10.2 Differentiable Functions.

A. The derivative.

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

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

3. **Theorem.** (10.2.2) If \( f \) is differentiable at \( x_0 \in D \) then \( f \) is continuous at \( x_0 \).
B. Computing $f'(x)$.

1. **Remark.** (a) If $f'(x) \in \mathcal{L}(\mathbb{E}^n, \mathbb{E}^m)$ 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 $f: \mathbb{E}^n \to \mathbb{E}^1$, that is a real-valued function of $n$ variables. Let us write $f(x) = f(x_1, x_2, ..., x_n)$. In this case, for a given $x_0 = (x_1^0, x_2^0, ..., x_n^0)$, $f'(x_0)$ is a linear transformation from $\mathbb{E}^n$ to $\mathbb{E}^1$ and hence can be written as

$$f'(x_0)h = a \cdot h$$

for $h \in \mathbb{E}^n$. And we have

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

(c) Since the limit exists, we can approach zero from any direction. By letting $h = h e_i$, we get $a \cdot h = a_i$, and writing the above limit in components we get

$$\lim_{h \to 0} \frac{f(x_1^0, ..., x_j^0 + h, ..., x_n^0) - f(x_1^0, ..., x_n^0)}{h} = a_i$$

But this is just the usual definition of the partial derivative. So we conclude

$$a = \left( \frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, ..., \frac{\partial f}{\partial x_n} \right) = \nabla f(x_0)$$
2. **Theorem (10.2.3)** Let \( f: D \to \mathbb{E}^m, \) \( D \) an open subset of \( \mathbb{E}^n, \) be differentiable at \( x \in D. \) Then the matrix of \( f'(x) \) with respect to the standard basis is given by

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

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

\[
f'(x)v = \lim_{t \to 0} \frac{f(x + tv) - f(x)}{t}
\]

which is defined as the *directional derivative of \( f \) in the direction \( v \) at \( x. \)*
3. **Remark.** (a) Differentiability of \( f \) at \( x \) implies that all of the partial derivatives of \( f \) exist at \( x \). However, the existence of all partial derivatives at \( x \) does not guarantee that \( f \) is differentiable at \( x \). This is in contrast to the case of real-valued functions of a single variable.

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

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