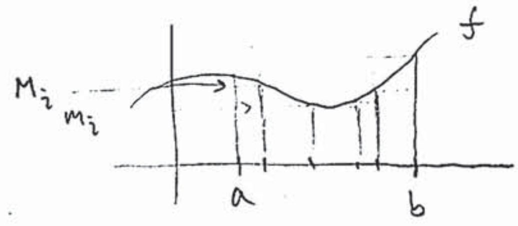
Integrability



$\{x_0\ x_1 \ldots x_n\} = P$

Riemann - Stieltjes ← $\alpha(x)$ monotonically increasing fcn

lower sum: $\sum m_i \overset{\Delta\alpha_i}{\cancel{\Delta x_i}} = L(P, f, \alpha)$

upper sum: $\sum M_i \Delta\alpha_i = U(P, f, \alpha)$

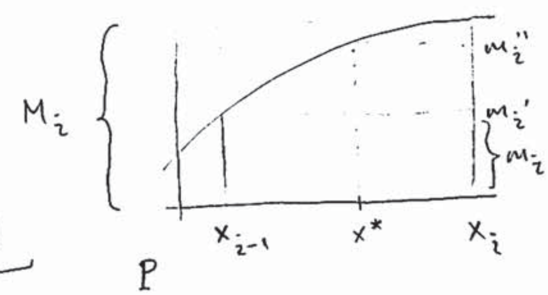
Let $\underline{\int} f\,d\alpha = \sup\limits_P L(P)$ and $\overline{\int} f\,d\alpha = \inf\limits_P U(P)$. If they are equal, we call f <u>R-S integrable</u>. We write $f \in R(\alpha)$ ← set of RS-integrable fcn's

Q: Which functions are in $R(\alpha)$?

Say P is a <u>partition</u>. Call P^* a refinement of P if $P^* \supseteq P$

<u>Claim:</u> $L(P) \le L(P^*)$ and $U(P) \ge U(P^*)$

<u>Proof:</u> Enough to show $P^* = P \cup \{x^*\}$ (then use induction). We compare: $L(P^*) - L(P) =$



$$\underbrace{m_i'[\alpha(x^*) - \alpha(x_{i-1})] - m_i[\quad]}_{\ge 0} + \underbrace{m_i''[\alpha(x_i) - \alpha(x^*)] - m_i[\quad]}_{\ge 0}$$

The expression ≥ 0 as desired. Similarly for $U(P) \ge U(P^*)$

<u>Thm:</u> $\underline{\int}_a^b f\,d\alpha \le \overline{\int}_a^b f\,d\alpha$

<u>Proof:</u> Given P_1, P_2 partitions. Let $P^* \supseteq P_1 \cup P_2$. Then,

$L(P_1) \le L(P^*) \le U(P^*) \le U(P_2)$ holds for any pair. Then,

$L(P_1) \le \inf\limits_{P_2} U(P_2) \leftarrow \overline{\int} f\,d\alpha$ and $\underline{\int} f\,d\alpha \rightarrow \sup\limits_{P_1} L(P_1) \le \inf\limits_{P_2} U(P_2)$, as desired.

Cauchy criterion for integrability:

Thm: A function $f \in R(\alpha)$ iff $\forall \varepsilon > 0$ $\exists P$ s.t. $U(P) - L(P) < \varepsilon$.

Proof: ($\Leftarrow$) Since $L(P) \le \underline{\int} f d\alpha \le \overline{\int} f d\alpha \le U(P)$ for any P,

$0 \le \overline{\int} f d\alpha - \underline{\int} f d\alpha < \varepsilon$, $\forall \varepsilon > 0$ so $\overline{\int} f d\alpha - \underline{\int} f d\alpha = 0$.

($\Rightarrow$) If $f \in R(\alpha)$, given $\varepsilon > 0$ $\exists P_1$ s.t. $U(P_1) - \int f d\alpha < \frac{\varepsilon}{2}$ by def'n of $\overline{\int}$ and $\exists P_2$ s.t. $\int f d\alpha - L(P_2) < \frac{\varepsilon}{2}$. We see

$$U(P_1) - L(P_2) < \varepsilon \quad \text{but} \quad L(P_2) \le \underbrace{L(P_1 \cup P_2) \le U(P_1 \cup P_2)}_{\text{differ by } < \varepsilon} \le U(P_1)$$

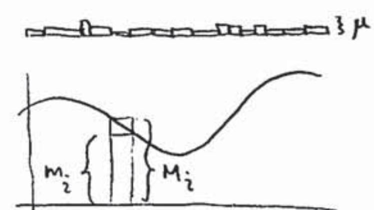
so $P_1 \cup P_2$ is the desired partition.

Ex: Dirichlet function $f(x) = \begin{cases} 1, & \text{if } x \text{ rational} \\ 0, & \text{otherwise} \end{cases}$

We see $L(P) = 0$ and $U(P) = 1$ are never close so by Cauchy criterion, it is not in R (Riemann integrable).

Recall: Cauchy criterion $f(x) \in R(\alpha) \iff \forall \varepsilon > 0$ $\exists P$ s.t. $U(P) - L(P) < \varepsilon$.

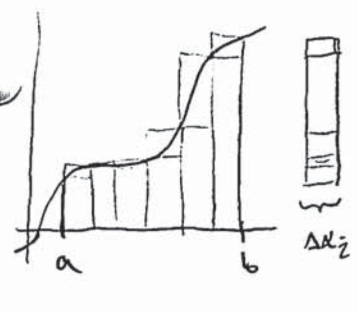
Thm: f continuous on $[a,b] \Rightarrow f \in R(\alpha)$ on $[a,b]$.



Notice $U(P) - L(P) = \sum_{i=1}^{n} (M_i - m_i) \Delta\alpha_i$

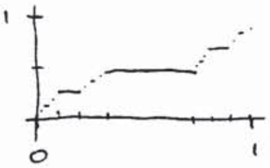
Proof: Given $\varepsilon > 0$, choose η s.t. $\eta [\Sigma \Delta\alpha_i] = \eta[\alpha(b) - \alpha(a)] < \varepsilon$. Since f is continuous on a compact set, f is uniformly continuous. So $\exists \delta > 0$ s.t. $|x - y| < \delta \Rightarrow |f(x) - f(y)| < \eta$. Choose P s.t. $\Delta x_i < \delta$. Then, $U(P) - L(P) = \Sigma (M_i - m_i) \Delta\alpha_i \le \Sigma \eta \Delta\alpha_i$ $= \eta \Sigma \Delta\alpha_i = \eta[\alpha(b) - \alpha(a)] < \varepsilon$. By the Cauchy criterion, $f \in R(\alpha)$.

Thm: f monotonic on $[a,b]$, α continuous on $[a,b] \Rightarrow f \in R(\alpha)$.



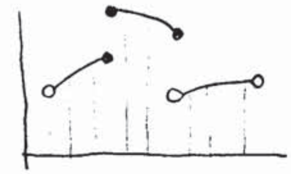
Proof: Let $n = \#$ intervals. Let $\Delta\alpha_i = \frac{\alpha(b) - \alpha(a)}{n}$. Choose n large enough so $\Delta\alpha_i < \frac{\varepsilon}{f(b) - f(a)}$. Then $U(P) - L(P)$ $= \Sigma (M_i - m_i) \Delta\alpha_i < \sum_i (M_i - m_i) \frac{\varepsilon}{f(b) - f(a)} = \varepsilon$. By the Cauchy criterion, $f \in R(\alpha)$.

Ex: Devil's Staircase (monotonic, not continuous)



rises from 0 to 1 on Cantor set

Thm: f is bounded on $[a,b]$ and has finite discontinuities. α is continuous where f is not. Then $f \in R(\alpha)$.



discontinuous at w_i

Idea: bound $\sum (M_i - m_i)\Delta\alpha_i$ in 2 parts:
where f continuous, make $M_i - m_i$ small
where f discontinuous, make $\Delta\alpha_i$ small

Proof: Given $\varepsilon > 0$, choose (u_i, v_i) containing w_i s.t. $\sum |\alpha(v_i) - \alpha(u_i)| < \varepsilon$. (We can do this b/c finitely many w_i) For continuous part of f, use uniform continuity to make $M_i - m_i$ small: $\exists\, \delta > 0$ s.t. $|x-y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon$. Choose P s.t. includes u_i, v_i's and $\Delta x_i < \delta$ on continuous part $\left([a,b] - \bigcup_i (u_i, v_i)\right)$. Then

$$U(P) - L(P) = \underbrace{\sum (M_i - m_i)\Delta\alpha_i}_{\text{intervals where } f \text{ cont.}} + \underbrace{\sum (\overbrace{M_i - m_i}^{f \text{ bounded by } K})\Delta\alpha_i}_{\text{other intervals}}$$

$$< \varepsilon[\alpha(b) - \alpha(a)] + 2K\varepsilon$$

As $\varepsilon \to 0$, $U(P) - L(P) \to 0$ so by the Cauchy criterion, $f \in R(\alpha)$.

Integration Properties

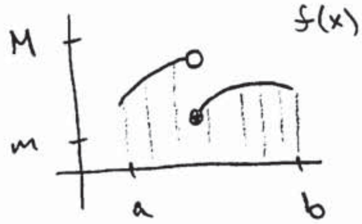
Compositions

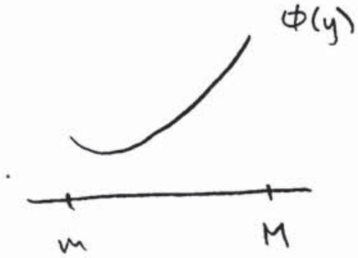
Ex: Let $D(x) = \begin{cases} 1; & x \text{ rational} \\ 0; & \text{else} \end{cases} \notin R$, $E(x) = \begin{cases} 1, & x \neq 0 \\ 0, & \text{else} \end{cases} \in R$, $F(x) = \begin{cases} 1/q, & x = p/q \text{ lowest term} \\ 0, & \text{else} \end{cases} \in R$

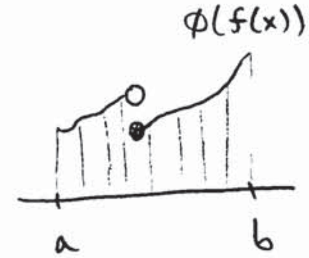
We see that $D(x) = E(F(x))$ so the composition of integrable functions is not necessarily integrable.

Thm: If $f \in R(\alpha)$ on $[a,b]$, $m \le f \le M$ (bounded) and ϕ is continuous on $[m, M]$, and let $h(x) = \phi(f(x))$ on $[a,b]$. Then $h \in R(\alpha)$.

Cor: (i) $f \in R(\alpha) \Rightarrow f^2 \in R(\alpha)$, b/c $\phi(y) = y^2$ is cont.
(ii) $f, g \in R(\alpha) \Rightarrow fg \in R(\alpha)$, use $fg = \frac{(f+g)^2 - f^2 - g^2}{2}$ (also need thm about sums)







Idea: bound $\sum (M_i^* - m_i^*) \Delta\alpha_i$

Consider $U(P,h) - L(P,h) = \underset{\substack{\text{short}\\\text{boxes}}}{\sum} (M_i^* - m_i^*)\Delta\alpha_i + \underset{\substack{\text{tall}\\\text{boxes}}}{\sum} (M_i^* - m_i^*)\Delta\alpha_i$

$$< \underset{\substack{\text{by cont.}\\\text{of } \phi}}{\varepsilon} \underbrace{\sum \Delta\alpha_i}_{< \alpha(b)-\alpha(a)} + 2 \sup_{[m,M]} \phi \underbrace{\sum \Delta\alpha_i}_{\substack{\text{hope to make}\\\text{this small}}}$$

Idea: Given $\varepsilon > 0$, by uniform continuity of ϕ, $\exists \delta_1$ s.t. $|s-t| < \delta_1$ $\Rightarrow |\phi(s) - \phi(t)| < \varepsilon$. Choose $\delta = \min(\delta_1, \varepsilon)$ so $\delta < \varepsilon$ needed later. Choose P s.t. $U(P,f) - L(P,f) < \delta^2$ (b/c $f \in R(\alpha)$). We have tall (use f): $M_i - m_i \geq \delta$ and short ("): $M_i - m_i < \delta$. Then,

$$\delta \sum \Delta\alpha_i < \sum_{\text{tall}} (M_i - m_i)\Delta\alpha_i < \delta^2 < \delta < \varepsilon.$$

Some "easy" theorems:

<u>Thm:</u> $f_1, f_2, f \in R(\alpha)$, $c \in \mathbb{R}$

(a) $f_1 + f_2 \in R(\alpha)$ and $\int_a^b (f_1 + f_2)\,d\alpha = \int_a^b f_1\,d\alpha + \int_a^b f_2\,d\alpha$
$cf \in R(\alpha)$ and $\int_a^b cf\,d\alpha = c\int_a^b f\,d\alpha$

(b) $f_1(x) \leq f_2(x) \Rightarrow \int f_1\,d\alpha \leq \int f_2\,d\alpha$

(c) $\int_a^c f\,d\alpha = \int_a^b f\,d\alpha + \int_b^c f\,d\alpha$, $a < b < c$

(d) $|f(x)| \leq M \Rightarrow \left|\int_a^b f\,d\alpha\right| < M(\alpha(b) - \alpha(a))$

(e) If $f \in R(\alpha_1)$, $f \in R(\alpha_2) \Rightarrow f \in R(\alpha_1 + \alpha_2)$
and $\int_a^b f\,d(\alpha_1 + \alpha_2) = \int_a^b f\,d\alpha_1 + \int_a^b f\,d\alpha_2$
If $c > 0$, $f \in R(c\alpha)$
and $\int_a^b f\,d(c\alpha) = c\int_a^b f\,d\alpha$.

<u>Thm:</u> $f, g \in R(\alpha) \Rightarrow fg \in R(\alpha)$. (by comp)

$f \in R(\alpha) \Rightarrow |f| \in R(\alpha)$ (by comp w/ $\phi(t) = |t|$)

$\left| \int_a^b f d\alpha \right| \leq \int_a^b |f| d\alpha$ (idea: choose $c = \pm 1$ s.t. $c \int f d\alpha \geq 0$. Then, $\left| \int f d\alpha \right| = c \int f d\alpha = \int c f d\alpha \leq \int |f| d\alpha$)

Define for $a < b$ $\int_b^a f d\alpha := -\int_a^b f d\alpha$.

