In Chapter 4 of the textbook we used some basic facts without proof: for example, that a continuous function has a maximum on a bounded, closed interval, or that a function whose derivative is positive on an interval is increasing on that interval.

From a geometric point of view, these facts seem obvious. If we draw the graph of a continuous function, starting at one end of a bounded, closed interval and going to the other, it seems obvious that we must pass a highest point on the way. And if the derivative of a function is positive, then its graph must be sloping up, so the function has to be increasing.

However, this sort of graphical reasoning is not a rigorous proof, for two reasons. First, no matter how many pictures we imagine, we can’t be sure we have covered all possibilities. Second, our pictures often depend on the theorems we are trying to prove.

A Continuous Function on a Closed Interval Has a Maximum

The Extreme Value Theorem

If \( f \) is continuous on the interval \([a, b]\), then \( f \) has a global maximum and a global minimum on that interval.

Our proof has two parts: The first is to show that \( f \) has an upper bound on \([a, b]\), the second is to show that if \( f \) has an upper bound then it has a global maximum on the interval. Here we prove the second part; the first part is proved in Problems 13 and 14. In Problem 4 we extend the result from maxima to minima.

Proof

We assume that \( f \) is continuous and has an upper bound on the interval \([a, b]\). This means, by the Completeness Axiom, that \( f \) has a least upper bound \( M \) on \([a, b]\). We are going to prove that there is a number \( c \) in the interval \([a, b]\) such that \( f(c) = M \).

Divide \([a, b]\) into two halves. Then, on one of the halves, the least upper bound for \( f \) is \( M \), for if it were less than \( M \) on both halves, it would be less than \( M \) on the whole. Choose a half on which the least upper bound is equal to \( M \). Continue bisecting and at each stage choose the half-interval where the least upper bound for \( f \) is \( M \). See Figure F.20. This results in a sequence of nested intervals. By the Nested Interval Theorem on page 10, there is a number \( c \) in \([a, b]\) which is contained in all these intervals.

Since \( M \) is the least upper bound for \( f \), we have \( f(c) \leq M \). We are next going to prove that \( f(c) \) is not less than \( M \). For if \( f(c) < M \), then \( f(c) < M_0 \) for some number \( M_0 < M \). (For example, we could take \( M_0 \) to be half-way between \( M \) and \( f(c) \).) But then, since \( f \) is continuous, there would be a \( \delta > 0 \) such that \( f(x) < M_0 \) for all \( x \) in \([a, b]\) with \( c - \delta < x < c + \delta \). (See Figure F.20: Successively choosing the half-interval where the least upper bound of \( f \) is \( M \).)
Problem 12.) Since the nested intervals we constructed above have width tending to zero, one of them would be contained in the interval $c - \delta < x < c + \delta$. Therefore, $f$ would be bounded above by $M_b$ on one of the nested intervals. However, we chose each nested interval so that the least upper bound for $f$ is $M$. This is a contradiction since $M_b < M$.

So it is not possible that $f(c) < M$; we must have $f(c) = M$. Thus, $M$ is the global maximum of $f$ on $[a, b]$, which is what we wanted to show.

The Extreme Value Theorem guarantees the existence of global maxima (and minima) on an interval. To actually find the global maxima, we look at all the local maxima. The following theorem tells us that inside an interval, local maxima only occur at critical points, where the derivative is either zero or undefined.

**Theorem: Local Extrema and Critical Points**

Suppose $f$ is defined on an interval and has a local maximum or minimum at the point $x = a$, which is not an endpoint of the interval. If $f$ is differentiable at $x = a$, then $f'(a) = 0$.

**Proof**

We start with the definition of the derivative:

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

Remember that this is a two-sided limit:

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

Suppose that $f$ has a local maximum at $x = a$. By the definition of local maximum, $f(a+h) \leq f(a)$ for all sufficiently small $h$. Thus $f(a+h) - f(a) \leq 0$ for sufficiently small $h$. The denominator, $h$, is positive when we take the limit from the right and negative when we take the limit from the left. Thus

$$\lim_{h \to 0^-} \frac{f(a+h) - f(a)}{h} \geq 0 \quad \text{and} \quad \lim_{h \to 0^+} \frac{f(a+h) - f(a)}{h} \leq 0.$$ 

Since both these limits are equal to $f'(a)$, we have $f'(a) \geq 0$ and $f'(a) \leq 0$, so we must have $f'(a) = 0$.

