Chapter VII: Derivatives of Combinations

We show that various combinations of differentiable functions (including compositions) are differentiable; we derive formulas for the derivatives of the combinations in terms of the derivatives of the functions separately. We apply our results to show that polynomials are differentiable and that rational functions are differentiable where they are defined.

1. Sums, Differences, Products and Quotients

We show that sums, differences, products and quotients of two differentiable functions are differentiable; in the process, we derive formulas for the derivatives of the combined functions in terms of the derivatives of the functions separately. We apply our results in the next section to show polynomials and rational functions are differentiable.

Theorem 7.1: Let $X \subset \mathsf{R}^1$, and let $f, g : X \to \mathsf{R}^1$ be functions. If f and g are each differentiable at p, then $f + g$ is differentiable at p and

$$(f+g)'(p) = f'(p) + g'(p).$$

Proof: Using Theorem 4.1 for the third equality below, we have

$$\lim_{h\to 0} \frac{(f+g)(p+h)-(f+g)(p)}{h} = \lim_{h\to 0} \frac{f(p+h)-f(p)+g(p+h)-g(p)}{h}$$
$$= \lim_{h\to 0}\left[\frac{f(p+h)-f(p)}{h} + \frac{g(p+h)-g(p)}{h}\right]$$
$$\overset{4.1}{=} \lim_{h\to 0} \frac{f(p+h)-f(p)}{h} + \lim_{h\to 0} \frac{g(p+h)-g(p)}{h} = f'(p) + g'(p).$$ ¥

Corollary 7.2: Let $X \subset \mathsf{R}^1$, and let $f_1, f_2, ..., f_n : X \to \mathsf{R}^1$ be finitely many functions. If each of the functions $f_1, f_2, ..., f_n$ is differentiable at p, then the sum function $f_1 + f_2 + \cdots + f_n$ is differentiable at p and

$$(f_1 + f_2 + \cdots + f_n)'(p) = f_1'(p) + f_2'(p) + \cdots + f_n'(p).$$

Proof: The corollary follows from Theorem 7.1 by a simple induction (much like the proof of Theorem 4.5). ¥

Theorem 7.3: Let $X \subset \mathsf{R}^1$, let $f, g : X \to \mathsf{R}^1$ be functions. If f and g are each differentiable at p, then $f - g$ is differentiable at p and

$$(f-g)'(p) = f'(p) - g'(p).$$

Proof: The proof is similar to the proof of Theorem 7.1 using Theorem 4.2 (instead of Theorem 4.1). ¥

We know from the previous two theorems that derivatives "distribute over" sums and differences. Furthermore, the proofs of the two theorems use nothing more than the corresponding results about limits. Therefore, since limits "distribute over" products (Theorem 4.9), it is natural to expect that derivatives would do the same; in other words, we should expect that if f and g are differentiable at p, then

$$(f \cdot g)'(p) = f'(p)g'(p).$$

So, let's try to verify the formula and see what happens:

$$(f \cdot g)'(p) = \lim_{h \to 0} \frac{(f \cdot g)(p+h) - (f \cdot g)(p)}{h} = \lim_{h \to 0} \frac{f(p+h)g(p+h) - f(p)g(p)}{h};$$

this might not look promising, but remember the trick we used in proving the limit theorem for products (Theorem 4.9)? We subtracted and added an expression that enabled us to isolate expressions that related directly to the assumptions in the theorem. Let's try that here. Since we want to isolate the difference quotients for f and g, let's subtract and add $f(p)g(p+h)$ to the numerator of the second difference quotient above. The limit then becomes

$$\lim_{h \to 0} \frac{f(p+h)g(p+h) - f(p)g(p+h) + f(p)g(p+h) - f(p)g(p)}{h}$$
$$= \lim_{h \to 0}[g(p+h)\frac{f(p+h) - f(p)}{h} + f(p)\frac{g(p+h) - g(p)}{h}]$$

This doesn't look at all like $f'(p)g'(p)$! In fact, since $\lim_{h \to 0} g(p+h) = g(p)$ by Theorem 6.14 (and Theorem 3.12), we have uncovered a completely unexpected formula:

$$(f \cdot g)'(p) = g(p)f'(p) + f(p)g'(p)$$

(for this step, we are using the sum and product theorems for limits (Theorems 4.1 and 4.9)).

Thus, even though our initial guess about a formula for the derivative of a product was wrong, we have discovered the following theorem:

Theorem 7.4: Let $X \subset \mathsf{R}^1$, and let $f, g : X \to \mathsf{R}^1$ be functions. If f and g are each differentiable at p, then $f \cdot g$ is differentiable at p and

$$(f \cdot g)'(p) = f(p)g'(p) + g(p)f'(p).$$

Proof: The proof is in the discussion above. ¥

It is an understatement to say that the formula in Theorem 7.3 is not intuitive. But, at the very least, could we have known that our original "formula" – $(f \cdot g)'(p) = f'(p)g'(p)$ – could not be true before we tried to prove it? Yes, if we had tried to apply our "formula" in any one of several simple cases, such as to the product $x \cdot x$ or even to the function x written as $1x$ (we already computed the relevent derivatives in Examples 6.2 and 6.3).

Do we now discard our false formula so no one will know we made such a silly mistake? No! We turn our mistake into a question: For what differentiable functions f and g on R^1 is it true that $(f \cdot g)'(x) = f'(x)g'(x)$ for all $x \in \mathsf{R}^1$? We return to this question later.

Next, we show that the quotient of two differentiable functions is differentiable and, at the same time, we derive a formula for the derivative of the quotient.

Note that a quotient $\frac{f}{g}$ can be viewed as the product $f \cdot \frac{1}{g}$. Therefore, to simplify the proof of our theorem about quotients, we first prove a lemma concerning reciprocals. We did the same thing when we proved the theorem on limits of quotients in section 4 of Chapter IV (Lemma 4.19 and Theorem 4.20).

Lemma 7.5: Let $X \subset \mathsf{R}^1$, and let $g : X \to \mathsf{R}^1$ be a function. If g is differentiable at p and $g(p) \neq 0$, then $\frac{1}{g}$ is differentiable at p and

$$(\tfrac{1}{g})'(p) = \tfrac{-g'(p)}{[g(p)]^2}.$$

Proof: We begin by trying to get a feeling for what is going on:

$$(\tfrac{1}{g})'(p) = \lim_{h\to 0} \frac{\frac{1}{g(p+h)} - \frac{1}{g(p)}}{h} = \lim_{h\to 0} \tfrac{1}{h}\tfrac{g(p)-g(p+h)}{g(p+h)g(p)};$$

hence, isolating the difference quotient for g from the rest, we have

(1) $(\frac{1}{g})'(p) = \lim_{h\to 0}[\frac{g(p+h)-g(p)}{h}\frac{-1}{g(p+h)g(p)}]$.

Now, we see what to do: We evaluate the limits of the two quotients on the right - hand side of (1) separately.

Since g is differentiable at p, we have that

(2) $\lim_{h\to 0} \frac{g(p+h)-g(p)}{h} = g'(p)$.

Since g is continuous at p by Theorem 6.14, $\lim_{h\to 0} g(p+h) = g(p)$ by Theorem 4.29; thus, since $g(p) \neq 0$, $\lim_{h\to 0} \frac{1}{g(p+h)} = \frac{1}{g(p)}$ by Lemma 4.19. Therefore, by the limit theorem on products (Theorem 4.9), we have

(3) $\lim_{h\to 0} \frac{-1}{g(p+h)g(p)} = \frac{-1}{[g(p)]^2}$.

