10.2 Differentiable Functions.

A. The derivative.

1. <u>Motivation</u>. (a) Recall that a function $f: \mathbb{E}^1 \to \mathbb{E}^1$ is differentiable at x_0 in its domain, with derivative $f'(x_0)$ if

$$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h} = f'(x_0)$$

Rewriting this in terms of the definition of the limit gives: For every $\epsilon > 0$ there is a $\delta > 0$ such that if $|h| < \delta$ then

$$\left| \frac{f(x_0 + h) - f(x_0)}{h} - f'(x_0) \right| < \epsilon$$

or, rewriting again

$$|f(x_0 + h) - f(x_0) - f'(x_0)h| < \epsilon h$$

(b) If we define the linear transformation $A \in \mathcal{L}(\mathbb{E}^1, \mathbb{E}^1)$ by $A(h) = f'(x_0)h$ for all $h \in \mathbb{E}^1$, Then we can rewrite above as

$$|f(x_0 + h) - f(x_0) - A(h)| < \epsilon h$$

(c) Now clearly, for any linear transformation $A \in \mathcal{L}(\mathbb{E}^1, \mathbb{E}^1)$, or equivalently any number m, the quantity $|f(x_0 + h) - f(x_0) - A(h)| \to 0$ as $h \to 0$ (assuming f is continuous at x_0). However, the definition of differentiability says that in fact,

$$\frac{|f(x_0 + h) - f(x_0) - A(h)|}{h} \to 0$$

or in other words that $|f(x_0 + h) - f(x_0) - A(h)|$ goes to zero *faster than* h. There is only one transformation A that satisfies this criterion.

(d) We conclude that (i) the derivative $f'(x_0)$ can be thought of as a linear transformation, (ii) this linear transformation has the property that the difference between it and $f(x_0 + h) - f(x_0)$ goes to zero faster than h goes to zero, and (iii) it is the only linear transformation that does so.

2. Definition. Let $\mathbb{f}: D \to \mathbb{E}^m$ for some $D \subseteq \mathbb{E}^n$ and let $\mathbb{x} \in D$ be a cluster point of D. Then $\mathbb{f}$ is *differentiable* at $\mathbb{x}$ with derivative $\mathbb{f}'(\mathbb{x}) \in \mathcal{L}(\mathbb{E}^n, \mathbb{E}^m)$ if

$$\lim_{\mathbb{h} \to \mathbb{0}} \frac{||\mathbb{f}(\mathbb{x} + \mathbb{h}) - \mathbb{f}(\mathbb{x}) - \mathbb{f}'(\mathbb{x})(\mathbb{h})||}{||\mathbb{h}||} = 0$$

3. Theorem. (10.2.2) If $\mathbb{f}$ is differentiable at $\mathbb{x}_0 \in D$ then $\mathbb{f}$ is continuous at $\mathbb{x}_0$.

B. Computing $\mathbb{f}'(\mathbb{x})$.

1. Remark. (a) If $\mathbb{f}'(\mathbb{x}) \in \mathcal{L}(\mathbb{E}^n, \mathbb{E}^m)$ then it has a representation as a $m \times n$ matrix with respect to the standard basis. What is that matrix?

   (b) Consider first a function $\mathbb{f}\colon \mathbb{E}^n \to \mathbb{E}^1$, that is a real-valued function of n variables. Let us write $\mathbb{f}(\mathbb{x}) = f(x_1, x_2, \dots, x_n)$. In this case, for a given $\mathbb{x}_0 = (x_1^0, x_2^0, \dots, x_n^0)$, $\mathbb{f}'(\mathbb{x}_0)$ is a linear transformation from $\mathbb{E}^n$ to $\mathbb{E}^1$ and hence can be written as
   $$\mathbb{f}'(\mathbb{x}_0)\mathbb{h} = \mathbb{a} \cdot \mathbb{h}$$
   for $\mathbb{h} \in \mathbb{E}^n$. And we have
   $$\lim_{\mathbb{h} \to \mathbb{0}} \frac{\|\mathbb{f}(\mathbb{x}_0 + \mathbb{h}) - \mathbb{f}(\mathbb{x}_0) - \mathbb{a} \cdot \mathbb{h}\|}{\|\mathbb{h}\|} = 0$$

   (c) Since the limit exists, we can approach zero from any direction. By letting $\mathbb{h} = h\mathbb{e}_i$, we get $\mathbb{a} \cdot \mathbb{h} = a_i$, and writing the above limit in components we get
   $$\lim_{h \to 0} \frac{f\left(x_1^0, \dots, x_j^0 + h, \dots, x_n^0\right) - f(x_1^0, \dots, x_n^0)}{h} = a_i$$
   But this is just the usual definition of the partial derivative. So we conclude
   $$\mathbb{a} = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_n}\right) = \nabla f(\mathbb{x}_0)$$

2. <u>Theorem</u> (10.2.3) Let $\mathbb{f}: D \to \mathbb{E}^m,\ D$ an open subset of $\mathbb{E}^n$, be differentiable at $\mathbb{x} \in D$. Then the matrix of $\mathbb{f}'(\mathbb{x})$ with respect to the standard basis is given by

$$\mathbb{f}'(\mathbb{x}) = \left[\frac{\partial f_i}{\partial x_j}\right]_{m\times n}$$

Moreover, for any $\mathbb{v} \in \mathbb{E}^n$,

$$\mathbb{f}'(\mathbb{x})\mathbb{v} = \lim_{t\to 0}\frac{\mathbb{f}(\mathbb{x}+t\mathbb{v}) - \mathbb{f}(\mathbb{x})}{t}$$

which is defined as the *directional derivative of* $\mathbb{f}$ *in the direction* $\mathbb{v}$ *at* $\mathbb{x}$.

3. Remark. (a) Differentiability of $\mathbb{f}$ at $\mathbb{x}$ implies that all of the partial derivatives of $\mathbb{f}$ exist at $\mathbb{x}$. However, the existence of all partial derivatives at $\mathbb{x}$ does not guarantee that $\mathbb{f}$ is differentiable at $\mathbb{x}$. This is in contrast to the case of real-valued functions of a single variable.

   (b) However, if all of the partials of $\mathbb{f}$ are *continuous* then the story is different.

4. Theorem (10.2.3) Let $\mathbb{f}: D \to \mathbb{E}^m,\ D$ an open subset of $\mathbb{E}^n$. Then $\mathbb{f} \in C^1(D, \mathbb{E}^n)$, that is, considering $\mathbb{f}'$ is continuous as a function $\mathbb{f}': D \to \mathcal{L}(\mathbb{E}^m, \mathbb{E}^n)$ between two normed linear spaces, if and only if every partial derivative of $\mathbb{f}$ is continuous on D.