Differentiation in one variable (Rudin ch 5 = McMullen §5)

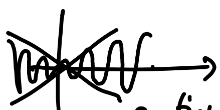
Def: $f : [a,b] \to \mathbb{R}$ is differentiable at x if $\lim_{t\to x} \frac{f(t)-f(x)}{t-x} =: f'(x)$ exists.

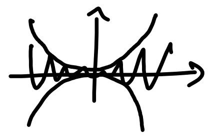
(ie. $\forall \varepsilon\ \exists \delta$ st. $0<|t-x|<\delta \Rightarrow \left|\frac{f(t)-f(x)}{t-x} - f'(x)\right| < \varepsilon$.)

- **Prop:** f differentiable at $x \Rightarrow f$ continuous at x. (The converse is false, eg. $|x|$ at 0).

  Pf: $f(t)-f(x) = \underbrace{\frac{f(t)-f(x)}{t-x}}_{\to f'(x) \text{ as } t\to x} . \underbrace{(t-x)}_{\to 0 \text{ as } t\to x}$ + multiplication is continuous $\Rightarrow f(t)-f(x) \to f'(x).0 = 0$.

- Usual rules of calculation hold: derivatives of $f+g$, fg, ... ; $(f\circ g)'(x) = f'(g(x)).g'(x)$ chain rule. (see Rudin p 104-105).

- **Ex:**
  - $\begin{cases} f(x) = x \sin\frac{1}{x} & (x\neq 0) \\ f(0) = 0 \end{cases}$  For $x\neq 0$, $f'(x) = \sin\left(\frac{1}{x}\right) - \frac{1}{x}\cos\left(\frac{1}{x}\right)$

    continuous but not differentiable at 0 $\left(\nexists \lim_{x\to 0} \frac{f(x)}{x}\right)$.
  - $\begin{cases} g(x) = x^2 \sin\frac{1}{x} \\ g(0) = 0 \end{cases} \Rightarrow$  differentiable $(g'(0)=0)$ but g' not continuous at 0.
  - $f(x) = \sum_{n=1}^{\infty} \frac{1}{n^2} \sin(n!\,x)$ continuous (series converges uniformly, since $\sum \frac{1}{n^2}$ conv.), nowhere differentiable! (See also Rudin 7.18 for a related example).

* **Mean value theorem:** $f : [a,b] \to \mathbb{R}$ differentiable $\Rightarrow \exists c\in(a,b)$ st. $f(b)-f(a) = f'(c).(b-a)$.

Follows logically from easier results:

(1) if $f : [a,b] \to \mathbb{R}$ has a local max (or min) at $x\in(a,b)$ (ie. max of $f_{|(x-\delta, x+\delta)}$) and f is differentiable at x, then $f'(x) = 0$.

(because $\frac{f(t)-f(x)}{t-x}$ is ≥ 0 for $t\in(x-\delta,x)$, ≤ 0 for $t\in(x,x+\delta)$ $\Rightarrow$ take lim. as $t\to x$ from left and from right.)

(2) if $f : [a,b] \to \mathbb{R}$ is differentiable and $f(a) = f(b)$ then $\exists c\in(a,b)$ st. $f'(c) = 0$.

clear if f is constant; otherwise look at max or min of f over $[a,b]$ & apply (1)

(3) mean val. thm = apply (2) to $g(x) = f(x) - \frac{f(b)-f(a)}{b-a}x$.

Corollary: mean value inequality: $m \leq f'(x) \leq M$ $\forall x\in(a,b) \Rightarrow m(b-a) \leq f(b)-f(a) \leq M(b-a)$.

* Generalization: **Taylor's theorem**:

$f : [a,b] \to \mathbb{R}$ n times differentiable. The deg.$(n-1)$ Taylor polynomial of f at a is:

$P(x) = \sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{k!}(x-a)^k$. Then $\exists c\in(a,b)$ st. $f(b) = P(b) + \frac{f^{(n)}(c)}{n!}(b-a)^n$.

Pf: • subtracting $P(x)$ from both sides, we can reduce to the case $f(a) = f'(a) = \ldots = f^{(n-1)}(a) = 0$. (and $P=0$).

• let $g(x) = f(x) - f(b)\frac{(x-a)^n}{(b-a)^n} \Rightarrow g(b) = g(a) = 0$ + still have $g'(a) = \ldots = g^{(n-1)}(a) = 0$.

- now: the mean value thm for g ; $g(a)=g(b)=0 \Rightarrow \exists x_1 \in (a,b)$ st. $g'(x_1)=0$.
  —— " —— g' : $g'(a)=g'(x_1)=0 \Rightarrow \exists x_2 \in (a,x_1)$ st. $g''(x_2)=0$

and so on until $\exists c = x_n \in (a, x_{n-1})$ st. $g^{(n)}(c)=0$. Ie: $f^{(n)}(c) - \dfrac{n!\, f(b)}{(b-a)^n} = 0$. □

Rmk:
- can compare $f(x)$ to $P(x)$ by applying thm. to $[a,x]$ instead!
- as with mean value inequality: a bound $|f^{(n)}| \leq M$ gives a bound $|f(x)-P(x)| \leq \dfrac{M(x-a)^n}{n!}$ over $[a,b]$.

