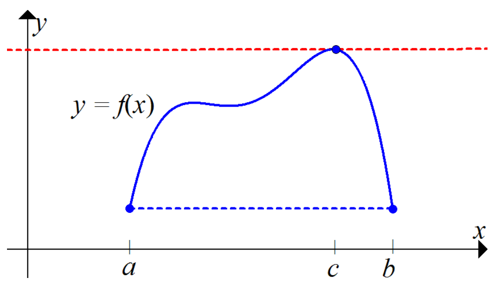
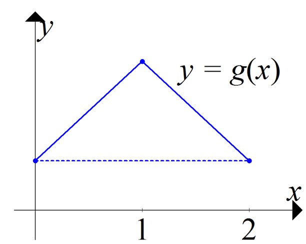
Mean Value Theorem

Vanishing derivatives. We will prove some basic theorems which relate the derivative of a function with the values of the function, culminating in the Uniqueness Theorem at the end. The first result is:

> *Rolle's Theorem:* If $f(x)$ is continuous on a closed interval $x \in [a,b]$ and differentiable on the open interval $x \in (a,b)$, and $f(a) = f(b)$, then there is some point $c \in (a,b)$ with $f'(c) = 0$.

Here $x \in [a,b]$ means $a \le x \le b$, and $x \in (a,b)$ means $a < x < b$. See the graph at left for an example: no matter how the curve wiggles, it must be horizontal somewhere.





Physically, imagine $f(t)$ represents the height of a rocket at time t, starting and finishing on its launch pad over the time interval $t \in [a,b]$. The theorem says there must be a pause in the motion where $f'(t) = 0$: this is the moment the rocket runs out of fuel and starts to fall.

Proof of Theorem. Assume $f(x)$ satisfies the hypotheses* of the Theorem. The Extremal Value Theorem (§3.1) guarantees that the continuous function $f(x)$ has at least one absolute maximum point $x = c_1 \in [a,b]$.

- If $c_1 \ne a, b$, then $c_1 \in (a,b)$, and the First Derivative Theorem (§3.1) says that $f'(c_1) = 0$.
- On the other hand, if $c_1 = a$ or b, then $f(c_1) = f(a) = f(b)$. Still, $f(x)$ also has an absolute minimum point $x = c_2$. If $c_2 \in (a,b)$, then $f'(c_2) = 0$ as before.
- The only case left is if $c_1 = a$ or b, and also $c_2 = a$ or b, so that $f(c_1) = f(c_2) = f(a) = f(b)$. Since the maximum and minimum values are the same, $f(x)$ cannot move above or below $f(a)$. Thus, $f(x)$ can only be a constant function, and $f'(c) = 0$ for *all* $c \in (a,b)$.

In every case, the conclusion† holds, Q.E.D.‡

Notes by Peter Magyar `magyar@math.msu.edu`

*In formal mathematics, *hypothesis* (plural *hypotheses*) means the "if" part of a theorem, the setup which is given or assumed. In our theorem, the three hypotheses are: $f(x)$ is continuous on $[a,b]$, $f(x)$ is differentiable on (a,b), and $f(a) = f(b)$.

†*Conclusion* means the "then" part of a theorem, the payoff which is to be deduced from the hypothesis: in our theorem, that $f'(c) = 0$.

‡Latin *quod erat demonstrandum*, "which was to be shown", the traditional end of a proof.

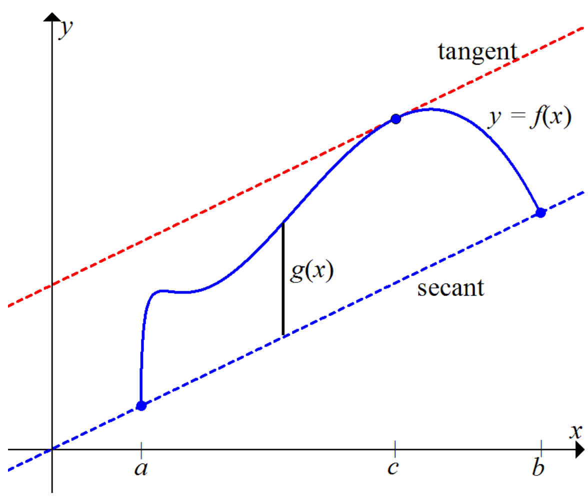
For Rolle's Theorem, as for most well-stated theorems, all the hypotheses are necessary to be sure of the conclusion. In the graph at right above, $y = g(x)$ has a corner and $g'(1)$ does not exist, so just one hypothesis fails at just one point. But already the conclusion is false: $g'(c) = 1$ for $c < 1$ and $g'(c) = -1$ for $c > 1$, but nowhere is $g'(c) = 0$. In physical terms, the velocity jumps instantaneously from 1 to -1 like an idealized ping-pong ball, and there is no well-defined velocity at the moment of impact.