By (1), (2) and (3), we can apply the limit theorem on products again to obtain that

$$(\tfrac{1}{g})'(p) = g'(p)\tfrac{-1}{[g(p)]^2} = \tfrac{-g'(p)}{[g(p)]^2}.$$

Have we proved the lemma? Yes, except for a technical detail: Even though $g(p) \neq 0$, there may be values h for which $g(p+h) = 0$, in which case the expression $\frac{1}{g(p+h)}$, which we used throughout the proof, does not make sense. However, this is easy to take care of: As already observed above (3),

$$\lim_{h\to 0} g(p+h) = g(p) \neq 0;$$

thus, there is a $\delta > 0$ such that

$$|g(p+h)| > \tfrac{|g(p)|}{2} \quad \text{when } p+h \in X \text{ and } |h| < \delta;$$

hence, $g(p+h) \neq 0$ for such h. Therefore, by stipulating at the beginning of the proof that all values h in the proof are restricted to those for which $|h| < \delta$, we take care of the matter. ¥

Theorem 7.6: Let $X \subset \mathsf{R}^1$, and let $f, g : X \to \mathsf{R}^1$ be functions. If f and g are each differentiable at p and $g(p) \neq 0$, then $\frac{f}{g}$ is differentiable at p and

$$(\frac{f}{g})'(p) = \frac{g(p)f'(p) - f(p)g'(p)}{[g(p)]^2}.$$

Proof: Since $\frac{f}{g} = f \cdot \frac{1}{g}$,

$$(\frac{f}{g})'(p) = (f \cdot \frac{1}{g})'(p) \overset{7.4}{=} f(p)(\frac{1}{g})'(p) + \frac{1}{g(p)}f'(p)$$

$$\overset{7.5}{=} f(p)\frac{-g'(p)}{[g(p)]^2} + \frac{1}{g(p)}f'(p) = \frac{-f(p)g'(p) + g(p)f'(p)}{[g(p)]^2}. \text{ ¥}$$

Exercise 7.7: Assume that $(f + g)(x) = x^3 + 5x - 3$, where f and g are differentiable functions and $f'(4) = 2$. Find $g'(4)$.

Exercise 7.8: Assume that $(f \cdot g)(x) = \frac{3x}{x^2+8}$, where f and g are differentiable functions such that $f(2) = 4$ and $f'(2) = 5$. Find $g'(2)$.

Exercise 7.9: Assume that $\frac{f}{g}(x) = x^2+2x$, where f and g are differentiable functions such that $f(2) = 2$ and $f'(2) = 3$. Find $g'(2)$.

Exercise 7.10: Let f and g be differentiable functions with $g(x) \neq 0$ for all x. Assume that the equation of the tangent line to the graph of f at $(2, f(2))$ is $3x - y - 5 = 0$ and that the equation of the tangent line to the graph of $\frac{f}{g}$ at $(2, \frac{f}{g}(2))$ is $2x + y + 4 = 0$. Find the equation of the tangent line to the graph of g at $(2, g(2))$.

2. Differentiating Polynomials and Rational Functions

In Chapter IV we showed that polynomials and rational functions are continuous. We now prove these types of functions are differentiable.

Our results will follow immediately from theorems in the preceding section once we prove a lemma.

Lemma 7.11: The function $f(x) = x^n$ is differentiable for each $n = 1, 2, ...$. In fact, for each $n = 1, 2, ...,$

$$(x^n)' = nx^{n-1}.$$

Proof: We prove the lemma by induction (Theorem 1.20).

We already know that $f(x) = x$ is differentiable and that $x' = 1 = 1x^0$ (Example 6.2); in other words, the lemma is true when $n = 1$.

Assume inductively that the lemma is true for a given natural number k.

We show using our inductive assumption that $(x^{k+1})' = (k + 1)x^k$.

Note that $x^{k+1} = xx^k$ and that, by our inductive assumption, $(x^k)' = kx^{k-1}$. Thus, since $x' = 1$, we can apply Theorem 7.3 on products to obtain

$$(x^{k+1})' = (xx^k)' = x(x^k)' + (x^k)x' = x(kx^{k-1}) + x^k$$

$$= kx^k + x^k = (k + 1)x^k.$$

The lemma now follows from the Induction Principle (Theorem 1.20). ¥

Theorem 7.12: Every polynomial is differentiable. Furthermore, if

$$f(x) = c_0 + c_1x + c_2x^2 + c_3x^3 + \cdots + c_nx^n,$$

then

$$f'(x) = c_1 + 2c_2x + 3c_3x^2 + \cdots + nc_nx^{n-1}.$$

Proof: By Theorem 7.4 on products and Lemma 7.11, $(cx^m)' = cmx^{m-1}$ for any constant c and any $m = 1, 2, \ldots$. Therefore, the theorem follows from Corollary 7.2. ¥

Theorem 7.13: Every rational function is differentiable on its domain.

Proof: Every point of the domain of a rational function is a limit point of its domain (you are asked to prove this in Exercise 7.14). Therefore, our theorem follows from Theorem 7.12 and Theorem 7.6. ¥

We close with a word of caution about computing derivatives. We know from Lemma 7.11 that $(x^4)' = 4x^3$. However, this does not say that $((2x)^4)' = 4(2x)^3$; in fact, since $(2x)^4 = 16x^4$, we see from Lemma 7.11 and Theorem 7.4 on products that $((2x)^4)' = (16x^4)' = 64x^3$. In other words, in general, if f is differentiable, Lemma 7.11 does not tell us how to differentiate $(f(x))^n$ or even whether $(f(x))^n$ is differentiable. We will learn about this in the next section.

Exercise 7.14: Prove the statement *every point of the domain of a rational function is a limit point of its domain*, which we used in the proof of Theorem 7.13. (No fair using that polynomials have only finitely many roots).

Exercise 7.15: $(x^n)' = nx^{n-1}$ for each $n = -1, -2, \ldots$.

Exercise 7.16: Find $f'(2)$ for each of the following functions f:

$$f(x) = -4x^5 + \frac{2}{x^3} - 7;\ f(x) = \frac{3x^2-2x+1}{(2x-1)^2};\ f(x) = \frac{x}{(4x-6)^3}.$$

Exercise 7.17: Let $f(x) = \frac{x}{(1+\frac{1}{x})^2}$. Find the equation of the tangent line to the graph of f at $(1, f(1))$.

Exercise 7.18: Find a function whose derivative is $3x^5 - 2x^2 + 1$.

Exercise 7.19: Find a function whose derivative is $\frac{1}{x^3} - (4x^2+1)^3$.

Exercise 7.20: Is there a polynomial of degree 3 that has horizontal tangent lines to its graph at three different points?

Exercise 7.21: Recall our discussion of the bogus formula $(f \cdot g)'(x) = f'(x)g'(x)$ following Theorem 7.4. When do polynomials f and g satisfy the formula?

3. The Chain Rule

We have proved that the composition of two continuous functions is continuous (Theorem 4.28). We now prove that the composition of two differentiable functions is differentiable and derive a formula for the derivative of the composition. The formula is called the Chain Rule (Theorem 7.23). The Chain Rule is useful in computing derivatives and has far-reaching theoretical consequences.

We will see applications of the Chain Rule in the next chapter (e.g., proof of Theorem 8.16) and in other chapters as well.

