

Chapter 1

Differentiation

1.1 Basic definitions

Remark 1.1.1. *We can now define derivatives analytically, using limits, in contrast to the geometric definition of derivatives, which uses tangents. The advantage of working analytically is that (a) we do not need to know the axioms of geometry, and (b) these definitions can be modified to handle functions of several variables, or functions whose values are vectors instead of scalar. Furthermore, one's geometric intuition becomes difficult to rely on once one has more than three dimensions in play. (Conversely, one can use one's experience in analytic rigour to extend one's geometric intuition to such abstract settings; as mentioned earlier, the two viewpoints complement rather than oppose each other.)*

Definition 1.1.2 (Differentiability at a point). *Let X be a subset of \mathbb{R}^n , and let $x_0 \in X$ be an element of X which is also a limit point of X . Let $f : X \rightarrow \mathbb{R}^m$ be a function. If the limit*

$$\lim_{x \rightarrow x_0; x \in X \setminus \{x_0\}} \frac{f(x) - f(x_0)}{x - x_0}$$

converges to some real number L , then we say that f is differentiable at x_0 on X with derivative L , and write $f'(x_0)L$. If the limit does not exist, or if x_0 is not an element of X or not a limit point of X , we leave $f'(x_0)$ undefined, and say that f is not differentiable at x_0 on X .

Remark 1.1.3. *Note that we need x_0 to be a limit point in order for x_0 to be adherent to $X \setminus \{x_0\}$, otherwise the limit*

$$\lim_{x \rightarrow x_0; x \in X \setminus \{x_0\}} \frac{f(x) - f(x_0)}{x - x_0}$$

would automatically be undefined. In particular, we do not define the derivative of a function at an isolated point; In practice, the domain X will almost always be an interval, and so by i.9.1.21 all elements x_0 of X will automatically be limit points and we will not have to care much about these issues.

Remark 1.1.4. *This point is trivial, but it is worth mentioning: if $f : X \rightarrow \mathbb{R}^m$ is differentiable at x_0 , and $g : X \rightarrow \mathbb{R}^m$ is equal to f (i.e., $g(x) = f(x)$ for all $x \in X$), then g is also differentiable*

at x_0 and $g'(x_0) = f'(x_0)$. However, if two functions f and g merely have the same value at x_0 , i.e., $g(x_0) = f(x_0)$, this does not imply that $g'(x_0) = f'(x_0)$. Thus, there is a big difference between two functions being equal on their whole domain, and merely being equal at one point.

Remark 1.1.5. One sometimes writes $\frac{df}{dx}$ instead of f' . This notation is of course very familiar and convenient, but one has to be a little careful, because it is only safe to use as long as x is the only variable used to represent the input for f ; otherwise one can get into all sorts of trouble. Because of this possible source of confusion, we will refrain from using the notation $\frac{df}{dx}$ whenever it could possibly lead to confusion. (This confusion becomes even worse in the calculus of several variables, and the standard notation of $\frac{\partial f}{\partial x}$ can lead to some serious ambiguities. There are ways to resolve these ambiguities, most notably by introducing the notion of differentiation along vector fields, but this is beyond the scope of this text.)

Example 1.1.6. Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be the function $f(x) = x$, and let $x_0 = 0$. To see whether f is differentiable at 0 on \mathbb{R} , we compute the limit

$$\lim_{x \rightarrow 0; x \in \mathbb{R} \setminus \{0\}} \frac{f(x) - f(0)}{x - 0} = \lim_{x \rightarrow 0; x \in \mathbb{R} \setminus \{0\}} \frac{x}{x}.$$

Now we take left limits and right limits. The right limit is

$$\lim_{x \rightarrow 0; x \in (0, \infty)} \frac{x}{x} = \lim_{x \rightarrow 0; x \in (0, \infty)} \frac{x}{x} = \lim_{x \rightarrow 0; x \in (0, \infty)} 1 = 1,$$

while the left limit is

$$\lim_{x \rightarrow 0; x \in (-\infty, 0)} \frac{x}{x} = \lim_{x \rightarrow 0; x \in (-\infty, 0)} \frac{-x}{x} = \lim_{x \rightarrow 0; x \in (-\infty, 0)} -1 = -1,$$

and these limits do not match. Thus, $\lim_{x \rightarrow 0; x \in (0, \infty)} \frac{x}{x}$ does not exist, and f is not differentiable at 0 on \mathbb{R} . However, if one restricts f to $[0, \infty)$, then the restricted function $f|_{[0, \infty)}$ is differentiable at 0 on $[0, \infty)$, with derivative 1:

$$\lim_{x \rightarrow 0; x \in [0, \infty) \setminus \{0\}} \frac{f(x) - f(0)}{x - 0} = \lim_{x \rightarrow 0; x \in (0, \infty)} \frac{x}{x} = 1.$$

Similarly, when one restricts f to $(-\infty, 0]$, the restricted function $f|_{(-\infty, 0]}$ is differentiable at 0 on $(-\infty, 0]$, with derivative -1 . Thus, even when a function is not differentiable, it is sometimes possible to restore the differentiability by restricting the domain of the function.