**A Relationship Between Local and Global: The Mean Value Theorem**

We often want to infer a global conclusion (for example, $f$ is increasing on an interval) from local information ($f'$ is positive.) The following theorem relates the average rate of change of a function on an interval (global information) to the instantaneous rate of change at a point in the interval (local information).

**The Mean Value Theorem**

If $f$ is continuous on $[a, b]$ and differentiable on $(a, b)$, then there exists a number $c$, with $a < c < b$, such that

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

In other words, $f(b) - f(a) = f'(c)(b - a)$. 

To understand what this theorem is saying geometrically, consider the graph in Figure F.21. Join the points on the curve where \( x = a \) and \( x = b \) with a line and observe that the slope of this secant line \( AB \) is given by

\[
m = \frac{f(b) - f(a)}{b - a}.
\]

Now consider the tangent line drawn to the curve at each point between \( x = a \) and \( x = b \). In general, these lines will have different slopes. For the curve shown in Figure F.21, the tangent line at \( x = a \) is flatter than the secant line from \( A \) to \( B \). Similarly, the tangent line at \( x = b \) is steeper than the secant line. However, there is at least one point between \( a \) and \( b \) where the slope of the tangent line to the curve is precisely the same as the slope of the secant line. Suppose this occurs at \( x = c \). Then

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

The Mean Value Theorem tells us that the point \( x = c \) exists, but it does not tell us how to find \( c \).

Problems 15 and 16 show how the Mean Value Theorem can be deduced from the Extreme Value Theorem.

\[\text{Figure F.21: The point } c \text{ with } f'(c) = \frac{f(b) - f(a)}{b - a}\]

The Increasing Function Theorem

We say that a function \( f \) is increasing on an interval if, for any two numbers \( x_1 \) and \( x_2 \) in the interval such that \( x_1 < x_2 \), we have \( f(x_1) < f(x_2) \). If instead we have \( f(x_1) \leq f(x_2) \), we say \( f \) is nondecreasing.

The Increasing Function Theorem

Suppose that \( f \) is continuous on \([a, b] \) and differentiable on \((a, b)\).

- If \( f'(x) > 0 \) on \((a, b)\), then \( f \) is increasing on \([a, b]\).
- If \( f'(x) \geq 0 \) on \((a, b)\), then \( f \) is nondecreasing on \([a, b]\).

\[\text{Proof: Suppose } a \leq x_1 < x_2 \leq b. \text{ By the Mean Value Theorem, there is a number } c, \text{ with } x_1 < c < x_2, \text{ such that}
\]

\[f(x_2) - f(x_1) = f'(c)(x_2 - x_1).
\]

If \( f'(c) > 0 \), this says \( f(x_2) - f(x_1) > 0 \), which means \( f \) is increasing. If \( f'(c) \geq 0 \), this says \( f(x_2) - f(x_1) \geq 0 \), which means \( f \) is nondecreasing.

It may seem that something as simple as the Increasing Function Theorem should follow immediately from the definition of the derivative, and that the use of the Mean Value Theorem (which in turn depends on the Extreme Value Theorem) is surprising. It is possible to give a proof which does not use the Mean Value Theorem, but not a simple one.
The Constant Function Theorem

If \( f \) is constant on an interval, then we know that \( f'(x) = 0 \) on the interval. The following theorem is the converse.

**The Constant Function Theorem**

Suppose that \( f \) is continuous on \([a, b]\) and differentiable on \((a, b)\). If \( f'(x) = 0 \) on \((a, b)\), then \( f \) is constant on \([a, b]\).

**Proof** The proof is the same as for the Increasing Function Theorem, only in this case \( f'(c) = 0 \) so \( f(x_2) - f(x_1) = 0 \). Thus \( f(x_2) = f(x_1) \) for \( a \leq x_1 < x_2 \leq b \), so \( f \) is constant.

A proof of the Constant Function Theorem using the Increasing Function Theorem is given in Problems 5 and 7.

The Racetrack Principle

**The Racetrack Principle**

Suppose that \( g \) and \( h \) are continuous on \([a, b]\) and differentiable on \((a, b)\), and that \( g'(x) \leq h'(x) \) for \( a < x < b \).

- If \( g(a) = h(a) \), then \( g(x) \leq h(x) \) for \( a \leq x \leq b \).
- If \( g(b) = h(b) \), then \( g(x) \geq h(x) \) for \( a \leq x \leq b \).

The Racetrack Principle has the following interpretation. We can think of \( g(x) \) and \( h(x) \) as the positions of two racehorses at time \( x \), with horse \( h \) always moving faster than horse \( g \). If they start together, horse \( h \) is ahead during the whole race. If they finish together, horse \( g \) was ahead during the whole race.