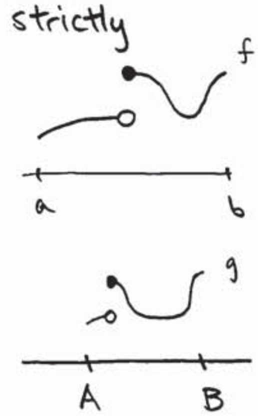
also $\int_a^\infty f(x) dx = \lim_{b \to \infty} \int_a^b f(x) dx$ (can define if f is not bounded (with care))

<u>Ex:</u> $f(x) = \frac{1}{\sqrt{x}}$ $\int_0^1 f(x) dx = \lim_{c \to 0^+} \int_c^1 f(x) dx$

Change of variable

<u>Thm:</u> Assume $\phi : [A, B] \to [a, b]$, $\phi(A) = a$, $\phi(B) = b$, ϕ is strictly increasing and continuous. Assume α on $[a, b]$, $f \in R(\alpha)$, β on $[A, B]$, $\beta = \alpha \circ \phi$. Let $g = f \circ \phi$. Then, $g \in R(\beta)$ on $[A, B]$ and

$$\int_A^B g d\beta = \int_a^b f d\alpha$$



See: U, L for g are same as for f (w/ transformed partitions)

Special case: $\int_a^b f(x) dx = \int_A^B f(\phi(y)) \phi'(y) dy$ (assuming ϕ differentiable, MVT gives $\Delta y_i = \phi'(y_0) \Delta x_i$ for some $y_0 \in \Delta y$)

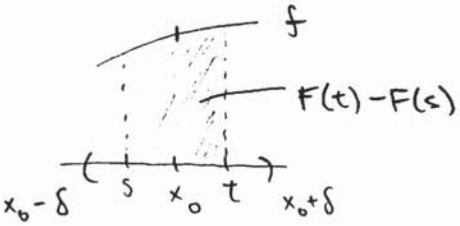
Fundamental Theorem of Calculus

<u>Thm:</u> If f bounded and $f \in R$, suppose for $x \in [a, b]$, $F(x) = \int_a^x f(t) dt$. Then F is continuous on $[a, b]$. Also, if f is continuous at $x_0 \in [a, b]$ then F is differentiable at x_0 and $F'(x_0) = f(x_0)$.

<u>Proof:</u> Say $|f(t)| \le M$ on (a,b). For $x<y$ in $[a,b]$

$$|F(y)-F(x)| = \left| \int_x^y f(t)\,dt \right| \le M(y-x)$$

For $\varepsilon>0$, choose $\delta = \varepsilon/M$. Then $|x-y|<\delta \Rightarrow |F(x)-F(y)|<\varepsilon$, so $F(x)$ is continuous. Now assume f is continuous at x_0, then given $\varepsilon>0$, $\exists\, \delta>0$ s.t. $|t-x_0|<\delta \Rightarrow |f(t)-f(x_0)|<\varepsilon$. If $s<t$ in $(x_0-\delta, x_0+\delta)$ then $\dfrac{F(t)-F(s)}{t-s} = \dfrac{1}{t-s}\displaystyle\int_s^t f(u)\,du$.



Idea: $\frac{\text{area}}{\text{width}} \approx \text{height} \to f(x_0)$

So, $$\left| \frac{F(t)-F(s)}{t-s} - f(x_0) \right| = \left| \frac{1}{t-s}\int_s^t [f(u)-f(x_0)]\,du \right|$$
$$\le \frac{1}{t-s}\,\varepsilon(t-s) = \varepsilon$$

So in the limit, $F'(x_0) = f(x_0)$.

<u>Thm:</u> (Fundamental Theorem of Calculus) If $f \in R$ on $[a,b]$ and $\exists F$ differentiable on $[a,b]$ s.t. $F' = f$ (called "anti-derivative"), then

$$\int_a^b f(x)\,dx = F(b)-F(a)$$

<u>Proof:</u> Given $\varepsilon>0$, choose P s.t. $U(P)-L(P)<\varepsilon$. Notice that $F(x_i)-F(x_{i-1}) =$ area of a rectangle. By summing over intervals, we can get $F(b)-F(a)$. We have $F(x_i)-F(x_{i-1}) \underset{MVT}{=} F'(t_i)\Delta x_i$, $t_i \in \Delta x_i$ so

$$F(b)-F(a) = \sum_i \underbrace{F'(t_i)}_{f(t_i)}\Delta x_i \quad \text{between } U(P) \text{ and } L(P)$$

But $\int_a^b f(x)\,dx$ is also between $U(P)$ and $L(P)$ so

$$\left| \int_a^b f(x)\,dx - (F(b)-F(a)) \right| < \varepsilon$$

for all $\varepsilon>0$. Therefore, they are equal.

Thm: (Integration by Parts) If on $[a,b]$, have $F' = f$ and $G' = g$ where $f, g \in R$, then

$$\int_a^b F(x)\,g(x)\,dx = F(b)G(b) - F(a)G(a) - \int_a^b f(x)\,G(x)\,dx$$

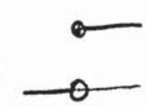
Proof idea: Product rule and FTC on $F(x)G(x)$.

Think of integration as a "functional" (a function on functions). Let $\mathcal{C}([a,b])$ denote the set of all continuous real-valued functions on $[a,b]$. Then $G: \mathcal{C}([a,b]) \to \mathbb{R}$ is a functional.

Ex:
- Integration: $G(f) = \int_a^b f\,dx$ (linear functional: $G(cf+g) = cG(f) + G(g)$)
- Evaluation at p: $G(f) = f(p)$

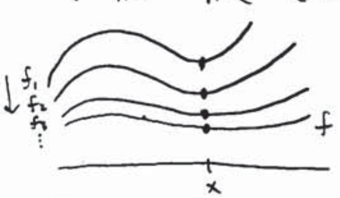
Riesz Representation Thm:
Let $G: \mathcal{C}[a,b] \to \mathbb{R}$ be a functional that is
- (i) positive ($\text{if } f \geq 0, \text{ then } G(f) \geq 0$)
- (ii) bounded ($|G(f)| \leq M \sup_{[a,b]} |f(x)|$)
- (iii) linear ($G(cf+g) = cG(f) + G(g)$)

then $\exists \alpha$ such that $G(f) = \int_a^b f\,d\alpha$.

Ex: Evaluation at p. Can be represented as $\int_a^b f\,d\alpha$ where $\alpha =$ 

Sequences of Functions

$f: \mathbb{R} \to \mathbb{R}$ (also $f: \mathbb{R}^n \to \mathbb{R}^m$ w/ $\|\vec{x}\| = \text{length}$) $\qquad f_1(x), f_2(x), \ldots \to\ ?$

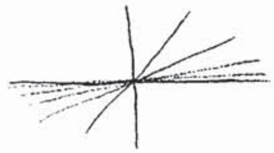


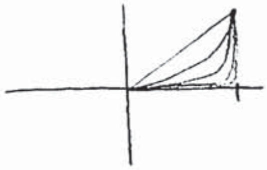
pointwise convergence: $f(x) = \lim_{n\to\infty} f_n(x)$ if exists

Ex: $f_n(x) = \cos(n!\,\pi x)$ 

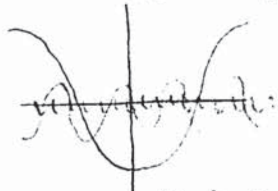
limit exists at $x = 0$ but not at other points

Ex: $f_n(x) = x - \frac{1}{n}$  $\longrightarrow f(x) = x$

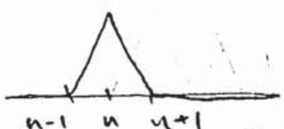
$f_n(x) = \frac{x}{n}$  $\longrightarrow f(x) = 0$

$f_n(x) = x^n$ on $[0,1]$  $\longrightarrow f(x) = \begin{cases} 1, & x = 1 \\ 0, & 0 \le x < 1 \end{cases}$

Note: limit not continuous

$f_n(x) = \frac{\sin(n^2 x)}{n}$  $\longrightarrow f(x) = 0$

Note: $f' \neq \lim_{n\to\infty} f_n'$

$f_n(x) =$ 

$\longrightarrow f(x) = 0$

Note: $\int f\,dx \neq \lim_{n\to\infty} \int f_n\,dx$

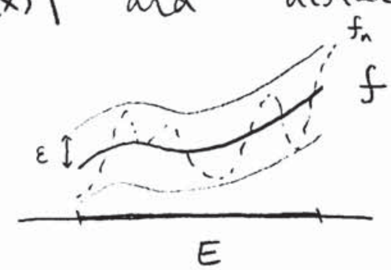
Q: When can lim and ($\int$ or $\frac{d}{dx}$) be switched? We want a notion of convergence that "plays nice" with limits. This is <u>uniform convergence</u> or "ribbon convergence".

<u>Def</u>: Say f_n <u>converges uniformly</u> to f on E if $\forall \varepsilon > 0, \exists N$ s.t.
$n \ge N \Rightarrow \underbrace{\forall x \in E, |f_n(x) - f(x)| < \varepsilon}_{\text{demand } f_n \text{ in } \varepsilon\text{-ribbon around } f \text{ "distance"} < \varepsilon}$.

Note: "uniform": Given x, $\exists N, \forall \varepsilon > 0, \ldots$
"pointwise": Given x, $\forall \varepsilon > 0, \exists N, \ldots$

<u>Def</u>: Let $\|f\| = \sup_{x\in E} |f(x)|$ and "distance" $d(f,g) = \|f - g\|$.

We write $f_n \xrightarrow{u} f$.



Q: How to tell if f_n converges uniformly? (Even when you don't know f)

<u>Thm</u>: (Cauchy criterion) $f_n \xrightarrow{u} f$ on $E \iff \forall \varepsilon > 0, \exists N$ s.t. $\forall n, m \geq N$, $d(f_n, f_m) < \varepsilon$ (or $\|f_n - f_m\| < \varepsilon$).

<u>Proof</u>: ($\Rightarrow$) Given $\varepsilon > 0$, by uniform convergence, $\exists N$ s.t. $n \geq N \Rightarrow$ $|f_n(x) - f(x)| < \varepsilon/2$. So for this N, $n, m \geq N$ implies

$$|f_n(x) - f_m(x)| \leq |f_n(x) - f(x)| + |f_m(x) - f(x)| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

as desired.

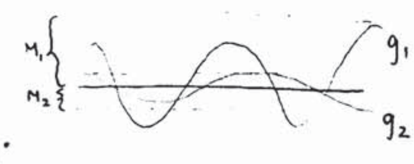
($\Leftarrow$) Let $f(x) = \lim_{n \to \infty} f_n(x)$. This pointwise limit exists because $\{f_n(x)\}$ is Cauchy for any fixed x by our hypothesis. So it converges because $\mathbb{R}$ is complete. Why does $f_n \xrightarrow{u} f$? Given $\varepsilon > 0$, choose N s.t. $n, m \geq N \Rightarrow |f_n(x) - f_m(x)| < \frac{\varepsilon}{2}$. Take $\lim_{m \to \infty}$ on both sides to get $|f_n(x) - f(x)| \leq \frac{\varepsilon}{2} < \varepsilon$ as desired.

We can talk about series of functions converging uniformly.

$$\sum_{n=1}^{\infty} g_n(x) := \lim_{k \to \infty} \sum_{n=1}^{k} g_n(x) \longleftarrow \text{partial sum } S_k(x)$$

To say $\sum g_n$ converges uniformly means the sequence $S_k(x)$ converges uniformly.

<u>Thm</u>: (Weierstrass M-test) $\{f_n\}$ on E, $|f_n(x)| \leq M_n$.
If $\sum_{n=1}^{\infty} M_n$ converges, then $\sum_{n=1}^{\infty} f_n(x)$ converge uniformly.



Pf idea: Let's show $S_k(x)$ converge uniformly.

$$|S_n(x) - S_m(x)| = \left|\sum_{i=m+1}^{n} f_i(x)\right| \underset{\Delta \text{ ineq.}}{\leq} \sum_{i=m+1}^{n} M_i \longleftarrow \text{Given } \varepsilon, \exists N \text{ s.t. this} < \varepsilon \text{ for } m, n \geq N.$$

So S_k converges uniformly.

Uniform convergence

We can use the Cauchy criterion (if we don't know f):

$$\forall \varepsilon > 0, \exists N \text{ s.t. } n, m \geq N \Rightarrow \|f_n - f_m\| < \varepsilon$$

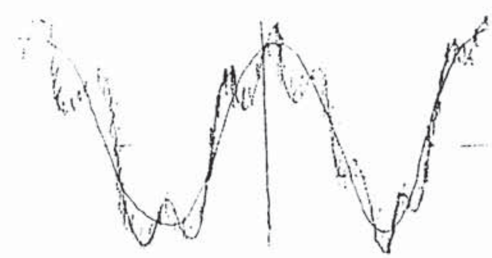
Thm: f_n continuous on E, $f_n \xrightarrow{u} f \Rightarrow f$ continuous on E

Proof: [$\varepsilon/3$ - argument] For $x \in E$, we want to show f is continuous at x. Given $\varepsilon > 0$, we choose f_n such that $|f_n(x) - f(x)| < \varepsilon/3$ by uniform convergence $f_n \to f$. Because f_n is continuous, $\exists\, \delta$ s.t. $|x-y| < \delta \Rightarrow |f_n(x) - f_n(y)| < \varepsilon/3$. For this δ, if $|x-y|<\delta$, then

$$|f(x) - f(y)| \le |f(x) - f_n(x)| + |f_n(x) - f_n(y)| + |f_n(y) - f(y)|$$
$$< \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon$$

as desired.

Ex: Weierstrass $\quad f(x) = \sum_{n=0}^{\infty} \left(\tfrac{3}{4}\right)^n \cos(9^n \pi x)$



We claim $f(x)$ is uniformly convergent by the M-test (using $M_n = (\tfrac{3}{4})^n$). Also, $f(x)$ is continuous because the partial sums are. It can be shown that f is not differentiable at any point. So f is "continuous every, differentiable nowhere".

Ex: Is there a curve which fills the entire square? $f: [0,1] \overset{\text{onto}}{\twoheadrightarrow} [0,1] \times [0,1]$

Yes, space-filling curves (Peano curves).

Idea:

$f_0 \qquad f_1 \qquad f_2 \qquad \ldots$

Construct a sequence f_k s.t.

① f_k is Cauchy sequence [then f is continuous].

② $\forall \delta > 0, \exists k$ s.t. $f_k([0,1])$ is a δ-net in $\square$, i.e. $\forall z \in \square, \exists x \in [0,1]$ s.t. $\|f_k(x) - z\| < \delta$. (Every point z is δ-close to the image of f_k) [then f is onto]

Proof: Let $z \in \square$, we find x s.t. $f(x) = z$. For $\delta = 1/n$, choose k_n by ②. Then $\exists x_n$ s.t. $\|f_{k_n}(x_n) - z\| < \delta$. Since x_n is a sequence in a compact space $[0,1]$, there is a convergence subsequence (rename to x_n for simplicity). Then $x_n \to$ some x. We claim $f(x) = z$. Note that $f(x) = \lim_{n\to\infty} f(x_n)$ by continuity of f. So $\exists K_1$ s.t. $n \ge K_1 \Rightarrow \|f(x) - f(x_n)\| < \varepsilon/3$. Since $f_{k_n} \xrightarrow{u} f$, $\exists K_2$ s.t. $n \ge K_2 \Rightarrow \|f(x_n) - f_{k_n}(x_n)\| < \varepsilon/3$. Also, $\exists K_3$ s.t. $n \ge K_3 \Rightarrow \frac{1}{n} < \varepsilon/3$ so $\|f_{k_n}(x_n) - z\| < \varepsilon/3$. Let K be the max of K_1, K_2, and K_3. For $n > K$,

$$\|f(x) - z\| \le \|f(x) - f(x_n)\| + \|f(x_n) - f_{k_n}(x_n)\| + \|f_{k_n}(x_n) - z\|$$
$$\le \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon$$

as desired.

