(57) Rather than proving THM 1.8.1 now, let's note that

- parts 1,2,3,4 are straightforward from defins & limit laws
- parts 5,6,7 are done in book (in fact part 5 reduces componentwise to the $n=1$ special case of part 7)

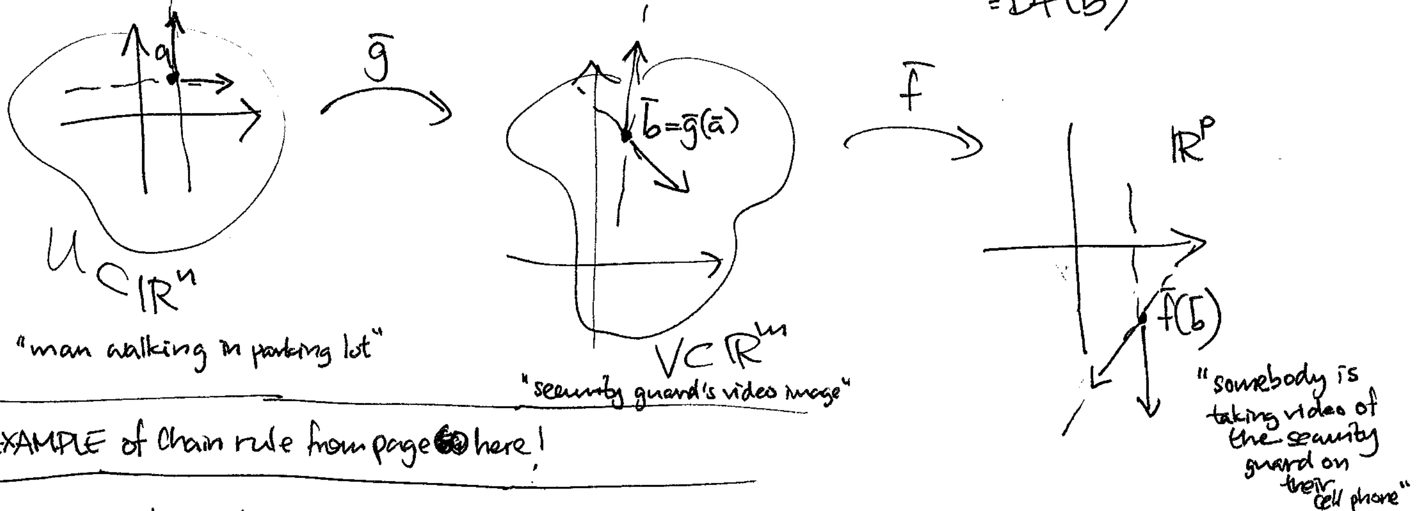
but we can deduce 5,6,7 from the <u>very</u> <u>important</u>...

THM 1.8.3 (Chain rule)

Assume
$$\begin{tikzcd} \overset{\text{open}}{\mathcal{U}} \arrow[r, "\bar{g}"] & \overset{\text{open}}{V} \arrow[r, "\bar{f}"] & \mathbb{R}^p \\ \mathbb{R}^n \arrow[u, phantom, "\cap"] & \mathbb{R}^m \arrow[u, phantom, "\cap"] & \end{tikzcd}$$

and $\bar{g}$ differentiable at $\underline{a}\in\mathcal{U}$, $\bar{f}$ differentiable at $\underline{b}=f(\underline{a})\in V$.

Then $\bar{f}\circ\bar{g}$ is differentiable at $\bar{a}$, with $[D(\bar{f}\circ\bar{g})(\bar{a})]=[\underbrace{D\bar{f}(\bar{g}(\bar{a}))}_{=D\bar{f}(\bar{b})}]\circ[D\bar{g}(\bar{a})]$



"man walking in parking lot"

"security guard's video image"

"somebody is taking video of the security guard on their cell phone"

Do EXAMPLE of chain rule from page [illegible] here!

Let's see how to prove part 7 from chain rule.

