Differentiation

Notes by Walter Noll, 1991

1 Differentiability, derivatives

Let I be a genuine real interval. Recall that the space of all functions with domain I is denoted by $\operatorname{Fun} I$ and the subspace of all continuous functions with domain I by $\operatorname{Cont} I$.

Let $f \in \operatorname{Fun} I$ and $t \in I$ be given. Consider the new function

$$\frac{(f \circ (t + \iota)) - f(t)}{\iota} : (-t + I) \backslash \{0\} \to \mathbb{R}. \tag{1.1}$$

Here, we have omitted the symbols that indicate necessary adjustments of domain or codomain. Also, we have denoted constants by their values. Had we not done so, this function would have been written

$$\frac{f \circ \left(t_{(-t+I)\to\mathbb{R}} + \iota\big|_{(-t+I)} \right)\Big|^{I} - f(t)_{(-t+I)\to\mathbb{R}}}{\iota\big|_{(-t+I)}} \ . \tag{1.2}$$

Clearly, it makes sense to ask whether this new function has a limit at 0.

Definition 1: *We say that $f : I \to \mathbb{R}$ is* **differentiable at** *a given $t \in I$ if the limit*

$$\partial_t f := \lim_0 \frac{f \circ (t + \iota) - f(t)}{\iota} \tag{1.3}$$

exists. If this is the case, then $\partial_t f$ is called the **derivative of f at t**.

We denote the set of all functions$f \in \operatorname{Fun} I$ that are differentiable at t by $\operatorname{Diff}_t I$.

We say that f is **differentiable** *if it is differentiable at every $t \in I$. If this is the case, the* **derivative** *$f^\bullet : I \to \mathbb{R}$ of f is defined by*

$$f^\bullet(t) := \partial_t f \text{ for all } t \in I. \tag{1.4}$$

We denote the set of all functions $f \in \operatorname{Fun} I$ that are differentiable by $\operatorname{Diff} I$.

We now list some basic facts of differential calculus. The first three are fairly easy to prove.

Proposition 1.1. *If $f \in \operatorname{Fun} I$ is differentiable at $t \in I$ then f is also continuous at t. If $f \in \operatorname{Fun} I$ is differentiable then f is also continuous .*

The converse if false: A function can be continuous at a point without being differentiable at that point. For example, the absolute-value function $(t \mapsto |t|) : \mathbb{R} \to \mathbb{R}$ is continuous but not differentiable at 0. One can even construct a function $f :]0, 1[\to \mathbb{R}$ which is continuous but not differentiable at *every* $t \in]0, 1[$. (The construction is very difficult.)

Proposition 1.2. *For every $t \in I$, $\operatorname{Diff}_t I$ is a subspace of $\operatorname{Fun} I$ and $\partial_t : \operatorname{Diff}_t I \longrightarrow \mathbb{R}$ is a linear functional.*

The mapping $(f \mapsto f^\bullet) : \operatorname{Diff} I \longleftarrow \operatorname{Cont} I$ is linear.

It is evident, from the definition, that all constants are differentiable and have the same derivative, namely the constant zero. The following converse is not as trivial and depends strongly on the assumption the I is a genuine interval.

Proposition 1.3. *Let a $f \in \mathrm{Diff\,I}$ be given. If $f^\bullet = 0$, then f is a constant.*

We say that a function $F \in \mathrm{Diff\,I}$ is an **anti-derivative** *of a given function $f \in \mathrm{FunI}$ if $F^\bullet = f$. If f has such an anti-derivative F, then*

$$\{F + c \mid c \text{ is a constant}\}$$

is the set of all anti-derivatives of f.

The following theorems will be presented with proofs.

Theorem 1.1. (product rule) *Let $t \in I$ and $f, g \in \mathrm{Diff_t\,I}$ be given. Then $fg \in \mathrm{Diff_t\,I}$ and*

$$\partial_t(fg) = f(t)\partial_t g + \partial_t f)g(t). \tag{1.5}$$

Given $f, g \in \mathrm{Diff\,I}$ we have $fg \in \mathrm{Diff_t\,I}$ and

$$(fg)^\bullet = f(g^\bullet) + (f^\bullet)g. \tag{1.6}$$

Proof. To be supplied. □

Theorem 1.2. (inversion rule) *Assume that $f : I \to J$ is invertible and differentiable, and that its derivative $f^\bullet : I \to \mathbb{R}$ is continuous. Let*

$$I' := I \backslash f^{\bullet <}(\{0\}) = \{t \in I \mid f^\bullet(t) \neq 0\}, \quad J' := f_>(I'). \tag{1.7}$$