Space of Functions

Let $\mathcal{C}_b(X)$ = all continuous bounded complex-valued functions on a metric space X

It has a norm ("size") $= \|f\| = \sup_{x \in X} |f(x)|$ (exists b/c bounded)

Key fact: Gives metric on $\mathcal{C}_b(x) = d(f,g) = \|f-g\|$ and convergence of f_n in this metric is uniform convergence of $f_n \to f$. Also, $\mathcal{C}_b(x)$ is complete with respect to this metric (b/c $\mathbb{C}$ is complete and continuous functions converge to a continuous function)

Ex: Define $\|f\|_2 = \left[\int_{-\infty}^{\infty} |f(x)|^2 dx \right]^{1/2}$ "the L^2-norm"

This arises from the inner product: $\langle f, g \rangle = \int_{-\infty}^{\infty} f(x) \overline{g(x)}\, dx$

We restrict our attention to <u>square-integrable functions</u> where $\int_{-\infty}^{\infty} |f(x)|^2 dx < \infty$. We refer to this metric space as the <u>L^2-space</u> or $L^2(\mathbb{C})$. Since it is a metric space, is complete, and has an inner product, it is a <u>Hilbert space</u>.

The Hölder inequality ($p=q=2$) gives $|\langle f, g \rangle| \le \|f\|_2 \|g\|_2$.

Uniform Convergence & Calculus

Q: Given $\{f_n\}$ converging pointwise to f, under what conditions do areas converge?

$f = 0$

f_n areas are 1
f area is 0

Thm: Say α monotonically increasing, $f_n \in R(\alpha)$ on $[a,b]$. Suppose $f_n \xrightarrow{u} f$ on $[a,b]$. Then $f \in R(\alpha)$ and $\int_a^b f\,d\alpha = \lim_{n\to\infty} \int_a^b f_n\,d\alpha$.

Proof: Let $\varepsilon_n = \| f_n - f \|$ so $f_n - \varepsilon_n \le f \le f_n + \varepsilon_n$ and

$$\int (f_n - \varepsilon_n)\,d\alpha \le \underline{\int} f\,d\alpha \le \overline{\int} f\,d\alpha \le \int (f_n + \varepsilon_n)\,d\alpha$$

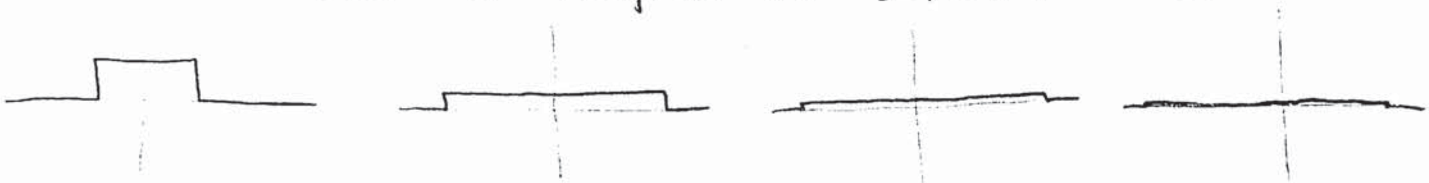
Note that

$$0 \le \overline{\int} f\,d\alpha - \underline{\int} f\,d\alpha \le \int_a^b 2\varepsilon_n\,d\alpha = 2\varepsilon_n\left[\alpha(b) - \alpha(a)\right]$$

As $\varepsilon_n \to 0$, we see that the upper and lower sums are equal.

Alternative proof: Consider the case where f_n is continuous. Let $G(f) = \int_a^b f\,d\alpha$ so G is a linear functional. This is a bounded linear functional: $|G(f)| \le M\|f\|$ so G is continuous:
$|G(f) - G(g)| = |G(f-g)| \le M\|f-g\|$. Then if $f_n \xrightarrow{u} f$, think of this as convergence of points in $\mathcal{C}_b([a,b])$. Then, $G(f_n) \to G(f)$, i.e., $\lim_{n\to\infty} \int f_n\,d\alpha = \int f\,d\alpha$.

Q: Is the thm true for integrals on $[a,\infty)$? No.



Q: If $f_n \xrightarrow{u} f$, it's not enough to ensure $f_n' \to f'$. What additional conditions are needed?

<u>Thm:</u> Suppose $\{f_n\}$ continuously (not needed) differentiable on $[a,b]$ and $\{f_n(x_0)\}$ converge for some $x_0 \in [a,b]$. If f_n' converge uniformly on $[a,b]$ then f_n converge uniformly to some f and

$$f'(x) = \lim_{n\to\infty} f_n'(x)$$

<u>Proof:</u> If f_n' continuous, then by FTC, consider $F_n(x) = \int_{x_0}^{x} f_n'(t)\,dt$. F_n is continuous, differentiable, and $F_n'(x) = f_n'(x)$. By FTC, $F_n(x) = f_n(x) - f_n(x_0) \Rightarrow f_n(x) = F_n(x) + f_n(x_0)$. We want to show f_n converge uniformly. Given $\varepsilon > 0$, because f_n' converge uniformly, $\exists K_1$ s.t. $n,m \geq K_1 \Rightarrow |f_n'(x) - f_m'(x)| < \frac{\varepsilon}{2(b-a)}$ so

$$|F_n(x) - F_m(x)| = \left| \int_{x_0}^{x} [f_n'(t) - f_m'(t)]\,dt \right| \leq \int_{x_0}^{x} |f_n'(t) - f_m'(t)|\,dt < \frac{\varepsilon}{2(b-a)} \int_{x_0}^{x} dt \leq \frac{\varepsilon}{2}$$

By hypothesis, $\exists K_2$ s.t. $n,m \geq K_2 \Rightarrow |f_n(x_0) - f_m(x_0)| < \frac{\varepsilon}{2}$. For $K = \max(K_1, K_2)$, if $n,m \geq K$, we have

$$|f_n(x) - f_m(x)| \leq |F_n(x) - F_m(x)| + |f_n(x_0) - f_m(x_0)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

so f_n converge uniformly to some f. We claim $f'(x)$ exists and is equal to $\lim_{n\to\infty} f_n'(x)$, which we call $L(x)$. Note that $L(x)$ is continuous by uniform convergence and continuity of f_n'. We define $F(x) = \lim_{n\to\infty} F_n(x) = \int_{x_0}^{x} \lim_{n\to\infty} f_n'(t)\,dt = \int_{x_0}^{x} L(t)\,dt$. Since $L(x)$ is continuous, $F'(x)$ exists and $F'(x) = L(x)$. We then see that

$$f_n(x) = \lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} F_n(x) + \lim_{n\to\infty} f_n(x_0)$$

so $f'(x) = F'(x) = L(x)$ as desired.

Equicontinuity

Recall in $\mathbb{R}^n$, a set is compact iff it is closed and bounded. In a general metric space, if a set is compact then it is closed and bounded.

The set $\mathcal{C}_b(\mathbb{R}) = \{\text{continuous bounded functions } \mathbb{R} \to \mathbb{R}\}$ is a metric space with

$$d(f,g) = \sup_{x \in \mathbb{R}} |f(x) - g(x)| = \|f - g\| \leftarrow \text{sup norm}$$

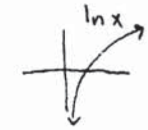
Convergence in $\mathcal{C}_b(\mathbb{R})$ is uniform convergence of functions.

Q: When does a set being closed and bounded imply that it is compact?

Recall a set S is <u>closed</u> when it contains all its limit points and is <u>bounded</u> if $\forall f \in S$, $\exists M$ s.t. $\|f\| \le M$.

<u>Def:</u> A family of functions $\mathcal{F}$ on a set E in a metric space X is said to be <u>equicontinuous</u> on E iff $\forall \varepsilon > 0$, $\exists \delta > 0$ s.t. $\forall x, y \in E$ and $\forall f \in \mathcal{F}$, $d(x,y) < \delta \Rightarrow |f(x) - f(y)| < \varepsilon$.

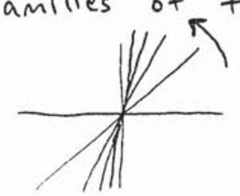
<u>Ex:</u> $f(x) = \ln x$ on $E = (0, \infty)$ in $\mathbb{R}$ is not uniformly continuous

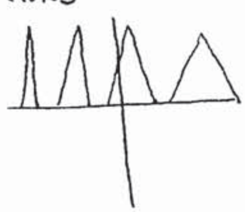


But on (a, ∞), $a > 0$, this f is uniformly continuous since we can use the smallest δ-ball.

A family of functions can't be equicontinuous if the functions inside are not uniformly continuous.

<u>Ex:</u> Families of functions that are uniformly continuous but not equicontinuous.





If $\mathcal{F}$ is finite and each $f \in \mathcal{F}$ is uniformly continuous then $\mathcal{F}$ is equicontinuous since we can use the smallest δ.

<u>Thm:</u> Let K be a compact metric space and $f_n \in \mathcal{C}_b(K)$. If $\{f_n\}$ converges uniformly on K, then $\{f_n\}$ is equicontinuous.

<u>Proof:</u> Given $\varepsilon > 0$. Each f_n is continuous on a compact set so each is uniformly continuous. So $\exists \delta_n$ s.t. $|x - y| < \delta_n \Rightarrow |f_n(x) - f_n(y)| < \varepsilon/3$. Since $\{f_n\}$ converges uniformly, $\exists N$ s.t. $\forall n \ge N$, $|f_n(x) - f_N(x)| < \varepsilon/3$, $\forall x \in K$. So if $n \ge N$ and $|x - y| < \delta_N$, then

$$|f_n(x) - f_n(y)| \le |f_n(x) - f_N(x)| + |f_N(x) - f_N(y)| + |f_N(y) - f_n(y)| < \varepsilon$$

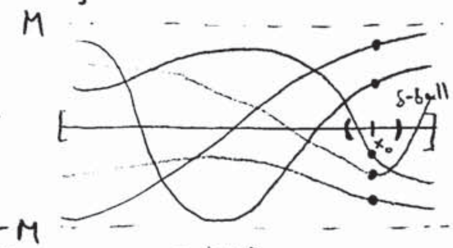
We use $\delta = \min\{\delta_1, \delta_2, \dots, \delta_N\}$. This works for all f_n, since $\forall n \in \mathbb{N}$, $|x-y| < \delta \Rightarrow |f_n(x) - f_n(y)| < \varepsilon/3$.

<u>Thm:</u> Let K be compact, $f_n \in \mathcal{C}(K)$, $\{f_n\}$ equicontinuous. Then $\{f_n\}$ is pointwise bounded implies $\{f_n\}$ is uniformly bounded.

<u>Proof:</u> By equicontinuity of $\{f_n\}$, $\exists\, \delta$ s.t. $|x-y| < \delta \Rightarrow |f_n(x) - f_n(y)| < \varepsilon$.
Choose an open cover of K by balls of radius δ. Since K is compact, there is a finite subcover by balls $B_\delta(x_i)$, $i = 1, \dots, m$.
Let M_i be a pointwise bound for x_i. Then $M_i + 1$ is a bound for $\{f_n\}$ on $B_\delta(x_i)$. Let $M = \max_i \{M_i + 1\}$. This is a uniform bound on $\{f_n\}$.

The Arzela-Ascoli Theorem

<u>Thm:</u> (Arzela-Ascoli) Suppose K is compact, $f_n \in \mathcal{C}(K)$, $\{f_n\}$ equicontinuous and (uniformly) bounded. Then $\{f_n\}$ contains a uniformly convergent subsequence.



<u>Idea:</u> Can get subsequence of $\{f_n\}$ to converge at x_0 & equicontinuity controls how much $f_n(x)$ changes in a δ-ball. We find N s.t.
$$n, m \geq N \Rightarrow |f_n(x) - f_m(x)| < \overbrace{|f_n(x) - f_n(x_0)|}^{\text{equicont}} + \overbrace{|f_n(x_0) - f_m(x_0)|}^{\text{conv. subseq.}} + \overbrace{|f_m(x_0) - f_m(x)|}^{\text{equicont.}} < \varepsilon.$$
But each δ-ball might involve a different subsequence.

<u>Lemma:</u> (countable selection) If $\{f_n\}$ is a pointwise bounded sequence of functions on a countable set E, then it has a convergent subsequence $\{f_{n_k}\}$ converging at all of E.

<u>Proof:</u> Let $\{x_i\}$ be the points of E. Since $\{f_n(x_1)\}$ is bounded in $\mathbb{R}$, $\exists$ subsequence $S_1 = \{f_{1,k}\}$ s.t. $\{f_{1,k}(x_1)\}$ converges as $k \to \infty$. By boundedness of x_2, $\exists$ subsequence of S_1, call it $S_2 = \{f_{2,k}\}$ s.t. it converges on x_2 as $k \to \infty$.
Continue to define $S_3, S_4, \dots$. We claim the subsequence $\{f_{k,k}\}$ converges on each x_i b/c past the i^{th} term, this is a subsequence of S_i.

$$\begin{array}{llll} S_1 & f_{1,1} & f_{1,2} & f_{1,3} \dots \text{conv on } x_1 \\ S_2 & f_{2,1} & f_{2,2} & f_{2,3} \dots \text{conv on } x_1, x_2 \\ S_3 & f_{3,1} & f_{3,2} & f_{3,3} \dots \text{conv on } x_1, x_2, x_3 \end{array}$$

<u>Proof:</u> Choose a countable dense subset E of K. (K is a compact metric space $\Rightarrow$ has countable basis of radius $1/n$ balls). By our lemma, $\exists$ subsequence $\{g_k\}$ that converges on E. We claim g_k converges uniformly. Given $\varepsilon > 0$, choose δ from equicontinuity s.t. $|x-y| < \delta \Rightarrow |g_k(x) - g_k(y)| < \varepsilon/3$. Notice $\{B_\delta(e)\}_{e \in E}$ covers K because every $k \in K$ is δ-close to a point of E. By compactness of K,

$\exists$ a finite subcover $\{B_\delta(e_i)\}_{i=1}^n$. Since $\{g_k\}$ converges on e_i, $\exists N_i$ s.t. $|g_n(e_i) - g_m(e_i)| < \varepsilon/3$, for all $n, m \geq N_i$. Let $N = \max N_i$. Given x, choose e_i s.t. $|x - e_i| < \delta$. Then, using Cauchy criterion, we see that

$$|g_n(x) - g_m(x)| \leq |g_n(x) - g_n(e_i)| + |g_n(e_i) - g_m(e_i)| + |g_m(e_i) - g_m(x)| < \varepsilon$$

Cor: Suppose $S \subseteq \mathcal{C}_b(K)$ for compact K. Then, set S compact $\Leftrightarrow$ S is closed, bounded, and equicontinuous.

