12: Differentiable Functions on $\mathbb{R}$

(8/4/23)

Alex Nita

Contents

1 Definition of Differentiability and its Basic Properties

Definition 1 Let $A \subseteq \mathbb{R}$ and let $f : A \to \mathbb{R}$. We say that f is **differentiable at a point** $a \in A$ if the functional limit

$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a} \stackrel{\text{def}}{=} f'(a)$$

exists in $\mathbb{R}$ (in which case it is called the **derivative** of f at a), that is,

$$\forall \varepsilon > 0,\ \exists \delta > 0,\ \left(0 < |x - a| < \delta \implies \left|\frac{f(x) - f(a)}{x - a} - f'(a)\right| < \varepsilon\right)$$

By a change-of-variable, letting $h \stackrel{\text{def}}{=} x - a$, so that $x = a + h$, we can write this limit in its other familiar form

$$\lim_{h \to 0} \frac{f(a + h) - f(a)}{h} \stackrel{\text{def}}{=} f'(a)$$

Expanded into its definitional ε-δ terms, this says

$$\forall \varepsilon > 0,\ \exists \delta > 0,\ \left(0 < |h| < \delta \implies \left|\frac{f(a + h) - f(a)}{h} - f'(a)\right| < \varepsilon\right)$$

We say f is **differentiable on** A if it is differentiable at every point in A. The **set of all differentiable functions on** A is denoted

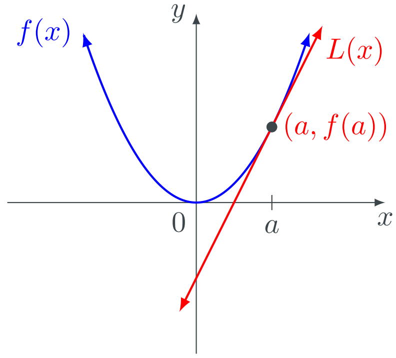
$$\mathcal{D}(A) \stackrel{\text{def}}{=} \textbf{all differentiable functions } f \textbf{ on } A$$ ■

Remark 2 *The derivative $f'(a) \in \mathbb{R}$, when it exists, is interpreted as the* instant rate of change of y with respect to x, *or as* slope of the tangent line to the graph of f at $(a, f(a))$. *Let us take this latter interpretation and find the* ***equation of the tangent line*** *to* graph f *at* $(a, f(a))$.

(1) *Use point-slope form with $P = (a, f(a))$ and $m = f'(a)$ to get an equation of the tangent line:*

$$y - f(a) = f'(a)(x-a) \quad \Longrightarrow \quad \boxed{\underbrace{y = f(a) + f'(a)(x-a)}_{\text{equation of tangent line, } L(x)}}$$

which some of you may remember from calculus as the ***linear approximation*** *to f or* ***local linearization*** *of f near $a \in A$, or, again, as the first Taylor polynomial of f at $a \in A$:*



$$\boxed{L(x) \stackrel{\text{def}}{=} f(a) + f'(a)(x-a)}$$

(2) *Use the definition of $f'(a)$ as a functional limit to clarify the idea of* ***approximation****: $\forall \varepsilon > 0$, $\exists \delta > 0$ such that $0 < |x-a| < \delta$ implies*

$$\left| \frac{f(x) - f(a)}{x-a} - f'(a) \right| < \varepsilon \iff -\varepsilon < \frac{f(x)-f(a)}{x-a} - f'(a) < \varepsilon$$
$$\iff -\varepsilon \cdot (x-a) < f(x) - \underbrace{\big(f(a) + f'(a)(x-a)\big)}_{= L(x)} < \varepsilon \cdot (x-a)$$

which shows that, if we take $\delta \le 1$ just as a precaution, we have $\forall x \in V_\delta(a)$

$$\boxed{|f(x) - L(x)| < \varepsilon |x-a| < \varepsilon \cdot \delta \le \varepsilon}$$

so that indeed near $x = a$

$$\boxed{f(x) \approx L(x) = f(a) + f'(a)(x-a)}$$

■

Notation 3 *The derivative of f at $a \in A$ is variously denoted as*

(Lagrange)	$f'(a)$
(Leibniz)	$\dfrac{df}{dx}(a)$ or $\dfrac{d\big(f(a)\big)}{dx}$ or $\left.\dfrac{df}{dx}\right\rvert_{x=a}$
(Euler)	$Df(a)$
(Newton)	$\dot{f}(a)$

Each has its virtues. We encounter $f'(a)$ and df/dx in calculus books, $Df(a)$ is multivariate analysis, and $\dot{f}(a)$ in classical physics texts. ■

Definition 4 If $A \subseteq \mathbb{R}$ and $f \in \mathcal{D}(A)$, we have a whole new function, the **derivative function**,

$$f' : A \to \mathbb{R}$$

$$f'(a) \stackrel{\text{def}}{=} \lim_{x \to a} \frac{f(x) - f(a)}{x - a}$$

If $f' \in \mathcal{D}(A)$, then we have the **second derivative function**,

$$f'' : A \to \mathbb{R}$$

$$f''(a) \stackrel{\text{def}}{=} (f')'(a) \stackrel{\text{def}}{=} \lim_{x \to a} \frac{f'(x) - f'(a)}{x - a}$$

Proceeding inductively, we may define the n**th derivative function** by

$$f^{(n)} : A \to \mathbb{R}$$

$$f^{(n)}(a) \stackrel{\text{def}}{=} (f^{(n-1)})'(a) \stackrel{\text{def}}{=} \lim_{x \to a} \frac{f^{(n-1)}(x) - f^{(n-1)}(a)}{x - a}$$

Let us denote the set of all k-times differentiable functions on A by

$$\mathcal{D}^k(A) \stackrel{\text{def}}{=} \{f : A \to \mathbb{R} \mid \exists f^{(i)} : A \to \mathbb{R}, \text{ for all } 0 \le i \le k\}$$ ■

