Differentiation

Definition. *Let $f : (a,b) \to \mathbb{R}$ be a function and $x_0 \in (a,b)$. The **derivative** of f at x_0, denoted by $f'(x_0)$, is defined as*

$$f'(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

whenever this limit exists. Equivalently, by setting $x = x_0 + h$, we can write

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

*If the limits above exist, then we say that the function is **differentiable** at x_0. If a function is differentiable at every point in its domain, then we say that it is a differentiable function, and we may think of the derivative as a function in its own right $f' : (a,b) \to \mathbb{R}$.*

Various notations are used to denote the derivative, for example,

$$f'(x_0) = \frac{df}{dx}(x_0) = \frac{df}{dx}\Big|_{x=x_0} = \frac{d}{dx}\Big(f(x)\Big)\Big|_{x=x_0}.$$

Theorem. *If a function $f : (a,b) \to \mathbb{R}$ is differentiable at $x_0 \in (a,b)$, then it is continuous at x_0.*

The converse does not hold, a function may be continuous at a point, but not differentiable at that point. For example, $f(x) = |x|$. Remarkably, there also exist functions which are continuous everywhere, but differentiable nowhere! Hence, we have an inclusion of sets:

$$\{\text{Differentiable functions}\} \subset \{\text{Continuous Functions}\} \subset \{\text{Functions}\}.$$

Theorem (Properties of the Derivative).

(i) Let $f : (a,b) \to \mathbb{R}$ be a differentiable function and $k \in \mathbb{R}$. Then $kf : (a,b) \to \mathbb{R}$ is differentiable and $(kf)'(x) = kf'(x)$.

(ii) Let $f, g : (a,b) \to \mathbb{R}$ be differentiable functions. Then $f + g : (a,b) \to \mathbb{R}$ is differentiable, and $(f+g)'(x) = f'(x) + g'(x)$.

(iii) Suppose $f : (a,b) \to \mathbb{R}$ is the constant function $f \equiv \lambda$ for some $\lambda \in \mathbb{R}$. Then f is differentiable, and $f'(x) \equiv 0$.

(iv) (Product Rule) If $f, g : (a,b) \to \mathbb{R}$ are differentiable functions, then so is fg, and $(fg)'(x) = f'(x)g(x) + f(x)g'(x)$.

(v) (Quotient Rule) If $f, g : (a,b) \to \mathbb{R}$ are differentiable functions, then f/g is differentiable whenever $g(x) \neq 0$, and

$$(f/g)'(x) = \frac{f'(x)g(x) + f(x)g'(x)}{(g(x))^2}.$$

Theorem (Chain Rule). *Suppose $f : (a,b) \to \mathbb{R}$ and $g : (c,d) \to \mathbb{R}$ are differentiable functions, and $f \circ g$ is defined. Then $f \circ g$ is differentiable and $(f \circ g)'(x) = f'(g(x))g'(x)$.*

Note that when a function is differentiable, we can ask whether the derivative is itself differentiable. If this is so, then we call the resulting function the second derivative. We can iterate this process to obtain higher and higher derivatives, provided the requisite limits exist. The usual notations for second derivatives are

$$f''(x_0) = \frac{d^2 f}{dx^2}(x_0) = \frac{d^2 f}{dx^2}\Big|_{x=x_0} = \frac{d^2}{dx^2}\Big(f(x)\Big)\Big|_{x_0}.$$

Similar notation is employed for higher derivatives. We use the notation $C^k\big((a,b),\mathbb{R}\big)$ to denote the set of all real-valued functions defined on (a,b) whose derivatives up to order k exist and are continuous, and $C^\infty\big((a,b),\mathbb{R}\big)$ to denote the set of all real-valued functions defined on (a,b) which are infinitely differentiable.

There are many different ways of interpreting the derivative. For example,

(i) (formal) The limit in the original definition.
(ii) (infinitesimal) The ratio of the infinitesimal change in the value of a function to the change in the function's argument.

(iii) (geometric) The slope of the tangent line to the function at a point.
(iv) (approximation) The best linear approximation to a function at a point.
(v) (symbolic) Thinking of the derivative purely in terms of the rules i.e. $x^n \mapsto nx^{n-1}$.
(vi) (rate of change) The instantaneous velocity of a particle whose position is described by the function, thinking of the argument as time.

Theorem. *If a differentiable function f has a local minimum or maximum at x_0, then $f'(x_0) = 0$. Hence, the vanishing of the derivative is a necessary (but not sufficient) condition for there to be a local minimum or maximum.*

Definition. *A point x_0 in the domain of f is called a **critical point** or **stationary point** of f if*

$$f'(x_0) = 0.$$

Theorem (Rolle's Theorem). *Let $f : [a, b] \to \mathbb{R}$ be a continuous function which is differentiable on (a, b). Suppose that $f(a) = f(b)$. Then there exists $x_0 \in (a, b)$ such that $f'(x_0) = 0$.*

Theorem (Mean Value Theorem). *Let $f : [a, b] \to \mathbb{R}$ be a continuous function which is differentiable on (a, b). Then there exists $x_0 \in (a, b)$ such that*

$$f'(x_0) = \frac{f(b) - f(a)}{b - a}.$$

Corollary. *Let $f : [a, b] \to \mathbb{R}$ be a continuous function which is differentiable on (a, b). Then*

$$|f(b) - f(a)| \leq \max_{a \leq x \leq b} |f'(x)||b - a|.$$

Theorem. *Suppose $f : (a, b) \to \mathbb{R}$ is a differentiable function. If $f'(x) = 0$ on some interval, then f is constant on that interval.*

