The Mean Value Theorem and Its Consequences

1 Minima and Maxima

The point \((M, f(M))\) is called an **absolute maximum** of \(f\) if \(f(x) \leq f(M)\) for every \(x\) in the domain of \(f\). The point \((m, f(m))\) is called an **absolute minimum** of \(f\) if \(f(x) \geq f(m)\) for every \(x\) in the domain of \(f\).

More than one absolute maximum or minimum may exist. For example, if \(f(x) = |x|\) for \(x \in [-1, 1]\) then \(f(x) \leq 1\) and there are absolute maxima at \((1, 1)\) and at \((-1, 1)\), but only one absolute minimum, at \((0,0)\).

Recall that if \(f\) is a continuous function with domain \([a,b]\) then there is some \(M \in [a,b]\) such that \(f(x) \leq f(M)\) for all \(x \in [a,b]\) and there is some \(m \in [a,b]\) such that \(f(x) \geq f(m)\). In fact, what we know is that the range of \(f\) is \([f(m), f(M)]\).

We want to see what additional information the differentiability of \(f\) gives us.

2 A convention to remember

If the domain of \(f\) is \([a,b]\) and

\[
\frac{f(x) - f(a)}{x - a}
\]

has a limit as \(x\) approaches \(a\) from above, then we say that \(f\) is differentiable at \(a\) and we write \(f'(a)\) for the derivative. Similarly, if

\[
\frac{f(x) - f(b)}{x - b}
\]

has a limit as \(x\) approaches \(b\) from below, then we say that \(f\) is differentiable at \(b\) and we write \(f'(b)\) for the derivative.

3 Derivatives, Maxima and Minima

The following is the important observation from which all the rest of our results will follow.

**Theorem 1** Suppose that the domain of \(f\) is \([a,b]\) and \(f'(c) = d \neq 0\). Then the following are true:

**Case I:** If \(a < c < b\) then \((c, f(c))\) is neither an absolute maximum nor an absolute minimum.

**Case IIA:** If \(c = a\) and \(f'(c) > 0\) then \((c, f(c))\) is not an absolute maximum.

**Case IIB:** If \(c = a\) and \(f'(c) < 0\) then \((c, f(c))\) is not an absolute minimum.

**Case IIIA:** If \(c = b\) and \(f'(c) > 0\) then \((c, f(c))\) is not an absolute minimum.

**Case IIIB:** If \(c = b\) and \(f'(c) < 0\) then \((c, f(c))\) is not an absolute maximum.

**Reason:** The arguments hinge on the following observation: If

\[
\lim_{x \to a} g(x) = G \neq 0
\]

then there is some \(d > 0\) so that if \(x\) is in the domain of \(g\) and \(0 < |x - u| < d\) then \(g(x)\) and \(G\) have the same sign. To see why, apply the definition of limit with the tolerance \(t = |G|/2\).

Now we work our way through the cases.

In **Case I** suppose that \(f'(c) > 0\) then there is some \(u > c\) so that

\[
\frac{f(u) - f(c)}{u - c} > 0.
\]

Since \(u - c > 0\) we know that \(f(u) - f(c) > 0\). This means \(f(u) > f(c)\) so \((c, f(c))\) is not an absolute maximum. There is also some \(v < c\) so that

\[
\frac{f(v) - f(c)}{v - c} > 0.
\]
Since \( v - c < 0 \) we know that \( f(v) - f(c) < 0 \). This means \( f(v) < f(c) \) so \((c, f(c))\) is not an absolute minimum.

The other possibility is that \( f'(c) < 0 \). Then there is some \( u > c \) so that

\[
\frac{f(u) - f(c)}{u - c} < 0.
\]

Since \( u - c > 0 \) we know that \( f(u) - f(c) < 0 \). This means \( f(u) < f(c) \) so \((c, f(c))\) is not an absolute minimum. There is also some \( v < c \) so that

\[
\frac{f(v) - f(c)}{v - c} < 0.
\]

Since \( v - c < 0 \) we know that \( f(v) - f(c) > 0 \). This means \( f(v) > f(c) \) so \((c, f(c))\) is not an absolute minimum.

