The Mean Value Theorem

We will begin with a classical theorem which serves in effect as a lemma in the proof of the more general and versatile mean value theorem. The following result was originally proven only for polynomials in 1691 by the French mathematician Michel Rolle, and the more general result given here continues to bear his name. Rolle was interested in finding methods for locating the roots of a polynomial—an area of research still very active today—and his proof is entirely algebraic. However, as knowledge spread of the infinitesimal calculus developed independently around the same time by the English scientist Isaac Newton and German polymath Gottfried Leibniz, it was recognized that the particular method employed by Rolle could be readily recast in the language of derivatives. Rolle's theorem in its modern form, as a statement belonging to the field of calculus, first appeared in 1868 in the work of French mathematician Joseph-Alfred Serret, who found inspiration in the approach taken by yet another French mathematician, Pierre-Ossian Bonnet, earlier in the nineteenth century. Both Serret and Bonnet are remembered today principally for their foundational work in the field of differential geometry.

(Cf. §3.2.3 of the course textbook.)

Rolle's theorem. *Let f be a real function which is continuous on a closed interval $[a, b]$ and differentiable on (a, b). If $f(a) = f(b)$ then there exists a point $\xi \in (a, b)$ such that $f'(\xi) = 0$.*

Proof. Let $f(a) = C = f(b)$. By the extreme value theorem, f attains a maximum M and minimum m on $[a, b]$. If $M = m = C$ then f is constant on this interval, and the result is immediate. (Its derivative is 0 everywhere on (a, b).) Suppose then that f attains an extremum different from C, necessarily at some point $\xi \in (a, b)$. By Fermat's theorem $f'(\xi) = 0$. $\square$

Example 1 (§3.2.5 Problem 16)**.** Use Rolle's theorem to show that $f(x) = 3x^4 - 6x^2$ is injective *as a fucntion* $[1, \infty) \to$.

A simple calculation shows that the derivative is

$$f'(x) = 12x^3 - 12x = 12x(x^2 - 1) = 12x(x + 1)(x - 1),$$

The roots of which are $x = 0, \pm 1$. Being a polynomial, f is continous on $[a, b]$ and differentiable on (a, b) for any $1 \le a < b$. Since no root of $f'(x)$ is in any such (a, b), the contrapositive of Rolle's theorem implies that $f(a) \ne f(b)$, from which it follows immediately that f is injective.

Now we are ready to state and prove one of the central results of univariate calculus. It and its direct consequences feature, for instance, in the proof of the fundamental theorem of calculus.

(Cf. §3.2.4 of the course textbook.)

Mean value theorem. *Let f be a real function which is continuous on a closed interval $[a, b]$ and differentiable on (a, b). Then there exists a point $\xi \in (a, b)$ such that*

$$f(b) - f(a) = f'(\xi)(b - a).$$

In other words there exists a point at which the instantaneous rate of change is equal to the average rate of change. Geometrically, there exists a point at which the tangent line to the graph is parallel to the secant line through the boundary points.

Remark 1. The preceding theorem also appears in the literature as the *Lagrange mean value theorem*, to distinguish it from the more general Cauchy mean value theorem (given below), and as the *finite growth theorem*, to distinguish it from the *(integral) mean value theorem*.

The mean value theorem was first proved in its modern form in 1823 by the French mathematician Augustin Louis Cauchy as a special case of the following more general eponymous result. The standard proof given today utilizes the modern form of Rolle's theorem. Cauchy was apparently unaware that Rolle's theorem applied to functions of which the derivative is not itself continuous, and his original proof, following on the work of the Franco-Italian mathematician Joseph-Louis Lagrange, takes a somewhat different approach. Cauchy is today considered as a "founding father" of the field of *mathematical analysis*, which includes calculus. Before him, Lagrange had expanded calculus into the rich field it has become and gives his name to *Lagrangian mechanics* in physics.

Cauchy mean value theorem. *Let the real-valued functions f and g be continous on $[a, b]$ and differentiable on (a, b). Then there exists a point $\xi \in (a, b)$ such that*

$$g'(\xi)(f(b) - f(a)) = f'(\xi)(g(b) - g(a)).$$

Geometrically, this theorem says that, for any curve given by a differentiable mapping $t \mapsto (f(t), g(t))$ for $t \in [a, b]$, there exists some $\xi \in (a, b)$ such that the tangent line to the curve at $(f(\xi), g(\xi))$ is parallel to the chord between the endpoints $(f(a), g(a))$ and $(f(b), g(b))$. In the special case where $g'(x) \neq 0$ for all $x \in (a, b)$, this form of the theorem can be used to verify *L'Hôpital's rule*.

Remark 2. This theorem may also be referred to with either of the words *generalized* or *extended* in place of the name *Cauchy*.

