Avez DC ch 4

Higher Order Derivatives

12/23/2023

1

Ch 4.1 As before, consider $f: \underset{E}{\mathcal{U}} \longrightarrow F$ then $Df_a : E \to F$ Bdd Linear

That is to say $Df_a \in \mathcal{L}(E \to F)$ and $Df : \mathcal{U} \longrightarrow \mathcal{L}(E \to F)$, $x \longmapsto Df_x(\cdot)$

then we'd have $D^2 f_a \in \mathcal{L}(E \to \mathcal{L}(E \to F))$ and $\underbrace{\mathcal{L}(E \to \mathcal{L}(E \to F)) \cong \mathcal{L}(E \times E \to F)}_{\text{abbrev. } \mathcal{L}^2(E \to F)}$ cannonical iso p.151

Here is that cannonical iso: $A \in \mathcal{L}(E \to \mathcal{L}_{EF})$ that means $A : E \longrightarrow \mathcal{L}_{EF}$, $k \longmapsto A^k : E \to F$ ← param; $h \mapsto A^k(h)$ linear here, linear here

So we can just write $\hat{A}(h,k) := A^k(h)$

$\hat{A} : E \times E \longrightarrow F$, $(h,k) \longmapsto A^k(h)$

Procede inductively... for the next one $A \in \mathcal{L}(E \to \mathcal{L}(E \times E \to F))$ $A : \ell \longmapsto A^\ell(h,k) = \hat{A}(h,k,\ell)$

Back to 2^{nd} deriv: Fix $h \in E$. Define $F^h(x) := Df_x(h)$ (h ← param) note "F" is not v.s. F! etc...

We seek a linear map $A^h_x(\cdot)$ (h ← linear) $\ni$ $\dfrac{\| F^h(x+k) - F^h(x) - A^h_x(k) \|_{F}}{\|k\|_E} \longrightarrow 0$ as $\|k\| \to 0$

i.e. $\dfrac{\| Df_{x+k}(h) - Df_x(h) - A_x(h,k) \|_{F}}{\|k\|_E} \longrightarrow 0$

First we compute Df_x and give it a direction h: $Df_x(h) =: F^h(x)$

Then compute $D(F^h)_x$ and give it a direction k: $D(F^h)_x(k)$

this is $D^2 f_x(h,k)$

$A_x(h,k)$ is bilinear, we name it $\boxed{D^2 f_x(h,k)}$

$D^2 f$ being cont (or $f \in C^2$) means $\|x-y\| < \delta \Rightarrow \| D^2 f_x - D^2 f_y \| < \varepsilon$

or $\|h\| < \delta \Rightarrow \| D^2 f_{x+h} - D^2 f_x \|_{op} < \varepsilon$

$$\| Df_{x+k}(\cdot) - Df_x(\cdot) - A_x(\cdot, k) \|_{\mathcal{L}_{EF}} \qquad \| Q(\cdot) \|_{\mathcal{L}} = \sup_{\|v\|=1} \| Q(v) \|_F$$

Ch 4.1.2 Calculation of $D^2 f_a$

From ch 1 p. 1-2 $\left.\dfrac{d}{dt} f(x+th)\right|_{t=0} = \lim_{t \to 0} \dfrac{1}{t}\left(f(x+th) - f(x)\right) = \delta f_x(h) = Df_x(h)$

The values are the same, provided the RHS exists.

Now for $F^h(x) := Df_x(h)$

$$\delta(F^h)_x(k) = \left.\frac{d}{ds} F^h(x+sk)\right|_{s=0} = \left.\frac{d}{ds}\left(Df_{x+sk}(h)\right)\right|_{s=0}$$

$$= \left.\frac{d}{ds}\left(\left.\frac{d}{dt} f(x+th+sk)\right|_{t=0}\right)\right|_{s=0} = D^2 f_x(h,k)$$

order doesn't matter since we will soon show $D^2 f_a$ is sym.

See my sheets later for an example of how this is used. sheets associated with showing $J : u \to u^{-1}$ is C^∞.

example 1 f is linear (continuous i.e. Bdd Linear) $f: E \longrightarrow F$

$x \longmapsto Lx$ where linear map L does not depend on x.

