The differential.

Let n be a positive integer.

Suppose $\mathbf{f}$ is functon whose domain is a subset of $\mathbf{R}^n$ and which has values in $\mathbf{R}^m$ for some positive integer m. For each $j = 1, \ldots, n$ the **partial derivative**

$$\partial_j \mathbf{f}$$

is, by definition, the set of ordered pairs $(\mathbf{a}, \mathbf{v})$ such that $\mathbf{a}$ is an interior point of the domain of $\mathbf{f}$ and

$$\mathbf{v} \;=\; \lim_{h \to 0} \frac{1}{h}\big[\mathbf{f}(\mathbf{a} + h\mathbf{e}_j) - \mathbf{f}(\mathbf{a})\big].$$

Note that for each $j = 1, \ldots, n$ the partial derivative $\partial_j \mathbf{f}$ is a function, possibly empty, with values in $\mathbf{R}^m$ whose domain is a subset of the domain of $\mathbf{f}$.

In case $n = 1$ we set

$$\mathbf{f}'(\mathbf{a}) \;=\; \partial_1 \mathbf{f}(\mathbf{a}).$$

We say **$\mathbf{f}$ is differentiable at $\mathbf{a}$** if $\mathbf{a}$ is in the domain of each of the partial derivatives $\partial_j \mathbf{f},\ j = 1, \ldots, n$ and if

$$\lim_{\mathbf{x} \to \mathbf{a}} \frac{1}{|\mathbf{x} - \mathbf{a}|}[\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{a}) - \sum_{j=1}^{n} (x_j - a_j)\partial_j \mathbf{f}(\mathbf{a})] \;=\; \mathbf{0}.$$

Note that if the j-th partial derivative of $\mathbf{f}$ exists at $\mathbf{a}$ if and only if the j-th partial derivative of each component $f_i,\ i = 1, \ldots, m$ exists at $\mathbf{a}$ in which case

$$\partial_j \mathbf{f}(\mathbf{a}) \;=\; \sum_{i=1}^{m} \partial_j f_i(\mathbf{a})\mathbf{e}_i.$$

Example. Define $\mathbf{f} : \mathbf{R}^2 \to \mathbf{R}^2$ by setting

$$\mathbf{f}(a, b) \;=\; (a^2 - b^2, 2ab) \qquad \text{whenever } (a, b) \in \mathbf{R}^{\mathbf{2}}.$$

We have

$$\partial_1 \mathbf{f}(a, b) \;=\; (2a, 2b), \qquad \partial_2 \mathbf{f}(a, b) \;=\; (-2b, 2a) \qquad \text{whenever } (a, b) \in \mathbf{R}^{\mathbf{2}}.$$

Linear functions.

Apart from the empty function and constant functions, the simplest kind of function carrying $\mathbf{R}^n$ into $\mathbf{R}^m$ is a *linear* function, which we now proceed to define. Suppose

$$\mathbf{l} : \mathbf{R}^n \to \mathbf{R}^m;$$

we say $\mathbf{l}$ is **linear** if

(1) $\mathbf{l}(c\mathbf{u}) = c\mathbf{l}(\mathbf{u})$ whenever $c \in \mathbf{R}$ and $\mathbf{u} \in \mathbf{R}^n$ and

(2) $\mathbf{l}(\mathbf{u} + \mathbf{v}) = \mathbf{l}(\mathbf{u}) + \mathbf{l}(\mathbf{v})$ whenever $\mathbf{u}, \mathbf{v} \in \mathbf{R}^n$.

If $n = 1$ then $\mathbf{l}$ is linear if and only if the graph of $\mathbf{l}$ is a line through $\mathbf{0}$ in $\mathbf{R} \times \mathbf{R}^m \equiv \mathbf{R}^{m+1}$. If $n = 2$ then $\mathbf{l}$ is linear if and only if the graph of $\mathbf{l}$ is a plane through $\mathbf{0}$ in $\mathbf{R}^2 \times \mathbf{R}^m \equiv \mathbf{R}^{m+2}$. Can you prove these assertions? At least in case $m = 1$? To succeed you will have to have a clear idea of what a line is and what a plane is.

