Definition

Let f be a real-valued function with domain D_f. Let x_0 be a cluster point of D_f. We define the number $f'(x_0)$ to be

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

provided the limit exists. If it does exist, we say that f is differentiable at x_0, and we call $f'(x_0)$ the derivative of f at x_0.

- Note that in the above definition, we only consider h for which $x_0 + h \in D_f$ since otherwise $f(x_0 + h)$ is not defined. This is always to be assumed; we won't mention it again.

- In Math 316, one of the hard things you will study is how to understand differentiability of functions $f : \mathbb{R}^n \to \mathbb{R}^m$ ("vector-valued functions of several variables").
- The definition of the derivative we've given above does not generalize to higher dimensions, so we ought to look for a definition that does generalize.
- The first thing we try to do is to formulate differentiability without any division; in some of the proofs (for example in proving the chain rule) it's hard to avoid dividing by 0 if we keep the definition of $f'(x_0)$ given above.

- To do this, let
$$\varepsilon := \frac{f(x_0+h)-f(x_0)}{h} - f'(x_0),$$
assuming that $f'(x_0)$ exists.
- Evidently ε is a function of h, and $\varepsilon(h)$ has the property that $\varepsilon(h)$ goes to 0 as h goes to 0.
- This can be rewritten as
$$f(x_0+h) - f(x_0) = f'(x_0)h + \varepsilon h, \quad \text{where } \varepsilon \to 0 \text{ as } h \to 0.$$
- Now suppose conversely that f satisfies the following:

  There exists a real number m and function ε of h such that $f(x_0+h) - f(x) = mh + \varepsilon h$, where $\varepsilon \to 0$ as $h \to 0$.

  Then we conjecture that the number m is unique, and it must be $f'(x_0)$.
- We summarize this in the following theorem.

Theorem (Linear Transformation Characterization of Differentiability)

Let f be a real-valued function with domain D_f. Let x_0 be a cluster point of D_f.

1. If f is differentiable at x_0, then there exists a function ε of h such that $f(x_0+h)-f(x_0)$ can be written in the form
$$f(x_0+h)-f(x_0)=f'(x_0)h+\varepsilon h$$
where $\varepsilon \to 0$ as $h \to 0$.
2. If there exists a real number m such that $f(x_0+h)-f(x_0)$ can be written in the form
$$f(x_0+h)-f(x_0)=mh+\varepsilon h$$
where ε is a function of h such that $\varepsilon \to 0$ as $h \to 0$, then f is differentiable at x_0 and $m=f'(x_0)$.

Exercise.

Write the proof of this theorem.

- We make use of this new way of looking at differentiability to prove the following theorem.

Theorem (Basic Properties of Differentiability)

Let f, g be functions with common domain D, let x_0 be a cluster point of D. Let c be a fixed real number. Suppose that f and g are differentiable at x_0.

a) (Differentiability implies continuity) Then f and g are continuous at x_0.

b) (Linearity) Then $f + g$ and $c\, f$ are both differentiable at x_0, and we have the formulas

$$(f+g)'(x_0) = f'(x_0) + g'(x_0)$$
$$(cf)'(x_0) = cf'(x_0).$$

c) (Product Rule) Then the product $f \cdot g$ is differentiable at x_0 and

$$(f \cdot g)'(x_0) = f'(x_0)g(x_0) + g'(x_0)f(x_0).$$

d) (Reciprocal Rule) If in addition $g(x_0) \neq 0$, then $1/g$ is differentiable at x_0 and

$$(1/g)'(x_0) = -\frac{g'(x_0)}{(g(x_0))^2}.$$

e) (Quotient Rule) If in addition $g(x_0) \neq 0$, then f/g is differentiable at x_0 and

$$(f/g)'(x_0) = \frac{f'(x_0)g(_0) - g'(x_0)f(x_0)}{(g(x_0))^2}.$$

Theorem (Basic Properties of Differentiability)

Let f, g be functions with common domain D, let x_0 be a cluster point of D. Let c be a fixed real number. Suppose that f and g are differentiable at x_0.

a) (Differentiability implies continuity) Then f and g are continuous at x_0.

Exercise.

Write the proof using the linear transformation characterization of differentiability.

Theorem (Basic Properties of Differentiability)

Let f, g be functions with common domain D, let x_0 be a cluster point of D. Let c be a fixed real number. Suppose that f and g are differentiable at x_0.

b) (Linearity) Then $f + g$ and $c\,f$ are both differentiable at x_0, and we have the formulas

$$(f+g)'(x_0) = f'(x_0) + g'(x_0)$$
$$(cf)'(x_0) = cf'(x_0).$$

Exercise.

Write the proof using the linear transformation characterization of differentiability.

Theorem (Basic Properties of Differentiability)

Let f, g be functions with common domain D, let x_0 be a cluster point of D. Let c be a fixed real number. Suppose that f and g are differentiable at x_0.

c) (Product Rule) Then the product $f \cdot g$ is differentiable at x_0 and

$$(f \cdot g)'(x_0) = f'(x_0)g(x_0) + g'(x_0)f(x_0).$$

Exercise.

Write the proof using the linear transformation characterization of differentiability.