Rmk: there exist nonzero functions whose Taylor polynomials are all zero!
eg. $f(x) = \exp(-\frac{1}{x^2})$, $f(0)=0$: $f \in C^\infty$ (all derivatives exist), $f^{(k)}(0)=0\ \forall k$
so the Taylor polynomials are all zero! The Taylor series of f at 0 converges but $\neq f$! (in other examples, it can also have $R=0$ ie. never converges for $x \neq a$).
Most C^∞ functions aren't analytic, ie. can't be written as power series.

Let $C^k([a,b],\mathbb{R}) = \{k\text{-times differentiable functions, } f^{(k)} \text{ continuous}\}$, with $\|f\|_{C^k} = \sum_{j=0}^{k} \|f^{(j)}\|_\infty$.

Thm: $f_n \in C^1$, $f_n \to f$ pointwise, $f_n' \to g$ uniformly $\Rightarrow f \in C^1$ and $f' = g$ (& $f_n \to f$ uniformly)

Pf:
* Fix $x \neq y \in [a,b]$, mean value theorem $\Rightarrow (*)\ \dfrac{f_n(y)-f_n(x)}{y-x} = f_n'(c_n)$ for some $c_n \in [x,y]$ or $[y,x]$.
  The left hand side $\to \dfrac{f(y)-f(x)}{y-x}$ as $n\to\infty$.
  For the right hand side: (c_n) has a subsequence (c_{n_k}) converging to some $c \in [x,y]$.
  Since f_n' is continuous, the uniform limit g is continuous; we claim $f_{n_k}'(c_{n_k}) \to g(c)$.
  Indeed: fix $\varepsilon>0$, let δ st. $|t-c|<\delta \Rightarrow |g(t)-g(c)|<\frac{\varepsilon}{2}$, and let N st.
  $n \geq N \Rightarrow \sup|f_n'-g| < \frac{\varepsilon}{2}$ and $n_k \geq N \Rightarrow |c_{n_k}-c|<\delta$. Then for $n_k \geq N$,
  $$|f_{n_k}'(c_{n_k}) - g(c)| \leq |f_{n_k}'(c_{n_k}) - g(c_{n_k})| + |g(c_{n_k})-g(c)| < \varepsilon.$$
  Hence: returning to $(*)$ and taking limit as $n\to\infty$: $\exists c \in [x,y]$ st. $\dfrac{f(y)-f(x)}{y-x} = g(c)$.
  We now take the limit as $y \to x$: the rhs. $\to g(x)$ using continuity of g and the fact that $|c-x| \leq |y-x|$ (check this!). Hence f is differentiable at x and $f'(x)=g(x)$. (+ since g is continuous, $f \in C^1$).
* Finally: mean value ineq. $\Rightarrow |f_n(x)-f(x)| \leq \underbrace{|f_n(a)-f(a)|}_{\to 0} + \underbrace{|x-a|}_{\leq (b-a)} \underbrace{\sup|f_n'-f'|}_{\to 0 \text{ since } f_n' \to g \text{ uniformly}}$
  gives a uniform bound so $\sup|f_n - f| \to 0$. □

Corollary: $C^k([a,b],\mathbb{R})$ is a complete metric space

Pf: Using completeness of C^0 (uniform top), (f_n) Cauchy in $C^1 \Rightarrow f_n, f_n'$ Cauchy in $C^0 \Rightarrow$
$\exists$ uniform limits $f, g \in C^0 \overset{\text{thm}}{\Rightarrow} f \in C^1$ and $f'=g$. Now $\left\{\begin{matrix} f_n \to f \\ f_n' \to f' \end{matrix}\right\}$ uniformly $\Rightarrow f_n \to f$ in C^1.
This proves the case $k=1$. Repeat same argument for successive derivatives for $k>1$. □

Corollary: $f(x) = \sum a_n x^n$ power series with radius of convergence $= R$
$\Rightarrow$ $f(x)$ is C^∞ over $(-R, R)$, and $f'(x) = \sum n a_n x^{n-1}$.