In Case IIA since \( c = a \) and \( f'(c) > 0 \) there is some \( u > c \) so that

\[
\frac{f(u) - f(c)}{u - c} > 0.
\]

Since \( u - c > 0 \) we have \( f(u) > f(c) \) so \((c, f(c))\) cannot be an absolute maximum.

In Case IIB since \( c = a \) and \( f'(c) < 0 \) there is some \( u > c \) so that

\[
\frac{f(u) - f(c)}{u - c} < 0.
\]

Since \( u - c > 0 \) we have \( f(u) < f(c) \) so \((c, f(c))\) cannot be an absolute minimum.

In Case IIIA since \( c = b \) and \( f'(c) > 0 \) there is some \( v < c \) so that

\[
\frac{f(v) - f(c)}{v - c} > 0.
\]

Since \( v - c < 0 \) we have \( f(v) < f(c) \) so \((c, f(c))\) cannot be an absolute minimum.

In Case IIIB since \( c = b \) and \( f'(c) < 0 \) there is some \( v < c \) so that

\[
\frac{f(v) - f(c)}{v - c} < 0.
\]

Since \( v - c < 0 \) we have \( f(v) > f(c) \) so \((c, f(c))\) cannot be an absolute maximum.

**QED**

**Corollary 1** Suppose that \( f \) is a continuous function with domain \([a, b]\). If \((M, f(M))\) is an absolute maximum, then one of the following is true.

- \( M = a \);
- \( M = b \);
- \( f \) is not differentiable at \( M \);
- \( f \) is differentiable at \( M \) and \( f'(M) = 0 \).

If \((m, f(m))\) is an absolute minimum, then one of the following is true.

- \( m = a \);
- \( m = b \);
- \( f \) is not differentiable at \( m \);
- \( f \) is differentiable at \( m \) and \( f'(m) = 0 \).

**Reason:** The only possibility not addressed is that \( f \) is differentiable at \( M \) or \( m \) with a non-zero derivative. In that case we have neither an absolute maximum nor an absolute minimum.

**QED**
3.1 A typical application

Suppose that \( f(x) = 2x^3 - 3x^2 + 2 \) with domain \([-2, 2]\). Locate the absolute maxima and absolute minima.

**Solution:** Since \( f \) is continuous and its domain is a closed interval, there is at least one absolute maximum and one absolute minimum. \( f \) is differentiable on its entire domain, with \( f'(x) = 6x^2 - 6x = 6x(x - 1) \). So the possible location of the absolute maxima and minima are

\[
\{(-2, f(-2)), (0, f(0)), (1, f(1)), (2, f(2))\} = \{(-2, -26), (0, 2), (1, 1), (2, 6)\}
\]

from which it is clear that the absolute maximum is at (2, 6) and the absolute minimum is at (−2, −26).

3.2 An important example

Suppose that \( f(x) = |x| \) with domain \([-2, 3]\). We know that the absolute maximum is (3, 3) and the absolute minimum is (0, 0). Notice that the minimum occurs at a place where the derivative is undefined.

3.3 A geometry problem

Consider the graph of \( y = x^2 \), for \( x \in [-2, 2] \). What point on the graph is closest to (0, 1) and which point is furthest away?

**Solution:** The distance from (0, 1) to \((x, x^2)\) is \( \sqrt{(x - 0)^2 + (x^2 - 1)^2} \). So we want to find the absolute maximum and minimum of the function \( f(x) = \sqrt{(x - 0)^2 + (x^2 - 1)^2} \) with domain \([-2, 2]\).

\( f \) is continuous, so there is an absolute maximum and an absolute minimum, while

\[
f'(x) = \frac{1}{2} (x^2 + (x^2 - 1)^2)^{-1/2} (2x + 2(x^2 - 1)(2x)) = \frac{x(2x^2 - 1)}{\sqrt{x^2 + (x^2 - 1)^2}}
\]

for every \( x \in [-2, 2] \). Therefore the only possible locations for the absolute maxima are

\[
\{(-2, f(-2)), (-1/\sqrt{2}, f(-1/\sqrt{2})), (0, f(0)), (1/\sqrt{2}, f(1/\sqrt{2})), (2, f(2))\}
\]

so we see the closest points are \((-1/\sqrt{2}, 1/2)\) and \((1/\sqrt{2}, 1/2)\), and the most distant points are \((-2, 4)\) and \((2, 4)\).

