

The "Differential of a function" →

102

DIFFERENTIATION

5-23. Give an example of a pair of functions f and g having finite derivatives in $(0, 1)$, such that

$$\lim_{x \rightarrow 0} \frac{f(x)}{g(x)} = 0,$$

but such that $\lim_{x \rightarrow 0} f'(x)/g'(x)$ does not exist, choosing g so that $g'(x)$ is never zero.

5-24. Prove the following theorem:

Let f and g be two functions having finite n th derivatives in (a, b) . For some interior point x_1 in (a, b) , assume that $f(x_1) = f'(x_1) = \dots = f^{(n-1)}(x_1) = 0$, and that $g(x_1) = g'(x_1) = \dots = g^{(n-1)}(x_1) = 0$, but that $g^{(n)}(x)$ is never zero in (a, b) . Show that

$$\lim_{x \rightarrow x_1} \frac{f(x)}{g(x)} = \frac{f^{(n)}(x_1)}{g^{(n)}(x_1)}.$$

NOTE. $f^{(n)}$ and $g^{(n)}$ are not assumed to be continuous at x_1 . [Hint: Let $F(x) = f(x) - (x - x_1)^{n-1}f^{(n)}(x_1)/(n-1)!$, define G similarly, and apply Theorem 5-15 to the functions F and G .]

5-25. Show that the formula in Taylor's theorem can also be written as follows:

$$f(x) = \sum_{k=0}^{n-1} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k + \frac{(x - x_0)(x - x_1)^{n-1}}{(n-1)!} f^{(n)}(x_1),$$

where x_1 is interior to the interval joining x and x_0 . Let $1 - \theta = (x - x_1)/(x - x_0)$. Show that $0 < \theta < 1$ and deduce the following form of the remainder term (due to Cauchy):

$$\frac{(1 - \theta)^{n-1}(x - x_0)^n}{(n-1)!} f^{(n)}[\theta x + (1 - \theta)x_0].$$

[Hint: Take $G(t) = g(t) = t$ in the proof of Theorem 5-15.]

Hobson, Ernest William. The Theory of Functions of a Real Variable & the Theory of Fourier Series

517, 5

H 6534 (2572)

Graves, Lawrence Murray. The Theory of Functions of a Real Variable

517, 5

G 7854 (2)

Goffman, Casper. Real Functions

517, 5

Taken from Hobson's The Theory of Functions of a Real Variable & the Theory of Fourier Series
by Hobson, Ernest William

CHAPTER 6

DIFFERENTIATION OF FUNCTIONS OF SEVERAL VARIABLES

6-1 Introduction. In Chapter 5 we considered derivatives of functions defined on subsets of the real line E_1 . We now wish to discuss differentiation of real-valued functions of several variables. Perhaps the simplest way to proceed is to reduce the discussion to the one-dimensional case by treating a function of several variables as a function of one variable at a time, holding the others fixed. This leads us to the concept of *partial derivative*, with which the reader is already somewhat familiar from his knowledge of elementary calculus.

If $\mathbf{x} = (x_1, \dots, x_n)$ is a point in E_n , and if $\mathbf{y} = (y_1, \dots, y_n)$ is another point all of whose coordinates except the k th are the same as those of \mathbf{x} , that is, $y_i = x_i$ if $i \neq k$ and $y_k \neq x_k$, then we can consider the limit

$$\lim_{y_k \rightarrow x_k} \frac{f(\mathbf{y}) - f(\mathbf{x})}{y_k - x_k}.$$

When this limit exists, it is called the partial derivative of f with respect to the k th coordinate and is denoted by $D_k f(\mathbf{x})$, or $f_k(\mathbf{x})$, or $\partial f(\mathbf{x})/\partial x_k$, or by a similar expression. We shall adhere to the notation $D_k f(\mathbf{x})$.

This process then yields, from a given function f , n further functions $D_1 f, D_2 f, \dots, D_n f$ defined at those points in E_n where the corresponding limits exist. One-sided and infinite partial derivatives could be defined as in the one-dimensional case, but we shall be interested only in finite derivatives which exist at interior points of certain open sets in E_n .