Theorem (Basic Properties of Differentiability)

Let f, g be functions with common domain D, let x_0 be a cluster point of D. Let c be a fixed real number. Suppose that f and g are differentiable at x_0.

d) (Reciprocal Rule) If in addition $g(x_0) \neq 0$, then $1/g$ is differentiable at x_0 and

$$(1/g)'(x_0) = -\frac{g'(x_0)}{(g(x_0))^2}.$$

Exercise.

Write the proof using the linear transformation characterization of differentiability.

Theorem (Basic Properties of Differentiability)

Let f, g be functions with common domain D, let x_0 be a cluster point of D. Let c be a fixed real number. Suppose that f and g are differentiable at x_0.

e) (Quotient Rule) If in addition $g(x_0) \neq 0$, then f/g is differentiable at x_0 and

$$(f/g)'(x_0) = \frac{f'(x_0)g(_0) - g'(x_0)f(x_0)}{(g(x_0))^2}.$$

Exercise.

Write the proof using the linear transformation characterization of differentiability.

We continue with applications of the Linear Characterization of Differentiability Theorem by proving the following theorem.

Theorem (Chain Rule)

Let g be a function with domain D_g and f a function whose domain is contained in the range of g. Let x_0 be a cluster point of D_g and $g(x_0)$ a cluster point of D_f. If g is differentiable at x_0 and f is differentiable at $g(x_0)$, then the composition $f \circ g$ is differentiable at x_0 and

$$(f \circ g)'(x_0) = f'(g(x_0))g'(x_0).$$

Exercise.

Write the proof.

Definition of Linear Transformation

A function $T : \mathbb{R}^n \to \mathbb{R}^m$ is called a linear transformation provided the following two things are true:

(i) $T(\vec{x} + \vec{y}) = T(\vec{x}) + T(\vec{y})$ for all $\vec{x}$ and $\vec{y}$ in $\mathbb{R}^n$;

(ii) $T(c\,\vec{x}) = c\,T(\vec{x})$ for all $\vec{x} \in \mathbb{R}^n$ and $c \in \mathbb{R}$.

Exercise

Show that the linear transformations $T : \mathbb{R} \to \mathbb{R}$ are precisely the functions of the form $T(x) = mx$ for some constant m.

- Recall that we've shown f is differentiable at x_0 if and only if there exists a number m such that $f(x_0+h)-f(x_0)$ can be written as $f(x_0+h)-f(x_0)=mh+\varepsilon h$, where ε is a function of h such that $\varepsilon \to 0$ as $h \to 0$.
- Thus f is differentiable at x_0 if and only if there exists a linear transformation $T:\mathbb{R}\to\mathbb{R}$ such that such that $f(x_0+h)-f(x_0)$ can be written as $f(x_0+h)-f(x_0)=T(h)+\varepsilon h$, where ε is a function of h such that $\varepsilon \to 0$ as $h \to 0$.
- The linear transformation T is unique, and instead of calling it T, it is commonly written as df_{x_0} and called the differential of f at x_0.
- Thus the differential df_{x_0} is viewed as a linear transformation from $\mathbb{R}$ to $\mathbb{R}$ whose value at h is $df_{x_0}(h)=f'(x_0)h$, just as in your basic calculus classes.

Why the name "Linear Transformation Characterization of Differentiability?"

- So to recap, f is differentiable at x_0 if and only if there exists a linear transformation df_{x_0} from $\mathbb{R}$ to $\mathbb{R}$ such that
$$f(x_0 + h) - f(x_0) = df_{x_0}(h) + \varepsilon(h)h,$$
where ε is a function of h such that $\varepsilon(h) \to 0$ as $h \to 0$.
- We can rewrite this by defining $\tilde{\varepsilon}(h) := \varepsilon(h)/h$. With this notation, f is differentiable at x_0 if and only if there exists a linear transformation df_{x_0} from $\mathbb{R}$ to $\mathbb{R}$ such that
$$f(x_0 + h) - f(x_0) = df_{x_0}(h) + \tilde{\varepsilon}(h),$$
where $\tilde{\varepsilon}$ is a function of h such that $\tilde{\varepsilon}(h)/h \to 0$ (equivalently, $|\tilde{\varepsilon}(h)|/|h| \to 0$ as $h \to 0$).
- This definition generalizes to functions $f : \mathbb{R}^n \to \mathbb{R}^m$.
- In this more general setting, h is a vector in $\mathbb{R}^n$, and $\tilde{\varepsilon} : \mathbb{R}^n \to \mathbb{R}^m$ such that $\dfrac{\|\tilde{\varepsilon}(h)\|}{\|h\|} \to 0$ as $\|h\| \to 0$.
The norm in the numerator of $\dfrac{\|\tilde{\varepsilon}(h)\|}{\|h\|}$ is in $\mathbb{R}^m$ and the norm in the denominator is in $\mathbb{R}^n$.
- It is a theorem that the linear functions from $\mathbb{R}^n$ to $\mathbb{R}^m$ are given by multiplication by an $m \times n$ matrix of real numbers (as suggested by the exercise on slide 12).
- The differential df_{x_0} is a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ for which the associated matrix turns out to be the so-called Jacobian matrix consisting of all of the partial derivatives of the various components of f. You will study this in detail in Math 316. The above suggests that linear algebra is an important tool in studying multivariable calculus.