Then, from ch 1, $Df_x = L$ and since there is no dependence on x,
$D^n f_x \equiv 0 \quad \forall\, n \geq 2$

example 2 f is bilinear $f: E_1 \times E_2 \longrightarrow F$

$(x_1, x_2) \longmapsto b(x_1, x_2) =: B(x)$ where $x = \langle x_1, x_2 \rangle$

Then from ch 1 $\underbrace{DB_x(h)}_{F^h(x)} = b(x_1, h_2) + b(h_1, x_2)$

To find $D^2 B_x$ compute:

$$F^h(x+k) - F^h(x) = b(x_1+k_1, h_2) + b(h_1, x_2+k_2) - [b(x_1,h_2) + b(h_1,x_2)]$$
$$= b(\cancel{x_1,h_2}) + b(k_1,h_2) + \cancel{b(h_1,x_2)} + b(h_1,k_2) - \cancel{b(x_1,h_2)} - \cancel{b(h_1,x_2)}$$
$$D^2B(h,k) = b(k_1,h_2) + b(h_1,k_2)$$

We can see this is linear in k. No dependence on x, so $D^3 B \equiv 0$, $D^4 B \equiv 0$, etc....
This shows bilinear b is C^∞ smooth.

Here is another way: $DB_x(h) = \underbrace{b(x_1, h_2)}_{P} + \underbrace{b(h_1, x_2)}_{Q}$

h is fixed, we have 2 linear maps in x: $P = P^{h_2} := b(\cdot, h_2)$
$Q = Q^{h_1} = b(h_1, \cdot)$

Then, just like for $x \longmapsto Lx$ and $DL_x(k) = Lk$, we have $DP_x(k) = Pk = b(k, h_2)$
$DQ_x(\ell) = Q\ell = b(h_1, \ell)$
and this is the same as above.

Now we could specialize to $E = \mathbb{R}^n$ $\quad b(x,y) = x^T A y$ for $n \times n$ matrix A

If now b is sym: $b(x,y) = b(y,x) \Rightarrow x^T A y = y^T A x$ ← This is a scalar, so it equals its own transpose
$= (y^T A x)^T = x^T A^T y$

$\Rightarrow A = A^T$ sym matrix

Now specialize to quadratic fcn: $y = x$ $\quad f(x) = b(x,x) = x^T A x$

$Df_x(h) = 2x^T A h$
$D^2 f_x(h,h) = b(h,h) + b(h,h) = 2b(h,h) = 2h^T A h$

Also in M&T VC ch 4.2 write-up sheets

so $D^2 f_x(\cdot,\cdot) = 2A$

Let's show this 2^{nd} deriv calculation for $f:\mathbb{R}^n\to\mathbb{R}$

[Avez doesn't do this until much later in chapter, and no details]

Recall: $f:\mathbb{R}^3\to\mathbb{R}$ why is $Df_x = [D_1 f \;\; D_2 f \;\; D_3 f]$?
$\qquad\qquad [f_x \;\; f_y \;\; f_z]$

f difb $\Rightarrow \exists\, A_x:\mathbb{R}^3\to\mathbb{R} \ni f(x+h)-f(x) = A_x h + o(h)$

Then for any basis vector $\hat{e}_j$ let $h = t\hat{e}_j$ $\quad \lim\limits_{t\to 0}\frac{1}{t}\left[f(x+h)-f(x)\right] = \lim\limits_{t\to 0}\frac{A_x(\cancel{t}\hat{e}_j)}{\cancel{t}} + \lim \cancel{\frac{o(t\hat{e}_j)}{t}}^{\,0}$

say $j=2$ $\hat{e}_2$ $\quad \lim\limits_{t\to 0}\frac{f(x,y+te_2,z)-f(x,y,z)}{t} = \left[a_1^{(1)}\; a_2^{(1)}\; a_3^{(1)}\right]\begin{bmatrix}0\\1\\0\end{bmatrix}$

$f_y(x,y,z) = a_2^{(1)}$

Thus we see $[D_1 f \;\; D_2 f \;\; D_3 f] = \left[a_1^{(1)}\; a_2^{(1)}\; a_3^{(1)}\right]$ and A is unique, so we know this is it.

