CHAPTER 2

Theoretical Considerations; Differentiation

§2.1. Limit Operations

In this section we shall go a little more deeply into the concept of limits than we did in chapter 1. Suppose that $y = f(x)$ is a function defined in an interval about the point x_0. Each value of x determines a value y using the rule represented by the function f. We say that y approaches a number L as x approaches x_0 if we can be sure that y is as close as we please to L just by taking x close enough to x_0. A little more precisely, if we allow an error $\varepsilon > 0$ in the calculation of L, we can find an error $\delta > 0$ for x_0 such that if x is within δ of x_0, then y is within ε of L. If the limit L is the number y_0 calculated from x_0 by f, then we say that f is *continuous* at x_0. That is the content of the following two definitions.

Definition 2.1 $\lim_{x\to x_0} f(x) = L$ *if, for any* $\varepsilon > 0$ *there is a* $\delta > 0$ *such that* $|x - x_0| < \delta$ *implies* $|f(x) - L| < \varepsilon$.

If the limit L is $f(x_0)$, then we say that f is continuous at x_0.

Definition 2.2 *A function f, defined in an interval about x_0 is continuous at x_0 if* $\lim_{x\to x_0} f(x) = f(x_0)$. *A function is said to be continuous if it is continuous at every point where it is defined.*

Example 2.1 Let $f(x) = x/|x|$. Then, $f(0) = 0$, for $x > 0$, $f(x) = 1$, and for $x < 0$, $f(x) = -1$. Thus, in any interval about 0, there are values of x for which $f(x) = 1$ and other values of x for which $f(x) = -1$. There is thus no number L such that both 1 and -1 are within .5 of L, so there can be no $\lim_{x\to 0} f(x)$.

Example 2.2 Let $f(x) = \cos(1/x)$ for $x \neq 0$. There is no value to assign to $f(0)$ to make this function continuous. For if $x = (2\pi n)^{-1}$, $f(x) = 1$ for n even, and $f(x) = -1$ for n odd, so we are in the same situation as that of example 1. However, for the function $g(x) = x\cos(1/x)$, we can define $g(0) = 0$ to get a continuous function. For $|g(x)| \leq |x|$ for every x, since the cosine is bounded by 1. Thus, for any $\varepsilon > 0$, if $|x| < \varepsilon$, we also have $|g(x)| < \varepsilon$.

Now we state the basic facts describing how limits behave under algebraic operations.

Proposition 2.1 *Suppose that f and g are functions defined in an interval about x_0 and that*

$$\lim_{x\to x_0} f(x) = L\,, \quad \lim_{x\to x_0} g(x) = M\,. \tag{2.1}$$

Then

a) $\lim_{x\to x_0} (f(x)+g(x)) = L+M$

b) $\lim_{x\to x_0} (f(x)\cdot g(x)) = L\cdot M$

If $M \neq 0$, then

c) $\lim_{x\to x_0} \dfrac{f(x)}{g(x)} = \dfrac{L}{M}$.

This proposition then tells us the following about continuous functions:

Proposition 2.2 *Suppose that f and g are defined in an interval about x_0, and are continuous at x_0. Then the sum $f+g$ and product $f\cdot g$ are also continuous at x_0. If $g(x_0)\neq 0$, the quotient f/g is also continuous at x_0.*

Now it is clear that a constant function is continuous: if $f(x) = C$ for all x, then the difference $|f(x) - C| = 0$ no matter what x is. Also, the function $f(x) = x$ is continuous everywhere: we can make $|f(x) - f(x_0)| < \varepsilon$ just by taking $|x - x_0| < \varepsilon$. Thus, by proposition 2.2, any function formed from constants and the function $f(x) = x$ by taking products and sums is continuous. But these are the polynomials.