Proposition 1.1.7 (Newton's approximation). *Let X be a subset of \mathbb{R} , let $x_0 \in X$ be a limit point of X , let $f : X \rightarrow \mathbb{R}$ be a function, and let L be a real number. Then the following statements are logically equivalent:*

1. f is differentiable at x_0 on X with derivative L .

2. For every $\varepsilon > 0$, there exists a $\delta > 0$ such that $f(x)$ is $\varepsilon x - x_0$ -close to $f(x_0) + L(x - x_0)$ whenever $x \in X$ is δ -close to x_0 , i.e., we have

$$f(x) - (f(x_0) + L(x - x_0)) \leq \varepsilon x - x_0$$

whenever $x \in X$ and $x - x_0 \leq \delta$.

Proof. We first show that the first statement implies the second statement. Since f is differentiable at x_0 on X with derivative L , by i:10.1.1 we have

$$\lim_{x \rightarrow x_0; x \in X \setminus x_0} \frac{f(x) - f(x_0)}{x - x_0} = L.$$

By i:9.3.6 this means

$$\forall \varepsilon \in ^+, \exists \delta \in ^+: \left(\forall x \in X \setminus x_0, x - x_0 < \delta \implies \frac{f(x) - f(x_0)}{x - x_0} - L \leq \varepsilon \right).$$

Thus, we have

$$\begin{aligned} & \forall x \in X \setminus x_0, x - x_0 \leq \delta/2 < \delta \\ \implies & \frac{f(x) - f(x_0)}{x - x_0} - L \leq \varepsilon \\ \implies & \frac{f(x) - f(x_0)}{x - x_0} - Lx - x_0 \leq \varepsilon x - x_0 \\ \implies & (f(x) - f(x_0)) - L(x - x_0) \leq \varepsilon x - x_0 \\ \implies & f(x) - (f(x_0) + L(x - x_0)) \leq \varepsilon x - x_0. \end{aligned}$$

If $x = x_0$, then we have

$$\begin{aligned} & 0 = x_0 - x_0 \leq \delta/2 < \delta \\ \implies & 0 = f(x_0) - (f(x_0) + L(x_0 - x_0)) \leq \varepsilon x_0 - x_0 = 0. \end{aligned}$$

Thus, we have $\forall \varepsilon \in ^+, \exists \delta \in ^+$ such that

$$\forall x \in X, x - x_0 \leq \delta/2 \implies f(x) - (f(x_0) + L(x - x_0)) \leq \varepsilon x - x_0.$$

Now we show that the second statement implies the first statement. By hypothesis we have $\forall \varepsilon \in ^+, \exists \delta \in ^+$ such that

$$\forall x \in X, x - x_0 \leq \delta \implies f(x) - (f(x_0) + L(x - x_0)) \leq \varepsilon x - x_0.$$

In particular, we have

$$\forall x \in X \setminus x_0, x - x_0 \leq \delta \implies f(x) - (f(x_0) + L(x - x_0)) \leq \varepsilon x - x_0.$$

Thus, we have

$$\begin{aligned}
& \forall x \in X \setminus x_0, x - x_0 \leq \delta \\
& \implies f(x) - (f(x_0) + L(x - x_0)) \leq \varepsilon x - x_0 \\
& \implies \frac{f(x) - (f(x_0) + L(x - x_0))}{x - x_0} \leq \varepsilon \\
& \implies \frac{f(x) - f(x_0)}{x - x_0} - L \leq \varepsilon.
\end{aligned}$$

By i:9.3.6 this means

$$\lim_{x \rightarrow x_0; x \in X \setminus x_0} \frac{f(x) - f(x_0)}{x - x_0} = L$$

and by i:10.1.1 we know that f is differentiable at x_0 on X with derivative L . \square

Remark 1.1.8. *Newton's approximation is of course named after the great scientist and mathematician Isaac Newton (1642–1727), one of the founders of differential and integral calculus.*