4 Rolle’s Theorem

Throughout this section we assume \( a < b \).

**Theorem 2 (Rolle’s Theorem)** Suppose that \( f \) is a continuous function with domain \([a, b]\), that \( f(a) = f(b) \) and \( f \) is differentiable at every \( x \in (a, b) \). Then there is at least one \( c \in (a, b) \) where \( f'(c) = 0 \).

**Reason:** There are two possibilities for for \( f \). The first is that \( f \) is constant. Then \( f'(c) = 0 \) for every \( c \in (a, b) \). The second possibility is that \( f \) is not constant. Then it has an absolute maximum \( (M, f(M)) \), an absolute minimum \( (m, f(m)) \) and \( f(m) < f(M) \). This means either \( M \in (a, b) \) or \( m \in (a, b) \). If the former, then \( f'(M) = 0 \) since \( f \) is differentiable at \( M \). If the latter, then \( f'(m) = 0 \) since \( f \) is differentiable at \( m \).

**QED**

Note: The point of Rolle’s theorem is that such a number \( c \) exists. We have no interest in the value of \( c \).

**Corollary 2 (Second Rolle’s Theorem)** Suppose that \( f \) is continuous on \([a, b]\), \( f(b) = f(a) = 0 \), and that \( f''(x) \) exists for each \( x \in (a, b) \). Suppose one of the following is true:

Case I: There is some \( c \in (a, b) \) such that \( f(c) = 0 \)

Case II: \( f'(a) = 0 \) and \( f' \) is continuous at \( a \).
Then there is some \( d \in (a,b) \) so that \( f'(d) = 0 \).

**Reason:** In Case I apply Rolle’s Theorem to \( f \) on \([a,c]\) and \([c,b]\) separately to infer that there are \( u \) and \( v \) with \( a < u < c < v < b \) and \( f'(u) = f'(v) = 0 \). Observe that Rolle’s Theorem now applies to \( f' \) on \([u,v]\).

In Case II apply Rolle’s Theorem on \([a,b]\) to infer that there is some \( u \in (a,b) \) so that \( f'(u) = 0 \). Observe that since \( f'(a) = 0 \) and \( f' \) is continuous on \([a,u]\) that Rolle’s Theorem applies to \( f' \) on \([a,u]\]. QED

## 5 The First Mean Value Theorem

**Theorem 3 (The First Mean Value Theorem)** If \( g \) is a continuous function with domain \([a,b]\) and \( g \) is differentiable at each \( x \in (a,b) \). Then there is at least one \( c \in (a,b) \) so that \( g(b) - g(a) = g'(c)(b-a) \).

**Reason:** Define \( f(x) = (g(x) - g(a))(b - a) - (x - a)(g(b) - g(a)) \). \( f \) is the sum of products of functions continuous on \([a,b]\) and differentiable on \((a,b)\), with

\[
f'(x) = (g'(x) - 0)(b - a) - (1 - 0)(g(b) - g(a)) = g'(x)(b - a) - (g(b) - g(a)).
\]

Furthermore, \( f(a) = f(b) = 0 \), so we may conclude from Rolle’s Theorem that there is some \( c \in (a,b) \) where

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

which is the same as \( g(b) - g(a) = g'(c)(b - a) \). QED

Again, the point of this theorem is that such a \( c \) exists, not to calculate it. For example:

**Corollary 3** Suppose that \( f \) is differentiable on \((a,b)\).

- If \( f'(x) = 0 \) for each \( x \in (A,B) \) then \( f \) is constant on \((A,B)\).
- If \( f'(x) > 0 \) for each \( x \in (A,B) \) then \( f \) is strictly increasing on \((A,B)\).
- If \( f'(x) \geq 0 \) for each \( x \in (A,B) \) then \( f \) is non-decreasing on \((A,B)\).
- If \( f'(x) < 0 \) for each \( x \in (A,B) \) then \( f \) is strictly decreasing on \((A,B)\).
- If \( f'(x) \leq 0 \) for each \( x \in (A,B) \) then \( f \) is non-increasing on \((A,B)\).