▷ Now lets apply this to 2^{nd} deriv:

Fix $h=\begin{bmatrix}h^1\\h^2\\h^3\end{bmatrix}$ Let $F^h(x) := Df_x(h) = [f_x(x)\;\; f_y(x)\;\; f_z(x)]\begin{bmatrix}h^{(1)}\\h^{(2)}\\h^{(3)}\end{bmatrix} = \sum^3 h_i D_i f(x) \in \mathbb{R}$

new "A"

$$F^h:\mathbb{R}^3\longrightarrow\mathbb{R}, \quad x\longmapsto \sum h_i D_i f(x) \qquad \text{Thus } A_x^h:\mathbb{R}^3\longrightarrow\mathbb{R}$$

$$A_x^h(k) = \left[a_1^h(x)\;\; a_2^h(x)\;\; a_3^h(x)\right]\begin{bmatrix}k_1\\k_2\\k_3\end{bmatrix} = \left[D_1F^h(x)\;\; D_2F^h(x)\;\; D_3F^h(x)\right]\begin{bmatrix}k_1\\k_2\\k_3\end{bmatrix}$$

$$= \left[D_1\left(\sum h_i D_i f(x)\right)\;\; D_2\left(\sum h_i D_i f\right)\;\; D_3\left(\sum h_i D_i f\right)\right]\begin{bmatrix}k_1\\k_2\\k_3\end{bmatrix}$$

$$= \left[\sum^3 h_i D_1 D_i f \;\;\; \sum h_i D_2 D_i f \;\;\; \sum h_i D_3 D_i f\right]\begin{bmatrix}k_1\\k_2\\k_3\end{bmatrix}$$

$$= \sum^3 h_i\left[D_1 D_i f\;\; D_2 D_i f\;\; D_3 D_i f\right]\begin{bmatrix}k\end{bmatrix} = \left(h_1[D_{11}\;D_{21}\;D_{31}] + h_2[D_{12}\;D_{22}\;D_{32}] + h_3[D_{13}\;D_{23}\;D_{33}]\right)[k]$$

$$= [h_1\; h_2\; h_3]\begin{bmatrix}D_{11}f & D_{21}f & D_{31}f\\ D_{12}f & D_{22}f & D_{32}f\\ D_{13}f & D_{23}f & D_{33}f\end{bmatrix}\begin{bmatrix}k_1\\k_2\\k_3\end{bmatrix} = D^2f_x(h,k)$$

□

Did I accidentally end up with transpose of matrix of partials? Usually no problem when $D_{ij} = D_{ji}f$

If we had $f:\mathbb{R}^n\to\mathbb{R}^m$, then this calculation would apply to each component $f^{(i)}$ and we'd end up with a stack of m of these

Say $n=2$ $m=2$

$$D^2f_x(h,k) = \begin{bmatrix} [h_1\; h_2]\begin{bmatrix}D_{11}f^{(1)} & D_{21}f^{(1)}\\ D_{12}f^{(2)} & D_{22}f^{(1)}\end{bmatrix}\begin{bmatrix}k_1\\k_2\end{bmatrix} \\ \\ [h_1\; h_2]\begin{bmatrix}D_{11}f^{(2)} & D_{21}f^{(2)}\\ D_{12}f^{(2)} & D_{22}f^{(2)}\end{bmatrix}\begin{bmatrix}k_1\\k_2\end{bmatrix}\end{bmatrix}$$

Thm (Schwartz) $D^2 f$ is Symmetric $D^2 f_x$ exists $\Rightarrow D^2 f_x(h,k) = D^2 f_x(k,h)$
It is not nec for $f \in C^2$ (see remarks after pf)

This is a hard and tricky thm to prove. Avez just gives a fragment of the pf. L&S p. 188 give the whole argument, but some details are very confusing. Richard Tapia: 'Appendix E: Differentiation in Abstract Spaces' p. 19 gave the clearest argument (online document).

pf. Fix x. We want to show $\| D^2 f_x(h,k) - D^2 f_x(k,h) \| < \text{const} \cdot \varepsilon$ ★

The only thing we have is: Given $\varepsilon > 0 \ \exists \ \delta > 0 \ni |s| < \delta \Rightarrow | Df_{x+s}(\cdot) - Df_x(\cdot) - D^2 f_x(\cdot, s) | < \varepsilon |s|$ call this $\exists D^2$

To get to ★ we use the '2nd diff' $\Delta^2 f_x(h,k) = f(x+h+k) - f(x+h) - f(x+k) + f(x)$

Note that $\Delta^2 f$ is sym in h, k

★ would become $\left| \left(\Delta^2 f(h,k) - D^2 f_x(h,k) \right) - \left(\Delta^2 f(k,h) - D^2 f_x(k,h) \right) \right| \le \underbrace{| \Delta^2 - D^2 |^{h,k}}_{a} + \underbrace{| \Delta^2 - D^2 |^{k,h}}_{b}$

we must bound each of a and b. Let's start with a:

Step 1 Define trick fcn $G_h : [0,1] \to F$

$t \longmapsto f(x+th+k) - f(x+th)$

Observe that $| \Delta^2_{hk} - D^2_{hk} | = | G_h(1) - G_h(0) - D^2 f_x(h,k) | \le \underbrace{| G_h(1) - G_h(0) - G_h'(0) |}_{\text{Bnd 1}} + \underbrace{| G_h'(0) - D^2 f_x(h,k) |}_{\text{Bnd 2}}$ Subtr & add

where $G_h(1) = f(x+h+k) - f(x+h)$
$G_h(0) = f(x+k) - f(x)$

$D(G_h)_t = G_h'(t) = Df_{x+th+k}(h) - Df_{x+th}(h)$
$G_h'(0) = Df_{x+k}(h) - Df_x(h)$

Step 2 To work on the bounds Bnd 1 and Bnd 2, define a quantity:

$G'(t) - D^2 f_x(h,k) = \left[Df_{x+th+k}(h) - Df_{x+th}(h) - Df_x(h) + Df_x(h) + D^2 f_x(h, th - th + k) \right]$

$= \left[Df_{x+th+k}(h) - Df_x(h) - D^2 f_x(h, th+k) \right] - \left[Df_{x+th}(h) - Df_x(h) - D^2 f_x(h, th) \right]$

To bound this, use the $\exists D^2$ cond: $0 \le t \le 1$ $|th+k| \le |h| + |k| < \delta$ if $|h| < \frac{1}{2}\delta$ and $|k| < \frac{1}{2}\delta$

plug h also into the open slot in $\exists D^2$ and get

$| Df_{x+th+k}(h) - Df_x(h) - D^2 f_x(h, th+k) | < \varepsilon |th+k| |h| \le \varepsilon (|h| + |k|) |h|$

$| Df_{x+th}(h) - Df_x(h) - D^2 f_x(h, th) | < \varepsilon |th| |h| \le \varepsilon (|h| + |k|) |h|$

Thus $| G'(t) - D^2 f_x(h,k) | \le 2\varepsilon (|h| + |k|) |h|$

Step 3 step 2 immediately gives us Bnd 2: $| G_h'(0) - D^2 f_x(h,k) | \le 2\varepsilon (|h| + |k|) |h|$

We can also bound $| G'(t) - G'(0) | \le | G'(t) - D^2 f_x(h,k) | + | G'(0) - D^2 f_x(h,k) | \le 4\varepsilon (|h| + |k|) |h|$

This will be used for Bnd 1 with MVT ($\mathbb{R} \to$ nls)

Define $g(t) := G(t) - G'(0) t$ for $0 \le t \le 1$ $\quad Dg_t = G'(t) - G'(0)$

Then $g(1) = G(1) - G'(0)$ $\quad g(0) = G(0)$

MVT $\Rightarrow | g(1) - g(0) | \le \sup_{0 \le t \le 1} | Dg_t | (1 - 0) = \sup | G'(t) - G'(0) | \le 4\varepsilon (|h| + |k|) |h|$

$\Rightarrow | G(1) - G(0) - G'(0) | \le 4\varepsilon (|h| + |k|) |h|$

Step 5 Plugging in for Bnd 1 and Bnd 2 in step 1, we see $| \Delta^2 f_x(h,k) - D^2 f_x(h,k) | \le 6\varepsilon (|h| + |k|) |h|$
$\le$ ~~6ε since $\delta < 1$ or we can take that.~~

step 6 By repeating from step 1 and switching h and k, we establish b and thus ★ $\longrightarrow$ next sheet
~~$D^2 f_x(h,k) = D^2 f_x(k,h)$ for small h, k. Since ε can be arb small~~

Step 7 This is not correct and requires additional argument on next page $\longrightarrow$

Schwartz Thm D^2f_x is Sym (cont'd):

First note that when we do the analogous arg for ⓑ, we'd in fact get

$$\left| \Delta^2 f_x(k,h) - D^2 f_x(k,h) \right| < 6\varepsilon(|k|+|h|)|k| \leq 6\varepsilon(|k|+|h|)(|k|+|h|)$$

This later bnd applies to ⓐ and ⓑ

So really $\left| D^2f_x(h,k) - D^2f_x(k,h) \right| < 12\varepsilon(|h|+|k|)^2$ provided $|h|, |k| < \frac{1}{2}\delta_\varepsilon$

