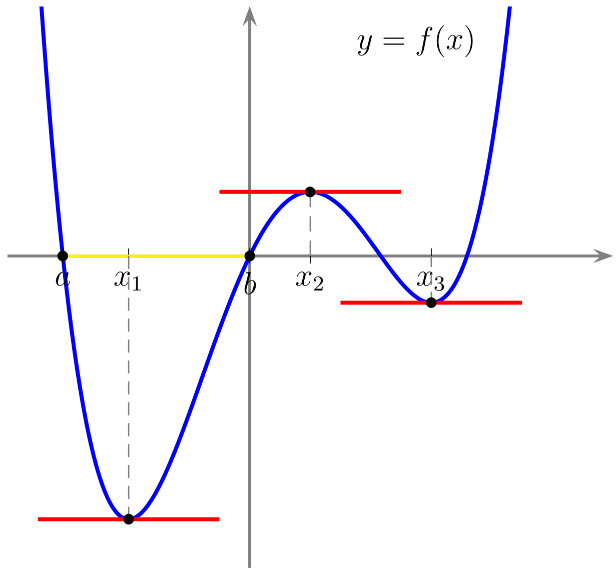
5.29 The Mean Value Theorem



We make the simple observation that at each of the indicated points in the sketch the tangent line has slope 0. If we focus on x_1, observe that $y = f(x)$ has a secant line through the points $(a, 0)$ and $(b, 0)$ that also has a slope of 0. Also notice that $a < x_1 < b$.
We can make this more precise with the following theorem.

Theorem 1. Rolle's Theorem
Suppose that f is continuous on the (closed) interval $[a, b]$ and differentiable on the (open) interval (a, b). Suppose also that $f(a) = f(b)$. Then there is a point $c \in (a, b)$ such that $f'(c) = 0$.

Remark. Notice the hypotheses in this theorem:

> Suppose that f is continuous on the (closed) interval $[a, b]$ and differentiable on the (open) interval (a, b).

This particular assumption will be very common for many of the exercises and theorems throughout the chapter.

Proof. There is no loss in generality in assuming that $f(a) = f(b) = 0$. If $f(x) = 0$ for all $x \in [a, b]$ there is nothing to prove for then $f'(x) = 0$ and we are done.
By Theorem 1 from section 4.1, f attains its minimum and maximum values on $[a, b]$. So there are points x_1 and x_2 such that

$$m = f(x_1) \leq f(x) \leq f(x_2) = M$$

for all $x \in [a, b]$.
If f is **not** a constant function then m and M are not both zero. In other words, at least one of the points x_1 or x_2 is an interior point. So f has a global (and hence local) extreme value at an interior point, c.
If $f(a) = f(b) = K \neq 0$ then we let $g(x) = f(x) - k$. Then g is is continuous on the (closed) interval $[a, b]$ and differentiable on the (open) interval (a, b). Also, $g(a) = g(b) = k - k = 0$. But $g'(x) = f'(x)$ so by the results above, there is a point $c \in (a, b)$ such that

$$0 = g'(c) = f'(c)$$

□

Example 1. Prove that $x^3 + x - 2 = 0$ has exactly one real root.

The condition that $f(a) = f(b)$ is a bit restrictive. What do we lose by dropping this assumption?

Theorem 2. The Mean Value Theorem
Suppose that f is continuous on the (closed) interval $[a, b]$ and differentiable on the (open) interval (a, b). Then there is a point $c \in (a, b)$ such that

$$f'(c) = \frac{f(b) - f(a)}{b - a} \tag{1}$$

Proof. We need to figure out a way to use the previous theorem. Let

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

Then

$$g(a) = f(a) + 0$$
$$g(b) = f(b) + \frac{f(a) - f(b)}{b - a}(b - a) = f(a)$$

and g is continuous on the (closed) interval $[a, b]$ and differentiable on the (open) interval (a, b). Hence g satisfies the hypotheses of Rolle's Theorem. So there is a $c \in (a, b)$ such

$$0 = g'(c) = f'(c) + \frac{f(a) - f(b)}{b - a}$$

Rearranging this yields

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

□

The Mean Value Inequality

Theorem 3. The Mean Value Inequality
Let $I = (\alpha, \beta)$ be an open interval on $\mathbb{R}$. Suppose that $a, b \in I$ with $a < b$ and let $K \geq 0$. If $f : I \to \mathbb{R}$ is differentiable with $f'(t) \leq K$ for all $t \in I$, then

$$f(b) - f(a) \leq K(b - a) \tag{2}$$

Proof. Let $\varepsilon > 0$. According to Exercise 1.3.8 from the text, it suffices to show that