Lemma 5 $\lim_{x\to a} f(x) = f(a) \quad \iff \quad \lim_{x\to a}\big(f(x) - f(a)\big) = 0.$

Proof: Suppose $\lim_{x\to a} f(x) = f(a)$, and observe that $\lim_{x\to a} f(a) = f(a)$ since $f(a)$ is a constant. Using the difference (functional) limit law we have

$$\begin{aligned}
\lim_{x\to a}\big(f(x) - f(a)\big) &= \big(\lim_{x\to a}\big(f(x)\big) - \big(\lim_{x\to a} f(a)\big) \\
&= f(a) - f(a) \\
&= 0
\end{aligned}$$

Conversely, if $\lim_{x\to a}\big(f(x) - f(a)\big) = 0$, then, since $\lim_{x\to a} f(a) = f(a)$, the sum (functional) limit law gives

$$\begin{aligned}
\lim_{x\to a} f(x) &= \lim_{x\to a}\big(f(x) - f(a) + f(a)\big) \\
&= \lim_{x\to a}\big[\big(f(x) - f(a)\big) + f(a)\big] \\
&= 0 + f(a) \\
&= f(a)
\end{aligned}$$

■

Theorem 6 *Differentiable functions are continuous,* $\mathcal{D}(A) \subseteq C(A)$.

Proof: $f \in \mathcal{D}(A)$ means $\forall a \in A$ the limit

$$f'(a) \overset{\text{def}}{=} \lim_{x\to a} \frac{f(x) - f(a)}{x - a}$$

exists in $\mathbb{R}$. Therefore, to show that $\lim_{x\to a} f(x) = f(a)$, we use this definition, the fact that all polynomials are continuous on $\mathbb{R}$ (Corollary 26, Lecture 10), and the product limit law to conclude

$$\begin{aligned}
\lim_{x\to a}\big(f(x) - f(a)\big) &= \lim_{x\to a}\big(f(x) - f(a)\big) \cdot \frac{x-a}{x-a} \\
&= \lim_{x\to a}\left[\left(\frac{f(x) - f(a)}{x - a}\right) \cdot (x - a)\right] \\
&= f'(a) \cdot 0 \\
&= 0
\end{aligned}$$

The lemma then ensures $\lim_{x\to a} f(x) = f(a)$. ■

Remark 7 *This is only a set inclusion, so far. Let us show that $\mathcal{D}(A)$ is in fact a vector subspace of $C(A)$. Indeed, it is a real associative subalgebra of $C(A)$.* ■

Theorem 8 (Generic Rules of Differentiation) *If $f, g \in \mathcal{D}(A)$ and $c \in \mathbb{R}$, then c (the constant function), cf, $f \pm g$, fg and $f/g \in \mathcal{D}(A)$ (the last wherever $g \neq 0$). Moreover, $\forall a \in A$ we have*

(1)	$c' = 0$	*(constant function rule)*
(2)	$(cf)'(a) = cf'(a)$	*(scalar multiple rule)*
(3)	$(f \pm g)'(a) = f'(a) \pm g'(a)$	*(sum/difference rules)*
(4)	$(fg)'(x_0) = f(a)g'(a) + f'(a)g(a)$	*(product rule)*
(5)	$\left(\frac{f}{g}\right)'(a) = \frac{f'(a)g(a) - f(a)g'(a)}{g^2(a)}$	*(quotient rule)*

Proof: Note that we can prove all of them by the combo of definition of differentiability + functional limit laws:

(1) Viewing c as the range of a function $f : A \to \mathbb{R}$, $f(x) \overset{\text{def}}{=} c$, we have

$$c' = \lim_{x\to a} \frac{c - c}{x - a} = \lim_{x\to a} 0 = 0$$

(2) Since $(cf)(x) \overset{\text{def}}{=} c \cdot f(x)$, we have

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

(3) Likewise $(f \pm g)(x) \overset{\text{def}}{=} f(x) \pm g(x)$, so

$$\begin{aligned}(f \pm g)'(a) &= \lim_{x\to a} \frac{(f \pm g)(x) - (f \pm g)(a)}{x - a} \\ &= \lim_{x\to a} \left[\frac{f(x) - f(a)}{x - a} \pm \frac{g(x) - g(a)}{x - a}\right] \\ &= f'(a) \pm g'(a)\end{aligned}$$

(4) Since $(fg)(x) \overset{\text{def}}{=} f(x)g(x)$, we have

$$\begin{aligned}(fg)'(a) &= \lim_{x\to a} \frac{(fg)(x) - (fg)(a)}{x - a} \\ &= \lim_{x\to a} \frac{f(x)g(x) - f(x)g(a) + f(x)g(a) - f(a)g(a)}{x - a} \\ &= \lim_{x\to a} \left[f(x) \cdot \frac{g(x) - g(a)}{x - a} + g(a) \cdot \frac{f(x) - f(a)}{x - a}\right] \\ &= f(a)g'(a) + f'(a)g(a)\end{aligned}$$

($f(x) \to f(a)$ because $f \in \mathcal{D}(A) \subseteq C(A)$)

(5) Finally, since $\left(\frac{f}{g}\right)(x) \stackrel{\text{def}}{=} \frac{f(x)}{g(x)}$, we have

$$\begin{aligned}
\left(\frac{f}{g}\right)'(a) &= \lim_{x\to a} \frac{\left(\frac{f}{g}\right)(x) - \left(fg\right)(a)}{x-a} = \lim_{x\to a} \frac{\frac{f(x)}{g(x)} - \frac{f(a)}{g(a)}}{x-a} \\
&= \lim_{x\to a} \frac{g(a)f(x) - f(a)g(x)}{(x-a)g(x)g(a)} \\
&= \lim_{x\to a} \frac{g(a)f(x) - g(a)f(a) + g(a)f(a) - f(a)g(x)}{(x-a)} \cdot \frac{1}{g(x)g(a)} \\
&= \lim_{x\to a} \left(\left[g(a) \cdot \frac{f(x)-f(a)}{(x-a)} - f(a) \cdot \frac{g(x)-g(a)}{(x-a)} \right] \cdot \frac{1}{g(x)g(a)} \right) \\
&= \frac{g(a)f'(a) - f(a)g'(a)}{g^2(a)}
\end{aligned}$$