Derivatives versus difference quotients. Throughout our theory, the derivative $f'(a)$ has been shadowed by the difference quotient $\frac{\Delta f}{\Delta x} = \frac{f(b)-f(a)}{b-a}$, over some interval $[a, b]$. Numerically, the difference quotient is an approximation to the derivative: $\frac{df}{dx} \approx \frac{\Delta f}{\Delta x}$, provided Δx is small. In physical terms, the difference quotient is the average rate of change of $f(x)$ over $x \in [a, b]$. Geometrically in terms of the graph $y = f(x)$, the difference quotient is the slope of the secant line cutting through the points $(a, f(a))$ and $(b, f(b))$.

Now we come to the most powerful result of this section, which says that the derivative is sometimes exactly equal to the difference quotient (the slope of one tangent is equal to the slope of the secant):

> *Mean Value Theorem (MVT):* If $f(x)$ is continuous on a closed interval $x \in [a, b]$ and differentiable on the open interval $x \in (a, b)$, then there is some point $c \in (a, b)$ with $f'(c) = \frac{f(b)-f(a)}{b-a}$.

See the picture below for an example: as the graph rises from $(a, f(a))$ to $(b, f(b))$, at some points the tangent line must be parallel to the secant line.



Note that Rolle's Theorem is the special case of MVT in which the secant line is horizontal. In fact, we will prove MVT for a general $f(x)$ by cooking up a new function $g(x)$ for which Rolle's Theorem applies, then translating Rolle's conclusion back in terms of $f(x)$.

Proof of MVT. Suppose $f(x)$ satisfies the hypotheses. Then define a new function $g(x)$, shown in the picture, which measures the height from the graph $y = f(x)$ down to the secant line $y = f(a) + \frac{f(b)-f(a)}{b-a}(x-a)$:

$$g(x) \;=\; f(x) - f(a) - \frac{f(b)-f(a)}{b-a}(x-a).$$

Then $g(x)$ is continuous on $[a, b]$ by the Limit Laws (§1.6), and differentiable on (a, b) by the Derivative Rules (§2.3). In fact,

$$g'(x) \;=\; f'(x) - 0 - \tfrac{f(b)-f(a)}{b-a}(1-0) \;=\; f'(x) - \tfrac{f(b)-f(a)}{b-a},$$

since $f(a)$ and $\frac{f(b)-f(a)}{b-a}$ are constants (having no x in them).

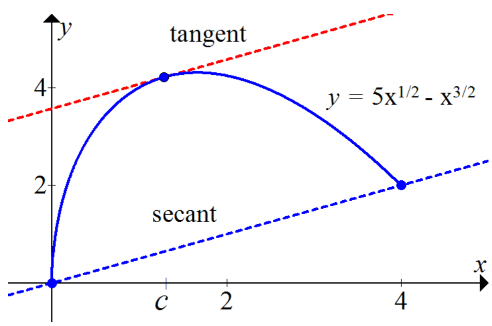
Also, we can easily compute that $g(a) = g(b) = 0$, so all the hypotheses of Rolle's Theorem hold for $g(x)$. Thus the conclusion of Rolle's Theorem also holds: there is some $c \in (a, b)$ with $g'(c) = 0$. That is,

$$g'(c) \;=\; f'(c) - \tfrac{f(b)-f(a)}{b-a} \;=\; 0,$$

which means $f'(c) = \frac{f(b)-f(a)}{b-a}$, Q.E.D.

The Mean Value Theorem does not give any way to find the particular $c \in (a, b)$ in the conclusion, so if we want this value in a particular case, we must solve for x in the equation $f'(x) = \frac{f(b)-f(a)}{b-a}$; however the Theorem will guarantee that there is some solution.

EXAMPLE: Let $f(x) = 5\sqrt{x} - x\sqrt{x}$ over the interval $[a, b] = [0, 4]$.