Assume that f and g are differentiable functions. Let us try to find out what the derivative of the composition $g \circ f$ should be. First, using the form in Exercise 6.10 for appearance sake only,

$$(g \circ f)'(p) = \lim_{x \to p} \tfrac{g(f(x)) - g(f(p))}{x-p} \text{ (if the limit exists).}$$

Next, as we have done on numerous occasions, we manipulate algebraically to obtain expressions that relate to our assumptions. Since we are assuming that f is differentiable, we want the difference quotient $\frac{f(x)-f(p)}{x-p}$ to appear as part of what we take the limit of to get $(g \circ f)'(p)$. We force this to happen by multiplying and dividing the expression $\frac{g(f(x))-g(f(p))}{x-p}$ by $f(x) - f(p)$, thereby obtaining

$$(g \circ f)'(p) = \lim_{x \to p} \tfrac{g(f(x)) - g(f(p))}{f(x)-f(p)} \tfrac{f(x)-f(p)}{x-p}.$$

Like the proverbial ostrich, we bury our head in the sand in order to believe that we have not divided by 0. Since $\lim_{x \to p}[f(x) - f(p)] = 0$ by Theorems 6.14 and 3.12, $\lim_{x \to p} \frac{g(f(x))-g(f(p))}{f(x)-f(p)}$ looks a lot like $g'(f(p))$. If the limit is $g'(f(p)$, then we can apply our theorem on limits of products (Theorem 4.9) to arrive at

$$(g \circ f)'(p) = \lim_{x \to p} \tfrac{g(f(x)) - g(f(p))}{f(x)-f(p)} \lim_{x \to p} \tfrac{f(x)-f(p)}{x-p} = g'(f(p))f'(p).$$

We have found a possible formula for $(g \circ f)'(p)$; we have not verified the formula (or even proved that $g \circ f$ is differentiable) since we may have divided by 0 in our computations. The following lemma overcomes this obstacle: the lemma will allow us to avoid limits of quotients with $f(x) - f(p)$ in the denominator, thereby verifying that the formula is indeed correct (Theorem 7.23).

Lemma 7.22: Let $X, Y, Z \subset \mathsf{R}^1$, and let $f : X \to Y$ and $g : Y \to Z$ be functions. Assume that f is continuous at p and that g is differentiable at $f(p) = q$. Define $G : Y \to \mathsf{R}^1$ by

$$G(y) = \begin{cases} \frac{g(y)-g(q)}{y-q} & \text{, if } y \neq q \\ g'(q) & \text{, if } y = q. \end{cases}$$

Then $\lim_{x \to p} G(f(x)) = g'(q)$ and

$$G(f(x)) \tfrac{f(x)-f(p)}{x-p} = \tfrac{g(f(x))-g(f(p))}{x-p}, \text{ all } x \in X - \{p\}.$$

Proof: Since g is differentiable at q, we see from Exercise 6.10 that

$$\lim_{y \to q} \tfrac{g(y)-g(q)}{y-q} = g'(q) = G(q);$$

hence, G is continuous at q by Theorem 3.12. Thus, since f is continuous at p and $f(p) = q$, $G \circ f$ is continuous at p by Theorem 4.28. Therefore, by Theorem 3.12,

$$\lim_{x\to p}(G\circ f)(x) = (G\circ f)(p) = G(f(p)) = G(q) = g'(q).$$

This proves the first part of the lemma.

To verify the equation in the second part of the lemma, let $x \in X - \{p\}$. Assume first that $f(x) \neq q$. Then, by the definition of G,

$$G(f(x)) = \tfrac{g(f(x))-g(q)}{f(x)-q};$$

thus, since $q = f(p)$,

$$G(f(x))\tfrac{f(x)-f(p)}{x-p} = \tfrac{g(f(x))-g(f(p))}{f(x)-f(p)}\tfrac{f(x)-f(p)}{x-p} = \tfrac{g(f(x))-g(f(p))}{x-p}.$$

This verifies the equation in the second part of the lemma when $f(x) \neq q$. Finally, when $f(x) = q$, we have $f(x) = f(p)$, so both sides of the equation are equal to 0. ¥

Theorem 7.23 (Chain Rule): Let $X, Y, Z \subset \mathsf{R}^1$, and let $f : X \to Y$ and $g : Y \to Z$ be functions. Assume that f is differentiable at p and that g is differentiable at $f(p) = q$. Then $g \circ f$ is differentiable at p and

$$(g\circ f)'(p) = g'(f(p))f'(p).$$

Proof: All the work is done (we can apply Lemma 7.22 below since f is continuous at p by Theorem 6.14):

$$(g\circ f)'(p) \overset{6.10}{=} \lim_{x\to p} \tfrac{g(f(x))-g(f(p))}{x-p} \overset{7.22}{=} \lim_{x\to p} G(f(x))\tfrac{f(x)-f(p)}{x-p};$$

also, $\lim_{x\to p} G(f(x)) = g'(q)$ (by Lemma 7.22) and $\lim_{x\to p} \frac{f(x)-f(p)}{x-p} = f'(p)$ (by Exercise 6.10). Therefore, we can apply Theorem 4.9 on limits of products to obtain

$$(g\circ f)'(p) = g'(q)f'(p) = g'(f(p))f'(p).$$ ¥

We conclude by illustrating how to use the Chain Rule in finding derivatives.

Example 7.24: Let $f(x) = (4x + 5)^{12}$. Note that $f = h \circ g$, where

$$g(x) = 4x + 5, \quad h(y) = y^{12}.$$

Hence, by the Chain Rule,

$$f'(x) = h'(g(x))g'(x) = 12(g(x))^{11}(4) = 48(4x + 5)^{11}.$$

Exercise 7.25: Find $f'(3)$ for each of the following functions f:

$$f(x) = \tfrac{1}{(1-x)^5};\ \ f(x) = \sqrt{x^6 + 3x^2 + 1};\ \ f(x) = [x + (x - x^3)^6]^7.$$

Exercise 7.26: Assume that $(g \circ f)(x) = \frac{x}{x+1}$, where f and g are differentiable functions such that $f(1) = 4$ and $g'(4) = 5$. Find $f'(1)$.

Exercise 7.27: Assume that $(g \circ f)(x) = x^4 + 3x$, where f and g are differentiable functions such that $f(2) = 3$ and $f'(2) = 5$. Find $g'(3)$.

Chapter VIII: The Inverse Function Theorem

The Inverse Function Theorem is concerned with one - to - one differentiable functions defined on an interval. The theorem tells us when the inverse of such a function is differentiable and provides a formula for the derivative.

After necessary preliminary results, we prove the Inverse Function Theorem in section 3. We apply the theorem in section 4 to show that rational powers of x are differentiable (where defined and for $x \neq 0$). We study the trigonometric functions in section 5: We show that the trigonometric functions are differentiable, and then we apply the Inverse Function Theorem to show that the inverse trigonometric functions are differentiable. We obtain formulas for the derivatives of rational powers, trigonometric functions and inverse trigonometric functions.

1. One - to - one Functions and Inverses

We recall some notions and notation from precalculus.

Let X and Y be sets. A function $f : X \to Y$ is said to be *one - to - one* provided that whenever $x_1, x_2 \in X$ and $x_1 \neq x_2$, then $f(x_1) \neq f(x_2)$.

Assume that $f : X \to Y$ is one - to - one. Then we can define a function $g : f(X) \to X$ as follows: For each $y \in f(X)$, $g(y)$ is the unique point in X that maps to y under f. In other words, $f(g(y)) = y$ for all $y \in f(X)$; in addition, $g(f(x)) = x$ for all $x \in X$. The function g is called the *inverse of* f, which we denote from now on by f^{-1}.

Do not confuse the notation f^{-1} with $\frac{1}{f}$; f^{-1} is the (unique) function such that $f \circ f^{-1}$ is the identity function on $f(X)$ and $f^{-1} \circ f$ is the identity function on X.