($g(x) \to g(a)$ because $g \in \mathcal{D}(A) \subseteq C(A)$). ■

Remark 9 *We could also prove (2) and (3) directly from the definition of $f'(a)$ as a functional limit: Choose $a \in A$ and $\varepsilon > 0$. Since $f \in \mathcal{D}(A)$, $\exists \delta > 0$ so that*

$$0 < |x-a| < \delta \implies \left| \frac{f(x)-f(a)}{x-a} - f'(a) \right| < \frac{\varepsilon}{|c|}$$

and therefore

$$\begin{aligned}
\left| \frac{cf(x)-cf(a)}{x-a} - cf'(a) \right| &= |c| \left| \frac{f(x)-f(a)}{x-a} - f'(a) \right| \\
&< |c| \cdot \frac{\varepsilon}{|c|} \\
&= \varepsilon
\end{aligned}$$

Similarly, if $f, g \in \mathcal{D}(A)$, then $\forall a \in A,\ \forall \varepsilon > 0,\ \exists \delta_1, \delta_2 > 0$ such that

$$0 < |x-a| < \delta \stackrel{\text{def}}{=} \min\{\delta_1, \delta_2\} \implies \begin{pmatrix} \left| \frac{f(x)-f(a)}{x-a} - f'(a) \right| < \frac{\varepsilon}{2} \\ \text{and} \\ \left| \frac{g(x)-g(a)}{x-a} - g'(a) \right| < \frac{\varepsilon}{2} \end{pmatrix}$$

so by a triangle inequality we have

$$\begin{aligned}
&\left| \frac{(f \pm g)(x) - (f \pm g)(a)}{x-a} - (f' \pm g')(a) \right| \\
&\quad = \left| \left(\frac{f(x)-f(a)}{x-a} - f'(a) \right) \pm \left(\frac{g(x)-g(a)}{x-a} - g'(a) \right) \right| \\
&\quad \leq \left| \frac{f(x)-f(a)}{x-a} - f'(a) \right| + \left| \frac{g(x)-g(a)}{x-a} - g'(a) \right| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon
\end{aligned}$$

The others could also be proven directly, I suppose, but the complications multiply with the product and quotient rules. We leave these as exercises! ■

Theorem 10 (Power Rule; Polynomials are Differentiable on $\mathbb{R}$)
For all $n \in \mathbb{N}$, the monomial $x^n \in \mathbb{R}[x]$ is differentiable everywhere, $x^n \in \mathcal{D}(\mathbb{R})$, and

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

As a result, all polynomials are differentiable everywhere,

$$\mathbb{R}[x] \subseteq \mathcal{D}(\mathbb{R}) \subseteq C(\mathbb{R})$$

and the derivative operator $\frac{d}{dx}$ restricts to a linear map on $\mathbb{R}[x]$:

$$\boxed{\left(\sum_{k=0}^{n} a_k x^k\right)' = \sum_{k=1}^{n} k a_k x^{k-1}}$$

Proof: Let us use the algebraic identity

$$x^n - a^n = (x-a)(x^{n-1} + x^{n-2}a + \cdots + xa^{n-2} + a^{n-1})$$

We have, $\forall a \in \mathbb{R}$,

$$\begin{aligned}
\frac{d(x^n)}{dx}(a) &\overset{\text{def}}{=} \lim_{x\to a} \frac{x^n - a^n}{x-a} \\
&= \lim_{x\to a} \frac{\cancel{(x-a)}(x^{n-1} + x^{n-2}a + \cdots + a^{n-1})}{\cancel{x-a}} \\
&= \lim_{x\to a} (x^{n-1} + x^{n-2}a + \cdots + a^{n-1}) \\
&= na^{n-1}
\end{aligned}$$

Since this is true for all $a \in \mathbb{R}$, we conclude that the derivative function is given by $(x^n)' = nx^{n-1}$. The action of $\frac{d}{dx}$ on $\mathcal{D}(\mathbb{R})$ is linear by (2)-(3) of the previous theorem, while the action on monomials produces other monomials, so $\frac{d}{dx}$ restricts to a linear operator on $\mathbb{R}[x]$, because polynomials are linear combinations of monomials. ■

We can enlarge this power rule to **all rational powers** p/q, in two steps, using the same algebraic identity as in the proof above:

Theorem 11 (qth Root Power Rule) *For all odd $q \in \mathbb{N}$ we have $x^{1/q} \in \mathcal{D}(\mathbb{R})$ and $x^{-1/q} \in \mathcal{D}(\mathbb{R}-\{0\})$, while for all even $q \in \mathbb{N}$ we have $x^{1/q} \in \mathcal{D}\big([0,\infty)\big)$ and $x^{-1/q} \in \mathcal{D}\big((0,\infty)\big)$, and their derivative functions are given by*

$$\left(x^{1/q}\right)' = \frac{1}{q}x^{(1/q)-1}$$
$$\left(x^{-1/q}\right)' = -\frac{1}{q}x^{-(1/q)-1}$$

