Chapter 4

Differentiable Functions

A differentiable function is a function that can be approximated locally by a linear function.

4.1. The derivative

Definition 4.1. Suppose that $f : (a, b) \to \mathbb{R}$ and $a < c < b$. Then f is differentiable at c with derivative $f'(c)$ if

$$\lim_{h \to 0} \left[\frac{f(c+h) - f(c)}{h} \right] = f'(c).$$

The domain of f' is the set of points $c \in (a, b)$ for which this limit exists. If the limit exists for every $c \in (a, b)$ then we say that f is differentiable on (a, b).

Graphically, this definition says that the derivative of f at c is the slope of the tangent line to $y = f(x)$ at c, which is the limit as $h \to 0$ of the slopes of the lines through $(c, f(c))$ and $(c + h, f(c + h))$.

We can also write

$$f'(c) = \lim_{x \to c} \left[\frac{f(x) - f(c)}{x - c} \right],$$

since if $x = c + h$, the conditions $0 < |x - c| < \delta$ and $0 < |h| < \delta$ in the definitions of the limits are equivalent. The ratio

$$\frac{f(x) - f(c)}{x - c}$$

is undefined $(0/0)$ at $x = c$, but it doesn't have to be defined in order for the limit as $x \to c$ to exist.

Like continuity, differentiability is a local property. That is, the differentiability of a function f at c and the value of the derivative, if it exists, depend only the values of f in a arbitrarily small neighborhood of c. In particular if $f : A \to \mathbb{R}$

where $A \subset \mathbb{R}$, then we can define the differentiability of f at any interior point $c \in A$ since there is an open interval $(a, b) \subset A$ with $c \in (a, b)$.

4.1.1. Examples of derivatives. Let us give a number of examples that illustrate differentiable and non-differentiable functions.

Example 4.2. The function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^2$ is differentiable on $\mathbb{R}$ with derivative $f'(x) = 2x$ since

$$\lim_{h\to 0}\left[\frac{(c+h)^2 - c^2}{h}\right] = \lim_{h\to 0}\frac{h(2c+h)}{h} = \lim_{h\to 0}(2c+h) = 2c.$$

Note that in computing the derivative, we first cancel by h, which is valid since $h \neq 0$ in the definition of the limit, and then set $h = 0$ to evaluate the limit. This procedure would be inconsistent if we didn't use limits.

Example 4.3. The function $f : \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} x^2 & \text{if } x > 0, \\ 0 & \text{if } x \leq 0. \end{cases}$$

is differentiable on $\mathbb{R}$ with derivative

$$f'(x) = \begin{cases} 2x & \text{if } x > 0, \\ 0 & \text{if } x \leq 0. \end{cases}$$

For $x > 0$, the derivative is $f'(x) = 2x$ as above, and for $x < 0$, we have $f'(x) = 0$. For 0,

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

The right limit is

$$\lim_{h\to 0^+}\frac{f(h)}{h} = \lim_{h\to 0} h = 0,$$

and the left limit is

$$\lim_{h\to 0^-}\frac{f(h)}{h} = 0.$$

Since the left and right limits exist and are equal, so does the limit

$$\lim_{h\to 0}\left[\frac{f(h) - f(0)}{h}\right] = 0,$$

and f is differentiable at 0 with $f'(0) = 0$.

Next, we consider some examples of non-differentiability at discontinuities, corners, and cusps.

Example 4.4. The function $f : \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} 1/x & \text{if } x \neq 0, \\ 0 & \text{if } x = 0, \end{cases}$$

is differentiable at $x \neq 0$ with derivative $f'(x) = -1/x^2$ since

$$\begin{aligned} \lim_{h\to 0}\left[\frac{f(c+h)-f(c)}{h}\right] &= \lim_{h\to 0}\left[\frac{1/(c+h)-1/c}{h}\right] \\ &= \lim_{h\to 0}\left[\frac{c-(c+h)}{hc(c+h)}\right] \\ &= -\lim_{h\to 0}\frac{1}{c(c+h)} \\ &= -\frac{1}{c^2}. \end{aligned}$$

However, f is not differentiable at 0 since the limit

$$\lim_{h\to 0}\left[\frac{f(h)-f(0)}{h}\right] = \lim_{h\to 0}\left[\frac{1/h-0}{h}\right] = \lim_{h\to 0}\frac{1}{h^2}$$

does not exist.

Example 4.5. The sign function $f(x) = \operatorname{sgn} x$, defined in Example 2.6, is differentiable at $x \neq 0$ with $f'(x) = 0$, since in that case $f(x+h) - f(x) = 0$ for all sufficiently small h. The sign function is not differentiable at 0 since

$$\lim_{h\to 0}\left[\frac{\operatorname{sgn} h - \operatorname{sgn} 0}{h}\right] = \lim_{h\to 0}\frac{\operatorname{sgn} h}{h}$$

and

$$\frac{\operatorname{sgn} h}{h} = \begin{cases} 1/h & \text{if } h > 0 \\ -1/h & \text{if } h < 0 \end{cases}$$

is unbounded in every neighborhood of 0, so its limit does not exist.

