CHAPTER IV
DIFFERENTIATION, LOCAL BEHAVIOR
$e^{i\pi} = -1$.

In this chapter we will finally see why $e^{i\pi}$ is -1. Along the way, we will give careful proofs of all the standard theorems of Differential Calculus, and in the process we will discover all the familiar facts about the trigonometric and exponential functions. At this point, we only know their definitions as power series functions. The fact that $\sin^2 + \cos^2 = 1$ or that $e^{x+y} = e^x e^y$ are not at all obvious. In fact, we haven't even yet defined what is meant by e^x for an arbitrary number x.

The main theorems of this chapter include:

(1) The **Chain Rule** (Theorem 4.7),
(2) The **Mean Value Theorem** (Theorem 4.9),
(3) The **Inverse Function Theorem** (Theorem 4.10),
(4) The **Laws of Exponents** (Corollary to Theorem 4.11 and Exercise 4.20), and
(5) **Taylor's Remainder Theorem** (Theorem 4.19).

THE LIMIT OF A FUNCTION

The concept of the derivative of a function is what most people think of as the beginning of calculus. However, before we can even define the derivative we must introduce a kind of generalization of the notion of continuity. That is, we must begin with the definition of the limit of a function.

DEFINITION. Let $f : S \to \mathbb{C}$ be a function, where $S \subseteq \mathbb{C}$, and let c be a limit point of S that is not necessarily an element of S. We say that f *has a limit* L *as* z *approaches* c, and we write

$$\lim_{z\to c} f(z) = L,$$

if for every $\epsilon > 0$ there exists a $\delta > 0$ such that if $z \in S$ and $0 < |z - c| < \delta$, then $|f(z) - L| < \epsilon$.

If the domain S is unbounded, we say that f *has a limit* L *as* z *approaches* ∞, and we write

$$L = \lim_{z\to\infty} f(z),$$

if for every $\epsilon > 0$ there exists a positive number B such that if $z \in S$ and $|z| \geq B$, then $|f(z) - L| < \epsilon$.

Analogously, if $S \subseteq \mathbb{R}$, we say $\lim_{x\to\infty} f(x) = L$ if for every $\epsilon > 0$ there exists a real number B such that if $x \in S$ and $x \geq B$, then $|f(x) - L| < \epsilon$. And we say that $\lim_{x\to-\infty} f(x) = L$ if for every $\epsilon > 0$ there exists a real number B such that if $x \in S$ and $x \leq B$, then $|f(x) - L| < \epsilon$.

Finally, for $f : (a, b) \to \mathbb{C}$ a function of a real variable, and for $c \in [a, b]$, we define the *one-sided* (left and right) limits of f at c. We say that f has a *left hand limit* of L at c, and we write $L = \lim_{x\to c-0} f(x)$, if for every $\epsilon > 0$ there exists a $\delta > 0$ such that if $x \in (a, b)$ and $0 < c - x < \delta$ then $|f(x) - L| < \epsilon$. We say that f has a *right hand limit* of L at c, and write $L = \lim_{x\to c+0} f(x)$, if for every $\epsilon > 0$ there exists a $\delta > 0$ such that if $x \in S$ and $0 < x - c < \delta$ then $|f(x) - L| < \epsilon$.

The first few results about limits of functions are not surprising. The analogy between functions having limits and functions being continuous is very close, so that for every elementary result about continuous functions there will be a companion result about limits of functions.