Remark 1.1.9. *We can phrase i:10.1.7 in a more informal way: if f is differentiable at x_0 , then one has the approximation $f(x) \approx f(x_0) + f'(x_0)(x - x_0)$, and conversely.*

Proposition 1.1.10 (Differentiability implies continuity). *Let X be a subset of \mathbb{R} , let $x_0 \in X$ be a limit point of X , and let $f : X \rightarrow \mathbb{R}$ be a function. If f is differentiable at x_0 , then f is also continuous at x_0 .*

Proof. Since f is differentiable at x_0 , by i:10.1.1 we have

$$L = \lim_{x \rightarrow x_0; x \in X \setminus x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

for some $L \in \mathbb{R}$. By i:10.1.7, we have $\forall \varepsilon \in^+, \exists \delta \in^+$ such that

$$\begin{aligned}
& \forall x \in X, x - x_0 \leq \delta \\
& \implies f(x) - (f(x_0) + L(x - x_0)) \leq \varepsilon x - x_0 \\
& \implies f(x) - (f(x_0) + L(x - x_0)) + L(x - x_0) \leq \varepsilon x - x_0 + L(x - x_0) \\
& \implies f(x) - (f(x_0) + L(x - x_0)) + L(x - x_0) \leq (\varepsilon + L)x - x_0 \\
& \implies f(x) - (f(x_0) + L(x - x_0)) + L(x - x_0) \\
& \quad \leq f(x) - (f(x_0) + L(x - x_0)) + L(x - x_0) \\
& \quad \leq (\varepsilon + L)x - x_0 \\
& \implies f(x) - f(x_0) \leq (\varepsilon + L)x - x_0.
\end{aligned}$$

Let $\delta' = \min(\delta, \varepsilon/(\varepsilon + L))$. Then we have

$$\begin{aligned}
& \forall x \in X, x - x_0 < \delta' \leq \delta \\
& \implies f(x) - f(x_0) \leq (\varepsilon + L)x - x_0 \\
& \implies f(x) - f(x_0) \leq (\varepsilon + L)x - x_0 \leq (\varepsilon + L)\delta' \leq (\varepsilon + L) \frac{\varepsilon}{\varepsilon + L} \\
& \implies f(x) - f(x_0) \leq \varepsilon.
\end{aligned}$$

Thus, by i:9.3.6 we have $\lim_{x \rightarrow x_0; x \in X} f(x) = f(x_0)$, and by i:9.4.1 f is continuous at x_0 . \square

Definition 1.1.11 (Differentiability on a domain). *Let X be a subset of \mathbb{R} , and let $f : X \rightarrow \mathbb{R}$ be a function. We say that f is differentiable on X if, for every limit point $x_0 \in X$, the function f is differentiable at x_0 on X .*

Corollary 1.1.12. *Let X be a subset of \mathbb{R} , and let $f : X \rightarrow \mathbb{R}$ be a function which is differentiable on X . Then f is also continuous on X .*

Proof. By i:9.1.11 we know that $\forall x_0 \in X$, x_0 is an adherent point. By i:ex:9.1.9 we know that x_0 is either a limit point or an isolated point. By i:10.1.10 and i:10.1.11 we know that if x_0 is a limit point then f is continuous at x_0 . So we only need to show that if x_0 is an isolated point, then f is also continuous at x_0 . Suppose that x_0 is an isolated point of X . By i:9.1.18 we know that $\exists \varepsilon' \in \mathbb{R}^+$ such that $x - x_0 > \varepsilon'$ for all $x \in X \setminus x_0$. To show that f is continuous at x_0 , by i:9.4.1 and i:9.3.6 we need to show that

$$\forall \varepsilon \in \mathbb{R}^+, \exists \delta \in \mathbb{R}^+ : (\forall x \in X, x - x_0 < \delta \implies f(x) - f(x_0) \leq \varepsilon).$$

Let $\delta = \varepsilon'$. Since x_0 is an isolated point, the only $x \in X$ satisfying $x - x_0 < \varepsilon'$ is x_0 . Thus, we have $0 = f(x_0) - f(x_0) \leq \varepsilon$ and $\lim_{x \rightarrow x_0; x \in X} f(x) = f(x_0)$. \square

Theorem 1.1.13 (Differential calculus). *Let X be a subset of \mathbb{R} , let $x_0 \in X$ be a limit point of X , and let $f : X \rightarrow \mathbb{R}$ and $g : X \rightarrow \mathbb{R}$ be functions.*