Example 4.6. The absolute value function $f(x) = |x|$ is differentiable at $x \neq 0$ with derivative $f'(x) = \operatorname{sgn} x$. It is not differentiable at 0, however, since

$$\lim_{h\to 0}\frac{f(h)-f(0)}{h} = \lim_{h\to 0}\frac{|h|}{h} = \lim_{h\to 0}\operatorname{sgn} h$$

does not exist.

Example 4.7. The function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^{1/3}$ is differentiable at $x \neq 0$ with

$$f'(x) = \frac{1}{3x^{2/3}}.$$

To prove this, we use the identity for the difference of cubes,

$$a^3 - b^3 = (a-b)(a^2 + ab + b^2),$$

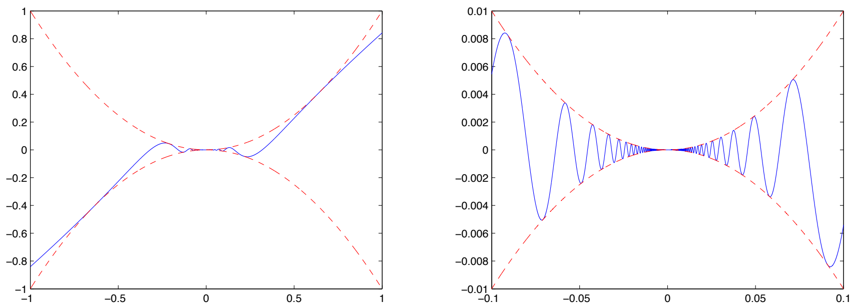

Figure 1. A plot of the function $y = x^2 \sin(1/x)$ and a detail near the origin with the parabolas $y = \pm x^2$ shown in red.

and get for $c \neq 0$ that

$$\begin{aligned}
\lim_{h\to 0}\left[\frac{f(c+h)-f(c)}{h}\right] &= \lim_{h\to 0}\frac{(c+h)^{1/3}-c^{1/3}}{h} \\
&= \lim_{h\to 0}\frac{(c+h)-c}{h\left[(c+h)^{2/3}+(c+h)^{1/3}c^{1/3}+c^{2/3}\right]} \\
&= \lim_{h\to 0}\frac{1}{(c+h)^{2/3}+(c+h)^{1/3}c^{1/3}+c^{2/3}} \\
&= \frac{1}{3c^{2/3}}.
\end{aligned}$$

However, f is not differentiable at 0, since

$$\lim_{h\to 0}\frac{f(h)-f(0)}{h} = \lim_{h\to 0}\frac{1}{h^{2/3}},$$

which does not exist.

Finally, we consider some examples of highly oscillatory functions.

Example 4.8. Define $f : \mathbb{R} \to \mathbb{R}$ by

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

It follows from the product and chain rules proved below that f is differentiable at $x \neq 0$ with derivative

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

However, f is not differentiable at 0, since

$$\lim_{h\to 0}\frac{f(h)-f(0)}{h} = \lim_{h\to 0}\sin\frac{1}{h},$$

which does not exist.

Example 4.9. Define $f : \mathbb{R} \to \mathbb{R}$ by

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

Then f is differentiable on $\mathbb{R}$. (See Figure 1.) It follows from the product and chain rules proved below that f is differentiable at $x \neq 0$ with derivative

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

Moreover, f is differentiable at 0 with $f'(0) = 0$, since

$$\lim_{h \to 0} \frac{f(h) - f(0)}{h} = \lim_{h \to 0} h \sin \frac{1}{h} = 0.$$

In this example, $\lim_{x \to 0} f'(x)$ does not exist, so although f is differentiable on $\mathbb{R}$, its derivative f' is not continuous at 0.

4.1.2. Derivatives as linear approximations. Another way to view Definition 4.1 is to write

$$f(c + h) = f(c) + f'(c)h + r(h)$$

as the sum of a linear approximation $f(c) + f'(c)h$ of $f(c+h)$ and a remainder $r(h)$. In general, the remainder also depends on c, but we don't show this explicitly since we're regarding c as fixed.

As we prove in the following proposition, the differentiability of f at c is equivalent to the condition

$$\lim_{h \to 0} \frac{r(h)}{h} = 0.$$

That is, the remainder $r(h)$ approaches 0 faster than h, so the linear terms in h provide a leading order approximation to $f(c + h)$ when h is small. We also write this condition on the remainder as

$$r(h) = o(h) \qquad \text{as } h \to 0,$$

pronounced "r is little-oh of h as $h \to 0$."

Graphically, this condition means that the graph of f near c is close the line through the point $(c, f(c))$ with slope $f'(c)$. Analytically, it means that the function

$$h \mapsto f(c + h) - f(c)$$

is approximated near c by the linear function

$$h \mapsto f'(c)h.$$

Thus, $f'(c)$ may be interpreted as a scaling factor by which a differentiable function f shrinks or stretches lengths near c.

If $|f'(c)| < 1$, then f shrinks the length of a small interval about c by (approximately) this factor; if $|f'(c)| > 1$, then f stretches the length of an interval by (approximately) this factor; if $f'(c) > 0$, then f preserves the orientation of the interval, meaning that it maps the left endpoint to the left endpoint of the image and the right endpoint to the right endpoints; if $f'(c) < 0$, then f reverses the orientation of the interval, meaning that it maps the left endpoint to the right endpoint of the image and visa-versa.