Note that for any $\mathbf{x} \in \mathbf{R}^n$ we have

$$\mathbf{l}(\mathbf{x}) = \mathbf{l}(\sum_{j=1}^n x_j \mathbf{e}_j) = \sum_{j=1}^n \mathbf{l}(x_j \mathbf{e}_j) = \sum_{j=1}^n x_j \mathbf{l}(\mathbf{e}_j);$$

thus $\mathbf{l}$ is completely determined by its values on $\mathbf{e}_j,\ j = 1, \ldots, n$.

On the other hand, if $\mathbf{v}_j \in \mathbf{R}^m,\ j = 1, \ldots, n$, and if $\mathbf{l} : \mathbf{R}^n \to \mathbf{R}^m$ is defined by setting

$$\mathbf{l}(\mathbf{x}) = \sum_{j=1}^n x_j \mathbf{v}_j \qquad \text{for each } \mathbf{x} \in \mathbf{R}^n$$

then it is easy to see that $\mathbf{l}$ is linear.

It is easy to verify, under appropriate hypotheses about domains, that a scalar multiple of a linear function is a linear function; that the sum of linear functions is linear; and that the composition of linear functions is linear.

Differentiability. Suppose $\mathbf{f}$ is functon whose domain is a subset of $\mathbf{R}^n$ and which has values in $\mathbf{R}^m$ for some positive integer m. We say **f is differentiable at a** if $\mathbf{a}$ is an interior point of the domain of $\mathbf{f}$ and if there is $\mathbf{l}$ such that $\mathbf{l} : \mathbf{R}^n \to \mathbf{R}^m$, $\mathbf{l}$ is linear and

$$\lim_{\mathbf{x} \to \mathbf{a}} \frac{1}{|\mathbf{x} - \mathbf{a}|} \big[\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{a}) - \mathbf{l}(\mathbf{x} - \mathbf{a})\big] = \mathbf{0}. \tag{1}$$

Note that $\mathbf{l}$ is uniquely determined by (1) because it implies that

$$\mathbf{l}(\mathbf{e}_j) = \partial_j \mathbf{f}(\mathbf{a}) \qquad \text{for } j = 1, \ldots, n;$$

we call $\mathbf{l}$ the **differential of f at a**. We let

$$d\mathbf{f}$$

be the set of ordered pairs $(\mathbf{a}, \mathbf{l})$ such that $\mathbf{f}$ is differentiable at $\mathbf{a}$ and $\mathbf{l}$ is the differential of $\mathbf{f}$ at $\mathbf{a}$. Note that $d\mathbf{f}$, which we call the **differential of f**, is a function whose domain is a subset of the domain of $\mathbf{f}$ and whose range is a subset of the set of linear functions carrying $\mathbf{R}^n$ into $\mathbf{R}^m$. Note also that

$$d\mathbf{f}(\mathbf{a}) = \lim_{h \to 0} \frac{1}{h} [\mathbf{f}(\mathbf{a} + h\mathbf{v}) - \mathbf{f}(\mathbf{a})]$$

whenever $\mathbf{f}$ is differentiable at $\mathbf{a}$ and $\mathbf{v} \in \mathbf{R}^n$; we call this vector the **derivative of f at a in the direction v**.

Make sure you understand that if m and n are both 1 then this amounts to the definition of differentiability in one variable calculus. You may wonder why $\mathbf{l}$ is required to be linear. The answer is that everything works under this hypothesis and that it is naturally verified in situations where one wishes to apply multivariable calculus; in this regard, study the proof of the chain rule.

Here a simple and very useful sufficient condition for differentiability.

Theorem. Suppose

(1) $\mathbf{a}$ is an interior point of the domain of each of the partial derivatives of $\mathbf{f}$ and

(2) each of the partial derivatives of $\mathbf{f}$ is continuous at $\mathbf{a}$.