Let $X \subset \mathsf{R}^1$, and let $f : X \to \mathsf{R}^1$ be one - to - one. Then the graph of f^{-1} is obtained by reflecting the graph of f about the line $y = x$ in the plane. The reason is quite simple: The reflection about $y = x$ changes a point $(x, f(x))$ to the point $(f(x), x)$, and $f^{-1}(f(x)) = x$.

The simple relation between the graphs of f and f^{-1} just mentioned can provide geometric intuition for the Inverse Function Theorem and for some results preceding it. In particular, examining the graphs of f and f^{-1} in the same picture can serve to motivate the results and provide insight. I leave it to the reader to draw pictures of continuous one - to - one functions on intervals, together with their inverses, and differentiable one - to - one functions on intervals, together with their inverses, before reading further – try to predict (from the pictures) a geometric characterization of one - to - one continuous functions on intervals, and try to determine what the formula should be for the derivative of f^{-1} in terms of the derivative of f.

2. Continuity of the Inverse Function

We prove that the inverse of a one - to - one continuous function on an interval is continuous. We will use this result in the next section to prove that the inverse of a differentiable function (on an interval) is differentiable. The pattern should be familiar from the preceding chapter: There we used the continuity of compositions (in the proof of Lemma 7.22) in proving the Chain Rule.

Our result about continuity of the inverse function is Theorem 8.6. We prove the result by first characterizing one-to-one continuous functions defined on intervals in a geometric way. The terminology for the characterization is as follows:

Definition: Let $X \subset \mathsf{R}^1$ and let $f : X \to \mathsf{R}^1$ be a function. We say that f is *increasing on* X provided that whenever $x_1, x_2 \in X$ such that $x_1 < x_2$, then $f(x_1) \leq f(x_2)$; f is *strictly increasing on* X provided that whenever $x_1, x_2 \in X$ such that $x_1 < x_2$, then $f(x_1) < f(x_2)$.

Similarly, f is *decreasing on* X (or *strictly decreasing on* X) provided that whenever $x_1, x_2 \in X$ such that $x_1 < x_2$, then $f(x_1) \geq f(x_2)$ (or $f(x_1) > f(x_2)$, respectively).

If $Y \subset X$, we say f is *increasing* (*strictly increasing*, etc.) *on* Y to mean $f|Y$ is increasing (strictly increasing, etc.) on Y.

Exercise 8.1: Let $X \subset \mathsf{R}^1$ and let $f : X \to \mathsf{R}^1$ be a function. If f is strictly increasing on X, then f^{-1} is strictly increasing on $f(X)$; if f is strictly decreasing on X, then f^{-1} is strictly decreasing on $f(X)$.

It is obvious that if a function f is either strictly increasing or strictly decreasing, then f is one-to-one. Our characterization theorem says that the converse is also true when f is continuous on an interval (Theorem 8.4). We first prove a lemma; we use the lemma in the proof of the characterization theorem and in the proof of the subsequent theorem about the continuity of the inverse function. The proof of the lemma uses the Intermediate Value Theorem and the Maximum-Minimum Theorem.

Lemma 8.2: Let $f : [a,b] \to \mathsf{R}^1$ be a one-to-one continuous function.

(1) If $f(a) < f(b)$, then f is strictly increasing on $[a,b]$, $f([a,b]) = [f(a), f(b)]$, and f^{-1} is strictly increasing on $[f(a), f(b)]$.

(2) If $f(a) > f(b)$, then f is strictly decreasing on $[a,b]$, $f([a,b]) = [f(b), f(a)]$, and f^{-1} is strictly decreasing on $[f(b), f(a)]$.

Proof: We prove part (1); part (2) follows easily from part (1) (Exercise 8.3).

Assume that $f(a) < f(b)$. Then, by the last part of Theorem 5.13,

$$f([a,b]) = [c,d] \text{ for some } c < d.$$

We show that $f(a) = c$ and $f(b) = d$. Since $f([a,b]) = [c,d]$, there exist $s, t \in [a,b]$ such that $f(s) = c$ and $f(t) = d$. Let J denote the closed interval with end points s and t (i.e., $J = [s,t]$ if $s < t$ and $J = [t,s]$ if $t < s$). Since $f(s) = c$ and $f(t) = d$ and since $f(J) \subset [c,d]$, we see by the Intermediate Value Theorem (Theorem 5.2) that $f(J) = [c,d]$. Thus, since $f(a), f(b) \in [c,d]$, there exist $p, q \in J$ such that $f(p) = f(a)$ and $f(q) = f(b)$. Now, since f is one-to-one on $[a,b]$, $p = a$ and $q = b$. Hence, $a, b \in J$. Thus, $J = [a,b]$. Therefore, $s = a$ or b, and $t = a$ or b; furthermore, if $s = b$, then $t = a$, and we have

$$f(b) = f(s) = c < d = f(t) = f(a),$$

which contradicts our assumption that $f(a) < f(b)$. Hence, $s = a$ and, consequently, $t = b$. Therefore, $f(a) = c$ and $f(b) = d$.

We have proved the following:

$$(*)\ f([a, b]) = [f(a), f(b)].$$

We use (*) to prove that f is strictly increasing on $[a, b]$. Suppose that f is not strictly increasing on $[a, b]$. Then, since f is one-to-one, there are points x_1 and x_2 such that

$$a \leq x_1 < x_2 \leq b \quad \text{and} \quad f(x_1) > f(x_2).$$

Furthermore, $x_1 > a$ (otherwise, $x_1 = a$ and, hence, $f(a) > f(x_2)$, which contradicts (*)); also, since f is one-to-one and $a \neq x_2$, we see from (*) that $f(a) < f(x_2)$. To summarize, we have that

$$a < x_1 \quad \text{and} \quad f(a) < f(x_2) < f(x_1).$$

Thus, by the Intermediate Value Theorem (Theorem 5.2), there exists a point $c \in (a, x_1)$ such that $f(c) = f(x_2)$. However, this contradicts that f is one-to-one (since $c < x_1 < x_2$). Therefore, we have proved that f is strictly increasing on $[a, b]$.

Finally, we have shown that f is strictly increasing on $[a, b]$ and that $f([a, b]) = [f(a), f(b)]$ (by (*)); therefore, by Exercise 8.1, f^{-1} is strictly increasing on $[f(a), f(b)]$.

This completes the proof of part (1) of the lemma; part (2) is left as Exercise 8.3. ¥

Exercise 8.3: Finish the proof of Lemma 8.2 by showing how part (2) follows quickly from part (1).

We are now ready to prove the characterization theorem.

Theorem 8.4: Let I be an interval, and let $f : I \to \mathsf{R}^1$ be a continuous function. Then f is one-to-one if and only if f is either strictly increasing on I or strictly decreasing on I.

Proof: If f is either strictly increasing on I or strictly decreasing on I, then it is clear that f is one-to-one. Therefore, we need only prove the converse.

Any interval can be written as a countable union of closed and bounded intervals $[a_n, b_n]$, $n = 1, 2, \ldots$, where $[a_n, b_n] \subset [a_{n+1}, b_{n+1}]$ for all n. For example, $(a, b) = \cup_{n=1}^{\infty}[a + \frac{b-a}{2n}, b - \frac{b-a}{2n}]$, $[a, b) = \cup_{n=1}^{\infty}[a, b - \frac{b-a}{2n}]$, $(a, \infty) = \cup_{n=1}^{\infty}[a + \frac{1}{n}, a + n]$, and so on. Thus, whatever kind of interval the interval I in our theorem is (excluding the trivial case when $I = [a, a]$), we have

$$I = \cup_{n=1}^{\infty}[a_n, b_n], \quad [a_n, b_n] \subset [a_{n+1}, b_{n+1}] \text{ for all } n, \quad a_1 < b_1.$$

Now, assume that $f : I \to \mathsf{R}^1$ is one-to-one. Then either $f(a_1) < f(b_1)$ or $f(b_1) < f(a_1)$.