Proof: Recall that for all real numbers a and b and all $p \in \mathbb{N}$ we have

$$a^q - b^q = (a-b)(a^{q-1} + a^{q-2}b + \cdots + ab^{q-2} + p^{q-1}) \tag{1}$$

(just foil out the right hand side). With $a = (x+h)^{1/q}$ and $b = x^{1/q}$, we get

$$\begin{aligned}
\left(x^{1/q}\right)' &= \lim_{h\to 0}\frac{(x+h)^{1/q} - x^{1/q}}{h}\cdot\frac{\big((x+h)^{1/q}\big)^{q-1} + \big((x+h)^{1/q}\big)^{q-2}x^{1/q} + \cdots + (x^{1/q})^{1-q}}{\big((x+h)^{1/q}\big)^{q-1} + \big((x+h)^{1/q}\big)^{q-2}x^{1/q} + \cdots + (x^{1/q})^{1-q}} \\
&= \lim_{h\to 0}\frac{\big((x+h)^{1/q}\big)^{q} - (x^{1/q})^q}{h\Big[\big((x+h)^{1/q}\big)^{q-1} + \big((x+h)^{1/q}\big)^{q-2}x^{1/q} + \cdots + (x^{1/q})^{1-q}\Big]} \\
&= \lim_{h\to 0}\frac{\not h}{\not h\Big[\big((x+h)^{1/q}\big)^{q-1} + \big((x+h)^{1/q}\big)^{q-2}x^{1/q} + \cdots + (x^{1/q})^{1-q}\Big]} \\
&= \frac{1}{qx^{(q-1)/q}} \\
&= \frac{1}{q}x^{(1/q)-1}
\end{aligned}$$

which proves the first result. For the second, we have, again by (1),

$$\begin{aligned}
\left(x^{-1/q}\right)' &= \lim_{h\to 0}\frac{(x+h)^{-1/q} - x^{-1/q}}{h} \\
&= \lim_{h\to 0}\frac{\frac{1}{(x+h)^{1/q}} - \frac{1}{x^{1/q}}}{h} \\
&= \lim_{h\to 0}\frac{\frac{x^{1/q}-(x+h)^{1/q}}{(x+h)^{1/q}x^{1/q}}}{h}\cdot\frac{(x^{1/q})^{q-1} + (x^{1/q})^{q-2}(x+h)^{1/q} + \cdots + \big((x+h)^{1/q}\big)^{q-1}}{(x^{1/q})^{q-1} + (x^{1/q})^{q-2}(x+h)^{1/q} + \cdots + \big((x+h)^{1/q}\big)^{q-1}} \\
&= \lim_{h\to 0}\frac{x-(x+h)}{h(x+h)^{1/q}x^{1/q}\Big[(x^{1/q})^{q-1} + (x^{1/q})^{q-2}(x+h)^{1/q} + \cdots + \big((x+h)^{1/q}\big)^{q-1}\Big]} \\
&= \frac{-\not h}{\not h(x+h)^{1/q}x^{1/q}\Big[(x^{1/q})^{q-1} + (x^{1/q})^{q-2}(x+h)^{1/q} + \cdots + \big((x+h)^{1/q}\big)^{q-1}\Big]} \\
&= -\frac{1}{qx^{2/q}x^{(q-1)/q}} \\
&= -\frac{1}{qx^{1+(1/q)}}
\end{aligned}$$

■

Theorem 12 (Rational Power Rule) *If $p, q \in \mathbb{Z}$, with $q > 0$, then we have $x^{p/q} \in \mathcal{D}(\mathbb{R} - \{0\})$ if q is odd, or $x^{p/q} \in \mathcal{D}\big((0,\infty)\big)$ if q is even (and including 0 if $p/q > 1$), and*

$$\left(x^{p/q}\right)' = \frac{p}{q}x^{p/q-1}$$

Proof: By the previous theorem, if $q \in \mathbb{N}$ we have $\frac{d}{dx}x^{1/q} = \frac{1}{q}x^{1/q-1} = \frac{1}{q}x^{(1-q)/q}$ so by another application of (1) we have

$$\begin{aligned}
\left(x^{p/q}\right)' &= \lim_{h\to 0}\frac{(x+h)^{p/q} - x^{p/q}}{h} \\
&= \lim_{h\to 0}\frac{\left((x+h)^{1/q}\right)^{p} - (x^{1/q})^{p}}{h} \\
&= \lim_{h\to 0}\frac{\left((x+h)^{1/q} - x^{1/q}\right)\Big(\left((x+h)^{1/q}\right)^{p-1} + \cdots + (x^{1/q})^{p-1}\Big)}{h} \\
&= \left[\lim_{h\to 0}\frac{(x+h)^{1/q} - x^{1/q}}{h}\right]\left[\lim_{h\to 0}\Big(\left((x+h)^{1/q}\right)^{p-1} + \cdots + (x^{1/q})^{p-1}\Big)\right] \\
&= \left[\frac{d}{dx}x^{1/q}\right]\left[(x^{1/q})^{p-1} + (x^{1/q})^{p-2}x^{1/q} + \cdots + (x^{1/q})^{p-1}\right] \\
&= \left[\frac{1}{q}x^{(1-q)/q}\right]\left[p(x^{1/q})^{p-1}\right] \\
&= \frac{p}{q}x^{(1-q)/q+(p-1)/q} \\
&= \frac{p}{q}x^{(p-q)/q} \\
&= \frac{p}{q}x^{p/q-1}
\end{aligned}$$

■

Theorem 13 (Chain Rule) *If $g \in \mathcal{D}(A)$ and $f \in \mathcal{D}\big(g(A)\big)$, then $f \circ g \in \mathcal{D}(A)$ and*

$$(f \circ g)'(x) = f'\big(g(x)\big)g'(x)$$

In Leibniz notation, and using $z = f(y)$ and $y = g(x)$, this becomes

$$\frac{dz}{dx} = \frac{dz}{dy}\cdot\frac{dy}{dx}$$

Proof: The intuitive idea is to use the differentiability of f at $g(a)$ and multiply and divide the difference quotient of $(f \circ g)(x)$ by $g(x) - g(a)$,

$$\frac{f(g(x)) - f(g(a))}{x-a}\cdot\frac{g(x)-g(a)}{g(x)-g(a)} = \frac{f(g(x)) - f(g(a))}{g(x)-g(a)}\cdot\frac{g(x)-g(a)}{x-a}$$