$$f(b) - f(a) \leq (K + \varepsilon)(b - a). \tag{3}$$

Suppose the conclusion is false. That is, suppose that $f(b) - f(a) > (K + \varepsilon)(b - a)$. Our plan is derive a contradiction using the technique from the proof of the Nested Interval Theorem and Cauchy Criterion. Write $K_\varepsilon = K + \varepsilon$ and let $a_0 = a$ and $b_0 = b$. Now bisect the interval $[a_0, b_0]$ to obtain the midpoint M_0. Observe that

$$\begin{aligned} 0 &< f(b_0) - f(a_0) - K_\varepsilon(b_0 - a_0) \\ &= f(b_0) - f(M_0) - K_\varepsilon(b_0 - M_0) + f(M_0) - f(a_0) - K_\varepsilon(M_0 - a_0). \end{aligned}$$

Now at least one of the following expressions is positive.

$$f(M_0) - f(a_0) - K_\varepsilon(M_0 - a_0)$$
$$f(b_0) - f(M_0) - K_\varepsilon(b_0 - M_0)$$

If the first expression is positive, let $a_1 = a_0$ and $b_1 = M_0$. Otherwise, let $a_1 = M_0$ and $b_1 = b_0$. In either case, we have created a new (nested) interval $I_1 = [a_1, b_1]$ such that

$$f(b_1) - f(a_1) > K_\varepsilon(b_1 - a_1) \quad \text{and} \quad I_1 \subset I_0 = [a_0, b_0]$$

Now continue the process to create a sequence of intervals $\{I_n\}$ with the following properties.

$$f(b_n) - f(a_n) > K_\varepsilon(b_n - a_n) \quad \text{and} \quad I_n \subset I_{n-1}. \tag{4}$$

Once again, notice that the a_n form a bounded increasing sequence and the b_n form a bounded decreasing sequence. So by the Monotone Convergence Theorem, both sequences converge. Now since $b_n - a_n \to 0$ as $n \to \infty$, we have

$$\lim_{n\to\infty} b_n = \lim_{n\to\infty} (b_n - a_n + a_n) = \lim_{n\to\infty} (b_n - a_n) + \lim_{n\to\infty} a_n = \lim_{n\to\infty} a_n$$

That is, both sequences have a common limit $c \in I$. Now by the continuity of f on I,

$$f(c) = \lim_{n\to\infty} f(a_n) = \lim_{n\to\infty} f(b_n)$$

Since f is differentiable at c, there is a $\delta > 0$ with $(c-\delta, c+\delta) \subset (a,b)$ such

$$\left| \frac{f(c+h)-f(c)}{h} - f'(c) \right| < \varepsilon/2$$

or

$$|f(c+h)-f(c)-f'(c)h| < |h|\varepsilon/2$$

provided $0 < |h| < \delta$. The last inequality is equivalent to

$$\frac{-|h|\varepsilon}{2} < f(c+h)-f(c)-f'(c)h < \frac{|h|\varepsilon}{2}$$

Now suppose that $0 < h < \delta$. Focusing on the right inequality, we see that

$$\begin{aligned} f(c+h)-f(c) &< f'(c)h + \frac{|h|\varepsilon}{2} \\ &= f'(c)h + \frac{h\varepsilon}{2} \\ &\le (K+\varepsilon/2)h \qquad (5) \end{aligned}$$

On the other hand, if $-\delta < h < 0$, we can use the left inequality to obtain

$$\frac{h\varepsilon}{2} = \frac{-|h|\varepsilon}{2} < f(c+h)-f(c)-f'(c)h$$

Rearranging yields

$$\begin{aligned} f(c)-f(c+h) &< (f'(c)+\varepsilon/2)(-h) \\ &\le (K+\varepsilon/2)(-h) \qquad (6) \end{aligned}$$

Now $a_n \nearrow c$ implies that there exists $N_1 \in \mathbb{N}$ such that $c - a_n < \delta$ for all $n \ge N_1$. Likewise, $b_n \searrow c$ implies that there exists $N_2 \in \mathbb{N}$ such that $n \ge N_2$ implies $b_n - c < \delta$. As usual let $N = \max\{N_1, N_2\}$. Then

$$f(c)-f(a_N) < (K+\varepsilon/2)(c-a_N)$$
$$f(b_N)-f(c) < (K+\varepsilon/2)(b_N-c)$$

are immediate consequences of (5) and (6), respectively. Thus

$$\begin{aligned} f(b_N)-f(a_N) &= f(b_N)-f(c)+f(c)-f(a_N) \\ &< (K+\varepsilon/2)(b_N-c)+(K+\varepsilon/2)(c-a_N) \\ &= (K+\varepsilon/2)(b_N-a_N) \end{aligned}$$

contrary to (4). The result follows. □