Assume first that $f(a_1) < f(b_1)$. Then, by part (1) of Lemma 8.2, f is strictly increasing on $[a_1, b_1]$. Assume inductively that f is strictly increasing on $[a_k, b_k]$ for some given k. Since f is one-to-one, either $f(a_{k+1}) < f(b_{k+1})$ or $f(b_{k+1}) < f(a_{k+1})$. If $f(b_{k+1}) < f(a_{k+1})$, then we see from part (2) of Lemma 8.2 that f is strictly decreasing on $[a_{k+1}, b_{k+1}]$, hence on $[a_k, b_k]$; this contradicts our inductive assumption that f is strictly increasing on $[a_k, b_k]$. Hence, $f(a_{k+1}) < f(b_{k+1})$. Therefore, by part (1) of Lemma 8.2, f is strictly increasing on $[a_{k+1}, b_{k+1}]$. Hence, by the Induction Principle (Theorem 1.20), we have proved that f is strictly increasing on $[a_n, b_n]$ for all n. Therefore, since $I = \cup_{n=1}^{\infty}[a_n, b_n]$, it follows easily that f is strictly increasing on I.

We leave the case when $f(a_1) > f(b_1)$ as an exercise (Exercise 8.5). ¥

Exercise 8.5: Finish the proof of Theorem 8.4 (by taking care of the case when $f(a_1) > f(b_1)$).

Finally, we prove our main theorem.

Theorem 8.6: Let I be an interval. If $f : I \to \mathsf{R}^1$ is a one-to-one continuous function, then f^{-1} is continuous on $f(I)$.

Proof: By Theorem 8.4, f is either strictly increasing on I or strictly decreasing on I. We assume that

(1) f is strictly increasing on I.

By (1) and Exercise 8.1, we have that

(2) f^{-1} is strictly increasing on $f(I)$.

Now, to prove that f^{-1} is continuous on $f(I)$, let $p \in f(I)$. We prove that $\lim_{y \to p} f^{-1}(y) = f^{-1}(p)$.

Let $\epsilon > 0$. Let $q = f^{-1}(p)$.

By the Intermediate Value Theorem, $f(I)$ is an interval. We take two cases:

Case 1*:* p is not an end point of $f(I)$. Then it follows from (2) that q is not an end point of I. Hence, we can assume that ϵ is small enough so that

$$[q - \epsilon, q + \epsilon] \subset I.$$

Thus, since f is strictly increasing on $[q - \epsilon, q + \epsilon]$ (by (1)), we have

(3) $f(q - \epsilon) < f(q) = p < f(q + \epsilon)$.

By (1), f is strictly increasing on $[q-\epsilon, q+\epsilon]$; hence, by Lemma 8.2, we have that

(4) $f([q - \epsilon, q + \epsilon]) = [f(q - \epsilon), f(q + \epsilon)]$.

Now, let

$$\delta = \min\{p - f(q - \epsilon), f(q + \epsilon) - p\}.$$

By (3), $\delta > 0$. Assume that $|y - p| < \delta$. Then $y \in (f(q-\epsilon), f(q+\epsilon))$ since

$$f(q-\epsilon) = p - [p - f(q-\epsilon)] \leq p - \delta < y < p + \delta$$
$$\leq p + [f(q+\epsilon) - p] = f(q+\epsilon).$$

Hence, by (4), $f^{-1}(y) \in (q-\epsilon, q+\epsilon)$; in other words, $\left|f^{-1}(y) - q\right| < \epsilon$. Therefore, since $q = f^{-1}(p)$, $\left|f^{-1}(y) - f^{-1}(p)\right| < \epsilon$. Thus, we have proved that

$$\lim_{y\to p} f^{-1}(y) = f^{-1}(p).$$

Therefore, f^{-1} is continuous at p by Theorem 3.12.

Case 2: p is an end point of $f(I)$. Then it follows from (2) that q is an end point of I, and it is easy to modify the argument for Case 1 to prove that f^{-1} is continuous at p (replace $[q-\epsilon, q+\epsilon]$ with either $[q, q+\epsilon]$ or $[q-\epsilon, q]$, and make the obvious adjustments in the rest of the proof for Case 1). ¥

3. The Inverse Function Theorem

We prove the main theorem of the chapter. The assumption in the theorem that $f'(p) \neq 0$ is unconditionally necessary (see Exercise 8.8).

Theorem 8.7 (Inverse Function Theorem): Let I be an interval, let $f : I \to \mathsf{R}^1$ be a one-to-one continuous function, and let $p \in I$. If f is differentiable at p and $f'(p) \neq 0$, then f^{-1} is differentiable at $f(p) = q$ and

$$(f^{-1})'(q) = \frac{1}{f'(p)} = \frac{1}{f'(f^{-1}(q))}.$$

Proof: We will use Lemma 7.22 with f in the lemma replaced by f^{-1} here and g in the lemma replaced by f here (thus, the roles of p and q in the lemma are switched here). The function F defined below is the function G in Lemma 7.22 with the replacements just mentioned:

$$F(x) = \begin{cases} \frac{f(x)-f(p)}{x-p} & \text{, if } x \neq p \\ f'(p) & \text{, if } x = p. \end{cases}$$

The assumptions of Lemma 7.22 are satisfied since f^{-1} is continuous at q (by Theorem 8.6) and f is differentiable at $f^{-1}(q) = p$ (by assumption in our theorem). Hence, by Lemma 7.22 (as adjusted here),

$$\lim_{y\to q} F(f^{-1}(y)) = f'(p).$$

Thus, since $f'(p) \neq 0$ (by assumption), $\lim_{y\to q} \frac{1}{F(f^{-1}(y))} = \frac{1}{f'(p)}$ by Lemma 4.19. Therefore, using the formula for F for the first equality below,

$$\frac{1}{f'(p)} = \lim_{y\to q} \frac{1}{\frac{f(f^{-1}(y))-f(p)}{f^{-1}(y)-p}} = \lim_{y\to q} \frac{f^{-1}(y)-p}{f(f^{-1}(y))-f(p)}$$
$$= \lim_{y\to q} \frac{f^{-1}(y)-f^{-1}(q)}{y-q} \overset{6.10}{=} (f^{-1})'(q). \text{ ¥}$$

Exercise 8.8: The assumption that $f'(p) \neq 0$ in Theorem 8.7 is absolutely necessary: *If I is an interval and $f : I \to R^1$ is a one-to-one differentiable function such that $f'(p) = 0$ for some point p, then f^{-1} is not differentiable at $f(p)$.* Prove this result, and explain why the result is to be expected from a picture of the graphs of f and f^{-1}.

Exercise 8.9: In the proof of Theorem 8.7, we proved that f^{-1} is differentiable at q by deriving the formula for $(f^{-1})'(q)$. If we had known beforehand that f^{-1} is differentiable at q, then we could have derived the formula for $(f^{-1})'(q)$ using the Chain Rule (Theorem 7.23). Show how to derive the formula for $(f^{-1})'(q)$ using the Chain Rule under the assumption that f^{-1} is differentiable at q (and the assumptions about f in Theorem 8.7).

Exercise 8.10: Find $(f^{-1})'(1)$ when $f(x) = x^7 + x^3 + x + 1$.

Exercise 8.11: Find $(f^{-1})'(6)$ when $f(x) = \sqrt{x^3 + 2x + 3}$.