then take the limit as $x \to a$. Of course, we know that wherever f and g are differentiable, they are also continuous, so, by Theorem 29, Lecture 10, their composition is also continuous. Hence, for any $a \in A$ we have $\lim_{x \to a} f(g(x)) = f(g(a))$. Yet, it may be the case that $g(x)$ is constant on a neighborhood of a, so that $f(g(x)) = f(g(a))$ on that neighborhood. In this case, we have, for $y \neq g(a)$,

$$f'\big(g(a)\big) = \lim_{y \to g(a)} \frac{f(y) - f\big(g(a)\big)}{y - g(a)}$$

but we cannot replace y with $g(x)$ without dividing by 0. To prevent this, we define a workaround function $h(y)$,

$$h(y) \stackrel{\text{def}}{=} \begin{cases} \dfrac{f(y) - f\big(g(a)\big)}{y - g(a)}, & \text{if } y \neq g(a) \\ f'\big(g(a)\big), & \text{if } y = g(a) \end{cases}$$

Then we can say that

$$h\big(g(x)\big)[g(x) - g(a)] = (f \circ g)(x) - (f \circ g)(a)$$

is continuous on a neighborhood of a even if $g(x) = g(a)$ nearby, and moreover $\lim_{x \to a} h\big(g(x)\big) = f'\big(g(x)\big)$. Then,

$$\begin{aligned} (f \circ g)'(a) &\stackrel{\text{def}}{=} \lim_{x \to a} \frac{(f \circ g)(x) - (f \circ g)(a)}{x - a} \\ &= \lim_{x \to a} h\big(g(x)\big) \frac{g(x) - g(a)}{x - a} \\ &= f'\big(g(a)\big) g'(a) \end{aligned}$$

■

Theorem 14 (Inverse Function Theorem) *If $f \in C([a, b]) \cap \mathcal{D}((a, b))$ is also injective on $[a, b]$, then*

$$f^{-1} \in \mathcal{D}\Big(f\big((a, b)\big) - f\big((f')^{-1}(0)\big)\Big)$$

(that is f^{-1} is differentiable wherever $f'(c) \neq 0$). Letting $y = f(c)$, in fact

$$(f^{-1})'(y) = \frac{1}{f'\big(f^{-1}(y)\big)}$$

When additionally $f \in C^1([a, b])$, if $\exists c \in (a, b)$ where $f'(c) \neq 0$, then $\exists V_\varepsilon(c) \subseteq (a, b)$ on which $f'(x) \neq 0$, and in that case $f^{-1} \in \mathcal{D}\big(V_\varepsilon(c)\big)$.

Proof: (Exercise 5.2.12, Abbott) Since $f \in C([a,b])$ is injective, it is bijective onto its range $f([a,b]) = [m, M]$, and therefore a homeomorphism and so either strictly increasing or strictly decreasing on $[a,b]$, according to Theorem 11, Lecture 11. Since $f'(c) \neq 0$, moreover, we have

$$\begin{aligned}(f^{-1})'\big(f(c)\big) &= \lim_{x\to c} \frac{f^{-1}(f(x)) - f^{-1}(f(c))}{f(x)-f(c)} \\ &= \lim_{x\to c} \frac{x-c}{f(x)-f(c)} \\ &= \lim_{x\to c} \frac{1}{\frac{f(x)-f(c)}{x-c}} \\ &= \frac{1}{f'(c)}\end{aligned}$$

This simultaneously shows that f^{-1} is differentiable at $f(c)$ and that its derivative is $1/f'(c)$. Finally, if $f \in C^1([a,b])$, then $f' \in C([a,b])$, so that if $f'(c) \neq 0$ for $c \in (a,b)$ then there is some open neighborhood I of c on which $f'(x) \neq 0$, whence the theorem holds for each x in this interval. ∎

2 Function Subspaces of $\mathcal{D}(A)$

Example 15 Let $A \subseteq \mathbb{R}$ and consider the following vector subspaces of the space

$$\mathcal{D}(A) \stackrel{\text{def}}{=} \{ f : A \to \mathbb{R} \mid f \text{ is differentiable on } A \}$$

(1) The space of ***k-times continuously differentiable functions***, those for which all derivatives up to and including order k exist and are continuous,

$$C^k(A) \stackrel{\text{def}}{=} \left\{ f : \mathbb{R} \to \mathbb{R} \;\middle|\; f^{(j)} \equiv \frac{d^j f}{dx^j} \in C(A), \text{ for all } 1 \leq j \leq k \right\}$$

We let $\boxed{C^0(A) \stackrel{\text{def}}{=} C(A)}$. Observe that by Theorem 6 $\boxed{\mathcal{D}^k(A) \subseteq C^{k-1}(A)}$.

(2) Functions which have continuous derivatives of *all* orders are called **smooth functions** or **C^∞-functions**,

$$\begin{aligned} C^\infty(A) &\stackrel{\text{def}}{=} \left\{ f : \mathbb{R} \to \mathbb{R} \;\middle|\; f^{(j)} \equiv \frac{d^j f}{dx^j} \in C(A) \text{ for all } j \in \mathbb{N} \right\} \\ &= \bigcap_{k=0}^{\infty} C^k(A) \end{aligned}$$

(3) Hence, we introduce the **(real) analytic functions**, at least in a neighborhood of a point $x = a$,

$$C^\omega(A) \stackrel{\text{def}}{=} \left\{ f \in C^\infty(\mathbb{R}) \;\middle|\; f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n, \text{ for all } a \in \mathbb{R} \right\}$$

(4) The most familiar functions which equal their own Taylor series are the **polynomials** $\mathbb{R}[x]$:

$$\mathbb{R}[x] \stackrel{\text{def}}{=} \left\{ p(x) \in C^\omega(\mathbb{R}) \;\middle|\; p(x) = \sum_{k=0}^{n} a_k x^k, \; n \in \mathbb{N} \right\}$$