Proposition 2.3 *All polynomials are continuous everywhere. A rational function (that is, a quotient of polynomials) is continuous everywhere where its denominator is non-zero.*

Example 2.3 That is not to say that a rational function is *not* continuous where the denominator is zero; perhaps it can be defined at those points so as to be continuous, For example, consider

$$f(x) = \frac{x^2-4x-5}{x-5}\,. \tag{2.2}$$

Since we cannot divide by zero, $f(x)$ is not defined for $x=5$. But the issue is, can we define $f(5)$ so that the function is continuous? Noting that $x^2-4x-5 = (x-5)(x+1)$, we see that for $x\neq 5$, $f(x) = x+1$. Thus by defining $f(5) = 6$, we get a continuous function.

Suppose now that g is defined in an interval around x_0 and f is a function defined on the range (set of values) of g. Then we can form the *composition* of the two functions, $f\circ g$, just by applying the rule defining f to the value of g : $f\circ g(x) = f(g(x))$.

Proposition 2.4 *Suppose that g is defined in an interval about the point x_0, $g(x_0) = y_0$ and f is defined in an interval about y_0. If g is continuous at x_0, and f is continuous at y_0, then $h = f\circ g$ is also continuous at x_0.*

To show this, we have to show that we can insure that $h(x)$ is within ε of $h(x_0)$ by taking x close enough to x_0. By the continuity of f we can be sure that $f(y)$ is within ε of $f(y_0)$ by taking y within some small number, η of y_0. But then, by the continuity of g, there is a δ such that, if x is within δ of x_0, $g(x)$ is within η of $g(x_0) = y_0$, and finally, $f(g(x))$ is within ε of $f(y_0) = f(g(x_0))$.

A useful technique is what is called the squeeze theorem. Suppose, in some interval containing the point a, the values of f lie between those of two other functions g and h. Suppose also that g and h have the same limit as x approaches a, then f also has that limit.

Proposition 2.5 (Squeeze Theorem) *Suppose that f, g, h are defined in a interval containing a and that $g(x) \leq f(x) \leq h(x)$. If*

$$\lim_{x\to a} g(x) = \lim_{x\to a} h(x) = L \ , \tag{2.3}$$

we also have $\lim_{x\to a} f(x) = L$.

Suppose an allowable error $\varepsilon > 0$ is specified. From the hypothesis, we know that there is a $\delta_1 > 0$ such that if x is within δ_1 of x_0, then $g(x) \geq L - \varepsilon$, and there is a $\delta_2 > 0$ such that if x is within δ_2 of x_0, then $h(x) \leq L + \varepsilon$. Then, so long as δ is less than both δ_1 and δ_2, we have

$$L - \varepsilon \leq g(x) \leq f(x) \leq h(x) \leq L + \varepsilon \tag{2.4}$$

which is to say that $f(x)$ is within ε of L.

Now suppose again that f is defined in a neighborhood of x_0 and continuous there. We now turn to the question of the differentiability of f at x_0.

Definition 2.3 *Let f be defined in a neighborhood of x_0. If the limit*

$$\lim_{x\to x_0} \frac{f(x) - f(x_0)}{x - x_0} \tag{2.5}$$

exists, it is denoted $f'(x_0)$, and is called the **derivative** *of f at x_0. f is said to be* **differentiable** *at x_0.*

Proposition 2.1c suggests that the limit does not exist since the denominator approaches 0. But we have to be careful: the numerator is also going to zero. In fact, as we saw by the division theorem of chapter 1, If f is a polynomial, then so is this difference quotient, and the limit is the value of that quotient at x_0. In fact, in general it is a necessary condition for differentiability that the limit of the numerator is zero - a fact we already used several times in chapter 1.

Proposition 2.6 *Let f be defined in a neighborhood of x_0. If f is differentiable at x_0, then it is continuous at x_0.*

Let $L = f'(x_0)$. The hypothesis tells us that we can be sure the difference quotient is within ε of L by taking x close enough to x_0. So, taking, for example, $\varepsilon = 1$, then if x is close enough to x_0,

$$-1 < \frac{f(x) - f(x_0)}{x - x_0} - L < +1, \tag{2.6}$$

from which we conclude that

$$(L-1)(x - x_0) < f(x) - f(x_0) < (L+1)(x - x_0) \ . \tag{2.7}$$

Now, the left and right hand sides tend to 0 as x approaches x_0, so, by the squeeze theorem, $\lim(f(x) - f(x_0)) = 0$. But that is the same as $\lim f(x) = f(x_0)$; that is, f is continuous at x_0.