Exercise 8.12: Let $f : \mathsf{R}^1 \to \mathsf{R}^1$ be a one-to-one differentiable function such that $f(3) = 4$ and $f'(3) = \frac{1}{4}$. Let $h = \frac{1}{f^{-1}}$. Find $h'(4)$.

4. Differentiating Rational Powers

We know that for all integers n, x^n is differentiable and $(x^n)' = nx^{n-1}$ (by Lemma 7.11 and Exercise 7.15). We extend the result to expressions of the form $x^{\frac{m}{n}}$, where m and n are integers ($n \neq 0$) and $x \neq 0$. The proof is an application of the Inverse Function Theorem (Theorem 8.7) and the Chain Rule (Theorem 7.23).

We begin by examining the function $f(x) = x^n$, where n is a natural number. We need to distinguish between the case when n is even and the case when n is odd; the reason will be apparent when we use the following lemma as a guide for defining the n^{th} root function.

Lemma 8.13: Let n be a natural number, and let $f(x) = x^n$ for all $x \in \mathsf{R}^1$.

(1) If n is even, then f is strictly increasing, hence one-to-one, on $[0, \infty)$ and $f([0, \infty)) = [0, \infty)$.

(2) If n is odd, then f is strictly increasing, hence one-to-one, on R^1 and $f(\mathsf{R}^1) = \mathsf{R}^1$.

Proof: All the numbers $0, 1, 2^n, 3^n, ..., k^n, ...$ are values of f; in addition, if n is odd, $n = 2m+1$, all the numbers $-k(-k)^{2m}$ for $k = 0, 1, 2, ...$ are values of f. Also, f is continuous by Theorem 4.16. Hence, it follows from the Intermediate Value Theorem (Theorem 5.2) and Lemma 1.21 that $f([0, \infty)) = [0, \infty)$ and, if n is odd, $f(\mathsf{R}^1) = \mathsf{R}^1$.

The fact that f is strictly increasing can be proved by induction; we leave this to the reader (Exercise 8.14).

Finally, f is one-to-one since a strictly increasing function is obviously one-to-one. ¥

Exercise 8.14: Finish the proof of Lemma 8.13 as indicated.

Lemma 8.13 provides a completely different proof of Theorem 1.25 and extends Theorem 1.25 to negative real numbers when n is odd. Which proof of Theorem 1.25 do you like better – the original proof or this proof?

Definition: Let n be a natural number, and let $f(x) = x^n$ for all $x \in \mathsf{R}^1$.

- With Lemma 8.13 in mind, we define the n^{th} *root function* to be the inverse of $f|[0,\infty)$ if n is even and to be the inverse of f if n is odd. Hence, the n^{th} root function has domain and range $[0,\infty)$ when n is even, and the n^{th} root function has domain and range R^1 when n is odd.
- The value of the n^{th} root function at x is denoted by $x^{\frac{1}{n}}$. Thus, we have *defined* $x^{\frac{1}{n}}$ for all $x \geq 0$ when n is even and for all real numbers x when n is odd; in other words, $x^{\frac{1}{n}}$ is defined to be the unique number such that

$$(x^{\frac{1}{n}})^n = x = (x^n)^{\frac{1}{n}}.$$

- For any integer m, $x^{\frac{m}{n}}$ is defined to be $(x^{\frac{1}{n}})^m$ for all x such that $x^{\frac{1}{n}}$ is defined; in addition, $x = 0$ is excluded if $m < 0$. Thus, except for $x = 0$ if $m < 0$, the function $h(x) = x^{\frac{m}{n}}$ is defined (only) on $[0,\infty)$ if n is even and on all of R^1 if n is odd.

The following theorem, which we use to prove our main theorem, is a consequence of the Inverse Function Theorem.

Theorem 8.15: Let g denote the n^{th} root function for some natural number n. Then g is differentiable at every point x in its domain except $x = 0$ and

$$g'(x) = (x^{\frac{1}{n}})' = \tfrac{1}{n}x^{\frac{1}{n}-1}.$$

Proof: Let $f(x) = x^n$, where f is restricted to $[0,\infty)$ if n is even. Note that $g = f^{-1}$. We will apply the Inverse Function Theorem (Theorem 8.7) to g. To know that we can do so, note the following: f is one-to-one (by Lemma 8.13), f is continuous (by Theorem 4.16), and $f'(x) = nx^{n-1}$ (by Lemma 7.11), hence $f'(x) \neq 0$ if $x \neq 0$. Therefore, by the Inverse Function Theorem, if $x \neq 0$,

$$g'(x) = \tfrac{1}{f'(g(x))} = \tfrac{1}{f'(x^{\frac{1}{n}})} = \tfrac{1}{n(x^{\frac{1}{n}})^{n-1}} = \tfrac{1}{nx^{1-\frac{1}{n}}} = \tfrac{1}{n}x^{\frac{1}{n}-1}.$$

Finally, since $f'(0) = 0$, we know that g is not differentiable at $x = 0$ by Exercise 8.8. ¥

We now prove our main theorem using Theorem 8.15 and the Chain Rule.

Theorem 8.16: Let n be a natural number, and let $m \neq 0$ be an integer. The function $h(x) = x^{\frac{m}{n}}$ is differentiable at every point x in its domain except $x = 0$ and

$$h'(x) = (x^{\frac{m}{n}})' = \tfrac{m}{n}x^{\frac{m}{n}-1}.$$

Proof: By the definition of $x^{\frac{m}{n}}$ (above Theorem 8.15), $h(x) = (x^{\frac{1}{n}})^m$. Hence, $h = f \circ g$, where $g(x) = x^{\frac{1}{n}}$ and $f(x) = x^m$. Thus, by the Chain Rule (Theorem 7.23), $h'(x) = f'(g(x))g'(x)$. Therefore, since $g'(x) = \frac{1}{n}x^{\frac{1}{n}-1}$ (by Theorem 8.15) and $f'(x) = mx^{m-1}$ (by Lemma 7.11 and Exercise 7.15), we have

$$h'(x) = f'(x^{\frac{1}{n}})g'(x) = m(x^{\frac{1}{n}})^{m-1}(\tfrac{1}{n}x^{\frac{1}{n}-1}) = \tfrac{m}{n}x^{\frac{m}{n}-1}. \text{ ¥}$$

It is natural to wonder if Theorem 8.16 holds for *all* powers of x rather than just for rational powers (when considering an irrational power of x, we assume that $x > 0$). However – we must first wonder what x^p means for a given irrational number p: What do we mean by $2^{\sqrt{2}}$, 3^{π}, etc.? Once we give an appropriate definition of x^p for any given irrational number p (and $x > 0$), we will see that $(x^p)' = px^{p-1}$ for any given real number p and all $x > 0$. The definition of x^p for irrational powers p awaits further developments, namely, the natural logarithm, which we define in Chapter XVI using the integral. The definition for x^p is above Exercise 16.20, and the result about the derivative of x^p is Theorem 16.31.

5. Differentiating Trigonometric Functions and Their Inverses

We first show that the trigonometric functions are differentiable. The fact that the inverse trigonometric functions are differentiable is then a consequence of the Inverse Function Theorem. As we have done in the past, we obtain formulas for all derivatives.

We assume that the reader is familiar with the definitions of the trigonometric functions and basic trigonometric identities. The independent variable, x, for a trigonometric function is a real number that is to be understood as the angle whose radian measure is x. Thus, when we write $\sin(x)$, $\cos(x)$ and so on, we assume x is radian measure; when we use degree measure, we will specifically write x° to mean x measured in degrees. We note the relationship between radian measure and degree measure: $1^\circ = \frac{\pi}{180}$ radians.