In other words, ***polynomials are those analytic functions whose Taylor series are finite!*** For example, $p(x) = 2x^2 - 3x + 1$ satisfies $p'(x) = 4x - 3$ and $p''(x) = 4$, with higher derivatives $p^{(n)}(x) = 0$, so that, centering the Taylor series at $x = 0$ we have

$$\begin{aligned} \sum_{n=0}^{\infty} \frac{p^{(n)}(0)}{n!}(x-0)^n &= p(0) + \frac{p'(0)}{1!}x + \frac{p''(0)}{2!}x^2 \\ &= 1 + \frac{-3}{1}x + \frac{4}{2}x^2 \\ &= 1 - 3x + 2x^2 \\ &= p(x) \end{aligned}$$

We may also filter the set of polynomials into subclasses, namely **polynomials of degree at most** n,

$$\boxed{\mathbb{R}_n[x] \stackrel{\text{def}}{=} \{p \in \mathbb{R}[x] \mid \deg(p) \leq n\}}$$

It is an easy exercise to prove that for any $p(x) = \sum_{k=0}^{n} a_k x^k$ all coefficients satisfy $a_k = p^{(k)}(0)/k!$, so that $\mathbb{R}_n[x] \subseteq C^\omega(\mathbb{R})$. ∎

Remark 16 *The above examples can be arranged in an* ***vector space inclusion diagram*** *(the hooked arrows denote inclusion):*

$$\mathbb{R}_0 = \mathbb{R} \hookrightarrow \mathbb{R}_1[x] \hookrightarrow \cdots \hookrightarrow \mathbb{R}_n[x] \hookrightarrow \cdots \hookrightarrow \mathbb{R}[x] \hookrightarrow C^\omega(\mathbb{R}) \hookrightarrow$$

$$\hookrightarrow C^\infty(\mathbb{R}) \hookrightarrow \cdots \hookrightarrow C^k(\mathbb{R}) \hookrightarrow \cdots \hookrightarrow C^2(\mathbb{R}) \hookrightarrow C^1(\mathbb{R}) \hookrightarrow \mathcal{D}(\mathbb{R}) \hookrightarrow$$

$$\hookrightarrow C(\mathbb{R}) \hookrightarrow \mathbb{R}^{\mathbb{R}}$$

Exercise 17 *Prove that $\mathbb{R}[x] \subseteq C^\omega(\mathbb{R})$, i.e. that all real polynomials are real analytic. This is most easily achieved by demonstrating that for any $p(x) = \sum_{k=0}^{n} a_k x^k \in \mathbb{R}[x]$, the coefficients satisfy $a_k = \frac{p^{(k)}(0)}{k!}$.* ∎

Exercise 18 *Show that all $C^k(A)$, as well as $C^\infty(A)$ and $C^\omega(A)$, are* ***real vector spaces****, and in fact* ***real associative algebras*** *(they are closed under (pointwise) scalar multiplication, sums and differences, and multiplication; even division whenever the denominator is not 0).* ∎

Exercise 19 *In the real case, the filtration above is strict, in the sense that there are functions $f \in C^k(\mathbb{R})$ which are not in $C^{k+1}(\mathbb{R})$. Show that,*

(1) *$f(x) = |x|$ lies in $C(\mathbb{R})$ by not in $C^1(\mathbb{R})$.*

(2) *$f(x) = x|x|$ lies in $C^1(\mathbb{R})$ but not in $C^2(\mathbb{R})$.*

(3) *$f(x) = x^k|x|$ lies in $C^k(\mathbb{R})$ but not in $C^{k+1}(\mathbb{R})$.*

(4) *$f(x) = \begin{cases} e^{-1/x}, & \text{if } x > 0, \\ 0, & \text{if } x \leq 0, \end{cases}$ lies in $C^\infty(\mathbb{R})$ but not in $C^\omega(\mathbb{R})$. This example stands in contrast to functions such as e^x and $\sin x$, which do equal their Taylor series everywhere (and so lie in $C^\omega(\mathbb{R})$; in fact, these functions are defined in terms of their series).*

(5) *Certainly $e^x = \sum_{n=0}^{\infty} \frac{1}{n!} x^n$ lies in $C^\omega(\mathbb{R})$ but not in $\mathbb{R}[x]$.* ∎

Remark 20 *The situation with complex differentiability is different, because by Cauchy's theorem a complex-differentiable function is automatically analytic, that is* $\mathcal{D}_{\mathbb{C}}(\mathbb{C}) = C^{\omega}(\mathbb{C})$. ■

Example 21 The following are subspaces of $C(A)$. We encourage the reader to explore the functional analysis and partial differential equations literature on these:

(1) **Schwartz functions**

$$\mathcal{S}(A) \overset{\text{def}}{=} \{f \in C^{\infty}(A) \mid \|f\|_{\alpha,\beta} < \infty,\ \alpha, \beta \in \mathbb{N}_0\}$$

where $\|f\|_{\alpha,\beta} \overset{\text{def}}{=} \sup_{x\in A} |x^{\alpha} f^{(\beta)}(x)|$.