**Proof** Consider the function \( f(x) = h(x) - g(x) \). Since \( f'(x) = h'(x) - g'(x) \geq 0 \), we know that \( f \) is nondecreasing by the Increasing Function Theorem. So \( f(x) \geq f(a) = h(a) - g(a) = 0 \). Thus \( g(x) \leq h(x) \) for \( a \leq x \leq b \). This proves the first part of the Racetrack Principle. Problem 6 asks for a proof of the second part.

---

**Example 1**

Explain graphically why \( e^x \geq 1 + x \) for all values of \( x \). Then use the Racetrack Principle to prove the inequality.

**Solution**

The graph of the function \( f(x) = e^x \) is concave up everywhere and the equation of its tangent line at the point \((0, 1)\) is \( y = x + 1 \). (See Figure F.22.) Since the graph always lies above its tangent, we have the inequality

\[ e^x \geq 1 + x. \]

---

\(^5\)Based on the Racetrack Principle in *Calculus&Mathematica*, by William Davis, Horacio Porta, Jerry Uhl (Reading: Addison Wesley, 1994).
Now we prove the inequality using the Racetrack Principle. Let \( g(x) = 1 + x \) and \( h(x) = e^x \). Then \( g(0) = h(0) = 1 \). Furthermore, \( g'(x) = 1 \) and \( h'(x) = e^x \). Hence \( g'(x) \leq h'(x) \) for \( x \geq 0 \).

So by the Racetrack Principle, with \( a = 0 \), we have \( g(x) \leq h(x) \), that is, \( 1 + x \leq e^x \).

For \( x \leq 0 \) we have \( h'(x) \leq g'(x) \). So by the Racetrack Principle, with \( b = 0 \), we have \( g(x) \leq h(x) \), that is, \( 1 + x \leq e^x \).

\[ \text{Figure F.22: Graph showing that } e^x \geq 1 + x \]

Problems for Section F

1. Use the Racetrack Principle and the fact that \( \sin 0 = 0 \) to show that \( \sin x \leq x \) for all \( x \geq 0 \).

2. Use the Racetrack Principle to show that \( \ln x \leq x - 1 \).

3. Suppose that the position of a particle moving along the \( x \)-axis is given by \( s = f(t) \), and that the initial position and velocity of the particle are \( f(0) = 3 \) and \( f'(0) = 4 \). Suppose that the acceleration is bounded by \( 5 \leq f''(t) \leq 7 \) for \( 0 \leq t \leq 2 \). What can we say about the position \( f(2) \) of the particle at \( t = 2 \)?

4. Show that if every continuous function on an interval \([a, b]\) has a global maximum, then every continuous function has a global minimum as well. [Hint: Consider \(-f\).]

5. State a Decreasing Function Theorem, analogous to the Increasing Function Theorem. Deduce your theorem from the Increasing Function Theorem. [Hint: Apply the Increasing Function Theorem to \(-f\).]

6. Suppose that \( g \) and \( h \) are continuous on \([a, b]\) and differentiable on \((a, b)\). Prove that if \( g'(x) \leq h'(x) \) for \( a < x < b \) and \( g(b) = h(b) \), then \( h(x) \leq g(x) \) for \( a \leq x \leq b \).

7. Deduce the Constant Function Theorem from the Increasing Function Theorem and the Decreasing Function Theorem (see problem 5).

8. Prove that if \( f'(x) = g'(x) \) for all \( x \) in \((a, b)\), then there is a constant \( C \) such that \( f(x) = g(x) + C \) on \((a, b)\). [Hint: Apply the Constant Function Theorem to \( h(x) = f(x) - g(x) \).]

9. Suppose that \( f'(x) = f(x) \) for all \( x \). Prove that \( f(x) = Ce^x \) for some constant \( C \). [Hint: Consider the function \( f(x)/e^x \).

10. Suppose that \( f \) is a continuous function on \([a, b]\) and differentiable on \((a, b)\) and that \( m \leq f'(x) \leq M \) on \((a, b)\). Use the Racetrack Principle to prove that \( f(x) - f(a) \leq M(x - a) \) for all \( x \) in \([a, b]\), and that \( m(x - a) \leq f(x) - f(a) \) for all \( x \) in \([a, b]\). Conclude that \( m \leq (f(b) - f(a))/(b - a) \leq M \). This is called the Mean Value Inequality. In words: If the instantaneous rate of change of \( f \) is between \( m \) and \( M \) at every point in an interval, so is the average rate of change of \( f \) over the interval.

