7. $f, g : U \longrightarrow \mathbb{R}^m$ both differentiable at a, then the dot product function $f \cdot g : U \longrightarrow \mathbb{R}$ is also differentiable and

$$[D(f \cdot g)(a)](v) = Df(a)(v) \cdot g(a) + f(a) \cdot Dg(a)(v)$$

<u>The Chain rule!</u>

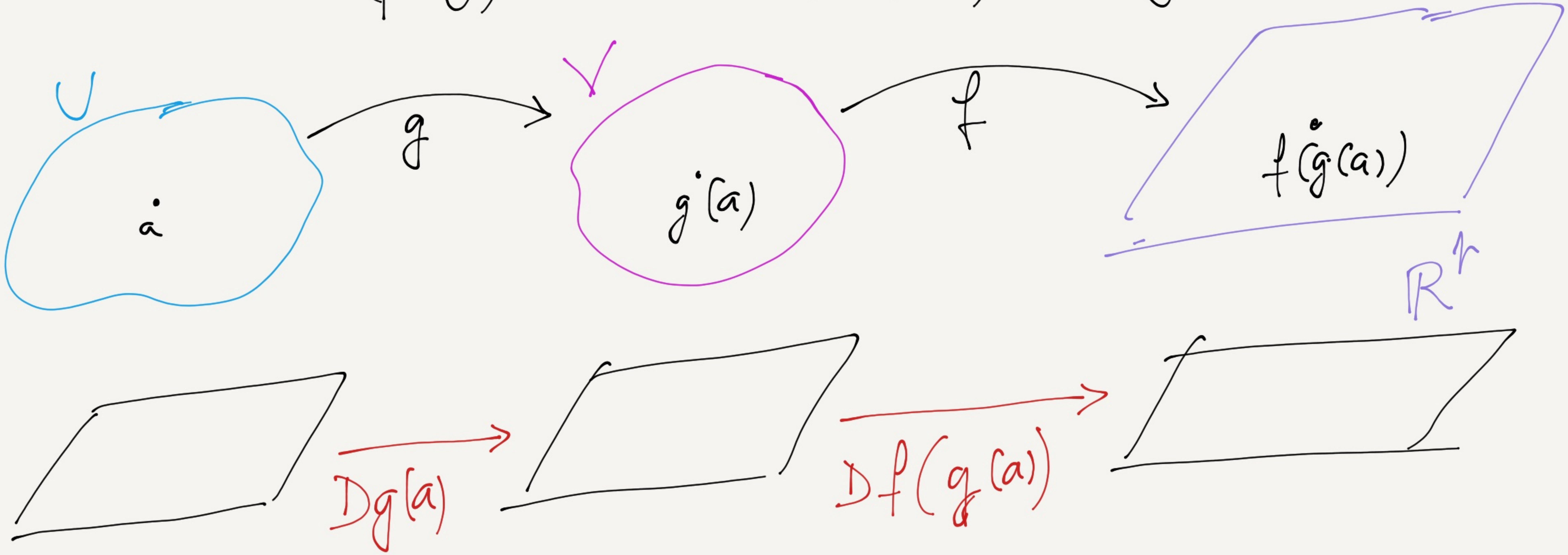
<u>Theorem 1.8.3</u>: $U \subset \mathbb{R}^n$, $V \subset \mathbb{R}^m$ open sets.

$g : U \longrightarrow V$, $f : V \longrightarrow \mathbb{R}^p$, $a \in U$.

If g is differentiable at a and f is differentiable at $g(a)$, then $f \circ g$ is differentiable at a and

$$[D(f \circ g)(a)](v) = [Df(g(a))]\big(Dg(a)(v)\big)$$

$$D(f\circ g)(a) = Df(g(a))\circ Dg(a)$$



$$\mathbb{R}^3 \xrightarrow{g} \mathbb{R}^2 \xrightarrow{f} \mathbb{R}$$

Example:

$$g\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x^2+y^2+z^2 \\ xyz \end{pmatrix} \qquad f(t,s) = st$$

$$a = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \qquad g(a) = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$$