In generalizing a concept from E_1 to E_n , we seek to preserve what we consider to be the important properties in the one-dimensional case. For example, the existence of the derivative at x implies continuity at x in the one-dimensional case. Therefore it seems desirable to have a notion of derivative for functions of several variables which will imply continuity. Partial derivatives do *not* do this. A function of n variables can have partial derivatives at a point with respect to each of the variables and yet not be continuous at the point. Consider the following example of a function of two variables:

$$f(x, y) = \begin{cases} x + y, & \text{if } x = 0 \text{ or } y = 0, \\ 1, & \text{otherwise.} \end{cases}$$

The partial derivatives $D_1f(0, 0)$ and $D_2f(0, 0)$ both exist. In fact,

$$D_1f(0, 0) = \lim_{x \rightarrow 0} \frac{f(x, 0) - f(0, 0)}{x - 0} = \lim_{x \rightarrow 0} \frac{x}{x} = 1$$

and, similarly, $D_2f(0, 0) = 1$. On the other hand, it is clear that this function is not continuous at $(0, 0)$.

The existence of the partial derivatives with respect to each variable separately implies continuity in each variable separately; but, as we have just seen, this does not necessarily imply continuity in all the variables simultaneously. The difficulty with partial derivatives is that by their very definition we are forced to consider only one variable at a time. Partial derivatives give us the rate of change of a function in the direction of each coordinate axis. It is natural to seek a more general concept of derivative which does not restrict our considerations to the special directions of the coordinate axes, but which allows us to study the rate of change in an arbitrary direction. The *directional derivative* serves this purpose.

Before we introduce the directional derivative, we wish to remark that in this chapter we shall restrict ourselves to functions defined on *open sets* S in E_n , so that with each point \mathbf{x} in S there will be a neighborhood $N(\mathbf{x}) \subset S$. Every point \mathbf{y} in $N(\mathbf{x})$ can then be expressed in the form $\mathbf{y} = \mathbf{x} + \lambda \mathbf{u}$, where \mathbf{u} is a unit vector; that is to say, $\mathbf{u} \in E_n$ and $|\mathbf{u}| = 1$. The number λ has an absolute value not exceeding the radius of the sphere $N(\mathbf{x})$.

6-2 The directional derivative.

6-1 DEFINITION. Let f be a real-valued function defined on an open set S in E_n and assume $\mathbf{x} \in S$. Let \mathbf{u} be a unit vector in E_n . We define the *directional derivative* of f at \mathbf{x} in the direction \mathbf{u} to be the number

$$D_{\mathbf{u}}f(\mathbf{x}) = \lim_{\lambda \rightarrow 0} \frac{f(\mathbf{x} + \lambda \mathbf{u}) - f(\mathbf{x})}{\lambda}$$

whenever the limit exists.

Observe that in the one-dimensional case, this reduces to the definition of $f'(\mathbf{x})$ if we take $\mathbf{u} = 1$. Also, this definition includes the k th partial derivative as a special case when the unit vector \mathbf{u} is taken to be the k th unit coordinate vector \mathbf{u}_k (that is, the vector having all components zero except the k th, which has the value 1). We then write D_k instead of $D_{\mathbf{u}_k}$. Observe also that if we introduce $F(\lambda) = f(\mathbf{x} + \lambda \mathbf{u})$, we have $D_{\mathbf{u}}f(\mathbf{x}) = F'(0)$.

If a function f defined in E_n has a directional derivative in *every* direction \mathbf{u} at a point \mathbf{x} , then, in particular, all partial derivatives D_1f, \dots, D_nf exist at \mathbf{x} . The converse is not true, however. For example, the function f considered above, which has the value $x + y$ at (x, y) if $x = 0$ or $y = 0$ and

has the value 1 otherwise, has finite partials $D_1f(0, 0)$ and $D_2f(0, 0)$. Nevertheless, if we consider any *other* direction $\mathbf{u} = (a_1, a_2)$, $a_1 \neq 0$, $a_2 \neq 0$, we have

$$\frac{f(\lambda a_1, \lambda a_2) - f(0, 0)}{\lambda} = \frac{1}{\lambda}$$

and this does not tend to a finite limit as $\lambda \rightarrow 0$.