(2) **Continuous, compactly supported functions** on A, also called **bump functions** or **test functions** in the context of distributions,

$$C_c(A) \overset{\text{def}}{=} \left\{f \in C(A) \mid \operatorname{supp}(f) \overset{\text{def}}{=} \overline{\{a \in A \mid f(a) \neq 0\}} \in \mathcal{K}_A\right\}$$

(3) **Functions vanishing at infinity**

$$C_0(A) \overset{\text{def}}{=} \{f \in C(A) \mid \forall \varepsilon > 0,\ \exists K \in \mathcal{K}_A,\ \sup_{x\in A-K} |f(x)| < \varepsilon\}$$

(4) **Bounded functions**

$$\mathcal{B}(A) \overset{\text{def}}{=} \{f : A \to \mathbb{R} \mid \exists M > 0,\ \forall a \in A,\ |f(a)| \leq M\}$$

(5) **Riemann integrable functions**

$$\mathcal{R}(A) \overset{\text{def}}{=} \left\{f : A \to \mathbb{R} \mid \int_A f\,dx \in \mathbb{R}\right\} \quad \text{(Riemann integral)}$$

(6) **Lebesgue square-integrable functions**

$$L^2(A) \overset{\text{def}}{=} \left\{f : A \to \mathbb{R} \mid \left(\int_A |f|^2\,d\mu\right)^{1/2} \in \mathbb{R}\right\} \quad \text{(Lebesgue integral)}$$

(7) **Sobolev spaces** $W^{k,p}(A)$ and $H^p(A)$, whose definition we leave to a specialized course.

Good places to look at these spaces are Reed and Simon, *Functional Analysis*; Kadison and Ringrose, *Fundamentals of the Theory of Operator Algebras, Volume I: Elementary Theory*; Duistermaat and Kolk, *Distributions: Theory and Applications*; Leoni, *A First Course in Sobolev Spaces*. The background required is a graduate real analysis course, e.g. Rudin's *Real and Complex Analysis*. ■

3 Mean Value Theorem and its Corollaries

Theorem 22 (Darboux/Fermat Theorem) *Let $A \subseteq \mathbb{R}$ and let $f \in \mathcal{D}(A)$. If $a \in A^\circ$ is a local extremum (min or max), then $f'(a) = 0$.*

Proof: If $f(a)$ is a local max on some neighborhood $V_\delta(a)$ of a, this means $f(x) \leq f(a)$ for all $x \in V_\delta$, and hence $f(x) - f(a) \leq 0$ for such x. On the other hand, $x - a < 0$ if $a - \delta < x < a$ and $x - a > 0$ if $a < x < a + \delta$. From this it follows that

$$L^- \stackrel{\text{def}}{=} \lim_{x \to a} \frac{f(x) - f(a)}{x - a} \geq 0$$
$$L^+ \stackrel{\text{def}}{=} \lim_{x \to a} \frac{f(x) - f(a)}{x - a} \leq 0$$

Since $'(a) = L^+ = L^-$ exists, we conclude that $f'(a) = L^+ = L^- = 0$. The local min case is proved similarly. ■

Theorem 23 (Rolle's Theorem) *Let $f \in C([a, b]) \cap \mathcal{D}\big((a, b)\big)$. If $f(a) = f(b)$, then $\exists c \in (a, b)$ such that $f'(c) = 0$.*

Proof: If f is a constant function, then any $c \in (a, b)$ will do. Otherwise, f must have an interior extremum $c \in (a, b)$ by EVT, at which $f'(c) = 0$ by the previous theorem. ■

Exercise 24 *If $f, g \in C([a, b]) \cap \mathcal{D}\big((a, b)\big)$, $f(a) = g(a)$ and $f(b) = g(b)$, then show that $\exists c \in (a, b)$ such that $f'(c) = g'(c)$.* ■

Theorem 25 (General Mean Value Theorem) *If $f, g \in C([a, b]) \cap \mathcal{D}\big((a, b)\big)$, then $\exists c \in (a, b)$ such that*

$$\big[f(b) - f(a)\big] g'(c) = \big[g(b) - g(a)\big] f'(c)$$

Proof: Let $h(x) = \big[f(b) - f(a)\big] g(x) - \big[g(b) - g(a)\big] f(x)$. Then,

$$\begin{aligned} h(a) &= f(b)g(a) + \cancel{f(a)g(a)} - g(b)f(a) + \cancel{g(a)f(a)} \\ &= f(b)g(a) - g(b)f(a) \end{aligned}$$

and

$$\begin{aligned} h(b) &= \cancel{f(b)g(b)} - f(a)g(b) - \cancel{g(b)f(b)} + g(a)f(b) \\ &= -f(a)g(b) + g(a)f(b) \end{aligned}$$

so that $h(a) = h(b)$. By Rolle's theorem, therefore, $\exists c \in (a, b)$ such that

$$h'(c) = f'(c)\big[g(b) - g(a)\big] - g'(c)\big[f(b) - f(a)\big] = 0$$

whence $\big[f(b) - f(a)\big] g'(c) = \big[g(b) - g(a)\big] f'(c)$. ■

Corollary 26 (Mean Value Theorem) *If $f \in C([a,b]) \cap \mathcal{D}\big((a,b)\big)$, then $\exists c \in (a,b)$*

$$f(b) - f(a) = f'(c)(b-a) \tag{2}$$

Proof: Let $g(x) = x$ in the General Mean Value Theorem. ■

Corollary 27 *For any $f \in C([a,b]) \cap \mathcal{D}\big((a,b)\big)$ we have*

$$\big(f'(x) \equiv 0 \text{ on } A \iff f \equiv c \text{ is constant on the interval}\big)$$

Proof: If $f \equiv c$ on A, then $f' \equiv 0$ on A by Theorem 8 above. If $f' \equiv 0$ on A, then by MVT, $\forall x, y \in (a,b),\ x < y \implies \exists \xi \in (x,y)$ such that $f(x) - f(y) = f'(\xi)(x-y) = 0(x-y) = 0$, implying that $f(x) = f(y)$. This is true for *all* $x, y \in (a,b)$, so f is constant on (a,b). Since f is continuous on $[a,b]$, it must have the same value at a and b as well. ■

Corollary 28 *Let $f \in \mathcal{D}\big((a,b)\big)$.*

(1) *If $f'(x) \geq 0$ on (a,b), then f is increasing on (a,b).*

(2) *If $f'(x) \leq 0$ on (a,b), then f is decreasing on (a,b).*

Moreover, if the inequalities are strict, then f is strictly increasing or strictly decreasing, respectively.