To check the hypotheses of MVT, note that $\sqrt{x}$ is continuous for all $x \geq 0$, and thus over $[0, 4]$. As for differentiability:

$$f'(x) \;=\; \left(5x^{1/2} - x^{3/2}\right)' \;=\; \tfrac{5}{2}x^{-1/2} - \tfrac{3}{2}x^{1/2}$$

is defined for $x > 0$, and hence over $x \in (0, 4)$: the hypothesis allows $f'(a) = f'(0)$ to be undefined. Thus we conclude there must be some $c \in (0, 4)$ with $f'(c) = \frac{f(b)-f(a)}{b-a} = \frac{2-0}{4-0} = \frac{1}{2}$. If we wish to find this c, we must solve:

$$f'(x) \;=\; \tfrac{5}{2}x^{-1/2} - \tfrac{3}{2}x^{1/2} \;=\; \tfrac{1}{2}\,,$$

which is equivalent to $3x + \sqrt{x} - 5 = 0$. Substituting the variable $u = \sqrt{x}$ gives $3u^2 + u - 5 = 0$, so the Quadratic Formula gives:

$$u \;=\; \sqrt{x} \;=\; \frac{-1 \pm \sqrt{1^2 - 4(3)(-5)}}{2(3)} \;=\; \frac{-1 \pm \sqrt{61}}{6}.$$

The negative solution is impossible, and the positive one gives $x = c = \left(\frac{\sqrt{61}-1}{6}\right)^2 \approx 1.29$, which agrees with the picture.

Mathematical and physical uniqueness. We come to the most important result of this section. Let $f(x)$ be a continuous function on an interval $x \in [p, q]$.

Uniqueness Theorem:

(a) If $f'(x) = 0$ for all $x \in (p, q)$, then $f(x) = C$, a constant function.

(b) If $f(x), g(x)$ have the same derivative $f'(x) = g'(x)$ for all $x \in (p, q)$, then $f(x) = g(x) + C$ for some constant C.

(c) If $f(x), g(x)$ have the same derivative $f'(x) = g'(x)$ for all $x \in (p, q)$, and the same initial value $f(c) = g(c)$ for some $c \in [p, q]$, then $f(x) = g(x)$.

Proof. (a) Assume the hypothesis $f'(x) = 0$ for all $x \in (p, q)$, and consider two outputs $f(a)$ and $f(b)$ for any $a < b$ in $[p, q]$. Applying the Mean Value Theorem to the smaller interval $[a, b]$, we get $\frac{f(b)-f(a)}{b-a} = f'(c) = 0$, since all derivatives $f'(x)$ are zero. Multiplying by $b-a$, we get $f(b)-f(a) = (b-a)\,0 = 0$, so $f(b) = f(a)$. That is, all the outputs of $f(x)$ are equal, and $f(x)$ is constant.

(b) Assume the hypothesis $f'(x) = g'(x)$ for all $x \in (a, b)$. Now the function $h(x) = f(x) - g(x)$ has $h'(x) = f'(x) - g'(x) = 0$, so we can apply part (a) to conclude that $h(x)$ is constant, $h(x) = f(x) - g(x) = C$, and $f(x) = g(x) + C$.

(c) In the situation of (b), we also assume $f(c) = g(c)$. By (b), we know $f(x) = g(x) + C$ for all x. In particular for $x = c$, we have $C = f(c) - g(c) = 0$, so $f(x) = g(x) + C = g(x)$, Q.E.D.

To see the significance of this theorem, recall from §2.7 the Ballistic Equation, which gives the height $s(t)$ of an object thrown straight up from initial height $s(0) = s_\circ$, with initial velocity $s'(0) = v_\circ$. The constant gravitational acceleration $-g$ means the velocity should steadily decrease with slope $-g$, so that $s'(t) = v_\circ - gt$. Now, the function

$$s(t) = s_\circ + v_\circ t - \tfrac{1}{2} g t^2$$

does indeed satisfy all these conditions.

But does this guarantee we have the correct function $s(t)$? What if there were some other function $\widetilde{s}(t)$ with the *same derivative* $\widetilde{s}'(t) = s'(t)$ and the *same initial value* $\widetilde{s}(0) = s(0)$? Then $\widetilde{s}(t)$ would be just as good a candidate to give the height of the object, and our mathematical theory would not produce a clear physical prediction. However, the Uniqueness Theorem (c) shows that $\widetilde{s}(t) = s(t)$: there is only one mathematical solution to the equation.