A rather surprising fact is that a function may have a finite directional derivative in every direction at some point but may fail to be continuous at that point. For example, consider the function of two variables defined by the formulas

$$f(x, y) = \begin{cases} xy^2/(x^2 + y^4), & \text{if } x \neq 0, \\ 0, & \text{if } x = 0. \end{cases}$$

Let $\mathbf{u} = (a_1, a_2)$ be an arbitrary unit vector in E_2 . Then we have

$$\frac{f(\lambda a_1, \lambda a_2) - f(0, 0)}{\lambda} = \frac{a_1 a_2^2}{a_1^2 + \lambda^2 a_2^4},$$

and hence $D_{\mathbf{u}}f(0, 0) = a_2^2/a_1$ if $a_1 \neq 0$. If $a_1 = 0$, we find $D_{\mathbf{u}}f(0, 0) = 0$. Therefore, $D_{\mathbf{u}}f(0, 0)$ is finite for all directions \mathbf{u} . On the other hand, the function f takes on the value $\frac{1}{2}$ at each point of the parabola $x = y^2$ (except at the origin), so that f is clearly not continuous at $(0, 0)$, since $f(0, 0) = 0$.

Thus we see that even the existence of *all* directional derivatives at a point fails to imply continuity at the point. For this reason directional, like partial derivatives, are a somewhat unsatisfactory extension of the one-dimensional concept of derivative. We now introduce a more suitable generalization which does imply continuity and, at the same time, permits us to extend the principal theorems of one-dimensional derivative theory to functions of several variables. The concept which seems to serve this purpose best is the notion of *differential*. We shall first discuss the one-dimensional case in detail before we define differentials in n -dimensions.

6-3 Differentials of functions of one real variable.

6-2 DEFINITION. Let f be a real-valued function defined on an open interval S in E_1 . Construct a new function g of two real variables as follows: For every point x in S such that $f'(x)$ exists (finite), and for every real number t , let

$$g(x; t) = f'(x)t.$$

The function g so defined is called the *differential* of f .

NOTE. We write $g(x; t)$ rather than $g(x, t)$ to place further emphasis on the different roles of x and t . The first point, x , must be a point where $f'(x)$ exists, whereas the second point, t , is an arbitrary point in E_1 . We sometimes say that $g(x; t)$ is the differential of f at x with increment t .

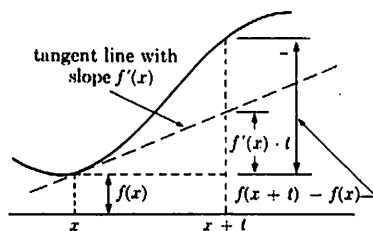


FIG. 6-1. Geometric interpretation of the differential in E_1 .

The differential can be given a geometric interpretation as indicated in Fig. 6-1. Observe in the figure that when t is "small," the difference $f(x+t) - f(x)$ and the differential $f'(x)t$ are nearly equal. This fact, which is of fundamental importance, is described precisely in the next theorem.

6-3 THEOREM. Let f have a finite derivative at x , and let $g(x; t) = f'(x)t$. Then for every $\epsilon > 0$, there exists a neighborhood $N(x)$ such that for every y in $N'(x)$ we have the inequality

$$|f(y) - f(x) - g(x; y - x)| < \epsilon |y - x|.$$

Proof. Given $\epsilon > 0$, there is a neighborhood $N(x)$ such that $y \in N'(x)$ implies

$$\left| \frac{f(y) - f(x)}{y - x} - f'(x) \right| < \epsilon.$$

Multiplication by $|y - x|$ gives the result.

The next property of differentials is an immediate consequence of the definition.

6-4 THEOREM. If $f'(x)$ exists and if $g(x; t) = f'(x)t$, then for all real numbers t, t', α , and α' , we have

$$g(x; \alpha t + \alpha' t') = \alpha g(x; t) + \alpha' g(x; t').$$

Proof. $g(x; \alpha t + \alpha' t') = f'(x)(\alpha t + \alpha' t') = \alpha[f'(x)t] + \alpha'[f'(x)t']$.

NOTE. We describe the property just proved by saying that g is linear in the second variable.