To prevent h,k from depending on δ, which depends on ε, fix 2 vectors $\bar{h}, \bar{k}$ (potentially large mag).

let $h := t\bar{h}$, $k := t\bar{k}$ $\quad |h| = |t\bar{h}| \overset{!}{<} \delta/2 \Rightarrow 0 \leq t \leq \min\left\{ \frac{1}{|\bar{h}|}\delta/2, \frac{1}{|\bar{k}|}\delta/2 \right\}$

Then we have $\left| D^2f_x(t\bar{h}, t\bar{k}) - D^2f_x(t\bar{k}, t\bar{h}) \right| < 12\varepsilon\left(|t\bar{h}| + |t\bar{k}| \right)^2 = 12t^2\varepsilon\underbrace{\left(|\bar{h}|+|\bar{k}| \right)^2}_{\text{no dependence on } \varepsilon}$

using bilinearity of D^2f_x

$$\cancel{t^2}\left| D^2f_x(\bar{h},\bar{k}) - D^2f_x(\bar{k},\bar{h}) \right| < \cancel{t^2}\, 12\varepsilon\left(|\bar{h}|+|\bar{k}|\right)^2$$

LHS holds for $\varepsilon > 0$ arb small, so we conclude $D^2f_x(\bar{h},\bar{k}) = D^2f_x(\bar{k},\bar{h})$ $\boxed{\text{QED}}$

▷ For what could have been a counter-example, I considered a $f: \mathbb{R}^2 \longrightarrow \mathbb{R}$ given below.
we know (cf M&T VC ch 2 sheet 15) f C^2 at pt $x \Rightarrow f_{xy}(x) = f_{yx}(x)$ $\quad x = \langle x, y \rangle$
But Rudin POMA prob 9.27 gives a pathalogical example where 2nd partials not cont at origin and $f_{xy}(0) \neq f_{yx}(0)$ $\left[0 = \langle 0,0 \rangle \right]$

Thus $D^2f_0(h,k) = \begin{bmatrix} k^1 & k^2 \end{bmatrix} \begin{bmatrix} f_{xx}(0) & f_{xy}(0) \\ f_{yx}(0) & f_{yy}(0) \end{bmatrix} \begin{bmatrix} h^1 \\ h^2 \end{bmatrix}$ would not be sym and thus contradicts thm.

The resolution is that, while the matrix of partials exists, it is not the Frechet deriv, which does not exist at 0.

$$f := \begin{cases} \dfrac{xy(x^2-y^2)}{x^2+y^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0) \end{cases}$$

This is C^∞ smooth for pts $x \neq 0$

[Note: f is invariant under the sym $(x,y) \xmapsto{S} (y,-x)$ and I discussed using this to help with the labor in my Math Symmetry sheets

$f(x) = f(S(x))$ $\quad Df_x = Df_{S(x)}S$ $\quad S = \begin{bmatrix} & 1 \\ -1 & \end{bmatrix}$

$D^2f_x = S^T D^2f_{S(x)} S$

ch 4.1.5 Generalizing Schwartz' Thm

Thm $D^n f_a$ is a sym n-linear map. i.e. Let σ be any perm of $\{1,2,\dots,n\}$

$$\boxed{D^n f_a(h_1,\dots,h_n) = D^n f_a(h_{\sigma(1)},\dots,h_{\sigma(n)})}$$

pf by induction – skipped

ch 4.1.6 Calculation of $D^n f_a$

If we know the n^{th} Frechet deriv exists, we can find it's value by generalizing the procedure mentioned before: $D^n f_a(h_1\dots h_n) = \frac{d}{dt_n}\cdots\frac{d}{dt_1} f\left(a+\sum^n t_i h_i\right)\Big|_{t_1=0,\ t_2=0\dots,\ t_n=0}$

Thm 4.3
$$\left.\begin{array}{l} \text{Gateaux } \delta^n f_x \text{ exists in open nbhd } \mathcal{U} \text{ of } a \\ \delta^n f_x \in \mathcal{L}^n(E\to F)\ \forall x\in\mathcal{U} \\ x \longmapsto \delta^n f_x \text{ is Cont} \end{array}\right\} \Rightarrow \begin{array}{l} f \text{ is class } C^n \\ \delta^n f_a = D^n f_a \end{array}$$