Pf: $f = \sum a_n x^n$ and $g = \sum n a_n x^{n-1}$ have the same radius of convergence, so the partial sums for both converge uniformly over compact subsets of $(-R,R)$, hence $f \in C^1$ and $f' = g$. Repeat for successive derivatives ($g \in C^1$ so $f \in C^2$, ...) ∎

Integration (Riemann $\int$, see Math 114 for Lebesgue integral and much more)

The definite integral of continuous functions is a linear operator $I_a^b : C^0([a,b]) \to \mathbb{R}$,
$f \longmapsto I_a^b(f) = \int_a^b f\,dx$

$\int_a^b (f+g)\,dx = \int_a^b f + \int_a^b g$
$\int_a^b cf\,dx = c\int_a^b f\,dx$

for each $a<b \in \mathbb{R}$, satisfying axioms:

1) If $f \geq 0$ then $\int_a^b f\,dx \geq 0$ $\quad$ ($\Rightarrow$ if $f \geq g$ then $\int_a^b f\,dx \geq \int_a^b g\,dx$).
2) If $a<c<b$ then $\int_a^b f\,dx = \int_a^c f\,dx + \int_c^b f\,dx$.
3) $\int_a^b 1\,dx = b-a$.

In fact, such a linear map is unique; the difference between different theories of integration is in how much more general functions we allow ourselves to integrate.

The Riemann integral starts from step functions: $s(x) : [a,b] \to \mathbb{R}$ such that $\exists\ a = x_0 < x_1 < \dots < x_n = b$ s.t. $s(x)$ is constant over each (x_{i-1}, x_i), $s(x) = s_i$. (the values at x_i don't matter). Then 2)+3) suggest we must have

$$I(s) = \int_a^b s(x)\,dx = \sum_{i=1}^n s_i\,(x_i - x_{i-1}).$$

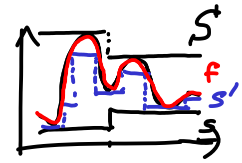
This definition of the integral for step functions satisfies the required axioms.

Next: if $s \leq f \leq S$ for s, S step functions, then $\int_a^b s\,dx \leq \int_a^b f\,dx \leq \int_a^b S\,dx$. $\quad (\star)$

In particular: $f : [a,b] \to \mathbb{R}$ bounded $\Rightarrow$ fixing $a = x_0 < x_1 < \dots < x_n = b$, we can take $s_i = \inf f([x_{i-1}, x_i])$ and $S_i = \sup f([x_{i-1}, x_i])$, giving the lower and upper Riemann sums of f for the given partition of $[a,b]$.

Refining (i.e. subdividing further) gives better bounds on f

$$\int s\,dx < \int s'\,dx < \int f\,dx < \int S'\,dx$$



Lower and upper Riemann integral:

$$I_-(f) = \sup\left\{ \int_a^b s\,dx \;\middle|\; \begin{array}{l} s \leq f \text{ on } [a,b] \\ s \text{ step function} \end{array} \right\}$$

$$I_+(f) = \inf\left\{ \int_a^b S\,dx \;\middle|\; \begin{array}{l} S \geq f \text{ on } [a,b] \\ S \text{ step function} \end{array} \right\}$$

$\forall f$ bounded $[a,b] \to \mathbb{R}$, $I_-(f) \leq I_+(f)$.

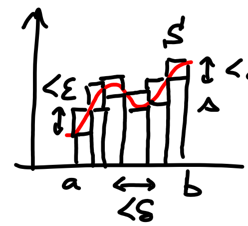
Def: f is Riemann integrable, $f \in \mathcal{R}([a,b])$, if $I_+(f) = I_-(f)$; we set $\int_a^b f\,dx = I_\pm(f)$.

Thm: Continuous functions are Riemann integrable.

Pf: The key ingredient is uniform continuity: $\forall \varepsilon > 0\ \exists \delta$ st. $x, y \in [a,b]$, $|x-y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon$.

(Recall: this is proved by applying the Lebesgue number lemma to the open cover $[a,b] \subset \bigcup_{c \in \mathbb{R}} f^{-1}((c, c+\varepsilon))$: $\exists \delta > 0$ st. $|x-y| = \operatorname{diam}(\{x,y\}) < \delta \Rightarrow \exists c$ st. $\{x,y\} \subset f^{-1}((c, c+\varepsilon))$)

Thus: given $\varepsilon > 0$, take δ as in uniform continuity, and split $a = x_0 < x_1 < \dots < x_n = b$ st. $x_{i+1} - x_i < \delta\ \forall i$. Then $s_i = \min f([x_i, x_{i+1}])$, $S_i = \max f([x_i, x_{i+1}])$ (attained) satisfy $S_i - s_i < \varepsilon\ \forall i$, and $s_i \le f \le S_i$ on $[x_i, x_{i+1}]$.



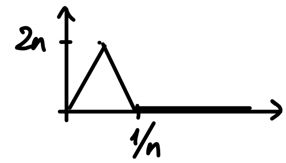
Let s, S = step functions taking values s_i, S_i on $[x_i, x_{i+1})$: $s \le f \le S$ on $[a,b]$, so $I(s) \le I_-(f)$, $I(S) \ge I_+(f)$;
moreover, $S_i - s_i < \varepsilon\ \forall i$ so $I(S) - I(s) < \varepsilon(b-a)$.