Proof: ($\Leftarrow$) By A-A, every sequence has a convergent subsequence. Then closed shows limit is in S.

($\Rightarrow$) To show S is equicontinuous, if not, $\exists \varepsilon > 0$ s.t. $\forall \delta, \exists f_\delta, x_\delta, y_\delta$ s.t. $|x_\delta - y_\delta| < \delta$ but $|f_\delta(x_\delta) - f_\delta(y_\delta)| > \varepsilon$. Use $\delta = 1/n$ and get a contradiction with compactness.

Integral Equations: In physical problems, we may need to solve given a, k

$$f(x) = a + \int_{y=0}^{x} k(x,y) f(y)\, dy$$

For example,

$$\frac{df}{dx} = f(x) \Leftrightarrow f(x) = a + \int_{y=0}^{x} f(y)\, dy$$

Does it have a solution? Strategy: on class of "approximate" solutions, use Arzela-Ascoli to find a uniformly convergent subsequence that converges to the solution.

Thm: If $\sup\limits_{x \in [0,b]} \int_{y=0}^{y=x} |k(x,y)|\, dy < 1$, then $f(x) = a + \int_{y=0}^{x} k(x,y) f(y)\, dy$ has a unique solution on $[0,b]$.

Proof: We'll use a contraction mapping on $\mathcal{C}_b([0,b])$. Define $T: \mathcal{C}_b([0,b]) \to \mathcal{C}_b([0,b])$ by $T(f)(x) = a + \int_0^x k(x,y) f(y)\, dy$ so a fixed point of T corresponds to a solution of our equation. We claim that T is a contraction.

$$\|T(f) - T(g)\| = \sup_{x \in [0,b]} |T(f)(x) - T(g)(x)| = \sup_{x \in [0,b]} \left| \int_0^x k(x,y)[f(y) - g(y)]\, dy \right|$$

$$\leq \sup_{x \in [0,b]} \int_0^x |k(x,y)| \underbrace{|f(y) - g(y)|}_{\leq \|f-g\|}\, dy \leq \lambda \|f - g\|$$

so T is a contraction because $\lambda < 1$. Recall the proof of the Contraction Mapping Theorem is constructive so we can iterate to find the solution.

Ex: $\frac{df}{dx} = f(x) \iff f(x) = \int_{y=0}^{x} f(y)\,dy + a$ w/ initial condition $f(0)=1$ so $a=1$.

Let $T(g)(x) = 1 + \int_0^x g(y)\,dy$. Start at $g=0$ function.

$T(0) = 1$, $T(1) = 1+x$, $T(1+x) = 1+x+\frac{x^2}{2}$, ...

This converges to the power series for e^x.

(Q.) When can a continuous function on some compact interval $[a,b]$ be uniformly approximated by polynomials?

(A.) Yes, we will show for $\mathcal{C}_b([0,1])$ using Bernstein polynomials.

Let $x \in [0,1]$ be the probability of HEADS for a coin. Given $f(x)$, think of f as payout that depends on the probability.

(Q.) What's the expected value of f after flipping n coins?

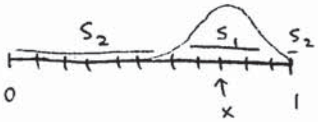
$$\underset{\text{polynomial of degree } n}{P_n(x)} = \sum_{k=0}^{n} f\left(\tfrac{k}{n}\right) \underbrace{\binom{n}{k} x^k (1-x)^{n-k}}_{b(x,n,k)} \xrightarrow[\text{claim it converges uniformly}]{} f(x)$$

Proof idea: Show $|f(x) - f_n(x)|$ is small:

$$\left| f(x) - \sum_k f\left(\tfrac{k}{n}\right) b(x,n,k) \right| \le \sum_k \left| f(x) - f\left(\tfrac{n}{k}\right)\right| b(x,n,k)$$

$\sum_k b(x,n,k) = 1$

$$\le \sum_{k \in S_1} + \sum_{k \in S_2}$$



S_1 = all k where $\left|\frac{k}{n} - x\right| < \delta$

S_2: all k where $\left|\frac{k}{n} - x\right| > \delta$

use combinatorial identities

$1 = \sum_k b(x,n,k)$

$nx = \sum_k k\, b(x,n,k)$

$n(n-1)x^2 = \sum_k k(k-1)\, b(x,n,k)$

We can rescale for $\mathcal{C}_b([a,b])$

Note that $\mathcal{C}_b([a,b])$ is <u>separable</u> or has a countable dense subset.

<u>Thm:</u> (Stone-Weierstrauss) Say X is a compact metric space. Suppose $\mathcal{A}$ is a subalgebra of $\mathcal{C}(X)$ s.t.

(i) $\mathcal{A}$ separates points in X, i.e. $\forall x,y \in X$, $\exists f \in \mathcal{A}$ s.t. $f(x) \neq f(y)$

(ii) $\mathcal{A}$ vanishes at no points of X, i.e. $\forall x \in X$, $\exists g \in \mathcal{A}$ s.t. $g(x) \neq 0$

Then $\mathcal{A}$ is dense in $\mathcal{C}(X)$. So functions in $\mathcal{C}(X)$ can be uniformly approximated by functions in $\mathcal{A}$.

Power Series

<u>Def:</u> A <u>power series</u> is a series of the form $\sum_{n=0}^{\infty} c_n (x-a)^n$.

We'll consider real series: $c_n, x, a \in \mathbb{R}$ with $a=0$ (since others are translates). We drop $\sum_{n=0}^{\infty}$ when it's understood.

Q: For what x does a power series converge? $\sum c_n x^n$

<u>Ex:</u> $\sum \frac{x^n}{n!}$ everywhere, $\sum \frac{x^n}{n}$ $[-1, 1)$, $\sum n! x^n$ $x=0$

<u>Def:</u> Given a sequence a_n, $\limsup a_n := \lim_{n\to\infty} \sup_{m \geq n} a_m$.

<u>Def:</u> For power series $\sum c_n x^n$, the radius of convergence R is defined by $\rho := \limsup |c_n|^{1/n}$ and $R = 1/\rho$. If $\rho = 0$, we say $R = \infty$. If $\rho = \infty$, we say $R = 0$.

<u>Ex:</u> $\sum x^n \Rightarrow \rho = \limsup (1)^{1/n} = 1 \Rightarrow R = 1 \Rightarrow$ converges on $(-1, 1)$

<u>Ex:</u> $\sum 2^n x^n \Rightarrow$ converges on $(-\frac{1}{2}, \frac{1}{2})$

<u>Thm:</u> (Cauchy-Hadamard) $\sum c_n x^n$ converges absolutely on $|x| < R$ and diverges on $|x| > R$.

<u>Proof:</u> If $0 < |x| < R$, $\exists b$ with $0 < b < 1$ s.t. $|x| < bR$. So $1/R < \frac{b}{|x|} \Rightarrow \rho < \frac{b}{|x|}$
$\Rightarrow \limsup |c_n|^{1/n} < \frac{b}{|x|} \Rightarrow |c_n|^{1/n} < \frac{b}{|x|}$ for large enough n. Thus, $|c_n x^n| < b^n$ for $b < 1$ and large enough $n \Rightarrow \sum |c_n x^n| < \sum b^n < \infty$ on it's tail so $\sum c_n x^n$ converges absolutely. If $|x| > R = 1/\rho$ then $\limsup |c_n|^{1/n} > \frac{1}{|x|} \Rightarrow |c_n|^{1/n} > \frac{1}{|x|}$ infinitely often $\Rightarrow |c_n x^n| > 1$ infinitely often so $\sum c_n x^n$ diverges by the term test.

In complex numbers, a power series converges on a disk centered at a with radius R. Anything is possible on the endpoints.

$\sum \frac{x^n}{n^2}$ $[-1, 1]$, $\sum \frac{x^n}{n}$ $[-1, 1)$, $\sum \frac{(-x)^n}{n}$ $(-1, 1]$, $\sum x^n$ $(-1, 1)$

Exercise: $R = \lim_{n\to\infty} \frac{|c_n|}{|c_{n+1}|}$ when limit exists

Q: Can I take derivatives of power series?

Thm: For any $\varepsilon > 0$, $\sum c_n x^n$ converges uniformly on the compact subset $[-R+\varepsilon, R-\varepsilon]$.

Proof: Let $b = \frac{R-\varepsilon/2}{R}$ and use previous proof. For big enough n, $|c_n x^n| \le \left(\frac{R-\varepsilon/2}{R}\right)^n$ by the M-test and we have uniform convergence.

Thm: $f(x) = \sum c_n x^n$ is continuous, differentiable on $(-R, R)$, and $f'(x) = \sum_{n=1}^{\infty} n c_n x^{n-1}$ on $|x| < R$.

Power series are ∞-differentiable since the radius of convergence remains the same. We say they are "C^∞ functions" or "smooth".

Proof: Let $f(x) = \sum c_n x^n$. Given $x_0 \in (-R, R)$, it's in some $[-R+\varepsilon, R-\varepsilon]$. Since f converges uniformly on this compact subset, f is continuous at x because it is a uniform limit of continuous functions. Note $\lim_{n\to\infty} n^{1/n} = 1 \Rightarrow \limsup |nc_n|^{1/n} = \limsup |c_n|^{1/n}$ so $\sum n c_n x^{n-1}$ has the same R as $\sum c_n x^n$. Thus, term-by-term differentiation works on $(-R, R)$.

Taylor Series

Recall a power series $\sum c_n (x-a)^n$ has a radius of convergence R, where $\frac{1}{R} = \limsup |c_n|^{1/n}$, so the series converges in $(a-R, a+R)$ and possibly the endpoints. It has uniform convergence on compact subsets, and it can be differentiated term-by-term with the same R.

Cor: $f(x) = \sum c_n x^n$ has infinitely many derivatives on $(-R, R)$ and

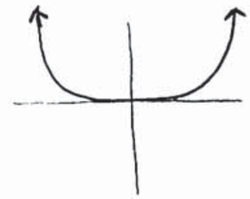
$$f^{(k)}(0) = k!\, c_k$$

Q: Are power series unique?

Thm: If $f(x) = \sum c_n x^n = \sum d_n x^n$ on $(-R, R)$, then $c_n = d_n$ for all n.

Why?: $f^{(k)}(0) = k!\,c_k = k!\,d_k$

Careful: $f(x) = \begin{cases} 0, & \text{if } x=0 \\ e^{-1/x^2}, & \text{else} \end{cases}$ Note that $f^{(k)}(0) = 0, \forall k$



The power series at 0 is 0 so it doesn't converge to $f(x)$ except at 0 but it converges everywhere.

Q: Can I integrate term-by-term?

Thm: Yes, over a compact interval. This follows from uniform convergence of power series on a compact set.

Thm: Suppose $f(x) = \sum c_n x^n$ has radius $R=1$ and $\sum c_n = c$. Then

$$\lim_{x \to 1^-} f(x) = c$$

Abel summation extends the notion of summation.

Ex: $1-1+1-1+1-1+\dots$ doesn't converge but it equals $\frac{1}{2}$ under Abel summation. Note that $\frac{1}{1+x} = 1 - x + x^2 - x^3 + \dots$ for $-1<x<1$. We let the sum be $\lim_{x\to 1^-} \frac{1}{1+x} = \frac{1}{2}$. If we rewrite the sum as $1+0-1+1+0-1+\dots$, we get that it is equal to $\frac{1-x^2}{1-x^3}$ which converges to $2/3$ by Abel summation.

We also have Cesaro summation. For $1-1+1-1+\dots$, if we take the sequence of the average of the partial sums $1, \frac{1}{2}, \frac{1}{3}, \dots$, it converges to $\frac{1}{2}$.

Q: When is a double sum "switchable"? $\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} a_{ij} \overset{?}{=} \sum_{j=1}^{\infty}\sum_{i=1}^{\infty} a_{ij}$

Thm: If $\sum_{j=1}^{\infty} |a_{ij}|$ converges (to say b_i) and $\sum_{i=1}^{\infty} b_i$ converge

then $\sum_i \sum_j a_{ij} = \sum_j \sum_i a_{ij}$

j \ i	1	2	3	4 ...
1	-1	0	0	0
2	1/2	-1	0	0
3	1/4	1/2	-1	0
4	1/8	1/4	1/2	-1
⋮				

$\sum_i \sum_j a_{ij} = 0$

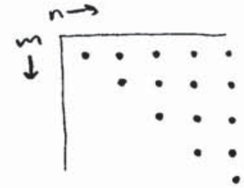
$\sum_j \sum_i a_{ij} = -2$

<u>Thm:</u> (Taylor's Theorem) If $f(x) = \sum_{n=0}^{\infty} c_n x^n$ on $|x| < R$, then

① If $a \in (-R, R)$, f has a power series centered at $x = a$ converging in $|x - a| < R - |a|$ (but possibly more).

② $f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n$

<u>Proof:</u> $f(x) = \sum_{n=0}^{\infty} c_n (x - a + a)^n = \sum_{n=0}^{\infty} c_n \sum_{m=0}^{n} \binom{n}{m} a^{n-m} (x-a)^m$



$$= \sum_{m=0}^{\infty} \underbrace{\left[\sum_{n=m}^{\infty} \binom{n}{m} c_n a^{n-m} \right]} (x-a)^m \quad \text{(by technical lemma)}$$

Check $\sum_n \sum_m \left| c_n \binom{n}{m} a^{n-m} (x-a)^m \right|$ converges on $|x-a| + |a| < R$

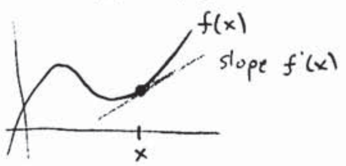
We claim this is the m^{th} derivative of $f(x)$ at $x = a$ divided by $m!$. Thus we have Taylor's Theorem.

Power series expansions at a must be identical in $(-R, R)$. If we start with some f that is not a power series, its Taylor series may not be f (e.g. e^{-1/x^2}).

Multivariable Functions

Ⓠ: Given $f: \mathbb{R}^n \to \mathbb{R}^m$, what is a "derivative"?