Pf – skipped

Ch 4.2 Rules for Calculation

I already showed linear and bilinear maps are C^∞ on sheet 1a

4.2.3. Leibniz bilinear prod rule: Consider some operation '$\star$' that is like a multiplication and bilinear in both its terms. see ch 1 sheet 7

$$f\star g : E \xrightarrow{\lambda} F_1\times F_2 \xrightarrow{b} G$$
$$x \longmapsto (f(x),g(x)) \longmapsto b(f(x),g(x))$$

In ch 1 sheet 7 $f\star g = b\circ\lambda$

$D(f\star g)_x = Db_{\lambda(x)} D\lambda_x$

$$\boxed{\text{Thm } f,g \text{ of class } C^k \Rightarrow f\star g \text{ of class } C^k}$$
(we can also replace 'C^k' by 'k-times difb': $C^{(k-1)+d}$)

Pf. This is essentially obvious because $f\star g = b(f,g)$ and bilinear b is C^∞

Say f,g are C^1

In ch 1 we found

$$D(f\star g)_x(h) = \underbrace{b}_{C^\infty}(\underbrace{Df_x(h)}_{C^0}, \underbrace{g(x)}_{C^1}) + b(f(x), Dg_x(h)) = Df_x(h)\star g + f(x)\star Dg_x(h)$$
$$= F^h_{(x)}\star g(x) + f(x)\star G^h(x)$$

The LHS ↑ is C^0 so $D(f\star g)$ is C^0

$\Rightarrow f\star g$ is C^1

We can continue in the same way:

$$D^2(f\star g)_x(h,k) = D\left(D(f\star g)_x(h)\right)_x(k) = D(F^h\star g + f\star G^h)$$
$$= b(\underbrace{DF^h(k)}_{C^0}, \underbrace{g}_{C^2}) + b(\underbrace{F^h}_{C^1}, \underbrace{Dg_x(k)}_{C^1}) + b(Df_x(k), G^h) + b(f, DG^h_x(k))$$

again all terms are at least C^0 so $D^2(f\star g)$ is C^0, so $f\star g$ is C^2

and so forth for higher derivatives... □

Ch 4.2.4 Higher order derivs of a product

avez just lists so results with limited discussion, so I shall be even more limited

consider first f, g scalar valued $f, g : E \longrightarrow K$ where $K = \mathbb{R}$ or $\mathbb{C}$

$$(fg)^{(n)} = \sum_{k=0}^{n} \binom{n}{k} f^{(n-k)} g^{(k)}$$

we can get a feel for this from

$(fg)' = f'g + fg'$

$(fg)'' = f''g + f'g' + f'g' + fg'' = \binom{2}{0}f''g + \binom{2}{1}f'g' + \binom{2}{2}fg''$

$$D^n(f \star g)_x(h_1 \dots h_n) = \sum_{k=0}^{n} \sum_{\text{'}\Omega\text{'}} b\left(D^k f_x(h_{i_1} \dots h_{i_k}),\ D^{n-k} g_x(h_{j_1} \dots h_{j_{n-k}})\right)$$

where 'Ω' means the sum over all partitions of $(h_1, \dots h_n)$ into 2 subsets with indices in ascending order: $i_1 < \dots < i_k$, $j_1 < \dots < j_{n-k}$

Define $\operatorname{Sym} A(h_1 \dots h_n) := \frac{1}{n!} \sum_{\text{all } \sigma} A(h_{\sigma(1)}, \dots h_{\sigma(n)})$ avg over all perms σ of $\{1, \dots, n\}$

So we could write $D^n(f \star g) = \operatorname{Sym}\left[\sum_{k=0}^{n} \binom{n}{k} b(D^{n-k} f, D^k g)\right]$

Ch 4.2.5 Derivs of Composite maps

Thm 4.4 $\left.\begin{array}{c} E \xrightarrow{g} F \xrightarrow{f} G \\ f, g \text{ class } C^k \end{array}\right\} \Rightarrow f \circ g$ is also class C^k

pf by induct – skipped

They give some horrible formulas, cumulating with

Avez has $g \circ f$

$$D^n(g \circ f) = \operatorname{Sym}\left[\sum_{q=1}^{n} \sum_{\Sigma r_j = n} \frac{n!}{r_1! \cdots r_q!} D^q(g \circ f)(D^{r_1} f, \dots, D^{r_q} f)\right]$$

I will carry out an extensive discussion of $D^2(f \circ g)_x$ →