**Reason:** Each case is similar.

- If \( A < a < b < B \) then \( f(b) - f(a) = f'(c)(b - a) = 0 \), so \( f(b) = f(a) \).
- If \( A < a < b < B \) then \( f(b) - f(a) = f'(c)(b - a) > 0 \), so \( f(b) > f(a) \).
- If \( A < a < b < B \) then \( f(b) - f(a) = f'(c)(b - a) \geq 0 \), so \( f(b) \geq f(a) \).
- If \( A < a < b < B \) then \( f(b) - f(a) = f'(c)(b - a) < 0 \), so \( f(b) < f(a) \).
- If \( A < a < b < B \) then \( f(b) - f(a) = f'(c)(b - a) \leq 0 \), so \( f(b) \leq f(a) \).

That is all we need to show. QED

Note that the first assertion says that if two function agree at one point of an interval and their derivatives agree at all points of the interval, then the functions agree on the interval. This extends the idea that parallel lines that cross are the same line, since parallel lines have the same slope, and the derivative gives the slope at a point.
5.1 Exponential Growth and Decay

We learned in algebra that in certain situations a function of the form \( f(t) = A \exp(rt) \) describes the phenomenon in question. For example, suppose that at time 0 a Petri dish has a colony consisting of 10 micrograms of Listeria bacteria, and 5 hours later there are 13 micrograms of this bacteria. How much bacteria is present at time \( t \) hours.

The underlying principle in these problems was that the instantaneous rate of change of the quantity in question was assumed to be proportional to the amount of the quantity present. We can now express this concisely using derivatives:

\[
\frac{d}{dt} Q(t) = r Q(t)
\]

where \( Q(t) \) is the amount of the quantity in question, and \( r \) is the proportionality constant. If we suppose that \( Q(0) \neq 0 \) is given then we can use the Mean Value Theorem to show that the only solution to our problem is \( Q(t) = Q(0) \exp(rt) \).

It is clear from the Chain Rule that \( Q(t) \) is one solution, since

\[
\frac{d}{dt} Q(t) = \frac{d}{dt} Q(0) \exp(rt) = Q(0) \frac{d}{dt} \exp(rt) = Q(0) \exp(rt) \frac{d}{dt}(rt) = Q(0) \exp(rt) \cdot r = r Q(t)
\]

and \( Q(0) = Q(0) \exp(0) \). Suppose now that \( f \) is another solution, that is \( f(0) = Q(0) \) and \( f'(t) = r f(t) \). We will show that \( f(t)/Q(t) = 1 \) for all \( t \). First, \( f(0)/Q(0) = Q(0)/Q(0) = 1 \). Second,

\[
\frac{d}{dt} \frac{f(t)}{Q(t)} = \frac{f'(t)Q(t) - Q'(t)f(t)}{Q(t)^2}.
\]

Since \( Q'(t) = r Q(t) \) and \( f'(t) = r f(t) \) we have

\[
\frac{d}{dt} \frac{f(t)}{Q(t)} = \frac{r f(t)Q(t) - r Q(t)f(t)}{Q(t)^2} = 0
\]

showing that \( f(t)/Q(t) = f(0)/Q(0) = 1 \).

5.2 Another look at Hooke’s Law

We examined the situation where we wanted to find a function \( f \) with the following properties:

\[
\begin{align*}
f''(t) &= -r^2 f(t) \\
f(0) &= A \\
f'(0) &= B
\end{align*}
\]

where \( r \neq 0 \).