Then $f^{\leftarrow}|_{J'}$ is differentiable and

$$\left(f^{\leftarrow}|_{J'}\right)^\bullet = \frac{1}{f^\bullet} \circ \left(f^{\leftarrow}|_{J'}^{I'}\right). \tag{1.8}$$

Proof. To be supplied. □

Theorem 1.3. Chain Rule: *Let genuine intervals I, J, and K and functions $f : I \to J$ and $g : J \to K$ be given.*

If f is differentiable at a given $t \in I$ and if g is differentiable at $f(t)$, then $g \circ f$ is differentiable at t and

$$\partial_t(g \circ f) = (\partial_{f(t)} g)(\partial_t f). \tag{1.9}$$

If f and g are both differentiable, so is $g \circ f$ and

$$(g \circ f)^\bullet = (g^\bullet \circ f) f^\bullet. \tag{1.10}$$

Proof: Assume that f is differentiable at $t \in I$ and g is differentiable at $z := g(t)$. The latter assumption means that the function $h : J \to \mathbb{R}$ defined by

$$h(x) := \left\{ \begin{array}{ccc} \dfrac{g(x) - g(z)}{x - z} & \text{for all} & x \in J \backslash \{z\} \\ \\ \partial_z g & \text{for} & x := z \end{array} \right\} \tag{1.11}$$

is continuous at z. Since the differentiability of f at t implies the continuity of f at t, it follows from Prop. 6,1 in *Limits and Continuity* that $h \circ f$ is continuous at t and hence, by Prop. 7.5 in *Limits and Continuity*, that

$$\lim_t (h \circ f)\Big|_{I \backslash \{t\}} = (h \circ f)(t) = h(z) = \partial_z g. \tag{1.12}$$

Noting the Dom $\left(\dfrac{f - f(t)}{\iota - t}\right) = I \backslash \{t\}$ and that $\dfrac{f - f(t)}{\iota - t}$ converges to $\partial_t f$ at t by assumption, it follows from Prop. 7.7 in *in Limits and Continuity*and from (1.12) that

$$\lim_{t}\left((h\circ f)|_{I\backslash\{t\}}\,\frac{f-f(t)}{\iota-t}\right)=\partial_z g\partial_t f,\quad z:=f(t). \tag{1.13}$$

Now let $s\in I\backslash\{t\}$ be given. We then have, by (1.11),

$$\left((h\circ f)|_{I\backslash\{t\}}\,\frac{f-f(t)}{\iota-t}\right)(s)=h(f(s))\frac{f(s)-f(t)}{s-t}$$

$$=\left\{\begin{array}{lll}\dfrac{g(f(s))-g(f(t))}{f(s)-f(t)}\cdot\dfrac{f(s)-f(t)}{s-t} & \text{if} & f(s)\neq f(t)\\[2ex] \partial_z g\cdot\dfrac{f(s)-f(t)}{s-t} & \text{if} & f(s)=f(t)\end{array}\right\}$$

$$=\left\{\begin{array}{lll}\dfrac{g(f(s))-g(f(t))}{s-t} & \text{if} & f(s)\neq f(t)\\[2ex] 0 & \text{if} & f(s)=f(t)\end{array}\right\}=\frac{g(f(s))-g(f(t))}{s-t}.$$

Hence, since $s\in I\backslash\{t\}$ was arbitrary, we have

$$(h\circ f)|_{I\backslash\{t\}}\,\frac{f-f(t)}{\iota-t}=\frac{g\circ f-(g\circ f)(t)}{\iota-t}.$$

Hence, by (1.13), we conclude that (1.9) is valid. ∎

Note that there are two statements in each of these theorems. The first asserts that a certain function is differentiable, the second gives a formula that tells how to obtain the derivative.

Definition 2: *We say that a function $f\in\operatorname{Diff}I$ is* **twice differentiable** *if its derivative $f^{\bullet}$ is differentiable and we define its* **second derivative** *$f^{\bullet\bullet};I\longrightarrow\mathbb{R}$ by $f^{\bullet\bullet}:=(f^{\bullet})^{\bullet}$. By recursion, given $n\in\mathbb{N}$, we can define what it means that $f\in\operatorname{Diff}I$ is n times differentiable and define its* **n'th derivative** *$f^{(n)}:I\longrightarrow\mathbb{R}$. We say that the function $f\in\operatorname{Fun}I$ is* **of class** *C^n and belongs to $C^n(I)$ if it is n times differentiable and $f^{(n)}$ is continuous. We say that the function $f\in\operatorname{Fun}\mathrm{I}$ is* **of class** *C^∞ and belongs to $C^\infty(I)$ if it is n times differentiable for all $n\in\mathbb{N}$.*

The sets $\mathrm{C}^n(I)$ with $n\in\mathbb{N}$ or $n:=\infty$ are all subspaces of $\operatorname{Cont}I$.