Now, in section 1 of chapter 1, no problems arose in calculating limits, since we were there dealing with polynomials (even in calculating derivatives). However, more generally questions about limits can become real issues. For example, when we turned to trigonometric functions and the square root function, we tacitly assumed their continuity. Since the continuity is intuitively clear (if we envision the graph of these functions), this was not an obstacle to finding derivatives. However, in more general contexts, the continuity is not at all clear. As preparation for this, we shall here reconsider the assumptions of continuity made in chapter 1. First, the square root.

Example 2.4 For $a \geq 0$, $\lim_{x\to a} \sqrt{x} = \sqrt{a}$.

We have to distinguish the cases $a \neq 0$ and $a = 0$. First, the case $a \neq 0$. We start with the identity

$$(\sqrt{x} - \sqrt{a})(\sqrt{x} + \sqrt{a}) = x - a \ , \tag{2.8}$$

which, for our purposes should be written as

$$\sqrt{x} - \sqrt{a} = \frac{x-a}{\sqrt{x} + \sqrt{a}} \ , \tag{2.9}$$

since it is the expression on the left we need to make small. Given $\varepsilon > 0$, choose $\delta > 0$ so that $\delta/\sqrt{a} < \varepsilon$. Then if $|x - a| < \delta$,

$$|\sqrt{x} - \sqrt{a}| = \frac{|x-a|}{\sqrt{x} + \sqrt{a}} < \frac{|x-a|}{\sqrt{a}} < \frac{\delta}{\sqrt{a}} < \varepsilon \ . \tag{2.10}$$

Since this argument fails if $a = 0$, we need another idea.

Proposition 2.7 (Archimedean principle) *For any positive real number M, there is an integer n such that $n > M$.*

Now, given $\varepsilon > 0$, choose the integer n so that $n > 1/\varepsilon^2$. Then $\sqrt{n} > 1/\varepsilon$, so $1/\sqrt{n} < \varepsilon$ which is what we need. For $x < 1/n$,

$$\sqrt{x} < \frac{1}{\sqrt{n}} < \varepsilon \ . \tag{2.11}$$

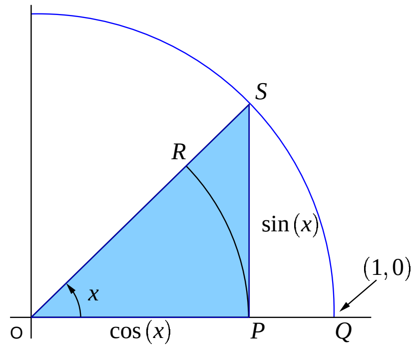
Now, in section 1.4 we derived the formulae for the derivatives of the sine and cosine functions, assuming that they were differentiable. Here we would like to justify that assumption. The crux of the matter is the following proposition (which we derived in section 1.4 from the formulae for differentiation).

Proposition 2.8

$$\lim_{x\to 0} \frac{\sin x}{x} = 1 \ , \qquad \lim_{x\to 0} \frac{\cos x - 1}{x} = 0 \ . \tag{2.12}$$

This is just the assertion that the sine and cosine functions are differentiable at $x = 0$. Then, using the addition formulae for these functions, we can prove their differentiability everywhere. Here is a geometric argument for proposition 2.8.