11. Suppose that \( f''(x) \geq 0 \) for all \( x \) in \((a, b)\). We will show the graph of \( f \) lies above the tangent line at \((c, f(c))\) for any \( c \) with \( a < c < b \).

(a) Use the Increasing Function Theorem to prove that \( f'(c) \leq f'(x) \) for \( c \leq x < b \) and that \( f'(c) \leq f'(x) \) for \( a < x \leq c \).

(b) Use (a) and the Racetrack Principle to conclude that \( f(c) + f'(c)(x - c) \leq f(x) \), for \( a < x < b \).

12. Suppose that \( f \) is a continuous function on \([a, b]\), and let \( c \) be in \([a, b]\). Show that if \( f(c) < M \), then there is a \( \delta \) such that \( f(x) < M \) for all \( x \) in \([a, b]\) such that \( c - \delta < x < c + \delta \). [Hint: Let \( \epsilon = M - f(c) \), and choose \( \delta \) such that \( |f(x) - f(c)| < \epsilon \) if \( |x - c| < \delta \).]

On page 29 we proved that a continuous function \( f \) has a global maximum on the interval \([a, b]\) under the assumption
that $f$ has an upper bound on $[a, b]$. Problems 13–14 give two different proofs of this assumption.

13. (a) Suppose that $f$ has no upper bound on $[a, b]$. Bisect $[a, b]$ into two halves. Deduce that $f$ has no upper bound on at least one of the halves. Call that half $[a_1, b_1]$.

(b) Continue bisecting so that at the $n$th stage you obtain an interval $[a_n, b_n]$ on which $f$ has no upper bound. By the Nested Interval Theorem on page 10, there is a point $c$ in all the intervals $[a_n, b_n]$.

(c) Use continuity of $f$ at $c$ to deduce that $f$ has an upper bound on $[a_n, b_n]$ for $n$ sufficiently large. This contradicts the original supposition, so $f$ must have an upper bound on $[a, b]$.

14. (a) Show that if $y \geq 0$, then $y/(1 + y) < 1$.

(b) Suppose that $f$ is continuous on $[a, b]$ and that $f(x) \geq 0$ on $[a, b]$. Define a function $g$ by $g(x) = f(x)/(1 + f(x))$. Show that $g$ is continuous and bounded on $[a, b]$. It follows from the partial proof of the Extreme Value Theorem on page 29 that $g$ has a global maximum on $[a, b]$ at some point $x = c$.

(c) Suppose that $y_1 \geq 0$ and $y_2 \geq 0$, and that $y_1/(1 + y_1) \leq y_2/(1 + y_2)$. Show that $y_1 \leq y_2$.

(d) Use parts (c) and (d) to show that $f$ has a global maximum at $x = c$.

(e) We have shown that if $f$ is continuous and non-negative on $[a, b]$, then it is bounded above on $[a, b]$. Now suppose that $f$ is continuous, but not necessarily non-negative. By applying the argument to $|f|$, deduce that $f$ is also bounded above.

15. In this problem we prove a special case of the Mean Value Theorem where $f(a) = f(b) = 0$. This special case is called Rolle’s Theorem: If $f$ is continuous on $[a, b]$ and differentiable on $(a, b)$, and if $f(a) = f(b) = 0$, then there is a number $c$, with $a < c < b$, such that

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

By the Extreme Value Theorem, $f$ has a global maximum and a global minimum on $[a, b]$.

16. Use Rolle’s Theorem to prove the Mean Value Theorem. Suppose that $f(x)$ is continuous on $[a, b]$ and differentiable on $(a, b)$.

(a) Prove Rolle’s theorem in the case that both the global maximum and the global minimum are at endpoints of $[a, b]$. [Hint: $f(x)$ must be a very simple function in this case.]

(b) Prove Rolle’s theorem in the case that either the global maximum or the global minimum is not at an endpoint. [Hint: Think about local maxima and minima.]

16. Use Rolle’s Theorem to prove the Mean Value Theorem. Suppose that $f(x)$ is continuous on $[a, b]$ and differentiable on $(a, b)$.

(a) Let $g(x)$ be the difference between $f(x)$ and the $y$-value on the secant line joining $(a, f(a))$ to $(b, f(b))$. See Figure F.23. Show that

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

(b) Use Rolle’s Theorem to show that there must be a point $c$ in $(a, b)$ such that $g'(c) = 0$.

(c) Show that if $c$ is the point in part (b), then

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