The function $\bar{h}(\bar{x})=\bar{f}(\bar{x})\bullet\bar{g}(\bar{x})$ (i.e. $\bar{h}=\bar{f}\bullet\bar{g}$ as in part 5)

can be thought of as a composition

$$\begin{tikzcd} \mathcal{U} \arrow[r, "\bar{k}=\begin{bmatrix}\bar{f}\\ \bar{g}\end{bmatrix}"] & \mathbb{R}^{2m} \arrow[r, "\ell"] & \mathbb{R} \\ \mathbb{R}^n \arrow[u, phantom, "\cap"] \quad x \arrow[r, mapsto] & \begin{pmatrix} f(\bar{x})\\ g(\bar{x})\end{pmatrix} & \\ & \begin{pmatrix}\bar{x}\\ \bar{y}\end{pmatrix} \arrow[r, mapsto] & \bar{x}\bullet\bar{y}=x_1y_1+\dots+x_my_m \end{tikzcd}$$

and we have $J\ell\begin{pmatrix}\bar{x}\\ \bar{y}\end{pmatrix}=\underbrace{[y_1\cdots y_m\ x_1\cdots x_m]}_{1\times 2m}$, $J\bar{k}(\bar{x})=\underbrace{\overbrace{\begin{bmatrix}J\bar{f}(\bar{x})\\ J\bar{g}(\bar{x})\end{bmatrix}}^{n}}_{2m\times n}\Big\}2m$

EXAMPLE of Chain rule:

$$h : \mathbb{R}^2 \to \mathbb{R}^2$$

$$\bar h\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}\sin(x^2y)\\ \cos(x^2y)\end{pmatrix} \text{ has } [J\bar h] = \begin{array}{c} \begin{array}{cc} D_1\sin(x^2y) & D_2\sin(x^2y) \\ \parallel & \parallel \end{array} \\ \begin{bmatrix} 2xy\cos(x^2y) & x^2\cos(x^2y) \\ -2xy\sin(x^2y) & -x^2\sin(x^2y) \end{bmatrix} \end{array}$$

but its also a composition $\bar h = \bar f \circ \bar g$:

$$\begin{tikzcd} \mathbb{R}^2 \arrow[r, "\bar g"] & \mathbb{R} \arrow[r, "\bar f"] & \mathbb{R}^2 \end{tikzcd}$$

$$\begin{pmatrix}x\\y\end{pmatrix} \longmapsto x^2y = \bar g\begin{pmatrix}x\\y\end{pmatrix}$$

$$x \longmapsto \begin{pmatrix}\sin(x)\\ \cos(x)\end{pmatrix} = \bar f(x)$$

Same!

and $[J\bar g] = [2xy \quad x^2]$

$$[J\bar f] = \begin{bmatrix}\cos(x)\\ -\sin(x)\end{bmatrix}$$

so

$$[J\bar h] = \underbrace{[J\bar f(\bar g(\bar x))][J\bar g(\bar x)]}_{\text{chain rule}} = \underbrace{\begin{bmatrix}\cos(x^2y)\\ -\sin(x^2y)\end{bmatrix}}_{2\times 1}\underbrace{[2xy \quad x^2]}_{1\times 2} = \underbrace{\begin{bmatrix} 2xy\cos(x^2y) & x^2\cos(x^2y) \\ -2xy\sin(x^2y) & -x^2\sin(x^2y)\end{bmatrix}}_{2\times 2}$$