Figure 2.1



In figure 2.1, let A be the area of the sector OPR, B the area of triangle OPS, and C the area of sector OQS. Then $A \leq B \leq C$. Using the formulae for these areas (measuring the angle x in radians), this gives us

$$\frac{1}{2} x \cos^2 x \leq \frac{1}{2} \cos x \sin x \leq \frac{1}{2} x (1)^2 \, . \tag{2.13}$$

Dividing by $x \cos x/2$, this gives us

$$\cos x \leq \frac{\sin x}{x} \leq \frac{1}{\cos x} \, . \tag{2.14}$$

But now, since $\lim_{x\to 0} \cos x = 1$, as is obvious from the figure, the first part of proposition 2.8 follows from the squeeze theorem. The second now follows from the first using the equalities:

$$\frac{\cos x - 1}{x} = \frac{\cos x - 1}{x} \frac{\cos x + 1}{\cos x + 1} = \frac{\cos^2 - 1}{x(\cos x + 1)} = \frac{\sin x}{x} \sin x \frac{1}{1 + \cos x} \to 0 \, , \tag{2.15}$$

since $\lim_{x\to 0} \sin x = 0$, as is clear from the figure.

Example 2.5 Find the limit as $x \to 0$ of $\sin(3x)/\sin(4x)$.

$$\lim_{x\to 0} \frac{\sin(3x)}{\sin(4x)} = \frac{3}{4} \lim_{x\to 0} \frac{\sin(3x)}{3x} \frac{4x}{\sin(4x)} = \frac{3}{4} \lim_{3x\to 0} \frac{\sin(3x)}{3x} \lim_{4x\to 0} \frac{4x}{\sin(4x)} = \frac{3}{4}(1)(1) = \frac{3}{4} \tag{2.16}$$

Example 2.6 Find the limit as $x \to \pi/2$ of $\cos x/(x - \pi/2)$.

Let $t = x - \pi/2$. Then $\cos x = \sin(\pi/2 - x) = -\sin(x - \pi/2) = -\sin t$, and $t \to 0$ as $x \to \pi/2$. Thus

$$\lim_{x\to \pi/2} \frac{\cos x}{x - \pi/2} = -\lim_{t\to 0} \frac{\sin t}{t} = -1 \, . \tag{2.17}$$

§2.2. Limits at Infinity

Suppose that f is defined for all positive numbers. We say that f has the limit L as $x \to +\infty$ if we can make f as close as we please to L by taking x large enough. For example

$$\lim_{x\to+\infty} \frac{1}{x} = 0, \tag{2.18}$$

since we can make $1/x < \varepsilon$ just by taking $x > 1/\varepsilon$.

Definition 2.4 *Suppose that $f(x)$ is defined for all $x > M_0$. We say that*

$$\lim_{x\to+\infty} f(x) = L \tag{2.19}$$

if, for every $\varepsilon > 0$, we can find an $M \geq M_0$ such that if $x > M$, then $|f(x) - L| < \varepsilon$. Suppose that $f(x)$ is defined for all $x < M_0$. We say that

$$\lim_{x\to-\infty} f(x) = L \tag{2.20}$$

if, for every $\varepsilon > 0$, we can find an $M \leq M_0$ such that if $x < M$, then $|f(x) - L| < \varepsilon$.

Example 8. $\lim_{x\to+\infty} \dfrac{x}{x+1} = 1.$

For, given $\varepsilon > 0$, choose $M = 1/\varepsilon$. Then, for $x > M$, we have

$$\left|\frac{x}{x+1} - 1\right| = \left|\frac{x-(x+1)}{x+1}\right| = \left|\frac{1}{x+1}\right| < \left|\frac{1}{x}\right| < \frac{1}{M} = \varepsilon. \tag{2.21}$$

Now, we define what it means to have $\pm\infty$ as a limit.

Definition 2.5 *Let f be defined for all x in an interval about a, except perhaps at a. We write*

$$\lim_{x\to a} f(x) = +\infty \tag{2.22}$$

if, for any $M > 0$, there is an $\varepsilon > 0$ such that for $|x - a| < \varepsilon$, we have $f(x) > M$. Similarly,

$$\lim_{x\to a} f(x) = -\infty \tag{2.23}$$

if, for any $M > 0$, there is an $\varepsilon > 0$ such that for $|x - a| < \varepsilon$, we have $f(x) < -M$.