We can use this description as a definition of the derivative.

Proposition 4.10. Suppose that $f : (a, b) \to \mathbb{R}$. Then f is differentiable at $c \in (a, b)$ if and only if there exists a constant $A \in \mathbb{R}$ and a function $r : (a-c, b-c) \to \mathbb{R}$ such that

$$f(c+h) = f(c) + Ah + r(h), \qquad \lim_{h\to 0} \frac{r(h)}{h} = 0.$$

In that case, $A = f'(c)$.

Proof. First suppose that f is differentiable at c, as in Definition 4.1, and define

$$r(h) = f(c+h) - f(c) - f'(c)h.$$

Then

$$\lim_{h\to 0} \frac{r(h)}{h} = \lim_{h\to 0} \left[\frac{f(c+h) - f(c)}{h} - f'(c)\right] = 0.$$

Conversely, suppose that $f(c+h) = f(c) + Ah + r(h)$ where $r(h)/h \to 0$ as $h \to 0$. Then

$$\lim_{h\to 0} \left[\frac{f(c+h) - f(c)}{h}\right] = \lim_{h\to 0} \left[A + \frac{r(h)}{h}\right] = A,$$

which proves that f is differentiable at c with $f'(c) = A$. □

Example 4.11. In Example 4.2 with $f(x) = x^2$,

$$(c+h)^2 = c^2 + 2ch + h^2,$$

and $r(h) = h^2$, which goes to zero at a quadratic rate as $h \to 0$.

Example 4.12. In Example 4.4 with $f(x) = 1/x$,

$$\frac{1}{c+h} = \frac{1}{c} - \frac{1}{c^2}h + r(h),$$

for $c \neq 0$, where the quadratically small remainder is

$$r(h) = \frac{h^2}{c^2(c+h)}.$$

4.1.3. Left and right derivatives. We can use left and right limits to define one-sided derivatives, for example at the endpoint of an interval, but for the most part we will consider only two-sided derivatives defined at an interior point of the domain of a function.

Definition 4.13. Suppose $f : [a, b] \to \mathbb{R}$. Then f is right-differentiable at $a \leq c < b$ with right derivative $f'(c^+)$ if

$$\lim_{h\to 0^+} \left[\frac{f(c+h) - f(c)}{h}\right] = f'(c^+)$$

exists, and f is left-differentiable at $a < c \leq b$ with left derivative $f'(c^-)$ if

$$\lim_{h\to 0^-} \left[\frac{f(c+h) - f(c)}{h}\right] = \lim_{h\to 0^+} \left[\frac{f(c) - f(c-h)}{h}\right] = f'(c^-).$$

A function is differentiable at $a < c < b$ if and only if the left and right derivatives exist at c and are equal.

Example 4.14. If $f : [0,1] \to \mathbb{R}$ is defined by $f(x) = x^2$, then

$$f'(0^+) = 0, \qquad f'(1^-) = 2.$$

These left and right derivatives remain the same if f is extended to a function defined on a larger domain, say

$$f(x) = \begin{cases} x^2 & \text{if } 0 \le x \le 1, \\ 0 & \text{if } x > 1, \\ 1/x & \text{if } x < 0. \end{cases}$$

For this extended function we have $f'(1^+) = 0$, which is not equal to $f'(1^-)$, and $f'(0^-)$ does not exist, so it is not differentiable at 0 or 1.

Example 4.15. The absolute value function $f(x) = |x|$ in Example 4.6 is left and right differentiable at 0 with left and right derivatives

$$f'(0^+) = 1, \qquad f'(0^-) = -1.$$

These are not equal, and f is not differentiable at 0.

4.2. Properties of the derivative

In this section, we prove some basic properties of differentiable functions.

4.2.1. Differentiability and continuity. First we discuss the relation between differentiability and continuity.

Theorem 4.16. If $f : (a,b) \to \mathbb{R}$ is differentiable at at $c \in (a,b)$, then f is continuous at c.

Proof. If f is differentiable at c, then

$$\begin{aligned} \lim_{h\to 0} f(c+h) - f(c) &= \lim_{h\to 0} \left[\frac{f(c+h) - f(c)}{h} \cdot h\right] \\ &= \lim_{h\to 0} \left[\frac{f(c+h) - f(c)}{h}\right] \cdot \lim_{h\to 0} h \\ &= f'(c) \cdot 0 \\ &= 0, \end{aligned}$$

which implies that f is continuous at c. □

For example, the sign function in Example 4.5 has a jump discontinuity at 0 so it cannot be differentiable at 0. The converse does not hold, and a continuous function needn't be differentiable. The functions in Examples 4.6, 4.7, 4.8 are continuous but not differentiable at 0. Example 5.24 describes a function that is continuous on $\mathbb{R}$ but not differentiable anywhere.

