiderivative. These two are of course linked by the *Fundamental theorem of calculus*, of which more will be said later.

For us, the study of the definite integral will start with an interval I which could be open, closed, or half-open, and a function $f : I \to \mathbf{R}$, and will lead us to a number $\int_I f$; we can write this integral as $\int_I f(x)\, dx$ (of course, we could replace x by any other dummy variable), or if I has endpoints a and b, we shall also write this integral as $\int_a^b f$ or $\int_a^b f(x)\, dx$.

To actually *define* this integral $\int_I f$ is somewhat delicate (especially if one does not want to assume any axioms concerning geometric notions such as area), and not all functions f are integrable. It turns out that there are at least two ways to define this integral: the *Riemann integral*, named after Georg Riemann (1826–1866), which we will do here and which suffices for most applications, and the *Lebesgue integral*, named after Henri Lebesgue (1875–1941), which supercedes the Riemann integral and works for a much larger class of functions. The Lebesgue integral will be constructed in Chapter 19. There is also the *Riemann-Steiltjes in-*

tegral $\int_I f(x)\, d\alpha(x)$, a generalization of the Riemann integral due to Thomas Stieltjes (1856–1894), which we will discuss in Section 11.8.

Our strategy in defining the Riemann integral is as follows. We begin by first defining a notion of integration on a very simple class of functions - the *piecewise constant* functions. These functions are quite primitive, but their advantage is that integration is very easy for these functions, as is verifying all the usual properties. Then, we handle more general functions by approximating them by piecewise constant functions.

11.1 Partitions

Before we can introduce the concept of an integral, we need to describe how one can partition a large interval into smaller intervals. In this chapter, all intervals will be bounded intervals (as opposed to the more general intervals defined in Definition 9.1.1).

Definition 11.1.1. Let X be a subset of $\mathbf{R}$. We say that X is *connected* iff the following property is true: whenever x, y are elements in X such that $x < y$, the bounded interval $[x, y]$ is a subset of X (i.e., every number between x and y is also in X).

Remark 11.1.2. Later on, in Section 13.4 we will define a more general notion of connectedness, which applies to any metric space.

Examples 11.1.3. The set $[1, 2]$ is connected, because if $x < y$ both lie in $[1, 2]$, then $1 \leq x < y \leq 2$, and so every element between x and y also lies in $[1, 2]$. A similar argument shows that the set $(1, 2)$ is connected. However, the set $[1, 2] \cup [3, 4]$ is not connected (why?). The real line is connected (why?). The empty set, as well as singleton sets such as $\{3\}$, are connected, but for rather trivial reasons (these sets do not contain two elements x, y for which $x < y$).

Lemma 11.1.4. *Let X be a subset of the real line. Then the following two statements are logically equivalent:*

(*a*) *X is bounded and connected.*

(*b*) *X is a bounded interval.*

Proof. See Exercise 11.1.1. □

Remark 11.1.5. Recall that intervals are allowed to be singleton points (e.g., the degenerate interval $[2,2] = \{2\}$), or even the empty set.

Corollary 11.1.6. *If I and J are bounded intervals, then the intersection $I \cap J$ is also a bounded interval.*

Proof. See Exercise 11.1.2. □

Example 11.1.7. The intersection of the bounded intervals $[2,4]$ and $[4,6]$ is $\{4\}$, which is also a bounded interval. The intersection of $(2,4)$ and $(4,6)$ is $\emptyset$.

We now give each bounded interval a length.

Definition 11.1.8 (Length of intervals). If I is a bounded interval, we define the *length* of I, denoted $|I|$ as follows. If I is one of the intervals $[a,b]$, (a,b), $[a,b)$, or $(a,b]$ for some real numbers $a < b$, then we define $|I| := b - a$. Otherwise, if I is a point or the empty set, we define $|I| = 0$.

Example 11.1.9. For instance, the length of $[3,5]$ is 2, as is the length of $(3,5)$; meanwhile, the length of $\{5\}$ or the empty set is 0.

Definition 11.1.10 (Partitions). Let I be a bounded interval. A *partition* of I is a finite set $\mathbf{P}$ of bounded intervals contained in I, such that every x in I lies in exactly one of the bounded intervals J in $\mathbf{P}$.

Remark 11.1.11. Note that a partition is a set of intervals, while each interval is itself a set of real numbers. Thus a partition is a set consisting of other sets.