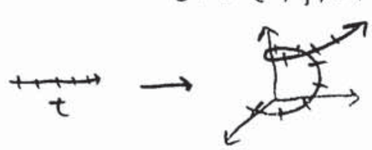
<u>Ex:</u> $f: \mathbb{R} \to \mathbb{R}$



graph

<u>Ex:</u> $f: \mathbb{R} \to \mathbb{R}^3$

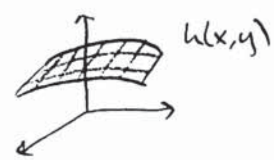
$t \mapsto (x, y, z)$



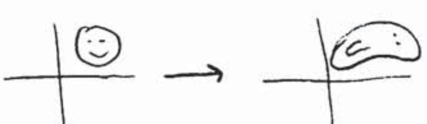
mapping

<u>Ex:</u> $h: \mathbb{R}^2 \to \mathbb{R}$

$(x, y) \mapsto z$

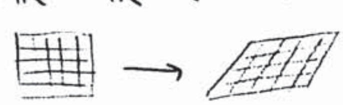


<u>Ex:</u> $j = \mathbb{R}^2 \to \mathbb{R}^2$



Derivative: "the best linear approximation"

Recall a linear functional $L: \mathbb{R}^n \to \mathbb{R}^m$ can be represented by a matrix $A_{m \times n} = (a_{ij})$



$[\;]_{m\times 1} = [A]_{m \times n} [\;]_{n \times 1}$

If $f: \mathbb{R}^n \to \mathbb{R}^m$ is differentiable (at $\vec{x}$) then locally (at $\vec{x}$) it behaves like a linear transformation. We expect the (total) derivative of f to be a linear map f' or $Df: \mathbb{R}^n \to \mathbb{R}^m$ such that

$$f(\vec{x}) - f(\vec{x}_0) = [Df](\vec{x} - \vec{x}_0)$$

<u>Def:</u> If $\exists$ linear map $A: \mathbb{R}^n \to \mathbb{R}^m$ s.t. for a given f,

$$\lim_{|h| \to 0} \frac{|f(x+h) - f(x) - Ah|_{\mathbb{R}^m}}{|h|_{\mathbb{R}^n}} = 0$$

then f is <u>differentiable</u> and $Df(x) = A$.

Alternatively, $f(x+h) - f(x) = Df(x) \cdot h + r(h)$ where $\lim_{|h| \to 0} \frac{|r(h)|}{|h|} = 0$.

Q: Is this well-defined? Suppose we have two linear maps A_1 and A_2. Let $B = A_1 - A_2$. Observe that $Bh = A_1 h - A_2 h =$ $\{f(x+h) - f(x) - A_2 h\} - [f(x+h) - f(x) - A_1 h]$. So, $\frac{|Bh|}{|h|} \le \frac{|\{\}|}{|h|} + \frac{|[]|}{|h|}$ then take $\lim_{|h| \to 0}$ yields $\lim_{|h| \to 0} \frac{|Bh|}{|h|} = 0$. Write $h = th_0$. Then, by linearity of B, $\frac{|Bth_0|}{|th_0|} = \frac{|t||Bh_0|}{|t||h_0|}$ so $\frac{|Bh_0|}{|h_0|} = 0$ for any h_0 so $B = 0$.

Let $L(\mathbb{R}^n, \mathbb{R}^m)$ = space of all linear transformations $\mathbb{R}^n \to \mathbb{R}^m$. In fact, this is a vector space so it has a norm. For $A \in L(\mathbb{R}^n, \mathbb{R}^m)$, define

max radius of image of unit ball under A → $\|A\| = \sup_{\underbrace{|x|_{\mathbb{R}^n} \le 1}_{\text{unit ball}}} |Ax|_{\mathbb{R}^m}$

Remarks:

1. $\forall w$, $|Aw| \le \|A\| |w|$ since $A \frac{w}{|w|} \le \|A\|$.
2. If $\forall x$, $|Ax| \le \lambda |x|$ then $\|A\| \le \lambda$ since $A \frac{w}{|w|} \le \lambda \cdot 1$ so $\underbrace{\sup_w \left| A \frac{w}{|w|} \right|}_{\|A\|} \le \lambda$ for all w.
3. $A \in L(\mathbb{R}^n, \mathbb{R}^m) \Rightarrow \|A\| < \infty$, A is uniformly continuous.

<u>Proof:</u> If $|\vec{x}| < 1$ then $\vec{x} = \sum c_i \vec{e}_i$ where $|c_i| \le 1$ and $\vec{e}_i$ are unit vectors. So $|A\vec{x}| = |\sum c_i A\vec{e}_i| \le \sum |c_i||A\vec{e}_i| \le \sum |A\vec{e}_i| < \infty$. So $\|A\|$ is finite. Also, $|A\vec{x} - A\vec{y}| = |A(\vec{x} - \vec{y})| \le \|A\| |\vec{x} - \vec{y}|$ so A uniformly continuous.

4. $\|\cdot\|$ is a norm: $\|A + B\| \le \|A\| + \|B\|$, $\|cA\| = |c| \|A\|$ so it induces a metric on L, $d(A,B) = \|A - B\|$

<u>Proof:</u> Follows from norm properties of $|\cdot|$ in $\mathbb{R}^n, \mathbb{R}^m$.

5. $A \in L(\mathbb{R}^n, \mathbb{R}^m)$, $B \in L(\mathbb{R}^m, \mathbb{R}^k) \Rightarrow \|BA\| \le \|B\| \|A\|$.

<u>Proof:</u> $\forall \vec{x}$, $|BA\vec{x}| \le \|B\| |A\vec{x}| \le \underbrace{\|B\| \|A\|}_{\lambda} |\vec{x}|$ so $\|BA\| \le \|B\| \|A\|$

⑥ <u>Thm</u>: If $A \in L(\mathbb{R}^n, \mathbb{R}^m)$ then $DA = A$.

<u>Proof</u>: Follows from $\frac{|A(x+h) - Ax - Ah|}{|h|} = 0$ ← by linearity, $=0$. so A satisfies the definition of the derivative.

For $f: \mathbb{R} \to \mathbb{R}$, if $A = [c]$ then $f(x) = cx$ and $Df = [c]$.

The Deriative Matrix

Recall: $f(\vec{x}+\vec{h}) - f(\vec{x}) = Df(\vec{x}) \cdot \vec{h} + r(\vec{h})$ where $\lim_{|\vec{h}| \to 0} \frac{|r(\vec{h})|}{|\vec{h}|} = 0$ (7)

We see that f differentiable $\Rightarrow f$ continuous because ε related to δ by $\|Df(x)\|$. Also, $Df(\vec{x}): \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation dependent on $\vec{x}$. We can think of Df as $Df: \mathbb{R}^n \to L(\mathbb{R}^n, \mathbb{R}^m)$. (8)

We claim if $f: \mathbb{R}^n \to \mathbb{R}^m$ is differentialle, then

$$Df = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix}_{m \times n}$$

where the basis for $\mathbb{R}^n$ is $\{e_j\}_{j=1}^n$, the basis for $\mathbb{R}^m$ is $\{u_i\}_{i=1}^m$, and

$$\frac{\partial f_i}{\partial x_j} = \lim_{t \to 0} \frac{f_i(x+te_j) - f_i(x)}{t}$$

<u>Thm</u>: Suppose $f: \mathbb{R}^n \to \mathbb{R}^m$ is differentialle at x. Then all $\frac{\partial f_i}{\partial x_j}$ exist and $Df(x) \cdot e_j = \sum_{i=1}^m \frac{\partial f_i}{\partial x_j}(x)\, u_i$

<u>Proof</u>: We note that by (7), $\underline{\frac{f(x+te_j) - f(x)}{t}} = \underline{\frac{Df(x) \cdot \cancel{t} e_j}{\cancel{t}}} + \underline{\frac{r(te_j)}{t}}$ which is

$$\sum_{i=1}^m \frac{f_i(x+te_j) - f_i(x)}{t} u_i = Df(x) \cdot e_j + \frac{r(te_j)}{t}$$

Taking $t \to 0$ gives

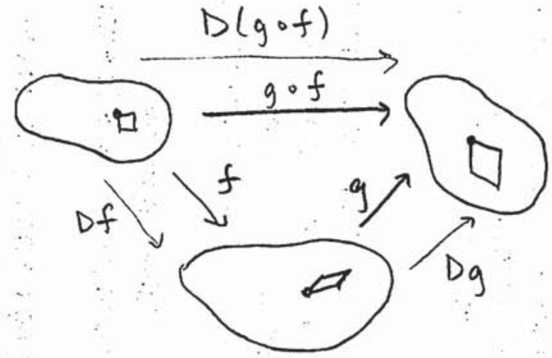
$$\sum_{i=1}^m \frac{\partial f_i}{\partial x_j}(x)\, u_i = Df(x) \cdot e_j$$

Chain Rule

The chain rule gives $D(g\circ f) = Dg \circ Df$

or $D(g\circ f)(x) = Dg(f(x)) \circ Df(x)$

(like $\frac{dT}{dx} = \frac{dT}{dW} \cdot \frac{dW}{dx}$)



$g(f(x)) = g\circ f\,(x)$.

Proof: Let $A = Df(x)$ so $f(x+h) = f(x) + Ah + u(h)$ (1).
Let $B = Dg(f(x))$ so $g(y+h) = g(y) + Bk + v(k)$ (2),
where $\frac{|u(h)|}{|h|} \to 0$ and $\frac{|v(k)|}{|k|} \to 0$. Let $y = f(x)$ and $F = g\circ f$.
Notice that

$$F(x+h) - F(x) - BAh = g(f(x+h)) - g(f(x)) - BAh$$

$$g(f(x) + \underbrace{Ah + u(h)}_{k}) = g(y) + B(Ah) + B(u(h)) + v(Ah+u(h))$$

$$= B(u(h)) + v(Ah + u(h))$$

so we have

$$\frac{|F(x+h) - F(x) - BAh|}{|h|} \le \frac{|Bu(h)|}{|h|} + \frac{|v(Ah+u(h))|}{|h|}$$

$$\le \|B\| \frac{|u(h)|}{|h|} + \frac{|v(Ah+u(h))|}{|Ah+u(h)|} \cdot \underbrace{\frac{|Ah+u(h)|}{|h|}}_{\text{need bounded } **}$$

as $|h| \to 0$, 0

$\uparrow$ 0, as long as *

* Show as $h \to 0$, $Ah + u(h) \to 0$

Proof: $|Ah + u(h)| \le \|A\|\,|h| + \frac{|u(h)|}{|h|}\,|h|$ $\to 0$

** Show $\frac{|Ah+u(h)|}{|h|}$ bounded

Proof: $\le \|A\| + \frac{|u(h)|}{|h|}$ $\to 0$

Chain Rule (Traditional)

$$\underset{DF}{[\ \circ\ ]} = \underset{Dg}{[\ \text{▭}\ ]}\ \underset{Df}{[\ \text{▯}\ ]} \qquad \frac{\partial F_i}{\partial x_j} = \sum_k \frac{\partial F_i}{\partial f_k} \frac{\partial f_k}{\partial x_j}$$

If $f: \mathbb{R}^n \to \mathbb{R}^m$, then $\det(Df)$ is the Jacobian (the "expansion factor" of volume locally)

Special case of chain rule:

$f: \mathbb{R}^n \to \mathbb{R}$ (pos. $\mapsto$ temp.) $x: \mathbb{R} \to \mathbb{R}^n$ (time $\mapsto$ pos.) (param. path) Let $F(t) = f(x(t))$ = temp of time t along path. Then,

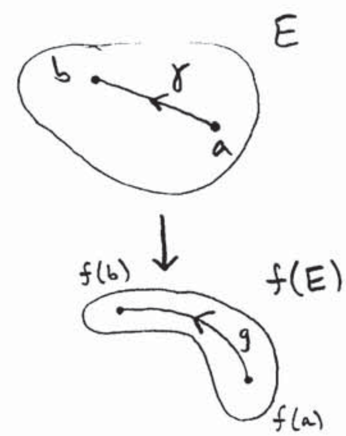
$$F'(t) = \underbrace{\begin{bmatrix} \; Df \; \end{bmatrix}_{1\times n}}_{\nabla f} \begin{bmatrix} Dx \end{bmatrix}_{n \times 1} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \frac{\partial x_i}{\partial t} = \nabla f(x(t)) \cdot x'(t)$$

If path has unit speed in direction $\vec{u}$ (so $x'(t) = \vec{u}$) then we get the <u>directional derivative</u> $D_{\vec{u}} f = \nabla f \cdot \vec{u}$.

$\mathcal{C}^1$ Functions

<u>Thm:</u> Let E be a convex open set in $\mathbb{R}^n$ and $f: E \to \mathbb{R}^m$ is differentiable on E. Suppose $\exists M$ s.t. $\|Df\| \le M$. Then,

$$|f(b) - f(a)| \le M|b-a|$$



<u>Proof:</u> Let $\vec{u}$ = unit vector in direction $f(a)$ to $f(b)$. Let $\gamma(t) = a + t(b-a)$ so $\gamma(0) = a$ and $\gamma(1) = b$. Let $g(t) = f(\gamma(t))$. Then consider $g(t)\cdot\vec{u}$, a function from $\mathbb{R}^1$ to $\mathbb{R}^1$. Apply MVT gives

$$g(1)\cdot\vec{u} - g(0)\cdot\vec{u} = \frac{d}{dt}(g(t)\cdot\vec{u})(1-0)$$
$$= Df(\gamma(t))\cdot\gamma'(t)$$

Taking the $|\cdot|$ on both sides,

$$|g(1) - g(0)||\vec{u}|\cos 0 = |Df||\gamma'(t)|\cos\theta$$
$$\le M|b-a|$$

Q: We saw if Df exists then $\frac{\partial f_i}{\partial x_j}$ exists. Is the converse true? No. See $f(x,y) = \frac{xy}{x^2+y^2}$ and 0 at $(0,0)$. Even if continuous? No. See $f(x,y) = \frac{xy^2}{x^2+y^2}$ and 0 at $(0,0)$.

<u>Def:</u> $f: \mathbb{R}^n \to \mathbb{R}^m$ is <u>continuously differentiable</u> in $E \subseteq \mathbb{R}^n$ if Df is continuous as the map $E \to L(\mathbb{R}^n, \mathbb{R}^m)$. We say f is $\mathcal{C}^1$ and we write $f \in \mathcal{C}^1(E)$ where $\mathcal{C}^k(E)$ = class of all functions with k derivatives all continuous.

Thm: $f \in \mathcal{C}'(E) \iff \frac{\partial f_i}{\partial x_j}$ all exist and are continuous.

Proof: ($\Rightarrow$) $\frac{\partial f_i}{\partial x_j}(x) - \frac{\partial f_i}{\partial x_j}(y) = \underbrace{[Df(x) - Df(y)]e_j}_{j^{th} \text{ row}} \cdot \underbrace{u_i}_{i^{th} \text{ col}}$

$$\leq \|[Df(x) - Df(y)]\| \, |e_j| \, |u_i|$$

This shows the forward direction.