Proof. Define the function h on $[a, b]$ by

$$h(x) = f(x)(g(b) - g(a)) - g(x)(f(b) - f(a)) + f(b)g(a) - f(a)g(b).$$

That h is continuous on $[a, b]$ and differentiable on (a, b) is a clear consequence of the same properties of f and g. Moreover, a simple calculation shows that $h(a) = h(b) = 0$. Hence we can invoke Rolle's theorem to deduce that there exists $\xi \in (a, b)$ such that $h'(\xi) = 0$. But in general

$$h'(x) = f'(x)(g(b) - g(a)) - g'(x)(f(b) - f(a)),$$

and consequently

$$0 = h'(\xi) = f'(\xi)(g(b) - g(a)) - g'(\xi)(f(b) - f(a))$$

$$\iff$$

$$g'(\xi)(f(b) - f(a)) = f'(\xi)(g(b) - g(a)).$$

$\square$

The "regular" mean value theorem proceeds from letting g be the identity function.

Example 2 (§3.2.5 Problem 23)**.** Use the mean value theorem to show that

$$\tfrac{1}{3} < \ln \tfrac{3}{2} < \tfrac{1}{2}.$$

The natural logarithm function is differentiable on $(0, \infty)$, from which it clearly follows that, in particular, it is continuous on $[2, 3]$ and differentiable on $(2, 3)$. Hence, by the mean value theorem, there exists some $\xi \in (2, 3)$ such that

$$\ln \tfrac{3}{2} = \ln 3 - \ln 2 = xx = \xi \ln x(3 - 2) = \tfrac{1}{\xi}(1) = \tfrac{1}{\xi}.$$

Now $2 < \xi < 3$ is equivalent to $\frac{1}{3} < \frac{1}{\xi} < \frac{1}{2}$, and the result follows.

The following corollaries of the mean value theorem are given in maximal generality. It is necessary that the student understand at least what they imply for a function defined on a single open interval.

Corollary 1. *If $f : D \subseteq \to$ is continuous on D and differentiable on the interior of D, then $f'(x) \equiv 0$ on the interior of D if and only if f is constant on each connected component of D.*

Proof. ($\Leftarrow$) Trivial.

($\Rightarrow$) Every connected component of D is necessarily an isolated point or an interval. In the former case, there is nothing to prove. As for the latter case, denote the interval I, and take any $a_1 < b_1 \in I$. By the mean value theorem, there exists $a_1 < p < b_1$ such that $f(b_1) - f(a_1) = f'(p)(b_1 - a_1)$. By hypothesis $f'(p) = 0$, so that $f(b_1) - f(a_1) = 0$, implying $f(b_1) = f(a_1)$. Hence f must be constant on I. □

Corollary 2. *Let $f, g : D \subseteq \to$ be continuous on D and differentiable on the interior of D. Then $f'(x) = g'(x)$ on the interior of D if and only if $f - g$ is constant on each connected component of D.*

Proof. ($\Leftarrow$) Trivial.

($\Rightarrow$) Define $h = f - g$, which is clearly continuous on D and differentiable on the interior of D. By the linearity of the derivative operator x, we have that $[h]x \equiv 0$ on the interior of D, hence h is constant on each connected component of D by the previous corollary. □

The following useful results also proceed from the mean value theorem.

Derivative extension theorem. *Let the real function f be differentiable on the open interval I except possibly at the single point $a \in I$. If its derivative function f' is such that*

$$\ell = \lim_{x \to a} f'(x)$$

exists and is finite, then $f'(a) = \ell$.

Proof. For any $x \in I$ with $x \neq a$, the mean value theorem guarantees the existence of some ξ strictly between a and x such that $f(x) - f(a) = f'(\xi)(x-a)$. Hence we can write

$$\frac{f(x) - f(a)}{x - a} = f'(\xi).$$

We take the limit of both sides as xa. Because ξ clearly approaches a itself as xa, this is equivalent on the right-hand side to taking the limit as ξa, hence

$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a} = \lim_{\xi \to a} f'(\xi) = \ell,$$

by hypothesis. Therefore, since ℓ is finite (i.e. some real number), the derivative of f at a is ℓ by definition. □

Using the preceding results, we can determine certain local characteristics of the graph of a function from the knowledge of the function's derivatives. These characteristics, which give an idea of the shape of the graph, and what the derivatives say about them are described below.

In the following let $I \subseteq$ be an arbitrary interval, including possibly $= (-\infty, \infty)$, and $f : I \to$. We denote by (I) the *interior* of I, which is the largest open interval contained in I; e.g. $[a, b] = [a, b) = (a, b] = (a, b) = (a, b)$.

Definition 1 (Monotonicity)**.** The function f is said to be ***increasing*** (or *montonically increasing* or *non-decreasing*), resp. ***decreasing*** (or *montonically decreasing* or *non-increasing*), if $f(x) \leq f(y)$, resp. $f(x) \geq f(y)$, whenever $x < y$ for all $x, y \in I$. A function is said to be ***monotonic*** if it is either increasing or decreasing.