In Example 4.9, the function is differentiable on $\mathbb{R}$, but the derivative f' is not continuous at 0. Thus, while a function f has to be continuous to be differentiable, if f is differentiable its derivative f' needn't be continuous. This leads to the following definition.

Definition 4.17. A function $f : (a, b) \to \mathbb{R}$ is continuously differentiable on (a, b), written $f \in C^1(a, b)$, if it is differentiable on (a, b) and $f' : (a, b) \to \mathbb{R}$ is continuous.

For example, the function $f(x) = x^2$ with derivative $f'(x) = 2x$ is continuously differentiable on any interval (a, b). As Example 4.9 illustrates, functions that are differentiable but not continuously differentiable may still behave in rather pathological ways. On the other hand, continuously differentiable functions, whose tangent lines vary continuously, are relatively well-behaved.

4.2.2. Algebraic properties of the derivative. Next, we state the linearity of the derivative and the product and quotient rules.

Theorem 4.18. If $f, g : (a, b) \to \mathbb{R}$ are differentiable at $c \in (a, b)$ and $k \in \mathbb{R}$, then kf, $f + g$, and fg are differentiable at c with

$$(kf)'(c) = kf'(c), \quad (f + g)'(c) = f'(c) + g'(c), \quad (fg)'(c) = f'(c)g(c) + f(c)g'(c).$$

Furthermore, if $g(c) \neq 0$, then f/g is differentiable at c with

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

Proof. The first two properties follow immediately from the linearity of limits stated in Theorem 2.22. For the product rule, we write

$$\begin{aligned}
(fg)'(c) &= \lim_{h\to 0} \left[\frac{f(c+h)g(c+h) - f(c)g(c)}{h}\right] \\
&= \lim_{h\to 0} \left[\frac{(f(c+h) - f(c))\, g(c+h) + f(c)\, (g(c+h) - g(c))}{h}\right] \\
&= \lim_{h\to 0} \left[\frac{f(c+h) - f(c)}{h}\right] \lim_{h\to 0} g(c+h) + f(c) \lim_{h\to 0} \left[\frac{g(c+h) - g(c)}{h}\right] \\
&= f'(c)g(c) + f(c)g'(c),
\end{aligned}$$

where we have used the properties of limits in Theorem 2.22 and Theorem 4.18, which implies that g is continuous at c. The quotient rule follows by a similar argument, or by combining the product rule with the chain rule, which implies that $(1/g)' = -g'/g^2$. (See Example 4.21 below.) □

Example 4.19. We have $1' = 0$ and $x' = 1$. Repeated application of the product rule implies that x^n is differentiable on $\mathbb{R}$ for every $n \in \mathbb{N}$ with

$$(x^n)' = nx^{n-1}.$$

Alternatively, we can prove this result by induction: The formula holds for $n = 1$. Assuming that it holds for some $n \in \mathbb{N}$, we get from the product rule that

$$(x^{n+1})' = (x \cdot x^n)' = 1 \cdot x^n + x \cdot nx^{n-1} = (n + 1)x^n,$$

and the result follows. It follows by linearity that every polynomial function is differentiable on $\mathbb{R}$, and from the quotient rule that every rational function is differentiable at every point where its denominator is nonzero. The derivatives are given by their usual formulae.

4.2.3. The chain rule. The chain rule states the differentiability of a composition of functions. The result is quite natural if one thinks in terms of derivatives as linear maps. If f is differentiable at c, it scales lengths by a factor $f'(c)$, and if g is differentiable at $f(c)$, it scales lengths by a factor $g'(f(c))$. Thus, the composition $g \circ f$ scales lengths at c by a factor $g'(f(c)) \cdot f'(c)$. Equivalently, the derivative of a composition is the composition of the derivatives. We will prove the chain rule by making this observation rigorous.

Theorem 4.20 (Chain rule)**.** Let $f : A \to \mathbb{R}$ and $g : B \to \mathbb{R}$ where $A \subset \mathbb{R}$ and $f(A) \subset B$, and suppose that c is an interior point of A and $f(c)$ is an interior point of B. If f is differentiable at c and g is differentiable at $f(c)$, then $g \circ f : A \to \mathbb{R}$ is differentiable at c and

$$(g \circ f)'(c) = g'(f(c))\, f'(c).$$

Proof. Since f is differentiable at c, there is a function $r(h)$ such that

$$f(c+h) = f(c) + f'(c)h + r(h), \qquad \lim_{h \to 0} \frac{r(h)}{h} = 0,$$

and since g is differentiable at $f(c)$, there is a function $s(k)$ such that

$$g(f(c) + k) = g(f(c)) + g'(f(c))\, k + s(k), \qquad \lim_{k \to 0} \frac{s(k)}{k} = 0.$$