1. *If f is a constant function, i.e., there exists a real number c such that $f(x) = c$ for all $x \in X$, then f is differentiable at x_0 and $f'(x_0) = 0$.*
2. *If f is the identity function, i.e., $f(x) = x$ for all $x \in X$, then f is differentiable at x_0 and $f'(x_0) = 1$.*
3. *(Sum rule) If f and g are differentiable at x_0 , then $f + g$ is also differentiable at x_0 , and $(f + g)'(x_0) = f'(x_0) + g'(x_0)$.*
4. *(Product rule) If f and g are differentiable at x_0 , then fg is also differentiable at x_0 , and $(fg)'(x_0) = f'(x_0)g(x_0) + f(x_0)g'(x_0)$.*
5. *If f is differentiable at x_0 and c is a real number, then cf is also differentiable at x_0 , and $(cf)'(x_0) = cf'(x_0)$.*
6. *(Difference rule) If f and g are differentiable at x_0 , then $f - g$ is also differentiable at x_0 , and $(f - g)'(x_0) = f'(x_0) - g'(x_0)$.*
7. *If g is differentiable at x_0 , and g is non-zero on X (i.e., $g(x) \neq 0$ for all $x \in X$), then $1/g$ is also differentiable at x_0 , and $(\frac{1}{g})'(x_0) = -\frac{g'(x_0)}{g(x_0)^2}$.*
8. *(Quotient rule) If f and g are differentiable at x_0 , and g is non-zero on X , then f/g is also differentiable at x_0 , and*

$$(\frac{f}{g})'(x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g(x_0)^2}.$$

Proof. (a) We have $\forall \varepsilon \in^+, \forall \delta \in^+$ such that

$$\forall x \in X \setminus x_0, x - x_0 < \delta \implies \frac{f(x) - f(x_0)}{x - x_0} - 0 = \frac{c - c}{x - x_0} - 0 = 0 \leq \varepsilon.$$

Thus, by i:9.3.6 we have

$$\lim_{x \rightarrow x_0; x \in X \setminus x_0} \frac{f(x) - f(x_0)}{x - x_0} = 0$$

and by i:10.1.1 we have $f'(x_0) = 0$. □

Proof. (b) We have $\forall \varepsilon \in^+, \forall \delta \in^+$ such that

$$\forall x \in X \setminus x_0, x - x_0 < \delta \implies \frac{f(x) - f(x_0)}{x - x_0} - 1 = \frac{x - x_0}{x - x_0} - 1 = 0 \leq \varepsilon.$$

Thus, by i:9.3.6 we have

$$\lim_{x \rightarrow x_0; x \in X \setminus x_0} \frac{f(x) - f(x_0)}{x - x_0} = 1$$

and by i:10.1.1 we have $f'(x_0) = 1$. □

Proof. (c) By i:10.1.1 and i:9.3.14 we have

$$\begin{aligned} & f'(x_0) + g'(x_0) \\ = & \lim_{x \rightarrow x_0; x \in X \setminus x_0} \frac{f(x) - f(x_0)}{x - x_0} + \lim_{x \rightarrow x_0; x \in X \setminus x_0} \frac{g(x) - g(x_0)}{x - x_0} \\ = & \lim_{x \rightarrow x_0; x \in X \setminus x_0} \frac{f(x) - f(x_0)}{x - x_0} + \frac{g(x) - g(x_0)}{x - x_0} \\ = & \lim_{x \rightarrow x_0; x \in X \setminus x_0} \frac{f(x) - f(x_0) + g(x) - g(x_0)}{x - x_0} \\ = & \lim_{x \rightarrow x_0; x \in X \setminus x_0} \frac{f(x) + g(x) - (f(x_0) + g(x_0))}{x - x_0} \\ = & \lim_{x \rightarrow x_0; x \in X \setminus x_0} \frac{(f + g)(x) - (f + g)(x_0)}{x - x_0} \\ = & (f + g)'(x_0). \end{aligned}$$

Thus, $f + g$ is differentiable at x_0 and $(f + g)'(x_0) = f'(x_0) + g'(x_0)$. □