We denote a trigonometric function raised to a power with a superscript directly after the function; for example $\sin^2(x)$ denotes $(\sin(x))^2$. As is consistent with our notation for inverse functions in general (section 1), we denote inverse trigonometric functions with a superscript of -1 directly after the function; for example, $\sin^{-1}(x)$ denotes the inverse sine of x, not $\frac{1}{\sin(x)}$ (which we denote by $(\sin(x))^{-1}$). The reader should *not*, for example, confuse $\sin^{-p}(x)$ with $(\sin^{-1}(x))^p$ when $p \neq 1$; $\sin^{-p}(x)$ for $p \neq 1$ always means $(\sin(x))^{-p} = \frac{1}{\sin^p(x)}$.

We use notation for the derivative of a trigonometric function and the derivative of its inverse that is consistent with our notation for derivatives in general: $\sin'$ or $\sin'(x)$ denotes the derivative of the sine function, $(\sin^{-1})'$ or $(\sin^{-1})'(x)$ denotes the derivative of the inverse sine function, and so forth.

We let S^1 denote the unit circle in the plane R^2 (i.e., S^1 is all points (x, y) in R^2 such that $\sqrt{x^2 + y^2} = 1$).

We note that the sine and cosine functions are continuous for use later:

Lemma 8.17: The sine function and the cosine function are continuous.

Proof: We do not belabor the proof that the sine and cosine functions are continuous. Their continuity is geometrically clear: Simply recognize that $(\cos(x), \sin(x))$ is the point on the unit circle S^1 corresponding to the angle whose radian measure is x, and observe that small changes in x result in small changes in the points $(\cos(x), \sin(x))$. ¥

To show that all trigonometric functions are differentiable, we focus on the sine function. Once we prove the sine function is differentiable, the differentiability of *all* trigonometric functions follows using elementary facts from trigonometry.

Let us see what is involved in showing that the sine function is differentiable: For a given x and for any $h \neq 0$,

$$\frac{\sin(x+h)-\sin(x)}{h} = \frac{\sin(x)\cos(h)+\sin(h)\cos(x)-\sin(x)}{h};$$

since $\lim_{h\to 0} \frac{\sin(x)}{h}$ and $\lim_{h\to 0} \frac{\cos(x)}{h}$ do not exist, we have no hope of proving that $\sin(x)$ is differentiable unless we put the expressions involving h together, obtaining

$$\frac{\sin(x+h)-\sin(x)}{h} = \sin(x)\frac{\cos(h)-1}{h} + \cos(x)\frac{\sin(h)}{h}.$$

Thus, we need to find two limits, $\lim_{h\to 0} \frac{1-\cos(h)}{h}$ and $\lim_{h\to 0} \frac{\sin(h)}{h}$, if the limits do, indeed, exist. The problem is not trivial, but can be solved with the aid of some elementary geometry:

Lemma 8.18: $\lim_{x\to 0} \frac{\sin(x)}{x} = 1$ and $\lim_{x\to 0} \frac{1-\cos(x)}{x} = 0$.

Proof: We first prove that

(1) $\lim_{x\to 0^+} \frac{\sin(x)}{x} = 1$.

Proof of (1)*:* Assume that $0 < x < \frac{\pi}{2}$. Referring to Figure 8.18 below, we see that

$$area(\Delta OAB) < area(\text{sector } OAC) < area(\Delta ODC).$$

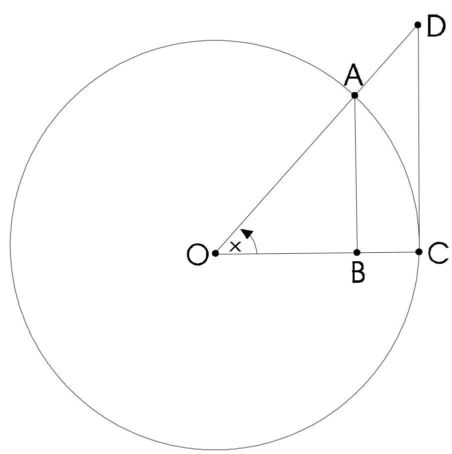

Figure 8.18

We write the inequalities above Figure 8.18 in terms of x (note: since a semicircle has area $\frac{\pi}{2}$ and is a sector with angle π, sector OAC with angle x has proportional area $\frac{x}{\pi}\frac{\pi}{2}$, which is $\frac{x}{2}$):

$$\tfrac{1}{2}\cos(x)\sin(x) < \tfrac{x}{2} < \tfrac{1}{2}\tan(x) = \tfrac{\sin(x)}{2\cos(x)}.$$

Since $0 < x < \frac{\pi}{2}$, $\sin(x) > 0$; hence, the inequalities remain in the same direction when we multiply through by $\frac{2}{\sin(x)}$, obtaining

$$\cos(x) < \tfrac{x}{\sin(x)} < \tfrac{1}{\cos(x)}.$$

Hence, taking reciprocals (thereby reversing inequalities), we have

$$\tfrac{1}{\cos(x)} > \tfrac{\sin(x)}{x} > \cos(x).$$

Moreover, by Lemma 8.17, Theorem 3.12, and Lemma 4.19,

$$\lim_{x\to 0^+} \cos(x) = 1 = \lim_{x\to 0^+} \tfrac{1}{\cos(x)}$$

Therefore, by the Squeeze Theorem (Theorem 4.34), $\lim_{x\to 0^+} \frac{\sin(x)}{x} = 1$. This proves (1).

Next, we prove that

(2) $\lim_{x\to 0^-} \frac{\sin(x)}{x} = 1$.

Proof of (2): Define $g : [0, \frac{\pi}{2}) \to \mathsf{R}^1$ by

$$g(x) = \begin{cases} \frac{\sin(x)}{x} & \text{, if } x > 0 \\ 1 & \text{, if } x = 0. \end{cases}$$

By (1) and Theorem 3.15, g is continuous. Define $f : (-\frac{\pi}{2}, 0] \to [0, \frac{\pi}{2})$ by $f(x) = -x$; obviously, f is continuous. Hence, by Theorem 4.28, $g \circ f$ is continuous. Thus, by Theorem 3.12,

$$\lim_{x\to 0^-}(g \circ f)(x) = (g \circ f)(0) = g(0) = 1;$$

furthermore, since $\sin(-x) = -\sin(x)$, we have that

$$(g \circ f)(x) = g(-x) = \tfrac{\sin(-x)}{-x} = \tfrac{-\sin(x)}{-x} = \tfrac{\sin(x)}{x}, \quad -\tfrac{\pi}{2} < x < 0.$$

Therefore,

$$\lim_{x\to 0^-} \tfrac{\sin(x)}{x} = \lim_{x\to 0^-}(g \circ f)(x) = 1.$$

This proves (2).

By (1), (2), and Theorem 3.16, $\lim_{x\to 0} \frac{\sin(x)}{x} = 1$. This proves that the first part of the lemma.

To prove the second part of the lemma, first observe that when $-\frac{\pi}{2} < x < \frac{\pi}{2}$ (to assure that $\cos(x) \neq -1$),

$$\frac{1-\cos(x)}{x} = \frac{1-\cos(x)}{x}\frac{1+\cos(x)}{1+\cos(x)} = \frac{\sin^2(x)}{x[1+\cos(x)]} = \frac{\sin(x)}{x}\frac{\sin(x)}{1+\cos(x)}.$$

Next, note that $\lim_{x\to 0}\frac{\sin(x)}{x} = 1$ (by the first part of the lemma) and that $\lim_{x\to 0}\frac{\sin(x)}{1+\cos(x)} = 0$ (by Lemma 8.17, Corollary 4.21 and Theorem 3.12). Therefore, by Theorem 4.9 on limits of products,

$$\lim_{x\to 0}\frac{1-\cos(x)}{x} = (\lim_{x\to 0}\frac{\sin(x)}{x})(\lim_{x\to 0}\frac{\sin(x)}{1+\cos(x)}) = (1)(0) = 0. \text{ ¥}$$

Exercise 8.19: Fix nonzero real numbers a and b. Find $\lim_{x\to 0}\frac{\sin(ax)}{bx}$ by making use of Theorem 3.15. Show all work carefully.