It follows that

$$\begin{aligned}(g \circ f)(c+h) &= g(f(c) + f'(c)h + r(h)) \\ &= g(f(c)) + g'(f(c))\,(f'(c)h + r(h)) + s(f'(c)h + r(h)) \\ &= g(f(c)) + g'(f(c))\, f'(c)h + t(h)\end{aligned}$$

where

$$t(h) = r(h) + s(\phi(h)), \qquad \phi(h) = f'(c)h + r(h).$$

Then, since $r(h)/h \to 0$ as $h \to 0$,

$$\lim_{h \to 0} \frac{t(h)}{h} = \lim_{h \to 0} \frac{s(\phi(h))}{h}.$$

We claim that this is limit is zero, and then it follows from Proposition 4.10 that $g \circ f$ is differentiable at c with

$$(g \circ f)'(c) = g'(f(c))\, f'(c).$$

To prove the claim, we use the facts that

$$\frac{\phi(h)}{h} \to f'(c) \quad \text{as } h \to 0, \qquad \frac{s(k)}{k} \to 0 \quad \text{as } k \to 0.$$

Roughly speaking, we have $\phi(h) \sim f'(c)h$ when h is small and therefore

$$\frac{s(\phi(h))}{h} \sim \frac{s(f'(c)h)}{h} \to 0 \qquad \text{as } h \to 0.$$

To prove this in detail, let $\epsilon > 0$ be given. We want to show that there exists $\delta > 0$ such that

$$\left| \frac{s(\phi(h))}{h} \right| < \epsilon \qquad \text{if } 0 < |h| < \delta.$$

Choose $\eta > 0$ so that

$$\left|\frac{s(k)}{k}\right| < \frac{\epsilon}{2|f'(c)|+1} \qquad \text{if } 0 < |k| < \eta.$$

(We include a "1" in the denominator to avoid a division by 0 if $f'(c) = 0$.) Next, choose $\delta_1 > 0$ such that

$$\left|\frac{r(h)}{h}\right| < |f'(c)| + 1 \qquad \text{if } 0 < |h| < \delta_1.$$

If $0 < |h| < \delta_1$, then

$$\begin{aligned} |\phi(h)| &\le |f'(c)|\,|h| + |r(h)| \\ &< |f'(c)|\,|h| + (|f'(c)| + 1)|h| \\ &< (2|f'(c)| + 1)\,|h|. \end{aligned}$$

Define $\delta_2 > 0$ by

$$\delta_2 = \frac{\eta}{2|f'(c)|+1},$$

and let $\delta = \min(\delta_1, \delta_2) > 0$. If $0 < |h| < \delta$, then $|\phi(h)| < \eta$ and

$$|\phi(h)| < (2|f'(c)| + 1)\,|h|.$$

It follows that for $0 < |h| < \delta$

$$|s(\phi(h))| < \frac{\epsilon|\phi(h)|}{2|f'(c)|+1} < \epsilon|h|.$$

(If $\phi(h) = 0$, then $s(\phi(h)) = 0$, so the inequality holds in that case also.) This proves that

$$\lim_{h\to 0} \frac{s(\phi(h))}{h} = 0.$$

□

Example 4.21. Suppose that f is differentiable at c and $f(c) \neq 0$. Then $g(y) = 1/y$ is differentiable at $f(c)$, with $g'(y) = -1/y^2$ (see Example 4.4). It follows that $1/f = g \circ f$ is differentiable at c with

$$\left(\frac{1}{f}\right)'(c) = -\frac{f'(c)}{f(c)^2}.$$

4.2.4. The derivative of inverse functions. The chain rule gives an expression for the derivative of an inverse function. In terms of linear approximations, it states that if f scales lengths at c by a nonzero factor $f'(c)$, then f^{-1} scales lengths at $f(c)$ by the factor $1/f'(c)$.

Proposition 4.22. Suppose that $f : A \to \mathbb{R}$ is a one-to-one function on $A \subset \mathbb{R}$ with inverse $f^{-1} : B \to \mathbb{R}$ where $B = f(A)$. If f is differentiable at an interior point $c \in A$ with $f'(c) \neq 0$, $f(c)$ is an interior point of B, and f^{-1} is differentiable at $f(c)$, then

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

Proof. The definition of the inverse implies that

$$f^{-1}(f(x)) = x.$$

Since f is differentiable at c and f^{-1} is differentiable at $f(c)$, the chain rule implies that

$$\left(f^{-1}\right)'(f(c))\, f'(c) = 1.$$

Dividing this equation by $f'(c) \neq 0$, we get the result. Moreover, it follows that f^{-1} cannot be differentiable at $f(c)$ if $f'(c) = 0$. □

Alternatively, setting $d = f(c)$, we can write the result as

$$(f^{-1})'(d) = \frac{1}{f'\left(f^{-1}(d)\right)}.$$

The following example illustrates the necessity of the condition $f'(c) \neq 0$ for the differentiability of the inverse.

Example 4.23. Define $f : \mathbb{R} \to \mathbb{R}$ by $f(x) = x^3$. Then f is strictly increasing, one-to-one, and onto with inverse $f^{-1} : \mathbb{R} \to \mathbb{R}$ given by

$$f^{-1}(y) = y^{1/3}.$$

Then $f'(0) = 0$ and f^{-1} is not differentiable at $f(0)= 0$. On the other hand, f^{-1} is differentiable at non-zero points of $\mathbb{R}$, with

$$(f^{-1})'(x^3) = \frac{1}{f'(x)} = \frac{1}{3x^2},$$

or, writing $y = x^3$,

$$(f^{-1})'(y) = \frac{1}{3y^{2/3}},$$

in agreement with Example 4.7.