6-4 Differentials of functions of several variables. Instead of defining a differential of a function of several variables by a formula, as we did in E_1 , we choose to define a differential in E_n by the properties we wish it to possess. The properties we want are the natural extensions of those embodied in Theorems 6-3 and 6-4.

6-5 DEFINITION. Let f be a real-valued function defined on an open set S in E_n , and assume $\mathbf{x} \in S$. We say that f has a differential at \mathbf{x} if there exists another function g which satisfies the following conditions:

- (a) g is a real-valued function of two n -dimensional variables, and the function values, denoted by $g(\mathbf{x}; t)$, are defined for the given point \mathbf{x} in S and for every point t in E_n .
- (b) g is linear in the second variable; that is, for every pair of points t and t' in E_n and for every pair of real numbers α and α' , we have

$$g(\mathbf{x}; \alpha t + \alpha' t') = \alpha g(\mathbf{x}; t) + \alpha' g(\mathbf{x}; t').$$

- (c) For every $\epsilon > 0$, there exists a neighborhood $N(\mathbf{x})$ such that $\mathbf{y} \in N'(\mathbf{x})$ implies

$$|f(\mathbf{y}) - f(\mathbf{x}) - g(\mathbf{x}; \mathbf{y} - \mathbf{x})| < \epsilon |\mathbf{y} - \mathbf{x}|.$$

NOTE. It is important to observe that we have no guarantee in advance that, for given f , any such function g exists. We shall prove later that when f is suitably restricted, there will be one and only one such function g .

Before we deal with the question of the existence of differentials, we shall prove some theorems on the assumption that a differential *does* exist. The first of these is a consequence of the linearity property and tells us that the value of a differential must be a linear combination of the components of the second variable.

6-6 THEOREM. Assume that f has a differential $g(\mathbf{x}; t)$ at \mathbf{x} , and write $t = (t_1, \dots, t_n)$. Then there exist n real numbers $a_1(\mathbf{x}), \dots, a_n(\mathbf{x})$ (depending on \mathbf{x} but independent of t) such that

$$g(\mathbf{x}; t) = \sum_{k=1}^n a_k(\mathbf{x}) t_k.$$

Proof. We can write $t = t_1 u_1 + \dots + t_n u_n$, where u_k is the k th unit

coordinate vector. From the linearity of g in the second variable, we then have

$$g(\mathbf{x}; \mathbf{t}) = g(\mathbf{x}; t_1 \mathbf{u}_1 + \cdots + t_n \mathbf{u}_n) = g(\mathbf{x}; \mathbf{u}_1) t_1 + \cdots + g(\mathbf{x}; \mathbf{u}_n) t_n.$$

This proves the theorem, if we take $a_k(\mathbf{x}) = g(\mathbf{x}; \mathbf{u}_k)$, $k = 1, 2, \dots, n$.

Theorem 6-6 will now be used to prove that if the differential exists at all, it is uniquely determined. In fact, we will show that the n numbers $a_1(\mathbf{x}), \dots, a_n(\mathbf{x})$ of Theorem 6-6 are simply the n partial derivatives $D_1 f(\mathbf{x}), \dots, D_n f(\mathbf{x})$.

6-7 THEOREM. (*Uniqueness theorem*). Assume that f has a differential $g(\mathbf{x}; \mathbf{t})$ at \mathbf{x} and write

$$g(\mathbf{x}; \mathbf{t}) = \sum_{k=1}^n a_k(\mathbf{x}) t_k,$$

in accordance with Theorem 6-6. Then each partial derivative $D_k f(\mathbf{x})$ exists and we have

$$a_k(\mathbf{x}) = D_k f(\mathbf{x}), \quad k = 1, 2, \dots, n.$$

Proof. By hypothesis, for every $\epsilon > 0$ there is a neighborhood $N(\mathbf{x})$ such that $\mathbf{y} \in N'(\mathbf{x})$ implies $|f(\mathbf{y}) - f(\mathbf{x}) - g(\mathbf{x}; \mathbf{y} - \mathbf{x})| < \epsilon |\mathbf{y} - \mathbf{x}|$. By Theorem 6-6, we also have

$$g(\mathbf{x}; \mathbf{y} - \mathbf{x}) = \sum_{r=1}^n a_r(\mathbf{x})(y_r - x_r).$$