($\Leftarrow$) Assume partials exist and are continuous. We want to show Df is in $\mathcal{C}'$. Recall

$$\frac{|f(x+h) - f(x) - \boxed{?}\, h|}{|h|} \to 0 \quad \text{as} \quad h \to 0$$

true if true for each component f_i so it's enough to check for $f: \mathbb{R}^n \to \mathbb{R}$. We claim

$$Df = \left[\frac{\partial f}{\partial x_1} \cdots \frac{\partial f}{\partial x_n}\right]^T$$

which is the map $h \to \sum \frac{\partial f}{\partial x_j} h_j$. We examine

$$\left|f(x+h) - f(x) - \sum_j h_j \frac{\partial f}{\partial x_j}(x)\right|$$

$$= \left|\sum h_j \frac{\partial f}{\partial x_j}(c_j) - \sum h_j \frac{\partial f}{\partial x_j}(x)\right| \quad \text{by MVT}$$

$$\leq \sum_j h_j \underbrace{\left|\frac{\partial f}{\partial x_j}(c_j) - \frac{\partial f}{\partial x_j}(x)\right|}_{\delta}$$

$$\leq |h| n \delta$$

so choose $\delta = \varepsilon/n$. Continuity of Df follows from continuity of partials.

The Inverse Function Theorem

Recall from linear algebra, to solve $A\vec{x} = \vec{y}$ (find $\vec{x}$ or given $\vec{y}$)

$$a_{11}x_1 + \dots + a_{1n}x_n = y_1$$
$$\vdots$$
$$a_{n1}x_1 + \dots + a_{nn}x_n = y_n$$

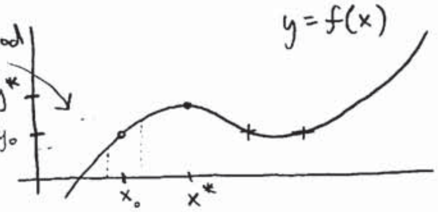
Key: We can solve uniquely if $A = (a_{ij})$ is invertible or $\det(A) \neq 0$.

Q: What about

$$f_1(x_1, \dots, x_n) = y_1$$
$$\vdots$$
$$f_n(x_1, \dots, x_n) = y_n$$

Ex: $f:\mathbb{R}\to\mathbb{R}$

∃ neighborhood V locally invertible b/c $f'(x_0)\neq 0$



Q: When can $f:\mathbb{R}^n\to\mathbb{R}^n$ be "locally inverted" near $f(x_0)=y_0$?

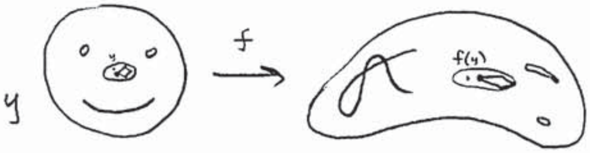
Inverse function theorem ... look at $\det(Df)$, want $f\in\mathcal{C}'$.

Thm: (Inverse Function Theorem) Suppose $f:\underset{\subseteq\mathbb{R}^n}{E}\to\mathbb{R}^n$ is $\mathcal{C}'$, $f(a)=b$, and $Df(a)$ is invertible. Then,

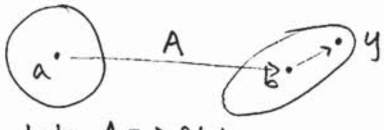
1) ∃ open U containing a and open V containing b s.t. f is one-to-one and onto on U and $f(U)=V$

2) Let $g=f^{-1}$ on V. Then $g\in\mathcal{C}'(V)$. (In fact if f is $\mathcal{C}^r$ then g is $\mathcal{C}^r$)

Idea: $f:\mathbb{R}^2\to\mathbb{R}^2$. The idea in Rudin is to construct a function φ_y to help find pre-image of y



?



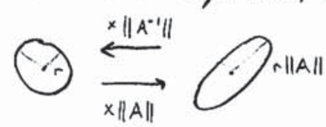
Let $A=Df(a)$

Approximate ? by $a+A^{-1}[y-f(a)]$.

Let $\varphi_y(x)=x+A^{-1}[y-f(x)]$. Note that the fixed point for φ_y is a preimage of y. Also, φ_y is a contraction if x is close to a. For $f:\mathbb{R}\to\mathbb{R}$ and $y=0$, finding preimage is finding roots so $\varphi_0(x)=x-\frac{f(x)}{f'(a)}$ which is like Newton's method.

Q: What neighborhood U to use? We'll choose U s.t. $\forall x\in U, \|Df(x)-A\|<\lambda$

We choose $\lambda=\frac{1}{2\|A^{-1}\|}$.



Step 1) f is 1-1 on U. Why? φ_y is a contraction on U, b/c
$D\varphi_y=I-A^{-1}\circ Df(x)=A^{-1}(A-Df(x))$ so $\|D\varphi_y\|\le\|A^{-1}\|\lambda=\frac{1}{2}$. By previous thm, $|\varphi_y(x)-\varphi_y(w)|\le\frac{1}{2}|x-w|$, a contraction.

Step 2) Let $V=f(U)$, we'll show V is open by taking $y_0\in V$ & finding open ball about y_0 still in V. Given y_0, $\exists x_0$ s.t. $f(x_0)=y_0$. Since U is open, choose ball B of radius r about x_0 s.t. $\overline{B}\subset U$. We'll show if $|y-y_0|<\lambda r$ then $y\in V$ (so V is open). Consider φ_y, claim $\varphi_y:\overline{B}\to\overline{B}$ (so φ_y contracts $\Rightarrow$ ∃ fixed pt for φ_y in $\overline{B}$, so $y\in f(\overline{B})\subset f(U)=V$) as desired.

If $x\in\overline{B}$, then
$$\begin{aligned}|\varphi_y(x)-x_0| &\le |\varphi_y(x)-\varphi_y(x_0)|+|\varphi_y(x_0)-x_0|\\ &\le \tfrac{1}{2}|x-x_0|+|A^{-1}(y-y_0)|\\ &\le \tfrac{1}{2}r+\|A^{-1}\|\,|y-y_0|\\ &\le \tfrac{1}{2}r+\tfrac{1}{2\lambda}\cdot\lambda r\le r\end{aligned}$$

So $\varphi_y(x)\in\overline{B}$

Recall: choose $\mathcal{U}$ to be an open ball in E s.t. $\forall x \in \mathcal{U}$

$$\|\underbrace{Df(x)}_{B} - \underbrace{Df(a)}_{A}\| < \lambda = \frac{1}{2\|A^{-1}\|}$$

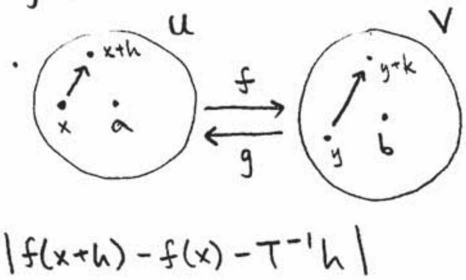
③ We will prove $B = Df(x)$ has an inverse, by showing if $w \neq 0$ then $Bw \neq 0$.
Recall A is invertible (nearby) so $|Aw| \le |(A-B)w| + |Bw|$

$$2\lambda|w| = 2\lambda|A^{-1}Aw| \le 2\lambda\|A^{-1}\||Aw| = |Aw| \le |(A-B)w| + |Bw| \le \lambda|w| + |Bw|$$

so $\lambda|w| \le |Bw|$ so $w \neq 0 \Rightarrow Bw \neq 0$ as desired.

Claim ④: By ① and ②, f has a local inverse, let $g = f^{-1}$ on V. We claim Dg exists on V, is continuous, and $Dg(y) = [Df(x)]^{-1}$

<u>Proof:</u> Consider $y, y+k$ with preimages under f: $x, x+h$.
By ③, let $T = [Df(x)]^{-1}$. Consider



(1) $|g(y+k) - g(y) - Tk| = |h - Tk| = |TT^{-1}h - Tk|$
$\le \|T\||k - T^{-1}h| = \|T\||f(x+h) - f(x) - T^{-1}h|$

Note that

$$|h - Tk| = |h - A^{-1}(f(x+h) - f(x))| = |\varphi_{y+k}(x+h) - \varphi_{y+k}(x)| \le \tfrac{1}{2}|x+h-x| = \tfrac{1}{2}|h|$$

so $|Tk| \ge \frac{1}{2}|h|$ so $|h| \le 2|Tk| = 2|A^{-1}k| \le 2\|A^{-1}\||k| = |k|/\lambda$. Then, looking at (1),

$$\frac{LHS}{|k|} \le \frac{RHS}{\lambda|h|}$$

so as $h \to 0$, $k \to 0$, so $\frac{LHS}{|k|} \to 0$ so Dg exists.

Claim ⑤: Dg is continuous. We know g is continuous (since it's differentiable) and Df is continuous (since f is $\mathcal{C}^1$). It is enough to show $T \to T^{-1}$ is continuous on $L(\mathbb{R}^n, \mathbb{R}^n)$ because $Dg(y) = [Df(g(y))]^{-1}$.
Why? $\|A^{-1} - B^{-1}\| = \|B^{-1}(A-B)A^{-1}\| \le \|B^{-1}\|\|A-B\|\|A^{-1}\|$

From ③, $\lambda|w| \le |Bw|$, $\forall w \neq 0$ so $\lambda|B^{-1}y| \le |y|$, $\forall y$ so $|B^{-1}y| \le \frac{1}{\lambda}|y| = 2\|A^{-1}\||y|$
so $\|B^{-1}\| \le 2\|A^{-1}\|$. Thus, $\|A^{-1} - B^{-1}\| \le 2\|A^{-1}\|^2\|A-B\|$ so the inverse is continuous.

Implicit Functions

Ex: $f: \mathbb{R}^2 \to \mathbb{R}$ s.t. $f(x,y) = 0$ so it defines x in terms of y implicitly.

$x^2 + y^2 - 1 = 0$

Q: When can x be explicitly solved in terms of y?

can't define y in terms of x locally

Q: In some neighborhood $\mathcal{U}$ of (a,b) where $f(a,b)=0$ is there some neighborhood W of b where each y has a unique x? If so, define a function $g(y)$ s.t. $g(b)=a$ and $f(g(y),y)=0$.

Note g may not exist where $\frac{\partial f}{\partial x}=0$.

The Implicit Function Theorem

Recall: $f:\mathbb{R}^2\to\mathbb{R}$, $f(x,y)=x^2+y^2-1$, $f(x,y)=0$ defines x in terms of y implicitly

Q: When can x be explicitly solved in terms of y in some neighborhood of (a,b) where $f(a,b)=0$? In some neighborhood $\mathcal{U}$ of (a,b), is there a neighborhood W of b where each y has unique x?

Ex: $f:\mathbb{R}^3\to\mathbb{R}^2$, $f(x_1,x_2,y)=(0,0)$

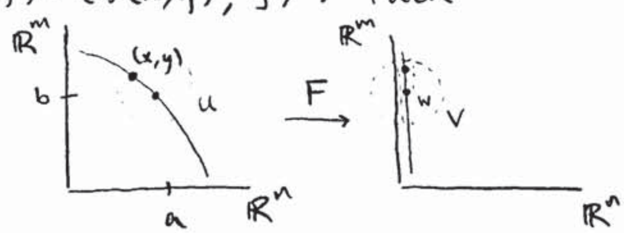
Q: When does $\exists g(y)$ s.t. $f(g(y),y)=(0,0)$ in a neighborhood of (a_1,a_2,b)?

Let $A=Df(a_1,a_2,b)=\Bigg[\underbrace{\begin{matrix}\frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2}\\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2}\end{matrix}}_{A_x}\ \underbrace{\begin{matrix}\frac{\partial f_1}{\partial y}\\ \frac{\partial f_2}{\partial y}\end{matrix}}_{A_y}\Bigg]$. We get problems when $\det(A_x)=0$.

Thm: (Implicit Function Theorem) Let $f:\underset{\subseteq\mathbb{R}^{n+m}}{E}\to\mathbb{R}^n$ be $\mathcal{C}^1$, and $f(\underset{\subseteq\mathbb{R}^n}{\vec{a}},\underset{\subseteq\mathbb{R}^m}{\vec{b}})=\underset{\subseteq\mathbb{R}^n}{\vec{0}}$ for $(\vec{a},\vec{b})\in E$. Let $A=Df(\vec{a},\vec{b})$, so $A=[A_x|A_y]$, and suppose $\det(A_x)\neq 0$. Then, $\exists\,\mathcal{U}\subseteq\mathbb{R}^{n+m}$ containing $(\vec{a},\vec{b})$ and $\exists W\in\mathbb{R}^m$ containing $\vec{b}$ s.t. $\forall\vec{y}\in W$, $\exists$ unique $\vec{x}$ s.t. $(\vec{x},\vec{y})\in\mathcal{U}$ and $f(\vec{x},\vec{y})=\vec{0}$. This defines g s.t. $\vec{x}=g(\vec{y})$. Then g is a $\mathcal{C}^1$ map $:W\to\mathbb{R}^n$ and $g(b)=a$, and $f(g(y),y)=\vec{0}$ and $Dg=-A_x^{-1}A_y$.

Proof idea: Apply Inverse Function Theorem to $F(x,y)=(\underset{\subseteq\mathbb{R}^n}{f(x,y)},\underset{\subseteq\mathbb{R}^m}{y})$. Then $F:\mathbb{R}^{n+m}\to\mathbb{R}^{n+m}$. Notice F is $\mathcal{C}^1$ b/c $F(x,y)=(\overset{\mathcal{C}^1}{f(x,y),0})+\overset{\mathcal{C}^\infty}{(0,y)}$. Also,

$$Df=\left[\begin{array}{c|c}\frac{\partial f_i}{\partial x_j} & \frac{\partial f_i}{\partial y_j}\\ \hline 0 & 1\end{array}\right]=\left[\begin{array}{c|c}A_x & A_y\\ \hline 0 & I\end{array}\right]_{(n+m)\times(n+m)}$$



which is invertible because A_x and I are invertible. By the Inverse Function Theorem, $\exists\,\mathcal{U}$ containing (a,b) & V containing $(0,b)$ s.t. F is a bijection between $\mathcal{U}$ and V. Let $W=\{y\in\mathbb{R}^m:(0,y)\in V\}$. Notice $b\in W$ and W is open in $\mathbb{R}^m$. Verify if $y\in W$, then $(0,y)\in V\Rightarrow$ local inverse (x,y) s.t. $F(x,y)=(0,y)$ so $f(x,y)=0$ as desired and given $y\in W$ $\exists$ unique $(x,y)\in V$ s.t. $f(x,y)=0$.

This defines g s.t. $x=g(y)$ for $y\in W$. Note

$$y \xrightarrow[\mathscr{C}^\infty]{} (0,y) \xrightarrow[\mathscr{C}^1]{F^{-1}} (g(y),y) \xrightarrow[\mathscr{C}^\infty]{} g(y)$$

so g is $\mathscr{C}^1$ as composition. By definition, $f(g(y),y)=0$. Take derivatives and use chain rule:

$$\underbrace{\left[Df(g(y),y)\right]}_{[A_x \mid A_y]} \cdot \left[\frac{Dg}{I}\right] = [0]$$

This gives $A_x Dg + A_y I = [0] \Rightarrow Dg = A_x^{-1}(-A_y) = -A_x^{-1}A_y$.