2 Important Theorems

In this section, we assume that a genuine interval I and functions $f, g \in \operatorname{Fun} I$ are given.

Theorem 2.1. Local Extremum Theorem: *Let $t \in I$ given such that t is not an endpoint of I and that f is differentiable at t. If there is $\delta \in \mathbb{P}^x$ such that $]t-\delta, t+\delta[\subset I$ and*

$$f(t) = \max f_>(]t-\delta, t+\delta[) \quad \text{or} \quad f(t) = \min f_>(]t-\delta, t+\delta[)$$

then $\partial_t f = 0$.

Proof: To say that f is differentiable at t, with derivative $\partial_t f$, means that the function $h : I \to \mathbb{R}$ defined by

$$h(s) := \left\{ \begin{array}{ccc} \dfrac{f(s)-f(t)}{s-t} & \text{for all} & s \in I \backslash \{t\} \\ \\ \partial_t f & \text{for} & s := t \end{array} \right\} \tag{2.1}$$

is continuous at t.

Now let $\delta \in \mathbb{P}^x$ such that $]t-\delta, t+\delta[\ \subset I$ be given. Suppose that $\partial_t f = h(t) > 0$. It follows from Prop.6.4 in *Limits and Continuity* that we can choose $s_1 \in\ ]t, t+\delta[$ and $s_2 \in\ ]t-\delta, t[$ such that $h(s_1) > 0$ and $h(s_2) > 0$. In view of (2.1), this means that

$$f(s_1) - f(t) > s_1 - t > 0 \quad \text{and} \quad f(t) - f(s_2) > t - s_2 > 0.$$

Hence $f(s_2) > f(t)$ and $f(s_2) < f(t)$, showing that $f(t)$ cannot be the maximum or the minimum of $f_>(]t-\delta, t+\delta[)$.

If $\partial_t f < 0$, we apply the argument above to $-f$ instead of f and arrive at the same conclusion. ■

Theorem 2.2. Rolle's Theorem: *Let $a, b \in I$ with $a \neq b$ be given. Assume that $f|_{[a,b]}$ is continuous and that $f|_{]a,b[}$ is differentiable. If $f(a) = f(b)$ then there is at least one $t \in]a, b[$ such that $\partial_t f = 0$.*

Proof: By the Theorem on Attainment of Extrema, $f|_{[a,b]}$ must attain a maximum and a minimum. Since $f(a) = f(b)$ at least one of these must be attained at a point $t \in]a,b[$. By the Local Extremeum Theorem, we have $\partial_t f = 0$. ■

Theorem 2.3. General Mean Value Theorem: *Let $a, b \in I$ be given. Assume that $f|_{[a,b]}$ and $g|_{[a,b]}$ are continuous and that $f|_{]a,b[}$ and $g|_{]a,b[}$ are differentiabole. Then there is $t \in]a,b[$ such that*

$$(f(b) - f(a))\,\partial_t g = (g(b) - g(a))\,\partial_t f. \tag{2.2}$$

Proof: If $f(b) = f(a)$, we can apply Rolle's Theorem and determine $t \in]a,b[$ such that $\partial_t f = 0$. Then (2.2) is valid because it reduces to $0 = 0$.

Assume now that $f(b) \neq f(a)$. We consider $h : I \to R$ defined by

$$h := g - Af \quad \text{with} \quad A := \frac{g(b) - g(a)}{f(b) - f(a)}. \tag{2.3}$$

Then

$$h(b) - h(a) = g(b) - g(a) - A(f(b) - f(a)) = 0$$

and hence $h(a) = h(b)$. Application of Rolle's Theorem to h shows that $\partial_t h = 0$ for some $t \in\ ]a,b[$ and hence, by (2.3), that $\partial_t g - A\partial_t f = 0$ for some $t \in\ ]a,b[$, which gives (2.2). ■

Theorem 2.4. Difference-Quotient Theorem: *Let $a, b \in I$ with $a \neq b$ be given. Assume that $f|_{[a,b]}$ is continuous and that $f|_{]a,b[}$ is differentiable. Then*

$$\frac{f(b) - f(a)}{b - a} \in \text{ Rng } \left(f\,|_{]a,b[}\right)^{\bullet}. \tag{2.4}$$

Proof: Apply the General Mean Value Theorem to the case when $g := \iota\,|_I$. ■