Proof: Let $x, y \in (a,b)$, with $x < y$. If $f'(x) \geq 0$ on (a,b), then by MVT applied to f on $[x,y]$, $\exists \xi \in (x,y)$ with

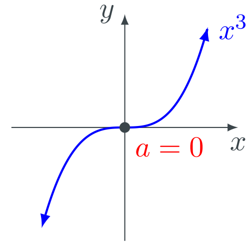
$$f(y) - f(x) = f'(\xi)(y-x) \geq 0$$

Since x and y were arbitrarily chosen, we see that f is increasing on (a,b). Similarly, if $f'(x) \leq 0$ on (a,b), we get

$$f(y) - f(x) = f'(\xi)(y-x) \leq 0$$

so f is decreasing on (a,b). If we replace the inequalities by strict inequalities, moreover, we see that f will become *strictly* increasing or decreasing, respectively. ■

Remark 29 *We have seen (Darboux/Fermat Theorem 22) that if c is a local interior extremum of a function $f \in \mathcal{D}(A)$, then $f'(c) = 0$. The converse, however, does not generally hold. Even if $f'(c) = 0$, it is not generally true that c is a local extremum. For example, if $f(x) = x^3$, then $f'(x) = 3x^2$, so $f'(0) = 0$, yet $x = 0$ is not an extremum.*



What, in addition to $f'(x) = 0$, do we need to determine whether x is a maximimum or a minimum of f? This is precisely what the first and second derivative tests tell us. ■

Definition 30 Let $f \in \mathcal{D}(A)$. Any zero of f', that is any $a \in A$ for which $f'(a) = 0$, is called a **critical point** or **critical number** of f. The set of all critical points is the preimage of 0 under f',

$$(f')^{-1}(0) \overset{\text{def}}{=} \textbf{all critical points of } f$$

The question is whether a given critical point a is

(1) a **local maximum** ($\exists \varepsilon > 0,\ \forall x \in V_\varepsilon(a),\ f(x) \leq f(a)$)

(2) a **local minimum** ($\exists \varepsilon > 0,\ \forall x \in V_\varepsilon(a),\ f(x) \geq f(a)$)

(3) a **saddle point** ($f(x) \leq f(a)$ and $f(x) \geq f(a)$ on $V_\varepsilon(a)$, the switch happening about a).

In order to determine this, we must **test** a given critical point. ■

Corollary 31 (First Derivative Test) *Let $f \in \mathcal{D}(A)$ and let $a \in (f')^{-1}(0)$ be a critical point of f. Suppose further that $a \in A^\circ$ and $\exists \varepsilon$ for which $V_\varepsilon(a) \subseteq A$.*

(1) *If $f'(x) > 0$ on $(a - \epsilon,\ a) \subsetneq V_\varepsilon(a)$ and $f'(x) < 0$ on $(a,\ a + \varepsilon) \subsetneq V_\varepsilon(a)$, then a is a* ***local maximum*** *of f.*

(2) *If $f'(x) < 0$ on $(a - \varepsilon,\ a)$ and $f'(x) > 0$ on $(a,\ a + \varepsilon)$, then a is a* ***local minimum*** *of f.*

(3) *If $f'(x) > 0$ on $V_\varepsilon(a) - \{a\}$ or $f'(x) < 0$ on $V_\varepsilon(a) - \{a\}$, then a is a* ***saddle point*** *of f.*

Proof: This follows from Corollary 28. If the first condition holds, then f is increasing on $(a-\varepsilon,\ a)$ and decreasing on $(a,\ a+\varepsilon)$. Hence for any $x \in V_\varepsilon(a) - \{a\}$ we must have $f(x) < f(a)$, so that a is a local max. Similarly a is a local min in the second case. The third case shows that either f is increasing on all of $V_\varepsilon(a) - \{a\}$ if $f'(x) > 0$ or decreasing on all of $V_\varepsilon(a) - \{a\}$ if $f'(x) < 0$, and hence reaches values both above and below $f(a)$. ■

Corollary 32 (Second Derivative Test) *Let $f \in \mathcal{D}^2(A)$ and consider an interior critical point $a \in A^\circ \cap (f')^{-1}(0)$.*

(1) *If $f''(a) > 0$, then a is a* ***local minimum*** *of f.*

(2) *If $f''(a) < 0$, then a is a* ***local maximum*** *of f.*

(3) *If $f''(a) = 0$, then* ***no conclusion*** *can be drawn, a may be a local minimum or a local maximum or a saddle point.*

Proof: Apply Corollary 28 to f'. If $f'' > 0$ on $V_\varepsilon(a)$, then f' is increasing on $V_\varepsilon(a)$. Since $f'(a) = 0$, we must have $f' < 0$ on $(a-\varepsilon,\ a)$ and $f' > 0$ on $(a,\ a+\varepsilon)$, so f has a local minimum by the First Derivative Test, Corollary 31. We sometimes say f is "concave up," meaning the (horizontal) tangent line at a is below the graph of f,

$$\underbrace{L(x) = f(a) + \overbrace{f'(a)(x-a)}^{=0} = \overbrace{f(a) \leq f(x)}^{a \text{ is a local min}}}_{L(x) \leq f(x)}$$

If $f'' < 0$ on $V_\varepsilon(a)$, then f' is decreasing on $V_\varepsilon(a)$. Since $f'(a) = 0$, we must have $f' > 0$ on $(a-\varepsilon,\ a)$ and $f' < 0$ on $(a,\ a+\varepsilon)$, so f has a local maximum by the First Derivative Test, Corollary 31. We sometimes say f is "concave down," meaning the tangent line at a is above the graph of f,

$$\underbrace{L(x) = f(a) + \overbrace{f'(a)(x-a)}^{=0} = \overbrace{f(a) \geq f(x)}^{a \text{ is a local max}}}_{L(x) \geq f(x)}$$

Otherwise, if $f''(a) = 0$, we cannot say which is the case. ■