We observed that \( f(t) = A \cos(rt) + B \sin(rt) \) fits the bill. We will now show that this is the only such function. Suppose that we also have

\[
\begin{align*}
g''(t) &= -r^2 g(t) \\
g(0) &= A \\
g'(0) &= B.
\end{align*}
\]

We will show that \( f(t) - g(t) = 0 \) for any \( t \). Let \( h(t) = f(t) - g(t) \). It is easy to check that

\[
\begin{align*}
h''(t) &= -r^2 h(t) \\
h(0) &= 0 \\
h'(0) &= 0.
\end{align*}
\]

Now put \( E(t) = (rh(t))^2 + (h'(t))^2 \). We have \( E(0) = 0^2 + 0^2 = 0 \) and

\[
\frac{d}{dt} E(t) = 2(rh(t))r h'(t) + 2 h'(t) h''(t) = 2 h'(t) (r^2 h(t) + h''(t)) = 0.
\]

Therefore \( E(t) = 0 \) for every \( t \). We then have \( 0 \leq (rh(t))^2 \leq E(t) = 0 \), so \( h(t) = 0 \) for all \( t \). Hence the only solution to our problem was \( f(t) = A \cos(rt) + B \sin(rt) \).
5.3 The Intermediate Value Theorem for Derivatives

Even though derivatives may not be continuous they still have the intermediate value property. **Hilfsatz** is the German word for **helping theorem**. A synonym is **lemma**.

**Hilfsatz 1** Suppose \( a < b \).

\[
A(x) = a + \frac{(x-a)^2}{b-a} \\
B(x) = b + \frac{(x-b)^2}{a-b}
\]

Then

\[
A(a) = B(a) = a \\
A(b) = B(b) = b \\
A(x) \in [a, b] \text{ for } x \in [a, b] \\
B(x) \in [a, b] \text{ for } x \in [a, b] \\
A'(a) = B'(b) = 0 \\
A'(b) = B'(a) = 2 \\
B(x) - A(x) = \frac{2(b-x)(x-a)}{b-a}
\]

**Reason:** This is all straightforward. Note that since \( B(a) - A(a) = B(b) - A(b) = 0 \) and \( B(x) - A(x) \) is quadratic, we know \( B(x) - A(x) = C(x-b)(x-a) \) for some constant \( C \). QED

**Hilfsatz 2** Suppose that \( f \) is continuous on \([a, b]\) and differentiable at \( a \) and at \( b \). Let \( d \) be any number lying between \( f'(a) \) and \( f'(b) \). Then there is some \( c \in (a, b) \) so that

\[
\frac{f(B(x)) - f(A(x))}{B(x) - A(x)} = d.
\]

**Reason:** Define a new function \( F \) by

\[
F(x) = \begin{cases} 
  f'(a) & \text{if } x = a \\
  \frac{f(B(x)) - f(A(x))}{B(x) - A(x)} & \text{if } x \in (a, b) \\
  f'(b) & \text{if } x = b
\end{cases}
\]

Once we show that \( F \) is continuous on \([a, b]\) then the conclusion follows from the Intermediate Value Theorem.

Since \( B(x) - A(x) > 0 \) for \( x \in (a, b) \) we know that \( F(x) \) is continuous on \((a, b)\). It remains to show that it is continuous at \( a \) and at \( b \). Observe that

\[
\frac{f(B(x)) - f(A(x))}{B(x) - A(x)} = f(B(x)) - f(A(x)) - (f(A(x)) - f(a))
\]

and recall the results of Hilfsatz 1. Keeping these fact in mind, if \( x \neq a \) and \( x \neq b \) we may write

\[
\frac{f(B(x)) - f(A(x))}{B(x) - A(x)} = \frac{b-a}{2(b-x)} \left( \frac{f(B(x)) - f(a)}{x-a} - \frac{f(A(x)) - f(a)}{x-a} \right)
\]

Therefore

\[
\lim_{x \to a^+} \frac{f(B(x)) - f(A(x))}{B(x) - A(x)} = \frac{b-a}{2(b-a)} (f'(a) \cdot B'(a) - f'(a) \cdot A'(a))
\]

\[
= \frac{1}{2} (f'(a) \cdot 2 - f'(a) \cdot 0)
\]

\[
= f'(a)
\]
so \( F \) is continuous at \( a \). Similarly,

\[
\frac{f(B(x)) - f(A(x))}{B(x) - A(x)} = \frac{b - a}{2(x - a)} \left( \frac{f(B(x)) - f(b)}{b - x} - \frac{f(A(x)) - f(b)}{b - x} \right)
\]