Proposition 4.22 is not entirely satisfactory because it assumes the differentiability of f^{-1} at $f(c)$. One can show that if $f : I \to \mathbb{R}$ is a continuous and one-to-one function on an interval I, then f is strictly monotonic and f^{-1} is also continuous and strictly monotonic. In that case, f^{-1} is differentiable at $f(c)$ if f is differentiable at c and $f'(c) \neq 0$. We omit the proof of these statements.

Another condition for the existence and differentiability of f^{-1}, which generalizes to functions of several variables, is given by the inverse function theorem: If f is differentiable in a neighborhood of c, $f'(c) \neq 0$, *and* f' is continuous at c, then f has a local inverse f^{-1} defined in a neighborhood of $f(c)$ and the inverse is differentiable at $f(c)$ with derivative given by Proposition 4.22.

4.3. Extreme values

Definition 4.24. Suppose that $f : A \to \mathbb{R}$. Then f has a global (or absolute) maximum at $c \in A$ if

$$f(x) \leq f(c) \qquad \text{for all } x \in A,$$

and f has a local (or relative) maximum at $c \in A$ if there is a neighborhood U of c such that

$$f(x) \leq f(c) \qquad \text{for all } x \in A \cap U.$$

Similarly, f has a global (or absolute) minimum at $c \in A$ if

$$f(x) \geq f(c) \qquad \text{for all } x \in A,$$

and f has a local (or relative) minimum at $c \in A$ if there is a neighborhood U of c such that

$$f(x) \geq f(c) \qquad \text{for all } x \in A \cap U.$$

If f has a (local or global) maximum or minimum at $c \in A$, then f is said to have a (local or global) extreme value at c.

Theorem 3.33 states that a continuous function on a compact set has a global maximum and minimum. The following fundamental result goes back to Fermat.

Theorem 4.25. Suppose that $f : A \to \mathbb{R}$ has a local extreme value at an interior point $c \in A$ and f is differentiable at c. Then $f'(c) = 0$.

Proof. If f has a local maximum at c, then $f(x) \leq f(c)$ for all x in a δ-neighborhood $(c - \delta, c + \delta)$ of c, so

$$\frac{f(c+h) - f(c)}{h} \leq 0 \qquad \text{for all } 0 < h < \delta,$$

which implies that

$$f'(c) = \lim_{h \to 0^+} \left[\frac{f(c+h) - f(c)}{h}\right] \leq 0.$$

Moreover,

$$\frac{f(c+h) - f(c)}{h} \geq 0 \qquad \text{for all } -\delta < h < 0,$$

which implies that

$$f'(c) = \lim_{h \to 0^-} \left[\frac{f(c+h) - f(c)}{h}\right] \geq 0.$$

It follows that $f'(c) = 0$. If f has a local minimum at c, then the signs in these inequalities are reversed and we also conclude that $f'(c) = 0$. □

For this result to hold, it is crucial that c is an interior point, since we look at the sign of the difference quotient of f on both sides of c. At an endpoint, we get an inequality condition on the derivative. If $f : [a, b] \to \mathbb{R}$, the right derivative of f exists at a, and f has a local maximum at a, then $f(x) \leq f(a)$ for $a \leq x < a + \delta$, so $f'(a^+) \leq 0$. Similarly, if the left derivative of f exists at b, and f has a local maximum at b, then $f(x) \leq f(b)$ for $b - \delta < x \leq b$, so $f'(b^-) \geq 0$. The signs are reversed for local minima at the endpoints.

Definition 4.26. Suppose that $f : A \to \mathbb{R}$. An interior point $c \in A$ such that f is not differentiable at c or $f'(c) = 0$ is called a critical point of f. An interior point where $f'(c) = 0$ is called a stationary point of f.

Theorem 4.25 limits the search for points where f has a maximum or minimum value on A to:

(1) Boundary points of A;

(2) Interior points where f is not differentiable;

(3) Stationary points of f.

4.4. The mean value theorem

We begin by proving a special case.

Theorem 4.27 (Rolle)**.** Suppose that $f : [a, b] \to \mathbb{R}$ is continuous on the closed, bounded interval $[a, b]$, differentiable on the open interval (a, b), and $f(a) = f(b)$. Then there exists $a < c < b$ such that $f'(c) = 0$.

Proof. By the Weierstrass extreme value theorem, Theorem 3.33, f attains its global maximum and minimum values on $[a, b]$. If these are both attained at the endpoints, then f is constant, and $f'(c) = 0$ for every $a < c < b$. Otherwise, f attains at least one of its global maximum or minimum values at an interior point $a < c < b$. Theorem 4.25 implies that $f'(c) = 0$. □

Note that we require continuity on the closed interval $[a, b]$ but differentiability only on the open interval (a, b). This proof is deceptively simple, but the result is not trivial. It relies on the extreme value theorem, which in turn relies on the completeness of $\mathbb{R}$. The theorem would not be true if we restricted attention to functions defined on the rationals $\mathbb{Q}$.

The mean value theorem is an immediate consequence of Rolle's theorem: for a general function f with $f(a) \neq f(b)$, we subtract off a linear function to make the values of the resulting function equal at the endpoints.