We will also say that $\lim_{x\to+\infty} f(x) = +\infty$ if we can make $f(x)$ as large as we please by taking x sufficiently large, and similarly, we define $\lim_{x\to+\infty} f(x) = -\infty, \lim_{x\to-\infty} f(x) = +\infty$, and so forth.

Proposition 2.9 *Let p be a polynomial of degree $n > 1$, with leading coefficient 1.*

a) If n is even, $\lim_{x\to\pm\infty} p(x) = +\infty$.

b) If n is odd $\lim_{x\to+\infty} p(x) = +\infty, \quad \lim_{x\to-\infty} p(x) = -\infty.$

To see this, write $p(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$. Then, by factoring out the highest power of x:

$$p(x) = x^n \left(1 + \frac{a_{n-1}}{x} + \cdots + \frac{a_1}{x^{n-1}} + \frac{a_0}{x^n}\right). \tag{2.24}$$

The term in parenthesis goes to 1 as $|x|$ becomes infinite. Now, since $|x^n| \geq |x|$ so long as $|x| \geq 1$, the term x^n approaches $+\infty$ as $|x|$ becomes large, except when n is odd, and $x \to -\infty$, in which case $x^n \to -\infty$.

Now, we can make the same kind of qualitative statements about quotients of polynomials (rational functions). Let $f(x) = p(x)/q(x)$ where p and q are polynomials with no common factors. Then f is defined and continuous at all points except those points a such that $q(a) = 0$. At such an a, the graph of $y = f(x)$ will go off the graph paper, either upwards or downwards, since the denominator is 0 at a. In this case we say that the graph has a *vertical asymptote* at $x = a$. To determine this behavior, we write $q(x) = (x-a)^n q_0(x)$ for some positive integer n and some polynomial q_0 such that $q_0(a) \neq 0$. Since p has no factor in common with q, $p(a) \neq 0$. Then

$$\lim_{x\to a} \frac{p(x)}{q(x)} = \lim_{x\to a} \frac{1}{(x-a)^n} \frac{p(x)}{q_0(x)}. \tag{2.25}$$

Since the second factor converges to $p(a)/q_0(a)$, the behavior of p/q near a is determined by the behavior of the first factor. For this, if n is odd, it depends upon whether we approach a from the right or the left, since $(x-a)^n$ is negative if $x < a$, and is positive if $x > a$. We summarize the result as

Proposition 2.10

a) If n is even, $\lim\limits_{x\to a} \dfrac{1}{(x-a)^n} = +\infty$.

b) If n is odd, $\lim\limits_{x\to a^-} \dfrac{1}{(x-a)^n} = -\infty, \quad \lim\limits_{x\to a^+} \dfrac{1}{(x-a)^n} = +\infty$.

Finally, we summarize the limits for rational functions as $x \to \pm\infty$.

Proposition 2.11 *Let $f(x) = p(x)/q(x)$, where p and q are polynomials of degree n and m respectively.*

a) If $n < m$, $\lim\limits_{x\to\pm\infty} f(x) = 0$.

b) If $n = m$, $\lim\limits_{x\to\pm\infty} f(x) = \dfrac{a_n}{b_n}$ *, where a_n, b_n are the leading coefficients of p and q respectively.*

c) If $n = m + d$, $\lim\limits_{x\to\pm\infty} |f(x) - Q(x)| = 0$ *where Q is the polynomial of degree d obtained by dividing the polynomial p by the polynomial q.*

In this last case, we say that the graph $y = f(x)$ approaches the graph $y = Q(x)$ *asymptotically*, and that the latter curve is an *asymptote* for $y = f(x)$.