Definition. *A function $f : (a, b) \to \mathbb{R}$ is **strictly increasing** on an interval (a, b) if $x_1 < x_2$ implies that $f(x_1) < f(x_2)$ for any two points $x_1, x_2 \in (a, b)$. Similarly, a function $f : (a, b) \to \mathbb{R}$ is **strictly decreasing** on an interval if $x_1 < x_2$ implies that $f(x_1) > f(x_2)$ for any two points $x_1, x_2 \in (a, b)$.*

Theorem. *Suppose $f : (a, b) \to \mathbb{R}$ is a differentiable function. If $f'(x) > 0$ (respectively, $f'(x) < 0$) on some interval, then $f(x)$ is strictly increasing (respectively, strictly decreasing) on that interval. This implies the following basic principle in calculus: if $f'(x_0) = 0$, and f' changes from positive to negative (respectively, negative to positive) at x_0, then f has a local maximum (respectively, local minimum) at x_0.*

Note that the converse to the above theorem does not hold. The cubic $y = x^3$ provides a counterexample.

Theorem. *Suppose $f : (a, b) \to \mathbb{R}$ is a twice differentiable function and $(c, d) \subset (a, b)$ is some interval. If $f''(x) > 0 \ \ \forall x \in (c, d)$, then over the interval (c, d):*

(i) the gradient of the function is strictly increasing,
(ii) the tangent lines lie below the graph (except at the point of tangency),
(iii) the line segment joining $(c, f(c))$ to $(d, f(d))$ lies strictly above the graph of f.

*Such a function is said to be **concave up** (or convex) on (c, d).*

Theorem. *Suppose $f : (a, b) \to \mathbb{R}$ is a twice differentiable function and $(c, d) \subset (a, b)$ is some interval. If $f''(x) < 0 \ \ \forall x \in (c, d)$, then over the interval (c, d):*

(i) the gradient of the function is strictly decreasing,
(ii) the tangent lines lie above the graph (except at the point of tangency),
(iii) the line segment joining $(c, f(c))$ to $(d, f(d))$ lies strictly below the graph of f.

*Such a function is said to be **concave down** (or concave) on (c, d).*

Definition. *Suppose $f : (a, b) \to \mathbb{R}$ is a twice differentiable function. A point $x = x_0$ is called a **point of inflection** if $f''(x_0) = 0$ and $f''(x)$ changes sign at $x = x_0$.*

Remark.

(i) At a point of inflection the graph of f changes from concave up to concave down, or vice versa.
(ii) At a point of inflection the tangent line passes through the graph of f.
(iii) A point of inflection of f corresponds to an isolated maximum or minimum for $f'(x)$.
(iv) Note that a point of inflection may be stationary (that is $f'(x_0) = 0$), or not stationary (and thus $f'(x_0) \neq 0$).

(v) The condition that f'' changes sign at a point of inflection is necessary, consider $f(x) = x^4$. At the origin this function has vanishing second derivative, but this isn't a point of inflection.

(vi) When sketching the graph of a function, in addition to considering the values of the first derivative (to determine the stationary points and where the function is increasing or decreasing), it is often useful to consider the values of the second derivative, in order to identify regions where it is concave up or concave down. Moreover, the second derivative can also help determine the nature of the stationary points (see below).

THEOREM (SECOND DERIVATIVE TEST). *Suppose $f : (a, b) \to \mathbb{R}$ is a twice differentiable function.*

(i) If $f'(x_0) = 0$ and $f''(x_0) > 0$, then f has a local minimum at $x = x_0$.
(ii) If $f'(x_0) = 0$ and $f''(x_0) < 0$, then f has a local maximum at $x = x_0$.
(iii) If $f'(x_0) = 0$ and $f''(x_0) = 0$, then the test is inconclusive.

DEFINITION. ***Antidifferentiation** is the inverse process to differentiation. Suppose that a function $F(x) : \mathbb{R} \to \mathbb{R}$ has derivative $f : \mathbb{R} \to \mathbb{R}$. This is written as $F'(x) = f(x)$, and we say that F is an **antiderivative** of f.*

Any other function $H : \mathbb{R} \to \mathbb{R}$ with derivative f will differ from F by a constant. Hence, any function whose derivative is f will be of the form $F(x) + c$ for some $c \in \mathbb{R}$. We call $F(x) + c$ the **general antiderivative** of f, or the **indefinite integral** of f, and write this as

$$\int f(x)\, dx = F(x) + c.$$

The symbol dx is simply notation to denote the extent of the integral, and the variable with respect to which we are antidifferentiating. A priori you are *not* permitted to algebraically manipulate it as if it were an independent variable.

The following properties of the indefinite integral are immediate consequences of the corresponding properties of derivatives.

THEOREM (PROPERTIES OF THE INDEFINITE INTEGRAL).

(i) $\int kf(x)\, dx = k \int f(x)\, dx$ *where* $k \in \mathbb{R}$.
(ii) $\int f(x) + g(x)\, dx = \int f(x)\, dx + \int g(x)\, dx$.