Writing $\mathbf{y} = \mathbf{x} + \lambda \mathbf{u}_k$, where $|\lambda|$ is less than the radius of $N(\mathbf{x})$ and \mathbf{u}_k is the k th unit coordinate vector, we have

$$0 < |\mathbf{y} - \mathbf{x}| = |\lambda|, \quad y_k - x_k = \lambda, \quad y_r - x_r = 0, \quad \text{if } r \neq k.$$

Therefore, the fundamental inequality becomes

$$|f(\mathbf{x} + \lambda \mathbf{u}_k) - f(\mathbf{x}) - \lambda a_k(\mathbf{x})| < \epsilon |\lambda|.$$

Dividing by $|\lambda|$, we find

$$\left| \frac{f(\mathbf{x} + \lambda \mathbf{u}_k) - f(\mathbf{x})}{\lambda} - a_k(\mathbf{x}) \right| < \epsilon,$$

and this implies that $D_k f(\mathbf{x})$ exists and has the value $a_k(\mathbf{x})$.

We have therefore proved that if a function f has a differential $g(\mathbf{x}; \mathbf{t})$ at \mathbf{x} , this differential is uniquely determined and must necessarily have the form

$$g(\mathbf{x}; \mathbf{t}) = \sum_{k=1}^n D_k f(\mathbf{x}) t_k, \quad \text{if } \mathbf{t} = (t_1, \dots, t_n).$$

It is customary to use the symbol df instead of g for the differential. It is also customary to use the symbols dx_1, \dots, dx_n instead of t_1, \dots, t_n for the components of \mathbf{t} , in which case the symbol $d\mathbf{x}$ is used in place of the vector \mathbf{t} . Theorem 6-7 then states that

$$df(\mathbf{x}; d\mathbf{x}) = \sum_{k=1}^n D_k f(\mathbf{x}) dx_k, \quad \text{if } d\mathbf{x} = (dx_1, \dots, dx_n).$$

This is sometimes expressed more briefly in the following notation:

$$df = \frac{\partial f}{\partial x_1} dx_1 + \cdots + \frac{\partial f}{\partial x_n} dx_n.$$

Indeed this last formula is often used as the definition of df and the properties given in our definition are then proved as theorems.

It is quite easy to see that the directional derivative $D_{\mathbf{u}} f(\mathbf{x})$ will exist in every direction \mathbf{u} if f has a differential at \mathbf{x} . In fact, the directional derivative is merely a special case of the differential.

6-8 THEOREM. Let f have a differential at a point \mathbf{x} of an open set S in E_n , and let \mathbf{u} be a unit vector in E_n . Then the directional derivative $D_{\mathbf{u}} f(\mathbf{x})$ exists and we have

$$D_{\mathbf{u}} f(\mathbf{x}) = df(\mathbf{x}; \mathbf{u}).$$

Proof. Given $\epsilon > 0$, there exists a neighborhood $N(\mathbf{x}; \delta)$ such that

$$|f(\mathbf{y}) - f(\mathbf{x}) - df(\mathbf{x}; \mathbf{y} - \mathbf{x})| < \epsilon |\mathbf{y} - \mathbf{x}|, \quad \text{if } \mathbf{y} \in N'(\mathbf{x}; \delta).$$

Given the unit vector \mathbf{u} , for every real $\lambda \neq 0$ such that $|\lambda| < \delta$, the point $\mathbf{x} + \lambda \mathbf{u}$ will be in $N'(\mathbf{x}; \delta)$. Taking $\mathbf{y} = \mathbf{x} + \lambda \mathbf{u}$ in the inequality, we obtain the relation

$$\left| \frac{f(\mathbf{x} + \lambda \mathbf{u}) - f(\mathbf{x})}{\lambda} - df(\mathbf{x}; \mathbf{u}) \right| < \epsilon, \quad \text{if } 0 < |\lambda| < \delta.$$

But this means that $D_{\mathbf{u}} f(\mathbf{x})$ exists and has the value $df(\mathbf{x}; \mathbf{u})$.