Definition 2. The function f is said to be ***strictly increasing*** or ***strictly decreasing*** if it is increasing or decreasing, resp., and is injective. This is equivalent to replacing "$\leq$" and "$\geq$" with "$<$" and "$>$", resp., in the preceding definition. A function which is either strictly increasing or strictly decreasing is said to be ***strictly monotone***.

Remark 3. To emphasize that the monotonicity of a function is not strict, one may use terms such as *weakly increasing* or *weakly monotone.*

The results of the next theorem are in effect an extension of the contrapositive statement of Rolle's theorem.

(Cf. §3.3 (introduction) of the course textbook.)

Proposition 1. *Let f be continuous on I and differentiable on (I). Then f is increasing, resp. decreasing, if and only if $f'(x) \geq 0$, resp. $f'(x) \leq 0$, for all $x \in (I)$. Moreover, f is* strictly *monotone if and only if, in addition, the set of points $x \in (I)$ such that $f'(x) = 0$ has empty interior, i.e. contains no (nonempty) open interval.*

Proof. Idea: The mean value theorem applies on every sub-interval $[a, b] \subset I$, and the results for the non-strict case are established by comparing signs in the equation $f(b) - f(a) = f'(\xi)(b-a)$. Proving the strict case when the derivative is everywhere positive or negative is entirely analogous; the more general statement requires a more complicated argument. □

First derivative test for local extrema. *Let $a \in (I)$ be a critical point of f, and assume f is continuous on I and differentiable on $I - a$.*

- *If there exists some $\delta > 0$ such that $f'(x) \geq 0$ for all $x \in (a-\delta, a)$ and $f'(x) \leq 0$ for all $x \in (a, a+\delta)$, then f attains a local maximum at a.*
- *If there exists some $\delta > 0$ such that $f'(x) \leq 0$ for all $x \in (a-\delta, a)$ and $f'(x) \geq 0$ for all $x \in (a, a+\delta)$, then f attains a local minimum at a.*
- *If there exists some $\delta > 0$ such that f' is either uniformly positive or uniformly negative on $x \in (a-\delta, a)(a, a+\delta)$, then f does not attain a local extremum at a.*

If none of the conditions above are met, then the test is inconclusive: no conclusion about $f(a)$ can be reached with this test.

Definition 3 (Concavity)**.** The function f is said to be ***concave up***, or ***convex***, if for any $a, b \in I$ with $a < b$ and $t \in [0, 1]$

$$f((1-t)a + tb) \leq (1-t)f(a) + tf(b).$$

If this instead holds with the relation reversed to "$\geq$", then f is said to be ***concave down***, or *concave*. These terms are the descriptions of the ***concavity*** of f. If the order relation is strict for all $t \in (0, 1)$, then the concavity is said to be strict, as in *strictly concave up*.

When f is differentiable and concave up, its graph will be locally above the tangent line at each point. Similarly, it will be locally below the tangent line when f is convave down.

Proposition 2. *Let I be open and f be twice differentiable. Then f is concave up, resp. concave down, if and only if $f''(x) \geq 0$, resp. $f''(x) \leq 0$, for all $x \in I$. The concavity is strict if and only if f'' is additionally nonzero on I.*

Definition 4 (Inflection point)**.** Let $I \subseteq$ be an open interval, $a \in I$ and $f : I \to$. If there exists some $R > 0$ such that f has uniform concavity on $(a-R, a)$ and on $(a, a+R)$ and the concavity on these two intervals are opposite, then a is called an ***inflection point*** of f.

Proposition 3. *Let $I \subseteq$ be an open interval with $a \in I$, and let $f : I \to$ be twice differentiable on $I - a$. Then a is an inflection point of f if and only if f'' changes sign at a, i.e. there exists some $R > 0$ such that the sign of f'' is uniform on each of the intervals $(a-R, a)$ and $(a, a+R)$ and the signs of f'' on these intervals are opposite.*

By combining knowledge of a function's first and second derivatives at a point, another test for local extrema arises. If the second derivative can be readily computed, this *second derivative test* will be more convenient than the first derivative test, as it dispenses with the need to consider the values of the derivative at all points on some neighborhood of the critical point in question.

Second derivative test. *Let $I \subseteq$ be an open interval with $a \in I$. Let $f : I \to$ be twice continuously differentiable, and suppose a is a critical point of f, i.e. $f'(a) = 0$.*

- *If $f''(a) > 0$ then f attains a local minimum at a.*

- *If* $f''(a) < 0$ *then* f *attains a local maximum at* a.

If $f''(a) = 0$ *then the test is inconclusive: no conclusion about* $f(a)$ *can be reached with this test.*

If the second derivative proves difficult to compute, then clearly this second test provides little improvement over the first. Another drawback is that it can reveal nothing when the second derivative is 0.

Definition 5 (Cusp)**.** ***cusp***