Example 2.8 Let $f(x) = (x^2 + x)/(x + 2)$. f is not defined at $x = -2$. To see how f behaves at -2 and at infinity, we do the division:

$$\frac{x^2 + x}{x + 2} = x - 1 + \frac{2}{x + 2} \tag{2.26}$$

Thus, for x very close to, but to the left of -2, $f(x)$ is negative; but for x very close to, but to the right of -2, $f(x)$ is positive. Thus

$$\lim_{x\to -2^-} \frac{x^2+x}{x+2} = -\infty\,, \qquad \lim_{x\to -2^+} \frac{x^2+x}{x+2} = \infty\,. \tag{2.27}$$

Finally, $y = f(x)$has the asymptote $y = x-1$ as x goes to infinity.

Example 2.9 Let $f(x) = \dfrac{x}{x-1}$. If we write

$$f(x) = x\frac{1}{x-1} \tag{2.28}$$

after observing that x is positive near 1, we see from proposition 2.10 that

$$\lim_{x\to 1^-} f(x) = -\infty\,, \qquad \lim_{x\to 1^+} f(x) = \infty\,. \tag{2.29}$$

Finally, to see what happens as x goes to infinity, write

$$f(x) = \frac{1}{1-\frac{1}{x}} \tag{2.30}$$

so that $y = 1$ is the asymptote.

Example 2.10 Let $f(x) = \dfrac{x}{(x-1)^2}$

If we write

$$f(x) = x\frac{1}{(x-1)^2} \tag{2.31}$$

after observing that x is positive near 1, we see from proposition 2.10 that $\lim_{x\to 1} f(x) = \infty$ from both sides. Since the degree of the denominator is greater than the degree of the numerator, the asymptote is $y = 0$, with $y = f(x)$ below the x-axis to the left, and above the x-axis to the right.

Example 2.11 Let $f(x) = \dfrac{x^3+2x^2}{x^2-3x+2}$.

First, we factor numerator and denominator as much as possible:

$$f(x) = \frac{x^2(x+2)}{(x-1)(x-2)}\,. \tag{2.32}$$

From this we see that f is defined except for the points $x = 1, 2$. We know that at these points $f(x)$ becomes infinite; we need only to determine whether the limit is $+\infty$ or $-\infty$. First we look near $x = 1$. To the left of 1, the negative terms are $x-1$ and $x-2$; since all others are positive, $f(x) > 0$. We conclude

$$\lim_{x\to 1^-} f(x) = +\infty\,. \tag{2.33}$$

But to the right of $x = 1$, the term $x-1$ changes sign, so now $f(x) < 0$, and we conclude

$$\lim_{x\to 1^+} f(x) = -\infty \,. \tag{2.34}$$

As we pass through $x = 2$, the only change is from $x-2 < 0$ to $x-2 > 0$, so we conclude

$$\lim_{x\to 2^-} f(x) = -\infty \,, \tag{2.35}$$

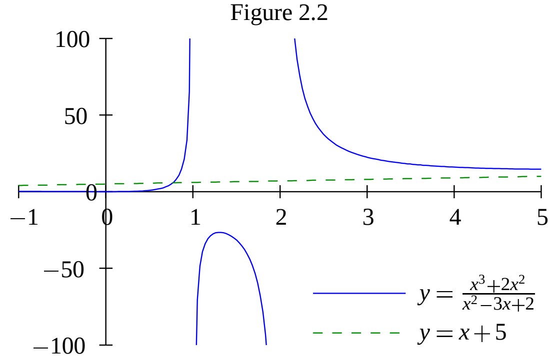
$$\lim_{x\to 2^+} f(x) = +\infty \,. \tag{2.36}$$

Finally we look for asymptotes as $x \to \pm\infty$. Long division gives

$$\frac{x^3+2x^2}{x^2-3x+2} = (x+5) + \frac{15x-10}{x^2-3x+2} \,, \tag{2.37}$$

so, as x approaches infinity, the graph approaches the asymptote $y = x+5$ (see figure 2.2).

Figure 2.2



§2.3. Some Basic Theorems

The preceding sections discussed the behavior of functions *locally*, that is, for x varying in a neighborhood of a particular point a. In this section we summarize more *global* results; that is the behavior of the function as x varies over an interval $[a, b]$. Most of these results are intuitively clear, and were taken as such by the founders of the Calculus.