$$Jg\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2x & 2y & 2z \\ yz & xz & xy \end{pmatrix}$$

$$Jg(a) = \begin{pmatrix} 2 & 2 & 2 \\ 1 & 1 & 1 \end{pmatrix}$$

$$Jf\begin{pmatrix} s \\ t \end{pmatrix} = \begin{pmatrix} t & s \end{pmatrix}$$

$$Jf(g(a)) = Jf\begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 3 \end{pmatrix}$$

$$J(f\circ g)(a) = Jf(g(a))\circ Jg(a) = \begin{pmatrix} 1 & 3 \end{pmatrix}\begin{pmatrix} 2 & 2 & 2 \\ 1 & 1 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 5 & 5 & 5 \end{pmatrix}$$

Proof of Theorem 1.9.8.

by Thm. 1.8.1 (3)

Since $f = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix} : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable iff $f_1, \dots, f_m$ are all differentiable, we can assume $f: \mathbb{R}^n \to \mathbb{R}$ is scalar valued.

We have to prove that, for all $a \in U$

$$\lim_{h \to 0} \frac{1}{|h|}\left(f(a+h) - f(a) - [Jf(a)]\, h\right) = 0$$

We only have information about the partials, so we should move only one variable at a time:

$$f\begin{pmatrix} a_1+h_1 \\ \vdots \\ a_n+h_n \end{pmatrix} - f\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = f\begin{pmatrix} a_1+h_1 \\ \vdots \\ a_n+h_n \end{pmatrix} - f\begin{pmatrix} a_1 \\ a_2+h_2 \\ \vdots \\ a_n+h_n \end{pmatrix}$$

$$+ f\begin{pmatrix} a_1 \\ a_2+h_2 \\ \vdots \\ a_n+h_n \end{pmatrix} - f\begin{pmatrix} a_1 \\ a_2 \\ a_3+h_3 \\ \vdots \\ a_n+h_n \end{pmatrix}$$

$$+ f\begin{pmatrix} a_1 \\ a_2 \\ a_3+h_3 \\ \vdots \\ a_n+h_n \end{pmatrix} - f\begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4+h_4 \\ \vdots \\ a_n+h_n \end{pmatrix}$$

$$\vdots$$

$$+ f\begin{pmatrix} a_1 \\ \vdots \\ a_{n-1} \\ a_n+h_n \end{pmatrix} - f\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

By the mean value theorem, in one variable (x_1 here) $\exists\, b_1 \in (a_1, a_1+h_1)$

s.t.
$$f\begin{pmatrix} a_1+h_1 \\ \vdots \\ a_n+h_n \end{pmatrix} - f\begin{pmatrix} a_1 \\ a_2+h_2 \\ \vdots \\ a_n+h_n \end{pmatrix} = h_1 \frac{\partial f}{\partial x_1}\begin{pmatrix} b_1 \\ a_2+h_2 \\ \vdots \\ a_n+h_n \end{pmatrix}$$

(works because h_i is very small, so $[a+h, a] \subset U$
and $[a+h_ie_i, a] \subset U$)

And similarly for the other variables.

So
$$f\begin{pmatrix} a_1+h_1 \\ \vdots \\ a_n+h_n \end{pmatrix} - f\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = h_1 \frac{\partial f}{\partial x_1}\begin{pmatrix} b_1 \\ a_2+h_2 \\ \vdots \\ a_n+h_n \end{pmatrix} + h_2 \frac{\partial f}{\partial x_2}\begin{pmatrix} a_1 \\ b_2 \\ a_3+h_3 \\ \vdots \\ a_n+h_n \end{pmatrix}$$
$$+ \cdots + h_n \frac{\partial f}{\partial x_n}\begin{pmatrix} a_1 \\ \vdots \\ a_{n-1} \\ b_n \end{pmatrix}$$