Theorem 4.28 (Mean value)**.** Suppose that $f : [a, b] \to \mathbb{R}$ is continuous on the closed, bounded interval $[a, b]$, and differentiable on the open interval (a, b). Then there exists $a < c < b$ such that

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

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

$$g(x) = f(x) - f(a) - \left[\frac{f(b) - f(a)}{b - a}\right](x - a)$$

is continuous on $[a, b]$ and differentiable on (a, b) with

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

Moreover, $g(a) = g(b) = 0$. Rolle's Theorem implies that there exists $a < c < b$ such that $g'(c) = 0$, which proves the result. □

Graphically, this result says that there is point $a < c < b$ at which the slope of the graph $y = f(x)$ is equal to the slope of the chord between the endpoints $(a, f(a))$ and $(b, f(b))$.

Analytically, the mean value theorem is a key result that connects the local behavior of a function, described by the derivative $f'(c)$, to its global behavior, described by the difference $f(b) - f(a)$. As a first application we prove a converse to the obvious fact that the derivative of a constant functions is zero.

Theorem 4.29. If $f : (a,b) \to \mathbb{R}$ is differentiable on (a,b) and $f'(x) = 0$ for every $a < x < b$, then f is constant on (a,b).

Proof. Fix $x_0 \in (a,b)$. The mean value theorem implies that for all $x \in (a,b)$ with $x \neq x_0$

$$f'(c) = \frac{f(x) - f(x_0)}{x - x_0}$$

for some c between x_0 and x. Since $f'(c) = 0$, it follows that $f(x) = f(x_0)$ for all $x \in (a,b)$, meaning that f is constant on (a,b). □

Corollary 4.30. If $f, g : (a,b) \to \mathbb{R}$ are differentiable on (a,b) and $f'(x) = g'(x)$ for every $a < x < b$, then $f(x) = g(x) + C$ for some constant C.

Proof. This follows from the previous theorem since $(f - g)' = 0$. □

We can also use the mean value theorem to relate the monotonicity of a differentiable function with the sign of its derivative.

Theorem 4.31. Suppose that $f : (a,b) \to \mathbb{R}$ is differentiable on (a,b). Then f is increasing if and only if $f'(x) \geq 0$ for every $a < x < b$, and decreasing if and only if $f'(x) \leq 0$ for every $a < x < b$. Furthermore, if $f'(x) > 0$ for every $a < x < b$ then f is strictly increasing, and if $f'(x) < 0$ for every $a < x < b$ then f is strictly decreasing.

Proof. If f is increasing, then

$$\frac{f(x+h) - f(x)}{h} \geq 0$$

for all sufficiently small h (positive or negative), so

$$f'(x) = \lim_{h \to 0} \left[\frac{f(x+h) - f(x)}{h} \right] \geq 0.$$

Conversely if $f' \geq 0$ and $a < x < y < b$, then by the mean value theorem

$$\frac{f(y) - f(x)}{y - x} = f'(c) \geq 0$$

for some $x < c < y$, which implies that $f(x) \leq f(y)$, so f is increasing. Moreover, if $f'(c) > 0$, we get $f(x) < f(y)$, so f is strictly increasing.

The results for a decreasing function f follow in a similar way, or we can apply of the previous results to the increasing function $-f$. □

Note that if f is strictly increasing, it does *not* follow that $f'(x) > 0$ for every $x \in (a,b)$.

Example 4.32. The function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^3$ is strictly increasing on $\mathbb{R}$, but $f'(0) = 0$.

If f is continuously differentiable and $f'(c) > 0$, then $f'(x) > 0$ for all x in a neighborhood of c and Theorem 4.31 implies that f is strictly increasing near c. This conclusion may fail if f is not continuously differentiable at c.

Example 4.33. The function

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

is differentiable, but not continuously differentiable, at 0 and $f'(0) = 1/2 > 0$. However, f is not increasing in any neighborhood of 0 since

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

is continuous for $x \neq 0$ and takes negative values in any neighborhood of 0, so f is strictly decreasing near those points.

4.5. Taylor's theorem

If $f : (a, b) \to \mathbb{R}$ is differentiable on (a, b) and $f' : (a, b) \to \mathbb{R}$ is differentiable, then we define the second derivative $f'' : (a, b) \to \mathbb{R}$ of f as the derivative of f'. We define higher-order derivatives similarly. If f has derivatives $f^{(n)} : (a, b) \to \mathbb{R}$ of all orders $n \in \mathbb{N}$, then we say that f is infinitely differentiable on (a, b).

Taylor's theorem gives an approximation for an $(n + 1)$-times differentiable function in terms of its Taylor polynomial of degree n.

Definition 4.34. Let $f : (a, b) \to \mathbb{R}$ and suppose that f has n derivatives $f', f'', \dots f^{(n)} : (a, b) \to \mathbb{R}$ on (a, b). The Taylor polynomial of degree n of f at $a < c < b$ is

$$P_n(x) = f(c) + f'(c)(x - c) + \frac{1}{2!} f''(c)(x - c)^2 + \dots + \frac{1}{n!} f^{(n)}(c)(x - c)^n.$$

Equivalently,

$$P_n(x) = \sum_{k=0}^{n} a_k (x - c)^k, \qquad a_k = \frac{1}{k!} f^{(k)}(c).$$

We call a_k the kth Taylor coefficient of f at c. The computation of the Taylor polynomials in the following examples are left as an exercise.