Proof. (d) By i:10.1.1 and i:9.3.14 we have

$$\begin{aligned}
& f'(x_0)g(x_0) + f(x_0)g'(x_0) \\
&= \left(\lim_{x \rightarrow x_0; x \in X \setminus x_0} \frac{f(x) - f(x_0)}{x - x_0} \right) \left(\lim_{x \rightarrow x_0; x \in X \setminus x_0} g(x) \right) \\
&\quad + f(x_0) \left(\lim_{x \rightarrow x_0; x \in X \setminus x_0} \frac{g(x) - g(x_0)}{x - x_0} \right) \\
&= \left(\lim_{x \rightarrow x_0; x \in X \setminus x_0} \frac{(f(x) - f(x_0))g(x)}{x - x_0} \right) + \left(\lim_{x \rightarrow x_0; x \in X \setminus x_0} \frac{f(x_0)(g(x) - g(x_0))}{x - x_0} \right) \\
&= \left(\lim_{x \rightarrow x_0; x \in X \setminus x_0} \frac{f(x)g(x) - f(x_0)g(x)}{x - x_0} \right) + \left(\lim_{x \rightarrow x_0; x \in X \setminus x_0} \frac{f(x_0)g(x) - f(x_0)g(x_0)}{x - x_0} \right) \\
&= \lim_{x \rightarrow x_0; x \in X \setminus x_0} \left(\frac{f(x)g(x) - f(x_0)g(x)}{x - x_0} + \frac{f(x_0)g(x) - f(x_0)g(x_0)}{x - x_0} \right) \\
&= \lim_{x \rightarrow x_0; x \in X \setminus x_0} \frac{f(x)g(x) - f(x_0)g(x) + f(x_0)g(x) - f(x_0)g(x_0)}{x - x_0} \\
&= \lim_{x \rightarrow x_0; x \in X \setminus x_0} \frac{f(x)g(x) - f(x_0)g(x_0)}{x - x_0} \\
&= \lim_{x \rightarrow x_0; x \in X \setminus x_0} \frac{(fg)(x) - (fg)(x_0)}{x - x_0} \\
&= (fg)'(x_0).
\end{aligned}$$

Thus, fg is differentiable at x_0 and $(fg)'(x_0) = f'(x_0)g(x_0) + f(x_0)g'(x_0)$. □

Proof. (e) By i:10.1.1 and i:9.3.14 we have

$$\begin{aligned}
& cf'(x_0) \\
&= c \left(\lim_{x \rightarrow x_0; x \in X \setminus x_0} \frac{f(x) - f(x_0)}{x - x_0} \right) \\
&= \lim_{x \rightarrow x_0; x \in X \setminus x_0} \left(c \frac{f(x) - f(x_0)}{x - x_0} \right) \\
&= \lim_{x \rightarrow x_0; x \in X \setminus x_0} \frac{cf(x) - cf(x_0)}{x - x_0} \\
&= \lim_{x \rightarrow x_0; x \in X \setminus x_0} \frac{(cf)(x) - (cf)(x_0)}{x - x_0} \\
&= (cf)'(x_0).
\end{aligned}$$

Thus, cf is differentiable at x_0 and $(cf)'(x_0) = cf'(x_0)$. □

Proof. (f)

$$\begin{aligned}
 & f'(x_0) - g'(x_0) \\
 &= f'(x_0) + (-g'(x_0)) \\
 &= f'(x_0) + ((-g)'(x_0)) && i : 10.1.13[e] \\
 &= (f + (-g))'(x_0) && i : 10.1.13[c] \\
 &= (f - g)'(x_0). && i : 9.2.1
 \end{aligned}$$

Thus, $f - g$ is differentiable at x_0 and $(f - g)'(x_0) = f'(x_0) - g'(x_0)$. □

Proof. (g) By i:10.1.1 and i:9.3.14 we have

$$\begin{aligned}
 & -\frac{g'(x_0)}{g(x_0)^2} \\
 &= \left(\lim_{x \rightarrow x_0; x \in X \setminus x_0} \frac{g(x) - g(x_0)}{x - x_0} \right) \left(\frac{-1}{g(x_0)^2} \right) \\
 &= \left(\lim_{x \rightarrow x_0; x \in X \setminus x_0} \frac{g(x) - g(x_0)}{x - x_0} \right) \left(\frac{-g(x_0)}{g(x_0)g(x_0)^2} \right) \\
 &= \left(\lim_{x \rightarrow x_0; x \in X \setminus x_0} \frac{g(x) - g(x_0)}{x - x_0} \right) \left(\lim_{x \rightarrow x_0; x \in X \setminus x_0} \frac{-g(x_0)}{g(x)g(x_0)^2} \right) \\
 &= \lim_{x \rightarrow x_0; x \in X \setminus x_0} \left(\left(\frac{g(x) - g(x_0)}{x - x_0} \right) \left(\frac{-g(x_0)}{g(x)g(x_0)^2} \right) \right) \\
 &= \lim_{x \rightarrow x_0; x \in X \setminus x_0} \frac{\frac{g(x_0)(g(x_0) - g(x))}{g(x)g(x_0)^2}}{x - x_0} \\
 &= \lim_{x \rightarrow x_0; x \in X \setminus x_0} \frac{\frac{g(x_0) - g(x)}{g(x)g(x_0)}}{x - x_0} \\
 &= \lim_{x \rightarrow x_0; x \in X \setminus x_0} \frac{\frac{1}{g(x)} - \frac{1}{g(x_0)}}{x - x_0} \\
 &= \lim_{x \rightarrow x_0; x \in X \setminus x_0} \frac{\frac{1}{g}}{x - x_0} \\
 &= \left(\frac{1}{g} \right)'(x_0).
 \end{aligned}$$