Theorem 2.5. L'Hôpital's Rule: *Let $f.g \in \operatorname{Fun} I$ be such that that $0 \notin \operatorname{Rng} f$ and $0 \notin \operatorname{Rng} f^{\bullet}$. Let a be an endpoint of I such that $a \notin I$. If f and g converge to 0 at a and if $\frac{g^{\bullet}}{f^{\bullet}}$ converges at a, then $\frac{g}{f}$ also converges at a and*

$$\lim_a \frac{g}{f} = \lim_a \frac{g^\bullet}{f^\bullet}. \tag{2.5}$$

Proof: We assume that a is a left endpoint of I and put $\overline{I} := \{a\} \cup I$. We assume that f and g converge to 0 at a. This means that the functions $\overline{f}, \overline{g} : \overline{I} \to \mathbb{R}$ defined by

$$\overline{g}(s) := \left\{ \begin{array}{lll} g(s) & \text{for all} & s \in I \\ & & \\ 0 & \text{for} & s := a \end{array} \right\}, \quad \overline{f}(s) := \left\{ \begin{array}{lll} f(s) & \text{for all} & s \in I \\ & & \\ 0 & \text{for} & s := a \end{array} \right\} \tag{2.6}$$

are continuous at a and hence continuous. Also, for every $t \in I$ the functions $\overline{g}\,|_{]a,t[} = g|_{]a,t[}$ and $\overline{f}|_{]a,t[} = f|_{]a,t[}$ are differentiable.

Now assume that $\dfrac{g^\bullet}{f^\bullet}$ converges at a and put

$$c := \lim_a \frac{g^\bullet}{f^\bullet} \tag{2.7}$$

Let $\varepsilon \in \mathbb{P}^x$ be given. We then may determine $\delta \in \mathbb{P}^x$ such that $a + \delta \in I$ and

$$\left| \frac{g^\bullet(s)}{f^\bullet(s)} - c \right| < \varepsilon \text{ for all } s \in \,]a, a+\delta[. \tag{2.8}$$

Now let $t \in \,]a, a+\delta[$ be given. By the General Mean Value Theorem, we can determine $s \in \,]a, a+t[\, \subset \,]a, a+\delta[$ such that

$$\frac{g(t)}{f(t)} = \frac{\overline{g}(t) - \overline{g}(a)}{\overline{f}(t) - \overline{f}(a)} = \frac{g^\bullet(s)}{f^\bullet(s)}.$$

By (2.8) we have

$$\left| \frac{g(t)}{f(t)} - c \right| < \varepsilon.$$

Since $t \in]a, a+\delta[$ was arbitary and since $\varepsilon \in \mathbb{P}^x$ was arbitrary, we conclude that $\dfrac{g}{f}$ converges to c at a.

The case when a is a right endpoint of I is easily reduced to the case considered above. ∎

Theorem 2.6. Taylor's Theorem: *Let $n \in \mathbb{N}^\times$ be given and assume that f is n times differentiable. Let $a, b \in I$ with $a \neq b$ be given. Then there is $t \in]a, b[$ such that*

$$f(b) = \sum_{k \in n^t} \frac{(b-a)^k}{k!} f^{(n)}(a) + \frac{(b-a)^n}{n!} f^{(n)}(t). \tag{2.9}$$

Proof: Consider the function $h : I \to \mathbb{R}$ defined by

$$h := \sum_{k \in n^t} \frac{(b-\iota)^k}{k!} f^{(k)}. \tag{2.10}$$

An easy calculation, using the formulas $((b-\iota)^k)^\bullet = k(b-\iota)^{k-1}$, $\left(f^{(h)}\right)^\bullet = f^{(k+1)}$ for all $k \in (n-1)^]$, the Sum-Rule, and the Product Rule, shows that h is differentiable and

$$h^\bullet = \frac{(b-\iota)^{n-1}}{(n-1)!} f^{(n)}. \tag{2.11}$$

Now consider the function $g : I \to \mathbb{R}$ defined by

$$g := h + \frac{(b-\iota)^n}{n!} A, \text{ with } A = \frac{n!}{(b-a)^n}(h(b) - h(a)) \tag{2.12}$$

An easy calculation shows that

$$g(b) = h(b) = g(a). \tag{2.11}$$

By (2.12) and (2.13), g is differentiable with

$$g^\bullet = h^\bullet - \frac{(b-\iota)^{n-1}}{(n-1)!} A = \frac{(b-\iota)^{n-1}}{(n-1)!}(f^{(n)} - A) \tag{2.14}$$

By (2.13) we may apply Rolle's Theorem to g and conclude that $g^\bullet(t) = 0$ for some $t \in]a, b[$. Hence, by (2.14), we have $A = f^{(n)}(t)$ for some $t \in]a, b[$. The desired result (2.9) follows from $(2.12)_2$ and (2.10). ∎