Example 4.35. If $P(x)$ is a polynomial of degree n, then $P_n(x) = P(x)$.

Example 4.36. The Taylor polynomial of degree n of e^x at $x = 0$ is

$$P_n(x) = 1 + x + \frac{1}{2!} x^2 \dots + \frac{1}{n!} x^n.$$

Example 4.37. The Taylor polynomial of degree $2n$ of $\cos x$ at $x = 0$ is

$$P_{2n}(x) = 1 - \frac{1}{2!} x^2 + \frac{1}{4!} x^4 - \dots + (-1)^n \frac{1}{(2n)!} x^{2n}.$$

We also have $P_{2n+1} = P_{2n}$.

Example 4.38. The Taylor polynomial of degree $2n + 1$ of $\sin x$ at $x = 0$ is

$$P_{2n+1}(x) = x - \frac{1}{3!} x^3 + \frac{1}{5!} x^5 - \dots + (-1)^n \frac{1}{(2n + 1)!} x^{2n+1}.$$

We also have $P_{2n+2} = P_{2n+1}$.

Example 4.39. The Taylor polynomial of degree n of $1/x$ at $x = 1$ is

$$P_n(x) = 1 - (x-1) + (x-1)^2 - \cdots + (-1)^n (x-1)^n.$$

Example 4.40. The Taylor polynomial of degree n of $\log x$ at $x = 1$ is

$$P_n(x) = (x-1) - \frac{1}{2}(x-1)^2 + \frac{1}{3}(x-1)^3 - \cdots + (-1)^{n+1}(x-1)^n.$$

We write

$$f(x) = P_n(x) + R_n(x).$$

where R_n is the error, or remainder, between f and its Taylor polynomial P_n. The next theorem is one version of Taylor's theorem, which gives an expression for the remainder due to Lagrange. It can be regarded as a generalization of the mean value theorem, which corresponds to the case $n = 0$.

The proof is a bit tricky, but the essential idea is to subtract a suitable polynomial from the function and apply Rolle's theorem, just as we proved the mean value theorem by subtracting a suitable linear function.

Theorem 4.41 (Taylor)**.** *Suppose $f : (a,b) \to \mathbb{R}$ has $n+1$ derivatives on (a,b) and let $a < c < b$. For every $a < x < b$, there exists ξ between c and x such that*

$$f(x) = f(c) + f'(c)(x-c) + \frac{1}{2!}f''(c)(x-c)^2 + \cdots + \frac{1}{n!}f^{(n)}(c)(x-c)^n + R_n(x)$$

where

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

Proof. Fix $x, c \in (a,b)$. For $t \in (a,b)$, let

$$g(t) = f(x) - f(t) - f'(t)(x-t) - \frac{1}{2!}f''(t)(x-t)^2 - \cdots - \frac{1}{n!}f^{(n)}(t)(x-t)^n.$$

Then $g(x) = 0$ and

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

Define

$$h(t) = g(t) - \left(\frac{x-t}{x-c}\right)^{n+1} g(c).$$

Then $h(c) = h(x) = 0$, so by Rolle's theorem, there exists a point ξ between c and x such that $h'(\xi) = 0$, which implies that

$$g'(\xi) + (n+1)\frac{(x-\xi)^n}{(x-c)^{n+1}}g(c) = 0.$$

It follows from the expression for g' that

$$\frac{1}{n!}f^{(n+1)}(\xi)(x-\xi)^n = (n+1)\frac{(x-\xi)^n}{(x-c)^{n+1}}g(c),$$

and using the expression for g in this equation, we get the result. □

Note that the remainder term

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

has the same form as the $(n+1)$th term in the Taylor polynomial of f, except that the derivative is evaluated at an (unknown) intermediate point ξ between c and x, instead of at c.

Example 4.42. Let us prove that

$$\lim_{x \to 0} \left(\frac{1 - \cos x}{x^2} \right) = \frac{1}{2}.$$

By Taylor's theorem,

$$\cos x = 1 - \frac{1}{2}x^2 + \frac{1}{4!}(\cos \xi)x^4$$

for some ξ between 0 and x. It follows that for $x \neq 0$,

$$\frac{1 - \cos x}{x^2} - \frac{1}{2} = -\frac{1}{4!}(\cos \xi)x^2.$$

Since $|\cos \xi| \leq 1$, we get

$$\left| \frac{1 - \cos x}{x^2} - \frac{1}{2} \right| \leq \frac{1}{4!}x^2,$$

which implies that

$$\lim_{x \to 0} \left| \frac{1 - \cos x}{x^2} - \frac{1}{2} \right| = 0.$$

Note that Taylor's theorem not only proves the limit, but it also gives an explicit upper bound for the difference between $(1 - \cos x)/x^2$ and its limit $1/2$.