Thus, $1/g$ is differentiable at x_0 and $(1/g)'(x_0) = -\frac{g'(x_0)}{g(x_0)^2}$. □

Proof. (h)

$$\begin{aligned}
& \left(\frac{f}{g}\right)'(x_0) \\
&= \left(f \cdot \frac{1}{g}\right)'(x_0) && i : 9.2.1 \\
&= f'(x_0) \frac{1}{g}(x_0) + f(x_0) \left(\frac{1}{g}\right)'(x_0) && i : 10.1.13[d] \\
&= \frac{f'(x_0)}{g(x_0)} + f(x_0) \left(\frac{1}{g}\right)'(x_0) && i : 9.2.1 \\
&= \frac{f'(x_0)}{g(x_0)} + f(x_0) \frac{-g'(x_0)}{g(x_0)^2} && i : 10.1.13[g] \\
&= \frac{f'(x_0)g(x_0)}{g(x_0)^2} - \frac{f(x_0)g'(x_0)}{g(x_0)^2} \\
&= \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g(x_0)^2}.
\end{aligned}$$

Thus, f/g is differentiable at x_0 and $(f/g)'(x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g(x_0)^2}$. \square

Remark 1.1.14. *The product rule is also known as the Leibniz rule, after Gottfried Leibniz (1646–1716), who was the other founder of differential and integral calculus besides Newton.*

Remark 1.1.15. *The trick of adding and subtracting an intermediate term is sometimes known as the “middle-man trick” and is very useful in analysis.*

Theorem 1.1.16 (Chain rule). *Let X, Y be subsets of \mathbb{R} , let $x_0 \in X$ be a limit point of X , and let $y_0 \in Y$ be a limit point of Y . Let $f : X \rightarrow Y$ be a function such that $f(x_0) = y_0$, and such that f is differentiable at x_0 . Suppose that $g : Y \rightarrow \mathbb{R}$ is a function which is differentiable at y_0 . Then the function $g \circ f : X \rightarrow \mathbb{R}$ is differentiable at x_0 , and*

$$(g \circ f)'(x_0) = g'(y_0)f'(x_0)$$

Proof. By i:10.1.7 we want to show that

$$\begin{aligned}
& \forall \varepsilon \in^+, \exists \delta \in^+ : \forall x \in X, x - x_0 \leq \delta \\
& \implies g(f(x)) - g(y_0) - g'(y_0)f'(x_0)(x - x_0) \leq \varepsilon x - x_0.
\end{aligned}$$

Since $g'(y_0)$ exists, by i:10.1.7 we have

$$\begin{aligned}
& \forall \varepsilon \in^+, \exists \delta_g \in^+ : \forall y \in Y, y - y_0 \leq \delta_g \\
& \implies g(y) - g(y_0) - g'(y_0)(y - y_0) \leq \varepsilon_g y - y_0
\end{aligned}$$

where

$$\varepsilon_g = \frac{-(f'(x_0) + g'(y_0)) + \sqrt{(f'(x_0) + g'(y_0))^2 + 4\varepsilon}}{2}.$$

Note that $\varepsilon_g \in^+$ since

$$\sqrt{(f'(x_0) + g'(y_0))^2 + 4\varepsilon} > \sqrt{(f'(x_0) + g'(y_0))^2} = f'(x_0) + g'(y_0).$$

Now fix such ε (and ε_g) and δ_g . Since $f'(x_0)$ exists, by i:10.1.12 f is continuous at x_0 and by i:9.4.1 $\lim_{x \rightarrow x_0; x \in X} f(x) = f(x_0)$. This means

(by i:9.4.7(d) and i:10.1.7)

$$\begin{aligned} &\implies f(x) - y_0 \leq \delta_g \\ &f(x) - y_0 - f'(x_0)(x - x_0) \leq \varepsilon_g x - x_0 \\ &f(x) - y_0 \leq \varepsilon_g x - x_0 + f'(x_0)(x - x_0) \\ &\implies f(x) - y_0 \leq \delta_g \\ &f(x) - y_0 - f'(x_0)(x - x_0) \leq \varepsilon_g x - x_0 \\ &f(x) - y_0 \leq (\varepsilon_g + f'(x_0))x - x_0 \\ &\exists \delta \in^+: \forall x \in X, x - x_0 \leq \delta \implies f(x) - y_0 \leq \delta_g f(x) - y_0 - f'(x_0)(x - x_0) \leq \varepsilon_g x - x_0 \quad (\text{by i:9.4.7(d) and } \end{aligned}$$