Therefore

\[
\lim_{x \to b^-} \frac{f(B(x)) - f(A(x))}{B(x) - A(x)} = \frac{b - a}{2(b - a)} (-f'(b) \cdot B'(b) + f'(b) \cdot A'(b)) = \frac{1}{2} (-f'(b) \cdot 0 + f'(b) \cdot 2) = f'(b)
\]

so \( F \) is continuous at \( b \). We have now shown that \( F \) is continuous on \([a, b]\). It now follows from the Intermediate Value Theorem that there is some \( c \in (a, b) \) so that \( F(c) = d \). \( \text{QED} \)

**Theorem 4 (Second Intermediate Value Theorem)** Suppose that \( f \) is differentiable on an interval \( I \). Then the range of \( f' \) is an interval.

Let \( a < b \) be elements of \( I \) and suppose that \( f'(a) \neq f'(b) \). Let \( d \) lie between \( f'(a) \) and \( f'(b) \). We have to show that there is some \( w \in I \) so that \( f(w) = d \).

According to Hilfsatz 2 there is some \( c \in (a, b) \) so that

\[
\frac{f(B(c)) - f(A(c))}{B(c) - A(c)} = d.
\]

Note that \( a \leq A(c) < B(c) \leq b \) so \([A(c), B(c)] \subset I\). This means that \( f \) is continuous on \([A(c), B(c)]\) and differentiable on \((A(c), B(c))\). Apply the First Mean Value Theorem to \( f \) on \([A(c), B(c)]\) to get that there is some \( w \in (A(c), B(c)) \) so that

\[
d = \frac{f(B(c)) - f(A(c))}{B(c) - A(c)} = f'(w),
\]

that is, \( d \) is in the range of \( f' \), as desired. \( \text{QED} \)

6 The Second Mean Value Theorem

**Theorem 5 (The Second Mean Value Theorem)** If \( h \) and \( g \) are continuous on \([a, b]\) and differentiable on \((a, b)\), then there is some \( c \in (a, b) \) so that

\[
h'(c)(g(b) - g(a)) = g'(c)(h(b) - h(a)).
\]

**Reason:** The reasoning is similar to that for the First Mean Value Theorem. Put

\[
f(x) = (g(x) - g(a))(h(b) - h(a)) - (h(x) - h(a))(g(b) - g(a))
\]

We have

\[
f'(x) = g'(x)(h(b) - h(a)) - h'(x)(g(b) - g(a)).
\]

We also have \( f(a) = f(b) = 0 \), so Rolle’s Theorem says that there is some \( c \in (a, b) \) so that

\[
0 = f'(c) = g'(c)(h(b) - h(a)) - h'(c)(g(b) - g(a))
\]

which is equivalent to the claim of the Second Mean Value Theorem. \( \text{QED} \)

Note that if you put \( h(x) = x \), the Second Mean Value Theorem reduces to the First Mean Value Theorem.

We now give some applications of the Second Mean Value Theorem.
6.1 L’Hopital’s Rule

So far we have considered two approaches to attacking the problem of limits of quotients when the limit of the numerator and of the denominator are both 0. The first approach is algebraic simplification, as in the case of
\[ \frac{x^3 - 27}{x^2 - 9} \]
when \( x \) approaches 3 or
\[ \frac{1 - \cos(t)}{t^2} \]
when \( t \) approaches 0.

The second approach is pinching, as in
\[ \frac{\sin(t)}{t} \]
\[ \frac{\exp(t) - 1}{t} \]
\[ \frac{\ln(1 + t)}{t} \]
as \( t \) approaches 0.