Differentiation of Integrals and Derivatives

(Q:) When is $\frac{d}{dt}\int_{x=a}^{b} \varphi(x,t)\,dx = \int_{x=a}^{b} \frac{\partial}{\partial t}\varphi(x,t)\,dx$

<u>Thm:</u> If ① $\varphi(x,t)$ defined in rectangle $x\in[a,b]$, $t\in[c,d]$; ② $\varphi(x,t_0)\in R$, $\forall t_0\in[c,d]$; and ③ $\frac{\partial\varphi}{\partial t}$ is continuous on rectangle, then for $s\in(c,d)$

$$\left[\frac{d}{dt}\int_a^b \varphi(x,t)\,dx\right]_{t=s} = \int_a^b \frac{\partial\varphi}{\partial t}(x,s)\,dx$$

<u>Proof:</u> Let $\Psi(x,t) = \dfrac{\varphi(x,t)-\varphi(x,s)}{t-s} \underset{MVT}{=} \dfrac{\partial\varphi}{\partial t}(x,u)$ for some $u\in(s,t)$. By ③,

$\forall \varepsilon>0, \exists\delta>0$ s.t. $|s-t|<\delta \Rightarrow \left|\Psi(x,t)-\frac{\partial\varphi}{\partial t}(x,s)\right|<\varepsilon$. So $\Psi(x,t_n)$ converges uniformly to $\frac{\partial\varphi}{\partial t}(x,s)$ as $t_n\to s$. Thus,

$$\int_a^b \Psi(x,t)\,dx \longrightarrow \int_a^b \frac{\partial\varphi}{\partial t}(x,s)\,dx \quad \text{as} \quad t\to s$$

Let $f(t)=\int_a^b \varphi(x,t)\,dx$. Then $\int_a^b \Psi(x,t)\,dx = \frac{f(t)-f(s)}{t-s} \to f'(s)$ as $t\to s$, as desired.

Higher Order Derivatives

$f:\mathbb{R}^n\to\mathbb{R}^m$

$Df:\mathbb{R}^n\to L(\mathbb{R}^n,\mathbb{R}^m)$ $\qquad Df(c):\mathbb{R}^n\to\mathbb{R}^m$

$D^2f = D(Df):\mathbb{R}^n\to L(\mathbb{R}^n, L(\mathbb{R}^n,\mathbb{R}^m))$ $\qquad D^2f(c):\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}^m$

If we restrict our attention to $f:\mathbb{R}^n\to\mathbb{R}$, then

$Df(c) = \left[\frac{\partial f}{\partial x_1} \;\dots\; \frac{\partial f}{\partial x_n}\right]$ and $D^2f(c) = \left[\frac{\partial^2 f}{\partial x_i \partial x_j}\right] = Hf(c)$ (Hessian matrix)

$D^2f(c):(\vec{y},\vec{z}) \mapsto \sum_{i,j} \frac{\partial^2 f}{\partial x_j \partial x_i}(c)\, y_i z_j$ or $\vec{y}^{\,T}[Hf]\vec{z}$

This gives a Taylor approximation to $f: \mathbb{R}^n \to \mathbb{R}$ (2^{nd} order)

$$f(\vec{x}) = f(\vec{a}) + Df(\vec{a}) \cdot (\vec{x} - \vec{a}) + \tfrac{1}{2}(\vec{x} - \vec{a})^T Hf(\vec{a})(\vec{x} \cdot \vec{a}).$$

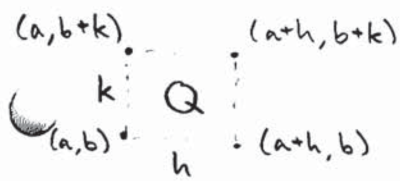
We see local maximums occur when $Df(\vec{a}) = 0$ and all eigenvalues of $Hf(\vec{a})$ are negative. (More generally, the number of positive and negative eigenvalues of Hf determine if the critical point is a max, min, or saddle point.)

Mixed Partials $\quad \frac{\partial}{\partial x_i}\left(\frac{\partial}{\partial x_j} f\right) = \frac{\partial^2}{\partial x_i \partial x_j} f = D_i(D_j f) = D_{ij} f$

<u>Thm:</u> $D_{ji} f = D_{ij} f$ for $f \in \mathcal{C}^2$. ($\Rightarrow Hf$ is symmetric $\Rightarrow$ eigenvectors form an orthogonal basis and eigenvalues are real)

<u>Cor:</u> $D_{ijk} f = D_{\sigma(ijk)} f$ for $f \in \mathcal{C}^3$.

<u>Proof idea:</u> Use MVT a lot.



Let $\Delta(f,Q) = [\overbrace{f(a+h,b+k) - f(a+h,b)}^{u(a+h)}] - [\overbrace{f(a,b+k) - f(a,b)}^{u(a)}]$.

Then $\Delta(f,Q) \underset{MVT}{=} hu'(x)$ for $x \in (a, a+h)$ $= hD_1 f(x,b+k) - hD_1 f(x,b)$

$\underset{MVT}{=} hk D_{21} f(x,y)$ for some $y \in (b, b+k)$

Repeat our argument to get $hk D_{12} f(x,y)$. We then get what we want when $h,k \to 0$.

Differential Forms

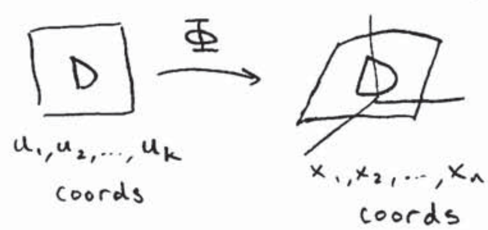
They are fundamental objects to integrate, geometric concepts are represented by forms, allow us to encode local "differential" info at each point, and there's a "cohomology" theory that reveals a topology of surfaces.

Q: What are they? Approaches to think about:

- I forms assign to each surface a value by (integrate)
- II forms assign to each point on a surface an alternating k-tensor (on tangent space)

<u>Def:</u> Let $\Phi : D_{cpt}^{\subset \mathbb{R}^k} \longrightarrow E_{open}^{\subset \mathbb{R}^n}$.

Φ is a k-surface



D is a parameter domain in coords $\{u_1, \dots, u_k\}$
often, $D = I^k$ the k-cell $[a_1, b_1] \times \dots \times [a_k, b_k]$
or $D = Q^k$ the k-simplex 
or built up from such pieces

Key idea: Any geometric concept can be represented by a form.

Let $x_1, \dots, x_n$ be standard coordinates in $\mathbb{R}^n$. Define

$$\Omega^* = \text{algebra}/_{\mathbb{R}} \text{ gen by symbols } dx_1, dx_2, \dots, dx_n$$

with wedge product $\wedge$ with relations:

i) $dx_i \wedge dx_i = 0$

ii) $dx_i \wedge dx_j = -dx_j \wedge dx_i$

As a vector space over $\mathbb{R}$, Ω^* has a basis: $\underset{\Omega^0}{1}, \underset{\Omega^1}{dx_i}, \underset{\Omega^2\ (i<j)}{dx_i \wedge dx_j}, \underset{\Omega^3\ (i<j<k)}{dx_i \wedge dx_j \wedge dx_k}, \dots$

We can define Ω^k generated by wedge products of k "basic" forms.

<u>Ex:</u> $\omega = 5dx - 2dy + \pi dz$ in $\mathbb{R}^3 \in \Omega^1$ "a 1-form" in $\mathbb{R}^3$
$v = 7dy \in \Omega^1$, $\omega \wedge v = 35\, dx \wedge dy - 7\pi\, dy \wedge dz$

<u>Def:</u> A <u>($\mathcal{C}^r$-) differential k-form</u> is a function $\omega : \mathbb{R}^n \to \Omega^k$.

<u>Ex:</u> $\omega|_{\vec{x}} = \sum_{I = \{i_1, \dots, i_k\}} a_{i_1, i_2, \dots, i_k}(\vec{x})\, dx_I$ where $dx_I = dx_{i_1} \wedge \dots \wedge dx_{i_k}$

an increasing k-index; $\mathcal{C}^r$-functions

We write $\omega(\vec{x}) = \sum_I a_I(\vec{x})\, dx_I \in \Omega^k(\mathbb{R}^n)$ the set of k-forms on $\mathbb{R}^n$

usually suppress writing $\vec{x}$, call it ω

<u>Ex:</u> 0-forms in $\mathbb{R}^3 \longleftrightarrow$ functions $\mathbb{R}^3 \to \mathbb{R}$

<u>Ex:</u> 1-forms in $\mathbb{R}^3 \longleftrightarrow a_1(\vec{x})dx_1 + a_2(\vec{x})dx_2 + a_3(\vec{x})dx_3$

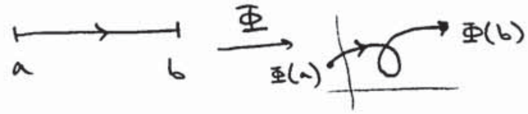
A k-form ω can be integrated over a k-surface Φ

$$\int_\Phi \omega := \int_D \sum_I a_I(\Phi(\vec{u}))\, \underbrace{\frac{\partial(\Phi_{i_1}, \dots, \Phi_{i_k})}{\partial(u_1, \dots, u_k)}}_{\text{Jacobian of } \Phi_I} du_1 \dots du_k$$

usual Riemann integration

$= \det$ (k×k matrix of partial derivatives evaluated at $\vec{u}$)

Ex: Φ path $[a,b] \to \mathbb{R}^2$.



$\omega = f(x,y)\,dx + g(x,y)\,dy$ is a general 1-form so

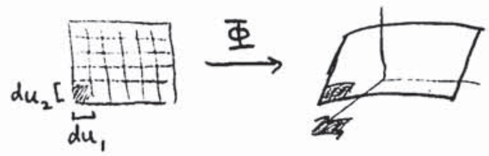
$$\int_{u=a}^{b} f(\underset{x(u)}{\Phi_1}, \underset{y(u)}{\Phi_2}) \underset{\frac{\partial x}{\partial u}}{\frac{\partial \Phi_1}{\partial u}} + g(x(u), y(u)) \frac{\partial y}{\partial u}\, du = \int_{\Phi} (f,g) \cdot d\vec{s} \quad \text{(line integral)}$$

Ex: $-y\,dx + x\,dy$ is the length form on unit circle in $\mathbb{R}^2$

Ex: $\eta = \frac{1}{2\pi}\left(\frac{-y}{x^2+y^2}\,dx + \frac{x}{x^2+y^2}\,dy\right)$ $\quad \int_{\Phi} \eta =$ winding number of path Φ around $(0,0)$ (counterclockwise)

Ex: $\int_{\text{path } \Phi} dx = \int_{t=a}^{b} \frac{\partial x}{\partial t}\, dt =$ (signed) length of projection of path on x-axis

Ex: $\omega = dx \wedge dy$ in $\mathbb{R}^3$, $\int_{\Phi} \omega = \int_D \frac{\partial(x,y)}{\partial(u_1,u_2)}\, du_1\, du_2$



= area of (signed) projection on xy-plane

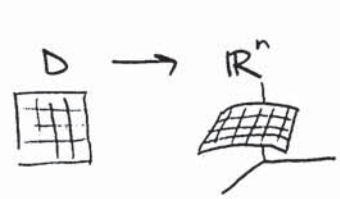
The Exterior Derivative

Recall: A differential form ω can be integrated - encodes geometric information! Write form in $\mathbb{R}^n$, in terms of basic forms $dx_1, dx_2, \ldots, dx_n$, basic k-forms $\underbrace{dx_{i_1} \wedge \ldots \wedge dx_{i_k}}_{dx_I}$, $\underbrace{i_1 < i_2 < \ldots < i_k}_{\text{index set } I \text{ chosen from } 1,\ldots,n}$.

General form: $\omega = \sum_I a_I(x)\, dx_I$

Ex: 0-form is a function $f(x)$ from $\mathbb{R}^n \to \mathbb{R}$.

Ex: $\int_C F_1\,dx + F_2\,dy = \int_C \vec{F} \cdot d\vec{s}$ $\quad (F_1, F_2)$



Recall: $\int_{\Phi} \omega := \iint \ldots \int_D \sum a_I(\Phi(\vec{u})) \frac{\partial(\Phi_{i_1}, \ldots, \Phi_{i_k})}{\partial(u_1, \ldots, u_k)}\, du_1\, du_2 \ldots du_k$

k-form; k-surface; Jacobian

Properties: $\int_{\Phi} \omega + \gamma = \int_{\Phi} \omega + \int_{\Phi} \gamma$, $\quad c\int_{\Phi} \omega = \int_{\Phi} c\omega$,

$\exists$ one basic n-form in $\mathbb{R}^n$: $dx_1 \wedge \ldots \wedge dx_n$ (volume form)

no k-forms in $\mathbb{R}^n$ for $k > n$

Wedge product of forms:

Basic forms: $\underbrace{dx_I}_{k\text{-form}} \wedge \underbrace{dx_J}_{\ell\text{-form}} = (dx_{i_1} \wedge dx_{i_2} \wedge \dots \wedge dx_{i_k}) \wedge (dx_{j_1} \wedge \dots \wedge dx_{j_\ell})$

$$= \begin{cases} 0 & \text{, if any index repeats} \\ (-1)^{\alpha} dx_{[I,J]} & \text{, else} \end{cases}$$

α = # differences $j_t - i_s$ that are neg.

$dx_{[I,J]}$: $(k+\ell)$-form; $[I,J]$: $I \cup J$, increasing order

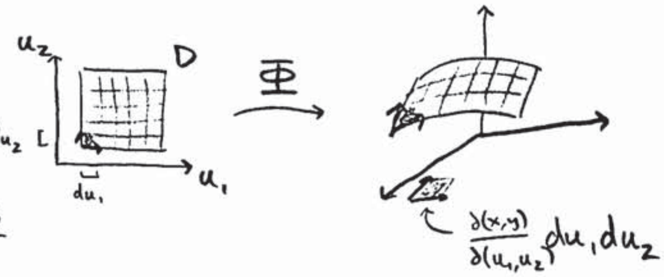
General forms: Say $\omega = \sum b_I dx_I$, $\lambda = \sum c_J dx_J$ are k & ℓ forms, then

$$\omega \wedge \lambda := \sum_{I,J} b_I c_J \, dx_I \wedge dx_J$$

Note if f is a 0-form, then $f \wedge \omega = f\omega$.

Recall: In $\mathbb{R}^3$, $\int_{\Phi} dx \wedge dy = \int_D \frac{\partial(x,y)}{\partial(u_1,u_2)} du_1 du_2$

= area of projection of Φ onto x-y plane



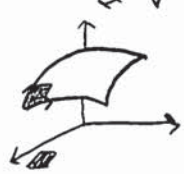
Alternate view of forms: A k-form ω specifies at each point p of a k-surface an alternating k-tensor $\omega|_p$

Ex: determinant $\omega(v_1, \dots, v_k) = -\omega(v_2, v_1, v_3, \dots, v_k)$

So in this view, a k-form locally "eats" k vectors, spits out number and integration chops up a surface by coordinates, the form ω produces a number, and the numbers get summed.