Hence: $I_+(f) - I_-(f) < \varepsilon(b-a)\ \forall \varepsilon > 0 \Rightarrow I_+(f) = I_-(f)$, $f \in \mathcal{R}([a,b])$. □

Rmk:
- piecewise continuous functions are also integrable ; and so do some stranger functions (see Rudin & see HW). However for example
$$f(x) = \begin{cases} 1 & \text{if } x \in \mathbb{Q} \\ 0 & x \notin \mathbb{Q} \end{cases} \quad \text{is not Riemann integrable} \quad \begin{pmatrix} I_-(f) = 0 \\ I_+(f) = b-a \end{pmatrix}.$$
The Lebesgue integral allows more general decompositions into "measurable" subsets (rather than just sub-intervals) & allows more general functions to be integrated (including unbounded functions, which are never Riemann integrable)
(eg for Riemann integration, $\int_0^x \frac{1}{\sqrt{t}}\, dt = \frac{1}{2}\sqrt{x}$ only makes sense as an "improper integral" ie. $\lim_{\varepsilon \to 0} \int_\varepsilon^x$, whereas Lebesgue can handle this & worse).
- In fact, Lebesgue gave a characterization of exactly which functions are Riemann-integrable: $f \in \mathcal{R}([a,b])$ iff f is bounded on $[a,b]$ and the set of points where f is discontinuous has Lebesgue measure 0, which means: $\forall \varepsilon > 0$ $\exists (I_i)$ at most countable collection of open intervals st $E \subset \bigcup I_i$ and $\sum \operatorname{length}(I_i) < \varepsilon$.
- It is easy to check (do it!) that $\mathcal{R}([a,b])$ is a vector space, $I : \mathcal{R}([a,b]) \to \mathbb{R}$ is linear and satisfies the above axioms.
- Fundamental thm of calculus: if f is continuous on $[a,b]$ then $F(x) = \int_a^x f(t)\,dt$ is differentiable and $F' = f$.

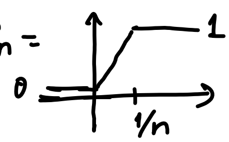
  Pf: $\frac{1}{h}(F(x+h) - F(h)) = \frac{1}{h}\int_x^{x+h} f(t)\,dt \xrightarrow[h \to 0]{} f(x)$ using continuity of f at x to estimate the integral for $h \to 0$. □

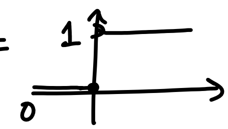
* **Thm:** $I : C^0([a,b]) \to \mathbb{R}$ is continuous with respect to the uniform topology: if $f_n \to f$ uniformly then $\int_a^b f_n\,dx \to \int_a^b f\,dx$.
In fact, $\left|\int f\,dx - \int g\,dx\right| \leq \int |f-g|\,dx \leq (b-a)\sup|f-g|$.

On the other hand, pointwise convergence isn't enough: $f_n =$ 

$f_n \to 0$ pointwise but $\int_0^1 f_n\,dx = 1 \not\to \int_0^1 0\,dx = 0$.

* Besides $\|f\|_\infty = \sup|f|$, we have other norms on the vector space $C^0([a,b],\mathbb{R})$: namely $\|f\|_1 = \int_a^b |f(x)|\,dx$, and also $\|f\|_p = \left(\int_a^b |f(x)|^p\,dx\right)^{1/p}$ $\forall p \geq 1$.
(Triangle inequality follows from Hölder's inequality, cf. homework).
These are called the L^p norms; since $\|f\|_p \leq (b-a)^{1/p}\|f\|_\infty$, balls for $\|\cdot\|_p$ contain balls for $\|\cdot\|_\infty$ and the topologies defined by these metrics are <u>coarser</u> than the uniform topology; $(C^0([a,b]), \|\cdot\|_p)$ isn't complete, its completion is the Lebesgue space $L^p([a,b])$ – see Math 114!

<u>Ex:</u> $f_n =$ 

is Cauchy in L^1 norm, in fact converges in L^1 to its pointwise limit $f =$ 

$\in \mathcal{R}$

$\left(\int_0^1 |f_n - f|\,dx = \frac{1}{2n} \to 0\right)$, but $f \notin C^0$.

L^1 is quite natural, but so is L^2, which comes from an inner product
$$\langle f, g\rangle_{L^2} = \int_a^b fg\,dx \qquad (\Rightarrow \|f\|_{L^2} = \sqrt{\langle f,f\rangle}).$$
(Cauchy-Schwarz: $\langle f,g\rangle \leq \|f\|_{L^2}\|g\|_{L^2}$ is a special case of Hölder's ineq.)