It is now easy to prove our result for the sine function:

Theorem 8.20: $\sin'(x) = \cos(x)$.

Proof: Fix $x \in \mathsf{R}^1$. Continuing from where we left off above Lemma 8.18,

$$\sin'(x) = \lim_{h\to 0}[\sin(x)\frac{\cos(h)-1}{h} + \cos(x)\frac{\sin(h)}{h}].$$

Therefore, by Lemma 8.18 and Theorem 4.1 on limits of sums,

$$\sin'(x) = \lim_{h\to 0}\sin(x)\frac{\cos(h)-1}{h} + \lim_{h\to 0}\cos(x)\frac{\sin(h)}{h} = \cos(x). \text{ ¥}$$

Corollary 8.21: $\cos'(x) = -\sin(x)$.

Proof: Since $\cos(x) = \sin(\frac{\pi}{2} - x)$ for all x, we see from Theorem 8.20 and the Chain Rule (Theorem 7.23) that

$$\cos'(x) = [\cos(\frac{\pi}{2} - x)][-1] = -\cos(\frac{\pi}{2} - x).$$

Therefore, since $\cos(\frac{\pi}{2} - x) = \sin(x)$, we have that $\cos'(x) = -\sin(x)$. ¥

Exercise 8.22: Using that all trigonometric functions can be expressed in terms of the sine and/or cosine functions, prove that the following formulas hold (for x in the domain of each function): $\tan'(x) = \sec^2(x)$, $\cot'(x) = -\csc^2(x)$, $\sec'(x) = \sec(x)\tan(x)$, and $\csc'(x) = -\csc(x)\cot(x)$.

Exercise 8.23: Would you expect the rate of change of a trigonometric function with respect to radian measure to be greater, smaller, or the same as the rate of change of the trigonometric function with respect to degree measure? Explain your answer intuitively, and prove your answer is correct.

Exercise 8.24: Direct computations using the Chain Rule (Theorem 7.23) and Theorem 8.20 give that

$$(\sin^2(x))' = 2\sin(x)\cos(x).$$

Thus, $(\sin^2(x))' = \sin(2x)$. Is this a coincidence, or can you explain why the result is to be expected from, say, a geometric point of view?

We turn our attention to derivatives of the inverse trigonometric functions.

The inverse sine function, $\sin^{-1}$, has domain $[-1,1]$ and range $[-\frac{\pi}{2}, \frac{\pi}{2}]$. Differentiating the inverse sine function is simply a matter of applying the Inverse Function Theorem (Theorem 8.7) in conjunction with Theorem 8.20:

Theorem 8.25: $(\sin^{-1})'(x) = \frac{1}{\sqrt{1-x^2}}$.

Proof: Fix $x \in [-1,1]$. Since $\sin' = \cos$ (Theorem 8.20), we see from the Inverse Function Theorem (Theorem 8.7) that

$$(\sin^{-1})'(x) = \frac{1}{\sin'(\sin^{-1}(x))} = \frac{1}{\cos(\sin^{-1}(x))} = \frac{1}{\sqrt{1-x^2}}. \;\text{¥}$$

Exercise 8.26: The inverse cosine function has domain $[-1,1]$ and range $[0,\pi]$. Prove that $(\cos^{-1})'(x) = \frac{-1}{\sqrt{1-x^2}}$.

Exercise 8.27: The inverse tangent has domain R^1 and range $(-\frac{\pi}{2}, \frac{\pi}{2})$. Prove that $(\tan^{-1})'(x) = \frac{1}{1+x^2}$.

Exercise 8.28: The inverse cotangent has domain R^1 and range $(0,\pi)$. Prove that $(\cot^{-1})'(x) = \frac{-1}{1+x^2}$.

Exercise 8.29: The inverse secant has domain $(-\infty,-1)\cup(1,\infty)$ and range $[0,\frac{\pi}{2})\cup(\frac{\pi}{2},\pi]$. Prove that $(\sec^{-1})'(x) = \frac{1}{|x|\sqrt{x^2-1}}$.

Exercise 8.30: The inverse cosecant has domain $(-\infty,-1)\cup(1,\infty)$ and range $[-\frac{\pi}{2},0)\cup(0,\frac{\pi}{2}]$. Prove that $(\csc^{-1})'(x) = \frac{-1}{|x|\sqrt{x^2-1}}$.

Chapter IX: Maxima, Minima and Derivatives

As a student in plane geometry, you may have seen the following problem: *If p and q are points on the same side of a line ℓ, find a point r on ℓ such that the sum of the distances pr and rq is a minimum.* The problem is solved easily by reflecting q across the line ℓ to the point q', and then observing that r must be the point on ℓ where the line from p to q' meets ℓ. What you may not have observed is that the minimum path from p to ℓ to q is the path for which the angles formed by pr and ℓ and by qr and ℓ are equal. Is this symmetry only a coincidence?

Is it merely a coincidence that the largest area enclosed by all curves in the plane of a given length is the area enclosed by the most symmetric of those curves (the circle)? And is it a coincidence that of all the rectangles having a given perimeter, the one with the largest area is the one that is most symmetric (the square)?

Surely, beauty in nature is intimately connected with symmetry, and it would appear that symmetry is connected with maxima and minima. Perhaps this is why maximum and minimum problems have been a constant theme throughout history. Leonhard Euler (1707 - 1783) articulated the importance of maxima and minima by saying that all interesting phenomena in this world can be explained in terms of maxima and minima.

We began our study of maxima and minima in Chapter V in the setting of continuous functions; there we proved that every continuous function on a closed and bounded interval has a maximum value and a minimum value (Theorem 5.13). We now localize the notions of maxima and minima and relate the local notions to derivatives. Our main result is Theorem 9.7, which lays the foundation for further study of maxima and minima. Theorem 9.7 sets the stage for the proof of the Mean Value Theorem (which we prove in the next chapter); the Mean Value Theorem is the essential ingredient for proving theorems that are used to classify local maxima and minima.

1. Neighborhoods

The following descriptive terminology will help us formulate statements concisely.

Definition: Let $X \subset \mathsf{R}^1$, and let $p \in X$. A *neighborhood of p in X* is the intersection of X with any open interval in R^1 containing p; in other words, if (a, b) is an open interval in R^1 such that $p \in (a, b)$, then $X \cap (a, b)$ is a neighborhood of p in X.

If $\epsilon > 0$, then $X \cap (p-\epsilon, p+\epsilon)$ is called the *ϵ - neighborhood of p in X*; thus, the ϵ - neighborhood of p in X is $\{x \in X : |p - x| < \epsilon\}$.

Example 9.1: $(-1, 1)$ is a neighborhood of 0 in $[-1, 1]$ (the 1 - neighborhood of 0 in $[-1, 1]$); $[0, \frac{1}{2})$ is a neighborhood of 0 in $[0, 1]$ (the $\frac{1}{2}$ - neighborhood of 0 in $[0, 1)$), but $[0, \frac{1}{2})$ is not a neighborhood of 0 in $[-1, 1]$; $(-\frac{1}{4}, \frac{1}{2})$ is a