Ex: $dx_i(\vec{v}) = v_i$ the i^{th} coordinate of v

Ex: $dx \wedge dy(\vec{v}, \vec{w})$ = area of projection of $\vec{v}, \vec{w}$ onto xy-plane (signed)



The Exterior Derivative

There's an operator $d : \underbrace{\Omega^k(\mathbb{R}^n)}_{k\text{-forms}} \to \underbrace{\Omega^{k+1}(\mathbb{R}^n)}_{(k+1)\text{-forms}}$ defined by

$$df := \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} dx_i \quad \text{for 0-forms}$$

and if $\omega = \sum_I b_I dx_I$, then

$$d\omega := \sum_I db_I \wedge dx_I$$

Ex: If $\omega = f(x,y,z)$ in $\mathbb{R}^3$, a 0-form, then

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz = \text{the "gradient" form}$$
$$= \nabla f \cdot (dx, dy, dz)$$

Note: $df(\vec{v}) = df(v_1, v_2, v_3) = \frac{\partial f}{\partial x}v_1 + \frac{\partial f}{\partial y}v_2 + \frac{\partial f}{\partial z}v_3 = \nabla f \cdot \vec{v}$ (directional derivative)

Ex: Let $\omega = f_1 dx + f_2 dy + f_3 dz$, a 1-form, then

$$\begin{aligned} d\omega &= df_1 \wedge dx + df_2 \wedge dy + df_3 \wedge dz \\ &= \left(\frac{\partial f_1}{\partial x}dx + \frac{\partial f_1}{\partial y}dy + \frac{\partial f_1}{\partial z}dz\right) \wedge dx + \dots \\ &= \left(\frac{\partial f_3}{\partial y} - \frac{\partial f_2}{\partial z}\right) dy \wedge dz - \left(\frac{\partial f_1}{\partial z} - \frac{\partial f_3}{\partial x}\right) dx \wedge dz + \left(\frac{\partial f_2}{\partial x} - \frac{\partial f_1}{\partial y}\right) dx \wedge dy \\ &= \operatorname{curl} \underset{(f_1, f_2, f_3)}{\vec{F}} \cdot d\vec{S} \end{aligned}$$

Ex: Check if ω is 2-form: $\omega = f_1 dy \wedge dz + f_2 dz \wedge dx + f_3 dx \wedge dy$, then $d\omega$ gives "divergence" form $\operatorname{div} \underset{(f_1,f_2,f_3)}{\vec{F}}\, dV$ ← volume form $dx \wedge dy \wedge dz$

We also see that $d(d\omega) = 0$.

Stokes' Theorem

Recall: The exterior derivative $d: \Omega^k(\mathbb{R}^n) \to \Omega^{k+1}(\mathbb{R}^n)$ defined as

$$df := \sum \frac{\partial f}{\partial x_i} dx_i \quad \text{for function } f$$
$$d\omega := \sum db_I \wedge dx_I \quad \text{for } \omega = \sum_I b_I dx_I$$

If ω is a k-form and λ an l-form then, check

- $d(\omega \wedge \lambda) = d\omega \wedge \lambda + (-1)^k \omega \wedge d\lambda$
- $d(d\omega) = 0$

Def: Given a k-form ω on V and $T: U \to V$ there's a natural k-form on U called the <u>pullback</u> of ω, denoted $T^*\omega$. For 0-forms, $T^*f = f \circ T$ is a 0-form on U. In general for $\omega = \sum_I b_I dx_I$ & $T = (T_1, T_2, \dots, T_n)$, define

$$T^*\omega\Big|_{\vec{x}} = \sum_I b_I(T(\vec{x}))\, dT_{i_1} \wedge \dots \wedge dT_{i_k}$$

Properties of Pullbacks

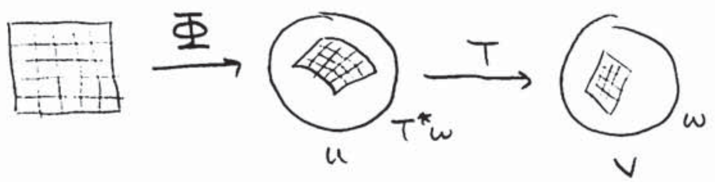
- commutes with $+, \wedge, d$ [need ω $\mathcal{C}^1$, T $\mathcal{C}^2$]

Check: $T^*(df) = d(T^*f)$

Compute $d(T^*f) = d(f(T(\vec{x}))) = \sum_i \frac{\partial(f\circ T)}{\partial x_i} dx_i = \sum_i \left[\sum_j \frac{\partial f}{\partial t_j}(T(\vec{x})) \cdot \frac{\partial T_j}{\partial x_i}(\vec{x}) \right] dx_i$

Compare $T^*(df) = T^*\left(\sum \frac{\partial f}{\partial t_j} dt_j\right) = \sum_j \frac{\partial f}{\partial t_j}(T(\vec{x}))\, dT_j = \sum_j \sum_i \frac{\partial f}{\partial t_j}(T(\vec{x})) \frac{\partial T_j}{\partial x_i} dx_i$

Now check for arbitrary $\omega = \sum b_I dx_I$ ...



$$\int_{\Phi} T^*\omega = \int_{T\circ\Phi} \omega \quad \text{"change of variable"}$$

Thm: (Stokes' Thm) If Ψ is a $\mathcal{C}^2$ oriented k-surface in open $V \in \mathbb{R}^m$ and ω is a $\mathcal{C}^1$ $(k-1)$-form on V then

$$\int_{\Psi} d\omega = \int_{\partial\Psi} \omega \quad \text{the "boundary" of } \Psi$$

Proof idea: Verify for standard simplex

$$\int_{T\sigma} d\omega = \int_{\sigma} T^*(d\omega) = \int_{\sigma} d(T^*\omega) \overset{\text{use FTC}}{=} \int_{\partial\sigma} T^*\omega = \int_{T(\partial\sigma)} \omega = \int_{\partial(T\sigma)} \omega$$

Def: If ω is a form such that $d\omega = 0$, call it a closed form.

$$\int_{\partial\Phi} \omega = \int_{\Phi} d\omega = 0$$

Def: If $\omega = d\lambda$ for some λ, call ω an exact form.

Exact forms are always closed, not necessarily vice versa.

Measure Theory

Def: A collection $\mathcal{M}$ of subsets of X is called a <u>σ-algebra</u> if $\mathcal{M}$ satisfies

① $X \in \mathcal{M}$
② $A \in \mathcal{M} \Rightarrow A^c \in \mathcal{M}$
③ $A_n \in \mathcal{M} \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{M}$

Consequently, $\emptyset \in \mathcal{M}$ and $\mathcal{M}$ is closed under countable intersections.

Ex: $\mathcal{M} = \{\emptyset, X\}$

Ex: $\mathcal{M} = 2^X$ = all subsets of X

Ex: Borel σ-algebra: generated by open sets and the three properties

Call $(X, \mathcal{M})$ a measurable space and the elements of $\mathcal{M}$ measurable sets.

Def: Say $f: X \to Y$ is a map of topological spaces (not necessarily continuous). Call f a <u>measurable function</u> if $\forall$ open sets $V \in Y$, the set $f^{-1}(V)$ is measurable.

If g is continuous and f is measurable, then $g \circ f$ is measurable.

$\{f_n\}$ measurable $\Rightarrow \sup f_n$, $|f_n|$, $\limsup f_n$, $\max\{f_1, f_2\}$, $f_1 + f_2$, $f_1 f_2$ measurable.

Def: A <u>measure</u> μ is a function $\mu: \mathcal{M} \to [0, \infty]$ that is countably additive (if $\{A_n\}$ is countable, disjoint sets in $\mathcal{M}$, then $\mu(\bigcup_{i=1}^{\infty} A_n) = \sum_{i=1}^{\infty} \mu(A_n)$). A <u>measure space</u>: $(X, \mathcal{M}, \mu)$.

Ex: zero measure: $\mu(E) = 0$, $\forall E \in \mathcal{M}$

Ex: counting measure: $\mu(E) = \begin{cases} \#\text{pts in } E & \text{, if finite} \\ \infty & \text{, else} \end{cases}$, $\mathcal{M} = 2^X$

Ex: Dirac measure: $\mathcal{M} = 2^X$, fix $x_0 \in X$, $\mu(E) = \begin{cases} 1, & \text{if } x_0 \in E \\ 0, & \text{else} \end{cases}$

Ex: probability measure: $\mathcal{M} = \{\text{measurable events}\}$, $\mu(E) = \text{prob}(E)$

Ex: Lebesgue measure: $X = \mathbb{R}^n$, $\mathcal{M} = ?$ with $\mu(E) =$ "volume" of E in $\mathbb{R}^n$. We demand $\mu(\text{box}) =$ product of side lengths. The "volume" idea can be extended to Borel sets and in fact to larger σ-algebra called the Lebesgue-measurable sets. Surprisingly, this is not true for all sets (Banach-Tarski paradox).

Def: A simple function is a function with a finite number of points in its range.

If $s: X \to [0, \infty)$ is a measurable simple function $s(x) = \sum a_i I_{A_i}(x)$ where $I_{A_i}(x) = \begin{cases} 1, & x \in A_i \\ 0, & \text{else} \end{cases}$, for $E \in \mathcal{M}$ define

$$\int_E s \, d\mu := \sum_{i=1}^{k} a_i \mu(E \cap A_i)$$

Given a measurable function $f: X \to [0, \infty)$, $\exists$ simple functions $\{s_n(x)\}$ s.t. $0 \le s_1 \le s_2 \le \dots \le f$ and $s_n(x) \to f(x)$ pointwise. Define

$$\int_E f \, d\mu := \sup_n \int s_n \, d\mu$$

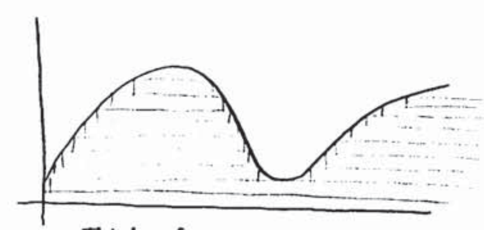
For general functions $f: X \to \mathbb{R}$, split f into f_+ and f_- such that $f = f_+ - (-f_-)$ and f_+ and $-f_-$ have nonnegative range. Then define

$$\int_E f \, d\mu := \int_E f_+ \, d\mu + \int_E f_- \, d\mu$$

Alternate definition for $f \ge 0$:

$$\int_{\mathbb{R}} f \, d\mu = \int_{t=0}^{\infty} \mu(x : f(x) > t) \, dt$$

Riemann integral



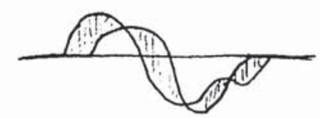
Think of it as horizontal integration.

Ex: Dirichlet function $f(x) = \begin{cases} 1, & x \text{ irrational} \\ 0, & \text{else} \end{cases}$ on $[0, 2]$

$$\int_{[0,2]} f \, d\mu = 2 \cdot 1$$ This is Lebesgue integrable but not Riemann integrable.

Let $\mathcal{C}_c(\mathbb{R})$ = continuous functions on $\mathbb{R}$ with compact support and has metric $d(f,g) = \int_{\mathbb{R}} |f-g|\,dx$. If we complete the space, we get a new space L^1 and $\mathcal{C}_c(\mathbb{R})$ is dense in L^1.

Riemann integration is uniformly continuous on $\mathcal{C}_c(\mathbb{R})$ so it can be extended to L^1. This is the Lebesgue integral w.r.t. the Lebegue measure.



Def: Call f and g equivalent or equal almost everywhere if they are equal except on a set of measure 0.

Thm: (Lebegue Monotone Convergence Thm) Let $E \in \mathcal{M}$, $\{f_n\}$ be a sequence of measurable functions on E. Suppose

① $0 \le f_1 \le f_2 \le \ldots < \infty$

② $f_n(x) \to f(x)$ as $n \to \infty$, $\forall x \in E$ (can be a.e. w.r.t. μ)

Then f is measurable and

$$\lim_{n\to\infty} \int_E f_n \, d\mu = \int_E f \, d\mu$$

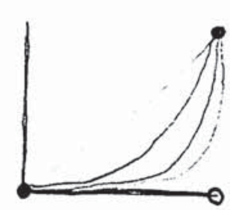
Thm: (Lebegue Dominated Convergence Thm) Let $E \in \mathcal{M}$, $\{f_n\}$ be a sequence of measurable functions on E. Suppose

① $f_n \to f$ pointwise

② $\exists g \in L^1(\mu)$ on E s.t. $|f_n(x)| \le g(x)$, $\forall x, n$

Then f is measurable and

$$\lim_{n\to\infty} \int_E f_n \, d\mu = \int_E f \, d\mu$$



Ex: $f_n(x) = x^n$ on $[0,1] \longrightarrow f(x) = \begin{cases} 1, & x=1 \\ 0, & \text{else} \end{cases}$

$f(x)$ is dominated by $g(x)=1$ so $\int_{[0,1]} f_n \, dx \to \int_{[0,1]} f \, dx$

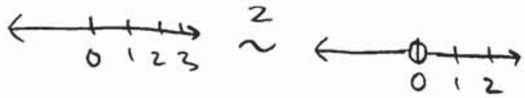
Non-Measurable Sets

We can't extend the idea of volume to all sets.

Banach-Tarski Paradox (1924): A solid ball in $\mathbb{R}^3$ can be partitioned into 5 pieces, that by rigid motions only, can reassemble into two solid balls, congruent to the original ball. We say this ball is <u>equidecomposable</u>.

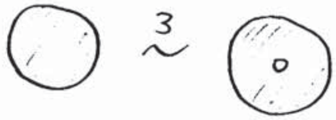
Another version says that a pea is equidecomposable into a sun.

<u>Ex:</u>

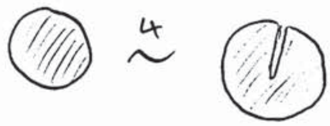


We do this by moving $k \to k+1$ for $k \in \mathbb{Z}$ and $k \geq 0$.

We do a similar operation and shift everything by 1. This works since the radius is irrational.



We first make a hole on the boundary and then shift it over to the center.



We do a similar operation on the radii.

Free group on 2 letters σ, τ, F_2: all words $\sigma, \tau, \sigma^{-1}, \tau^{-1}$ (Ex: $1, \sigma^2, \sigma\tau^{-5}\sigma^{-1}$). We claim F_2 is paradoxical using F_2 as the action, i.e. $F_2 \overset{4}{\sim} F_2 + F_2$

<u>Thm:</u> If G has a paradox and acts on X without fixed points then X has a paradox into G.