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 \( a \leq x \leq b \), 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.

**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 \( c_1 \neq a, b \), 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

*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: \( \frac{df}{dx} \approx \frac{\Delta f}{\Delta x} \). 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.

**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,

\[
g'(x) = f'(x) - 0 - \frac{f(b) - f(a)}{b-a} (1-0) = f'(x) - \frac{f(b) - f(a)}{b-a},
\]

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,

\[
g'(c) = f'(c) - \frac{f(b) - f(a)}{b-a} = 0,
\]

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:

\[
f'(x) = \left(5x^{1/2} - x^{3/2}\right)' = \frac{5}{2} \cdot \frac{1}{2} x^{-1/2} - \frac{3}{2} \cdot \frac{1}{2} x^{1/2}
\]

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:

\[
f'(x) = \frac{5}{2} \cdot \frac{1}{2} x^{-1/2} - \frac{3}{2} \cdot \frac{1}{2} x^{1/2} = \frac{1}{2} ,
\]

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 \( x = c = \left(\frac{-1 + \sqrt{61}}{6}\right)^2 \approx 1.29 \), 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 \( f(x) \) has derivative \( f'(x) = 0 \) for all \( x \in (p, q) \), then \( f(x) = C \) is 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 the same, and \( f(x) \) is a constant function.

(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 a constant function, \( h(x) = f(x) - g(x) = C \), meaning \( f(x) = g(x) + C \) for some constant \( C \).

(c) In the situation of (b), we also assume \( f(c) = g(c) \). By (b), we know \( f(x) = g(x) + C \) and \( C = f(x) - g(x) \) 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 \( s(t) = s_0 + v_0 t - \frac{1}{2}gt^2 \), which gives the height \( s(t) \) of an object thrown straight up from initial height \( s(0) = s_0 \), and an initial velocity \( v_0 \) together with constant gravitational acceleration \(-g\) produce the derivative \( s'(t) = v_0 - gt \).

But does this guarantee we have the correct function \( s(t) \)? What if there were some other function \( \tilde{s}(t) \) with the same derivative \( \tilde{s}'(t) = s'(t) \) and the same initial value \( \tilde{s}(0) = s(0) \)? Then \( \tilde{s}(t) \) 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 \( \tilde{s}(t) = s(t) \): 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 \( \tilde{s}(t) \) in other experiments. This is what we mean by physical law. Our Theorem shows the mathematical solution has the same uniqueness as experimental results.\(^5\)

\(^5\)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, but to probability distributions on all possible positions.
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 $|f''(x)| \leq B$ over an interval $x \in (a-\delta, a+\delta)$, then:

$$|f(x) - L_a(x)| \leq \frac{B}{2}|x-a|^2 \text{ for } x \in (a-\delta, a+\delta).$$

**Proof.** We will only establish a weaker inequality with the constant $B$ instead of $\frac{B}{2}$.

In our formulas, we consider $x$ as a variable and $a$ as an unspecified constant, but this is merely a point of view. We will 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 $t$, recalling that $x$ is a constant:

$$\varepsilon(t) = f(x) - L_t(x) = f(x) - f(t) - f'(t)(x-t)$$

$$\varepsilon'(t) = 0 - f'(t) - f'(t)(-1) - f''(t)(x-t) = -f''(t)(x-t).$$

Now apply MVT to the interval $[a,x]$:

there is some $t \in (a,x)$ with:

$$\frac{\varepsilon(x) - \varepsilon(a)}{x-a} = \varepsilon'(t).$$

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

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

Finally, we use the hypothesis $|f''(t)| < B$ and $|x-t| < |x-a|$ to get:

$$|f(x) - L_a(x)| = |\varepsilon(a)| < |f''(t)| \cdot |x-t| \cdot |x-a| < B |x-a|^2.$$