We know that

$$\begin{aligned} &\exists \delta \in^+: \forall x \in X, x - x_0 \leq \delta \\ &\implies f(x) - y_0 \leq \delta_g \\ &\implies g(f(x)) - g(y_0) - g'(y_0)(f(x_0) - y_0) \leq \varepsilon_g f(x) - y_0 \\ &\implies g(f(x)) - g(y_0) - g'(y_0)f'(x_0)(x - x_0) - g'(y_0)(f(x) - y_0 - f'(x_0)(x - x_0)) \\ &\quad \leq \varepsilon_g f(x) - y_0 \\ &\implies g(f(x)) - g(y_0) - g'(y_0)f'(x_0)(x - x_0) \\ &\quad \leq \varepsilon_g f(x) - y_0 + g'(y_0)(f(x) - y_0 - f'(x_0)(x - x_0)) \\ &\implies g(f(x)) - g(y_0) - g'(y_0)f'(x_0)(x - x_0) \\ &\quad \leq \varepsilon_g f(x) - y_0 + g'(y_0)f(x_0) - y_0 - f'(x_0)(x - x_0) \\ &\implies g(f(x)) - g(y_0) - g'(y_0)f'(x_0)(x - x_0) \\ &\quad \leq \varepsilon_g f(x) - y_0 + \varepsilon_g g'(y_0)x - x_0 \\ &\implies g(f(x)) - g(y_0) - g'(y_0)f'(x_0)(x - x_0) \\ &\quad \leq \varepsilon_g(\varepsilon_g + f'(x_0))x - x_0 + \varepsilon_g g'(y_0)x - x_0 \\ &\implies g(f(x)) - g(y_0) - g'(y_0)f'(x_0)(x - x_0) \\ &\quad \leq \varepsilon_g(\varepsilon_g + f'(x_0) + g'(y_0))x - x_0. \end{aligned}$$

Expanding ε_g we have

$$\begin{aligned}
 & \varepsilon_g(\varepsilon_g + f'(x_0) + g'(y_0)) \\
 &= \frac{-(f'(x_0) + g'(y_0)) + \sqrt{(f'(x_0) + g'(y_0))^2 + 4\varepsilon}}{2} \\
 & \quad \times \frac{(f'(x_0) + g'(y_0)) + \sqrt{(f'(x_0) + g'(y_0))^2 + 4\varepsilon}}{2} \\
 &= \frac{(f'(x_0) + g'(y_0))^2 + 4\varepsilon - (f'(x_0) + g'(y_0))^2}{4} \\
 &= \varepsilon.
 \end{aligned}$$

Thus, we conclude that

$$\begin{aligned}
 & \forall \varepsilon \in^+, \exists \delta \in^+ : \forall x \in X, x - x_0 \leq \delta \\
 & \implies g(f(x)) - g(y_0) - g'(y_0)f'(x_0)(x - x_0) \leq \varepsilon x - x_0
 \end{aligned}$$

and by i:10.1.7 we have $(g \circ f)'(x_0) = g'(y_0)f'(x_0)$. □

Remark 1.1.17. *If one writes y for $f(x)$, and z for $g(y)$, then the chain rule can be written in the more visually appealing manner $\frac{dz}{dx} = \frac{dz}{dy} \frac{dy}{dx}$. However, this notation can be misleading (for instance it blurs the distinction between dependent variable and independent variable, especially for y), and leads one to believe that the quantities dz, dy, dx can be manipulated like real numbers. However, these quantities are not real numbers (in fact, we have not assigned any meaning to them at all), and treating them as such can lead to problems in the future. For instance, if f depends on x_1 and x_2 , which depend on t , then chain rule for several variables asserts that $\frac{df}{dt} = \frac{\partial f}{\partial x_1} \frac{dx_1}{dt} + \frac{\partial f}{\partial x_2} \frac{dx_2}{dt}$, but this rule might seem suspect if one treated df, dt , etc. as real numbers. It is possible to think of dy, dx , etc. as “infinitesimal real numbers” if one knows what one is doing, but for those just starting out in analysis, I would not recommend this approach, especially if one wishes to work rigorously. (There is a way to make all of this rigorous, even for the calculus of several variables, but it requires the notion of a tangent vector, and the derivative map, both of which are beyond the scope of this text.)*