Experiment shows that objects launched in exactly the same way always fly the same way, not according to $s(t)$ in some experiments and a different $\widetilde{s}(t)$ in other experiments. This is what we mean by *physical law.* Our Theorem shows the mathematical solution has the same uniqueness as experimental results.[§]

[§]The theory of quantum mechanics, however, which explains atomic-scale phenomena, goes beyond the framework of deterministic laws, incorporating randomness not just as error (experiments are never perfectly controlled), but as an essential part of the setup. It requires a yet higher mathematical theory, in which we apply calculus not to specific positions of objects, but to probability distributions on all possible positions.

Derivative controls direction. We can similarly use the MVT to prove the expected ways that the sign of the derivative $f'(x)$ should control the increasing/decreasing behavior of the function $f(x)$. For example:

Theorem: A function with positive derivative is increasing. That is, if $f'(x) > 0$ for all x in an interval, then $f(a) < f(b)$ for any $a < b$ within the interval.

Proof: Assume $f'(x) > 0$ for all $x \in (a, b)$. By MVT, $f'(c) = \frac{f(b)-f(a)}{b-a}$ for some $c \in (a, b)$, so $f(b) - f(a) = (b-a)\, f'(c) > 0$ by assumption, and $f(b) > f(a)$.

Cauchy Mean Value Theorem. This is a generalized form of the MVT involving two functions $f(x)$ and $g(x)$, stating the quotient $\frac{\Delta f}{\Delta g} = \frac{f(b)-f(a)}{g(b)-g(a)}$ is equal to $\frac{f'(c)}{g'(c)}$ at some point c, provided the denominators are non-zero. Or in multiplied-out form:

Theorem. If $f(x), g(x)$ are continuous on $[a, b]$, differentiable on (a, b), then there is some $c \in (a, b)$ with

$$(f(b)-f(a))\, g'(c) \;=\; (g(b)-g(a))\, f'(c).$$

Proof. Apply Rolle's Theorem to $h(x) = (f(b)-f(a))g(x) - (g(b)-g(a))f(x)$, which has $h(a) = h(b)$. Then $h'(c) = 0$ gives the formula.

Proof of Linear Approximation Error Estimate. As a final application of MVT, consider as in §2.9 the linear approximation of a function $f(x)$ centered at $x = a$:

$$f(x) \;\approx\; L_a(x) \;=\; f(a) + f'(a)(x-a).$$

According to the Linear Approximation Theorem, the error is controlled by the second derivative: if $|f''(x)| \le B$ over an interval $x \in (a-\delta, a+\delta)$, then:

$$|f(x) - L_a(x)| \;\le \tfrac{1}{2} B\, |x-a|^2 \;\text{ for }\; x \in (a-\delta, a+\delta).$$

Proof. In our formulas, we consider x as a variable and a as an unspecified constant, but this is merely a point of view. We may instead hold x as a fixed value and allow a to vary, indicating this by replacing a with the variable t. Then we can take the derivative of the error function with respect to t, while keeping x constant:

$$\varepsilon(t) \;=\; f(x) - L_t(x) \;=\; f(x) - f(t) - f'(t)(x-t)$$
$$\varepsilon'(t) \;=\; 0 - f'(t) - \big(f'(t)(-1) + f''(t)(x-t)\big) \;=\; -f''(t)(x-t).$$

Now apply MVT to the interval $[a, x]$:[¶] there is some $t = c \in (a, x)$ with:

$$\frac{\varepsilon(x) - \varepsilon(a)}{x - a} \;=\; \varepsilon'(c).$$

Considering that $\varepsilon(x) = 0$, we find that:

$$\varepsilon(a) \;=\; -\varepsilon'(c)\,(x-a) \;=\; f''(c)\,(x-t)(x-a).$$

Finally, we use the hypothesis $|f''(c)| < B$ along with $|x-c| < |x-a|$ to get:

$$|f(x) - L_a(x)| = |\varepsilon(a)| < |f''(c)| \cdot |x-c| \cdot |x-a| < B\,|x-a|^2.$$

This is a slightly weaker upper bound than desired, since it is missing the factor of $\frac{1}{2}$.

To get the strong upper bound, we apply the Cauchy Mean Value Theorem to the functions $\varepsilon(t)$ and $g(x) = (x-t)^2$. There is some $c \in (a, x)$ with $\varepsilon'(c)/g'(c) = (\varepsilon(x) - \varepsilon(a))/(g(x) - g(a))$, so that $\varepsilon(a) = -\varepsilon'(c)(x-a)^2/2(x-c) = \frac{1}{2} f''(c)(x-a)^2$.

[¶] This assumes $a < x$, but the argument adapts easily to the other case $x < a$.