Theorem 2.1 (The Intermediate Value Theorem) *Suppose that f is a continuous function on the interval $[a, b]$. Then, for every number w between $f(a)$ and $f(b)$, there is a c between a and b such that $f(c) = w$.*

Intuitively, this says that as you draw the graph of the function $y = f(x)$, your pencil point never leaves the paper.

Theorem 2.2 (Maxima and Minima) *Suppose that f is a continuous function on the interval $[a.b]$. There are points c, C in $[a,b]$ such that $f(c)$ is the minimum value of f on the interval, and $f(C)$ is the maximum value of f on the interval.*

Theorem 2.3 (Rolle's Theorem) *Let f be continuous on $[a,b]$ and differentiable in (a,b), and suppose that $f(a) = f(b)$. Then there is a point c in (a,b) at which $f'(c) = 0$.*

We can derive Rolle's theorem from theorems 2.1 and 2.2. First of all, if f is constant, then $f'(c) = 0$ for all c. If f is nonconstant, there is a point c in the interval (a,b) at which f has either a maximum or a minimum. Suppose it is a maximum. Then, for all other x in (a,b), $f(x) \le f(c)$. In particular, for $x < c$, $f(x) - f(c)$ and $x - c$ have the same sign, so

$$\frac{f(x) - f(c)}{x - c} \ge 0\,. \tag{2.38}$$

Taking the limit as $x \to c$, we conclude $f'(c) \ge 0$. But now if $x > c$, the denominator changes sign, but the numerator doesn't, so, in this case

$$\frac{f(x) - f(c)}{x - c} \le 0\,, \tag{2.39}$$

from which we conclude $f'(c) \le 0$. Thus $f'(c) = 0$.

Notice, that the above argument incidentally shows that at a maximum or minimum point x_0 of a differentiable function, we must have $f'\left(x_0\right) = 0$. In the next chapter we shall see that this provides a method for finding maxima and minima.

Theorem 2.4 (The Mean Value Theorem) *Let f be continuous on $[a,b]$ and differentiable in (a,b). There is a c in (a,b) such that*

$$f'(c) = \frac{f(b) - f(a)}{b - a}\,. \tag{2.40}$$

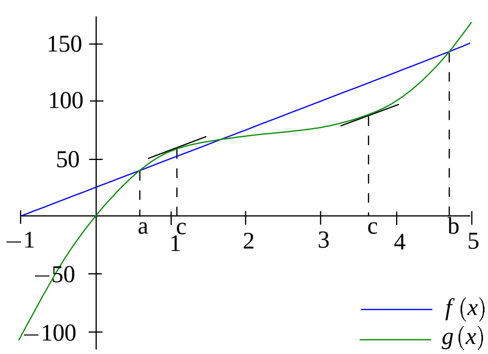


Figure 2.3

To see why this is true, consult figure 2.3. Here, $y = g(x)$ is the line joining $(a, f(a))$ to $(b, (f(b))$. The slope of this line is the right hand side of 2.40, which is also $g'(x)$ for any x in (a,b). Now the function $f(x) - g(x)$ satisfies the hypotheses of Rolle's theorem, so there is a c in (a,b) at which the derivative is zero, that is $f'(c) = g'(c)$. But this is the same as equation 2.40.

The point of this section is to demonstrate that Newton's uniqueness hypothesis (that a function with derivative zero everywhere is constant) follows from basic intuitive facts.

Theorem 2.5 *Suppose that f is differentiable in an interval, and has derivative zero everywhere. Then f is constant.*

Let $a,\ b$ be different points in the interval. By the Mean Value Theorem, there is a point c between a and b such that

$$f'(c) = \frac{f(b) - f(a)}{b-a}\ . \tag{2.41}$$

But the hypothesis is that $f'(c) = 0$, so we must have $f(b) - f(a) = 0$, or $f(a) = f(b)$. This is true for any two points a and b, so f is constant.