Now:

$$\lim_{h\to 0} \frac{1}{|h|}\left(\sum_{k=1}^{n} h_k \frac{\partial f}{\partial x_k}\begin{pmatrix} a_1 \\ \vdots \\ b_k \\ \vdots \\ a_n+h_n \end{pmatrix} - Jf\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}\begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix}\right)$$

$$Jf\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \left(\frac{\partial f}{\partial x_1}\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \quad \cdots \quad \frac{\partial f}{\partial x_n}\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}\right)$$

$$Jf\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}\begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} = \sum_{k=1}^{n} h_k \frac{\partial f}{\partial x_k}\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

$$\lim_{h\to 0} \frac{1}{|h|} \sum_{k=1}^{n} h_k \left[\frac{\partial f}{\partial x_k}\begin{pmatrix} a_1 \\ \vdots \\ b_k \\ \vdots \\ a_n+h_n \end{pmatrix} - \frac{\partial f}{\partial x_k}\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}\right]$$

The length is less than or equal to

$$\frac{1}{|h|}\sum_{k=1}^{n}|h_k|\left|\frac{\partial f}{\partial x_k}\begin{pmatrix}a_1\\ \vdots\\ b_k\\ \vdots\\ a_n+h_n\end{pmatrix}-\frac{\partial f}{\partial x_k}\begin{pmatrix}a_1\\ \vdots\\ \vdots\\ a_n\end{pmatrix}\right|$$

$$\leq \sum_{k=1}^{n}\left|\frac{\partial f}{\partial x_k}\begin{pmatrix}a_1\\ \vdots\\ b_k\\ \vdots\\ a_n+h_n\end{pmatrix}-\frac{\partial f}{\partial x_k}\begin{pmatrix}a_1\\ \vdots\\ \vdots\\ a_n\end{pmatrix}\right|$$

<u>Claim</u>: all the terms of the sum go to 0 as $h \to 0$

<u>e.g.</u>: $k=1$

$$\left|\frac{\partial f}{\partial x_1}\begin{pmatrix}b_1\\ a_2+h_2\\ \vdots\\ a_n+h_n\end{pmatrix}-\frac{\partial f}{\partial x_1}\begin{pmatrix}a_1\\ \vdots\\ \vdots\\ a_n\end{pmatrix}\right|$$

goes to 0 because $a_1 < b_1 < a_1+h_1$, so $\begin{pmatrix}b_1\\ a_2+h_2\\ \vdots\\ a_n+h_n\end{pmatrix} \longrightarrow \begin{pmatrix}a_1\\ \vdots\\ a_n\end{pmatrix}$

as the partials are continuous, this term goes to 0.

$\square$

Proof of: $\left(D\left(\frac{1}{f}\right)(a)\right)(v) = -\frac{(Df(a))(v)}{(f(a))^2}$

(part of Theorem 1.8.1)

$$\lim_{h\to 0}\frac{1}{|h|}\left(\frac{1}{f(a+h)}-\frac{1}{f(a)}+\frac{(Df(a))(h)}{(f(a))^2}\right)$$

$$=\lim_{h\to 0}\frac{1}{|h|}\left(\frac{f(a)-f(a+h)}{f(a)\,f(a+h)}+\frac{(Df(a))(h)}{f(a)^2}\right)$$

$$=\lim_{h\to 0}\frac{1}{|h|}\left[\frac{f(a)-f(a+h)+Df(a)(h)}{f(a)^2}-\frac{f(a)-f(a+h)}{f(a)^2}+\frac{f(a)-f(a+h)}{f(a)f(a+h)}\right]$$

First note $\lim_{h\to 0}\frac{1}{|h|}\left(f(a)-f(a+h)+Df(a)(h)\right)=0$ by the definition of $Df(a)$

So we are left with

$$\lim_{h\to 0} \frac{1}{f(a)}\left(\frac{f(a+h)-f(a)}{|h|}\right)\left(\frac{1}{f(a)}-\frac{1}{f(a+h)}\right)$$

We saw in homework (exercise 1.7.14) that when f is differentiable,
$\frac{f(a+h)-f(a)}{|h|}$ is bounded in a neighborhood of $h=0$.

Also, because f is continuous and $f(a)\neq 0$, $\frac{1}{f}$ is continuous at a.

So $\lim_{h\to 0}\left(\frac{1}{f(a)}-\frac{1}{f(a+h)}\right)=0$.

Hence the limit above is 0.

$\square$