Then $\mathbf{f}$ is differentiable at $\mathbf{a}$.

Proof. It's in the book for the case $n = 2$ and $m = 1$ and its a straightforward matter to extend the proof given there to other m and n. $\square$

Example. Let $\mathbf{f}$ be as in the previous example. Note that the partial derivatives are continuous everywhere, so $\mathbf{f}$ is differentiable everywhere. Let's show directly from the definition that this is the case. Fix $\mathbf{a} = (a, b) \in \mathbf{R}^2$ and define $\mathbf{l} : \mathbf{R}^2 \to \mathbf{R}^2$ by setting

$$\mathbf{l}(u, v) \;=\; u\partial_1\mathbf{f}(a, b) + v\partial_2\mathbf{f}(a, b) \;=\; u(2a, 2b) + v(-2b, 2a) \;=\; (2au - 2bv, 2bu + 2av) \qquad \text{for } (a, b) \text{ in } \mathbf{R^2}.$$

Next set

$$\epsilon(x, y) \;=\; \mathbf{f}(x, y) - \mathbf{f}(a, b) - \mathbf{l}(\mathbf{x} - \mathbf{a}) \qquad \text{for } (a, b) \text{ in } \mathbf{R^2}.$$

Note that

$$\begin{aligned} \epsilon(x, y) \;&=\; (x^2 - y^2 - a^2 + b^2 - 2a(x - a) + 2b(y - b), 2xy - 2ab - 2b(x - a) - 2a(y - b)) \\ &=\; ((x - a)^2 - (y - b)^2, 2(x - a)(y - b)) \end{aligned}$$

for (a, b) in $\mathbf{R^2}$. To show $\mathbf{f}$ is differentiable at (a, b) is to show that

$$\lim_{\mathbf{x} - \mathbf{a}} \epsilon(\mathbf{x}) \;=\; \mathbf{0}.$$

But this is clearly the case as

$$|\epsilon(\mathbf{x})| \;\le\; |x - a|^2 + |y - b|^2 + 2|x - a||y - b| \;=\; |\mathbf{x} - \mathbf{a}|^2$$

whenever $\mathbf{x} = (x, y) \in \mathbf{R}^2$; we used the triangle inequality to obtain the inequality.

A very important fact about differentiation of vector functions is the following.

The Chain Rule. Suppose

(1) $\mathbf{f}$ is a vector function whose domain is a subset of $\mathbf{R}^n$, whose range is a subset of $\mathbf{R}^m$ and which is differentiable at $\mathbf{a}$;

(2) $\mathbf{g}$ is a vector function whose domain is a subset of $\mathbf{R}^m$, whose range is a subset of $\mathbf{R}^l$ and which is differentiable at $\mathbf{f}(\mathbf{a})$;

Then $\mathbf{g} \circ \mathbf{f}$ is differentiable at $\mathbf{a}$ and

$$d(\mathbf{g} \circ \mathbf{f})(\mathbf{a}) \;=\; d\mathbf{g}(\mathbf{f}(\mathbf{a})) \circ d\mathbf{f}(\mathbf{a}).$$

Proof. See any good book on several variable calculus. Let me know when you find one. $\square$

Remark. Note that the chain rule implies

$$\partial_j(\mathbf{g} \circ \mathbf{f})(\mathbf{a}) \;=\; \sum_{i=1}^{m} \partial_j f_i(\mathbf{a})\partial_i\mathbf{g}(\mathbf{f}(\mathbf{a})),\; j = 1, \ldots, n,$$

and

$$\partial_j(g_k \circ \mathbf{f})(\mathbf{a}) \;=\; \sum_{i=1}^{m} \partial_j f_i(\mathbf{a})\partial_i g_k(\mathbf{f}(\mathbf{a})),\; j = 1, \ldots, n,\; k = 1, \ldots, m.$$