Exercises

- 10.1.1** Suppose that X is a subset of \mathbb{R} , x_0 is a limit point of X , and $f : X \rightarrow \mathbb{R}$ is a function which is differentiable at x_0 . Let $Y \subseteq X$ be such that $x_0 \in Y$, and x_0 is also a limit point of Y . Prove that the restricted function $f|_Y : Y \rightarrow \mathbb{R}$ is also differentiable at x_0 , and has the same derivative as f at x_0 . Explain why this does not contradict the discussion in i:10.1.2.

i:ex:10.1.1. Since f is differentiable at x_0 , by Newton's approximation (i:10.1.7) we have $\forall \varepsilon \in ^+, \exists \delta \in ^+$ such that

$$\forall x \in X, x - x_0 \leq \delta \implies f(x) - (f(x_0) + f'(x_0)(x - x_0)) \leq \varepsilon x - x_0.$$

Since $Y \subseteq X$, we have

$$\begin{aligned} & \forall x \in Y, x - x_0 \leq \delta \\ \implies & (x \in X) \wedge (x - x_0 \leq \delta) \\ \implies & f(x) - (f(x_0) + f'(x_0)(x - x_0)) \leq \varepsilon x - x_0 \\ \implies & f|_Y(x) - (f|_Y(x_0) + f'(x_0)(x - x_0)) \leq \varepsilon x - x_0. \end{aligned}$$

Thus, by Newton's approximation (i:10.1.7) we know that $f|'_Y(x_0) = f'(x_0)$. This does not contradict to i:10.1.2 since x_0 is a limit point of Y implies x_0 is also a limit point of X . \square

10.1.2 Prove i:10.1.7.

i:ex:10.1.2. See i:10.1.7. \square

10.1.3 Prove i:10.1.10.

i:ex:10.1.3. See i:10.1.10. \square

10.1.4 Prove i:10.1.13.

i:ex:10.1.4. See i:10.1.13. \square

10.1.5 Let n be a natural number, and let $f : \rightarrow$ be the function $f(x)x^n$. Show that f is differentiable on and $f'(x) = nx^{n-1}$ for all $x \in$ with the convention that $nx^{n-1} = 0$ when $n = 0$.

i:ex:10.1.5. We induct on n to show that $\forall n \in, f_n(x) = x^n$ is differentiable on and $f'_n(x) = nx^{n-1}$. For $n = 0$, we have $f_0(x) = x^0 = 1$. By i:10.1.13(a) we know that f_0 is differentiable on and $f'_0(x) = 0$ for every $x \in X$. Thus, (by convention) the base case holds. Suppose inductively that for some $n \geq 0$ we have $f_n(x) = x^n$ is differentiable on and $f'_n(x) = nx^{n-1}$. Then for $n + 1$, we have $f_{n+1}(x) = x^{n+1} = x^n \cdot x^1 = f_n(x)f_1(x) = (f_n \cdot f_1)(x)$ and

$$\begin{aligned} & (f_n \cdot f_1)'(x) \\ = & f'_n(x)f_1(x) + f_n(x)f'_1(x) && i : 10.1.13[d] \\ = & (nx^{n-1})(x^1) + f_n(x)f'_1(x) \\ = & (nx^{n-1})(x^1) + (x^n)(1x^0) && i : 10.1.13[b] \\ = & nx^n + x^n \\ = & (n + 1)x^n. \end{aligned}$$

This closes the induction. \square

10.1.6 Let n be a *negative* integer, and let $f : \setminus 0 \rightarrow$ be the function $f(x)x^n$. Show that f is differentiable on $\setminus 0$ and $f'(x) = nx^{n-1}$ for all $x \in \setminus 0$.

i:ex:10.1.6. Let $x \in \setminus 0$. Since $n \in^-$, $-n \in^+$. Then we have $(1/f)(x) = 1/x^n = x^{-n}$ and $f(x) = (1/(1/f))(x)$. Thus, f is differentiable at x and

$$\begin{aligned} f'(x) &= (1/(1/f))'(x) \\ &= -\frac{(1/f)'(x)}{((1/f)(x))^2} && i : 10.1.13[g] \\ &= -\frac{(-n)x^{-n-1}}{x^{-2n}} && i : ex : 10.1.5 \\ &= nx^{n-1}. \end{aligned}$$

□

10.1.7 Prove i:10.1.15.

i:ex:10.1.7. See i:10.1.15.

□