Hence $J\bar{h}(\bar{a}) = J(l \circ \bar{k})(\bar{a}) = [Jl(\bar{k}(\bar{a}))][J\bar{k}(\bar{a})]$

$$= \cancel{\begin{bmatrix} g_1(x) \\ g_m(x) \\ f_1(x) \\ f_m(x) \end{bmatrix}} [g_1(\bar{a}) \cdots g_m(\bar{a})\ f_1(\bar{a}) \cdots f_m(\bar{a})] \begin{bmatrix} J\bar{f}(\bar{a}) \\ J\bar{g}(\bar{a}) \end{bmatrix}$$

$$[D\bar{h}(\bar{a})]\bar{v} = [\bar{g}(\bar{a})\ \bar{f}(\bar{a})] \begin{bmatrix} J\bar{f}(\bar{a}) \\ J\bar{g}(\bar{a}) \end{bmatrix} \bar{v} = [\bar{g}(\bar{a})\ \bar{f}(\bar{a})] \begin{bmatrix} [J\bar{f}(\bar{a})]\bar{v} \\ [J\bar{g}(\bar{a})]\bar{v} \end{bmatrix}$$

$$= \bar{g}(\bar{a}) \cdot [J\bar{f}(\bar{a})]\bar{v} + \bar{f}(\bar{a}) \cdot [J\bar{g}(\bar{a})]\bar{v}$$

$$= D\bar{f}(\bar{a})\bar{v} \cdot \bar{g}(\bar{a}) + \bar{f}(\bar{a}) \cdot D\bar{g}(\bar{a})\bar{v}$$

10/19/2016 >

proof of Chain rule:

We're assuming $\bar{r}(\bar{h}) := \bar{g}(\bar{a}+\bar{h}) - \bar{g}(\bar{a}) - [D\bar{g}(\bar{a})]\bar{h}$ has $\lim\limits_{\bar{h}\to\bar{0}} \frac{\bar{r}(\bar{h})}{|\bar{h}|} = \bar{0}$

and $\bar{s}(\bar{k}) := \bar{f}(g(\bar{a})+\bar{k}) - \bar{f}(\bar{g}(\bar{a})) - [D\bar{f}(g(\bar{a}))]\bar{k}$ has $\lim\limits_{\bar{k}\to\bar{0}} \frac{\bar{s}(\bar{k})}{|\bar{k}|} = \bar{0}$

So we analyze

$$\bar{f}(\bar{g}(\bar{a}+\bar{h})) \overset{\text{defn of } \bar{r}(\bar{h})}{=} \bar{f}(g(\bar{a}) + \underbrace{[D\bar{g}(\bar{a})]\bar{h} + \bar{r}(\bar{h})}_{\text{call this } \bar{k} :=})$$

$$= \bar{f}(\bar{g}(\bar{a}) + \bar{k})$$

$$\overset{\text{defn of } \bar{s}(\bar{k})}{=} \bar{f}(\bar{g}(\bar{a})) + [D\bar{f}(\bar{g}(\bar{a}))]\bar{k} + \bar{s}(\bar{k})$$

$$= \bar{f}(\bar{g}(\bar{a})) + \underbrace{[D\bar{f}(\bar{g}(\bar{a}))][D\bar{g}(\bar{a})]\bar{h}}_{\text{what we wanted to approximate } f(g(\bar{a}+\bar{h}))-\bar{f}(\bar{g}(\bar{a}))!} + \underbrace{[D\bar{f}(\bar{g}(\bar{a}))]\bar{r}(\bar{h}) + \bar{s}(\bar{k})}_{\text{error term?}}$$

Thus we need to show the error term has

$$\lim_{\bar{h}\to\bar{0}} \frac{[D\bar{f}(\bar{g}(\bar{a}))]\bar{r}(\bar{h}) + \bar{s}(\bar{k})}{|\bar{h}|} = \bar{0}$$

Easy to see $\lim\limits_{\bar{h}\to\bar{0}} \frac{[D\bar{f}(\bar{g}(\bar{a}))]\bar{r}(\bar{h})}{|\bar{h}|} = \bar{0}$ since $\left|[D\bar{f}(\bar{g}(\bar{a}))] \frac{\bar{r}(\bar{h})}{|\bar{h}|}\right| \le \underbrace{\overset{\text{matrix norm}}{|[D\bar{f}(\bar{g}(\bar{a}))]|}}_{\text{fixed!}} \underbrace{\frac{|\bar{r}(\bar{h})|}{|\bar{h}|}}_{\to 0 \text{ as } \bar{h}\to 0}$