Neither of these approaches works on
\[ \frac{\tan(t) - t}{t^3} \]
\[ \frac{\sin(t) - t}{t^3} \]
as \( t \) approaches 0. The Second Mean Value Theorem gives a third approach, called L’Hopital’s Rule. This is a dangerous method, and you need to pay close attention to what it says. We will start with the example
\[ \frac{\sin(t) - t}{t^3} \]
as \( t \) approaches 0. Suppose we were to apply the Second Mean Value Theorem with \( a = 0 \) and \( b = t \) to \( g(t) = t^3 \) and \( h(t) = \sin(t) - t \). It says that for some \( c_t \) lying strictly between 0 and \( t \) we have
\[ h'(c_t)(g(t) - g(0)) = g'(c_t)(h(t) - h(0)) \]
\[ (\cos(c_t) - 1)(t^3) = 3c_t^2(\sin(t) - t) \]
\[ \frac{1}{3} \times \frac{1 - \cos(c_t)}{c_t^2} = \frac{\sin(t) - t}{t^4} \] (1)

We write \( c_t \) instead of \( c \) to emphasize that as you change \( t \) you may have to change \( c \). In any event, though, as \( t \) approaches 0, so does \( c_t \) since \( 0 < |c_t| < |t| \).

Now, we know that
\[ \lim_{c \to 0} \frac{1 - \cos(c)}{c^2} = \frac{1}{2} \]
so by our Theorem on limits of compositions of functions, we know that the limit as \( t \) approaches 0 of the lefthand side of (1) is \(-1/6\), so the limit as \( t \) approaches 0 of the righthand side of (1) is the same thing, that is
\[ \lim_{t \to 0} \frac{\sin(t) - t}{t^3} = -\frac{1}{6}. \]

This same reasoning allows us to prove

**Theorem 6 (L’Hopital’s Rule)** Suppose that \( f \) and \( g \) are differentiable in \( (a, b) \) and \( g'(x) \neq 0 \) for any \( x \in (a, b) \), and suppose that
\[ \lim_{x \to a^+} g'(x) = L. \]

8
If both \( f \) and \( g \) have a limit of 0 as \( x \) approaches \( a \) from above, or if \( g \) approaches \( +\infty \) as \( a \) approaches \( a \) from above, then
\[
\lim_{x \to a^+} \frac{f(x)}{g(x)} = L
\]
as well. The same conclusion holds if we consider \( x \) approaching \( b \) from below everywhere, or if \( g(x) \) approaches \( -\infty \) instead of \( +\infty \).

It is reasonable to wonder why one would consider such limits. Here is one possibility. Since we know that \( \sin(x)/x \) converges to 1 as \( x \) approaches 0 we know that \( \sin(x)/x \approx 1 \) when \( x \) is near to 0. In other words, \( \sin(x) - x \approx 0 \). Since \( \sin(x) - x \) is an odd function, we think it behaves like \( x^3 \) near 0. To investigate this conjecture we would consider the behavior of
\[
\frac{\sin(x) - x}{x^3}
\]
as \( x \) approaches 0.

6.2 Error Estimate in the Best Linear Approximation

**Theorem 7** Suppose that \( f \) is continuous on \([a, b]\), \( f' \) is continuous on \([a, b]\) and \( f''(x) \) exists for each \( x \in (a, b) \). Then there is some \( c \in (a, b) \) so that
\[
f(b) = f(a) + f'(a)(b - a) + \frac{1}{2} f''(c)(b - a)^2.
\]

**Reason:** Define
\[
F(x) = \frac{1}{2} (f(x) - f(a) - f'(a)(x - a))(b - a)^2 - \frac{1}{2} (f(b) - f(a) - f'(a)(b - a))(x - a)^2
\]
It is easy to check that
\[
F(a) = 0,
\]
\[
F'(b) = 0,
\]
\[
F'(x) = \frac{1}{2} (f'(x) - f'(a))(b - a)^2 - (f(b) - f(a) - f'(a)(b - a))(x - a)
\]
\[
F'(a) = 0,
\]
\[
F''(x) = \frac{1}{2} f''(x)(b - a)^2 - (f(b) - f(a) - f'(a)(b - a))
\]
so it follows from the Second Rolle’s Theorem that there is some \( c \in (a, b) \) so that
\[
0 = F''(c)
\]
\[
= \frac{1}{2} f''(c)(b - a)^2 - (f(b) - f(a) - f'(a)(b - a))
\]
\[
f(b) = f(a) + f'(a)(b - a) + \frac{1}{2} f''(c)(b - a)^2
\]
as desired. **QED**

Note that the same conclusion holds if \([a, b]\) is replaced by \([b, a]\).

For example, suppose \( f(x) = \ln(1 + x) \). Then for some \( y \) between 0 and \( x \) we have \( f(0) = 0, f'(0) = 1, f''(y) = -(1 + y)^{-2} \) and
\[
\ln(1 + x) = x - \frac{1}{2} \frac{1}{(1 + y)^2} x^2.
\]
For example, suppose we want an estimate of \( \ln(1.1) \). For \( 0 < y < 0.1 \) we have
\[
\frac{1}{1.1^2} \leq \frac{1}{(1 + y)^2} \leq 1
\]
so
\[
0.004 < \frac{1}{242} \leq 0.1 - \ln(1.1) \leq \frac{1}{200} = 0.005
\]
or,
\[
0.095 < \ln(1.1) < 0.096.
\]
Note that to 10 digits \( \ln(1.1) \approx 0.0531017980 \).
7 Newton’s Method

Suppose that we are to solve the equation \( f(x) = 0 \), and there is no algebraic solution. We know that if \( f \) is continuous then we know that we can use the bisection method. If \( f \) is differentiable and either convex or concave, we can do better. The idea is to find some \( a \) so that \( f(a) \approx 0 \). Then instead of trying to solve \( f(x) = 0 \) we solve \( f(a) + f'(a)(x-a) = 0 \), that is we look to see where the tangent line at \((a, f(a))\) crosses the horizontal axis. We then repeat this procedure. Although the following is true if convexity replaces our assumption about a second derivative, we will content ourselves with demonstrating:

**Theorem 8 (Newton’s Method)** Suppose \( a < c < b \), that \( f(a) = 0 \), and that \( f \) is twice differentiable on \( (a,b) \) and continuous at \( a \). Assume \( f'(x) > 0 \) and \( f''(x) \geq 0 \). Define the following infinite sequence recursively:

\[
x_0 = c \\
x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}
\]

Then the sequence \( x_n \) is monotone decreasing and it has a limit of \( a \).

**Reason:** Our hypotheses guarantee that \( x_{n+1} \leq x_n \). Now we will show by induction that \( x_n \geq a \). We assumed that \( x_0 > a \). Now, suppose that \( x_n \geq a \). We want to show that

\[ a \leq x_n - \frac{f(x_n)}{f'(x_n)}. \]

This is equivalent to showing

\[
\begin{align*}
f'(x_n)(a - x_n) &\leq -f(x_n) \\
f(x_n) + f'(x_n)(a - x_n) &\leq 0 \\
-f(x_n) - f'(x_n)(a - x_n) &\geq 0
\end{align*}
\]

Since \( f(a) = 0 \) this last inequality is equivalent to

\[ f(a) - f(x_n) - f'(x_n)(a - x_n) \geq 0 \]

If we apply our error estimate for the linear approximation we know that there is some \( z \in (a, x_n) \) so that

\[ f(a) - f(x_n) - f'(x_n)(a - x_n) = \frac{1}{2} f''(z)(a - x_n) \geq 0 \]

which is what we wanted to do. Now we know that

\[ \lim_{n \to \infty} x_n = L \geq a. \]

We need to show that \( L = a \). Suppose \( L > a \). We know that

\[ f'(x_n)(x_{n+1} - x_n) = f(x_n). \]

Take the limit of both sides of this equation as \( n \) tends to infinity. Since \( f \) and \( f' \) are continuous on \([L, c]\) we have \( 0 = f'(L)(L - L) = f(L) \), which is impossible since \( f \) is strictly increasing on \([a, c]\). This means that \( L = a \). \( \text{QED} \)

For example, suppose that we want to solve \( x^2 - 5 = 0 \) for \( x > 0 \). We let \( a = \sqrt{5} \) and \( b = 10 \). Take \( c = 3 \). Then since \( f'(x) = 2x \) we have

\[
x_0 = 3 \\
x_{n+1} = x_n - \frac{x_n^2 - 5}{2x_n} = \frac{x_n^2 + 5}{2x_n}
\]