THEOREM 4.1. *Let c be a complex number. Let $f : S \to \mathbb{C}$ and $g : S \to \mathbb{C}$ be functions. Assume that both f and g have limits as x approaches c. Then:*

(1) *There exists a $\delta > 0$ and a positive number M such that if $z \in S$ and $0 < |z-c| < \delta$ then $|f(z)| < M$. That is, if f has a limit as z approaches c, then f is bounded near c.*

(2)
$$\lim_{z\to c}(f(z)+g(z)) = \lim_{z\to c} f(z) + \lim_{z\to c} g(z).$$

(3)
$$\lim_{z\to c}(f(z)g(z)) = \lim_{z\to c} f(z) \lim_{z\to c} g(z).$$

(4) *If $\lim_{z\to c} g(z) \neq 0$, then*
$$\lim_{z\to c} \frac{f(z)}{g(z)} = \frac{\lim_{z\to c} f(z)}{\lim_{z\to c} g(z)},$$

(5) *If u and v are the real and imaginary parts of a complex-valued function f, then u and v have limits as z approaches c if and only if f has a limit as z approaches c. And,*
$$\lim_{z\to c} f(z) = \lim_{z\to c} u(z) + i \lim_{z\to c} v(z).$$

Exercise 4.1. (a) Prove Theorem 4.1.
HINT: Compare with Theorem 3.2.

(b) Prove that $\lim_{x\to c} f(x) = L$ if and only if, for every sequence $\{x_n\}$ of elements of S that converges to c, we have $\lim f(x_n) = L$.
HINT: Compare with Theorem 3.4.

(c) Prove the analog of Theorem 4.1 replacing the limit as z approaches c by the limit as z approaches ∞.

Exercise 4.2. (a) Prove that a function $f : S \to \mathbb{C}$ is continuous at a point c of S if and only if $\lim_{x\to c} f(x) = f(c)$.
HINT: Carefully write down both definitions, and observe that they are verbetim the same.

(b) Let f be a function with domain S, and let c be a limit point of S that is not in S. Suppose g is a function with domain $S \cup \{c\}$, that $f(x) = g(x)$ for all $x \in S$, and that g is continuous at c. Prove that $\lim_{x\to c} f(x) = g(c)$.

Exercise 4.3. Prove that the following functions f have the specified limits L at the given points c.

(a) $f(x) = (x^3 - 8)/(x^2 - 4)$, $c = 2$, and $L = 3$.
(b) $f(x) = (x^2 + 1)/(x^3 + 1)$, $c = 1$, and $L = 1$.
(c) $f(x) = (x^8 - 1)/(x^6 + 1)$, $c = i$, and $L = -4/3$.
(d) $f(x) = (\sin(x) + \cos(x) - \exp(x))/(x^2)$, $c = 0$, and $L = -1$.

Exercise 4.4. Define f on the set S of all nonzero real numbers by $f(x) = c$ if $x < 0$ and $f(x) = d$ if $x > 0$. Show that $\lim_{x\to 0} f(x)$ exists if and only if $c = d$.

(b) Let $f : (a, b) \to \mathbb{C}$ be a complex-valued function on the open interval (a, b). Suppose c is a point of (a, b). Prove that $\lim_{x\to c} f(x)$ exists if and only if the two one-sided limits $\lim_{x\to c-0} f(x)$ and $\lim_{x\to c+0} f(x)$ exist and are equal.

Exercise 4.5. (Change of variable in a limit) Suppose $f : S \to \mathbb{C}$ is a function, and that $\lim_{x\to c} f(x) = L$. Define a function g by $g(y) = f(y + c)$.

(a) What is the domain of g?

(b) Show that 0 is a limit point of the domain of g and that $\lim_{y\to 0} g(y) = \lim_{x\to c} f(x)$.

(c) Suppose $T \subseteq \mathbb{C}$, that $h : T \to S$, and that $\lim_{y\to d} h(y) = c$. Prove that

$$\lim_{y\to d} f(h(y)) = \lim_{x\to c} f(x) = L.$$

REMARK. When we use the word " interior" in connection with a set S, it is obviously important to understand the context; i.e., is S being thought of as a set of real numbers or as a set of complex numbers. A point c is in the interior of a set S of complex numbers if the entire disk $B_\epsilon(c)$ of radius ϵ around c is contained in S. While, a point c belongs to the interior of a set S of real numbers if the entire interval $(c - \epsilon, c + \epsilon)$ is contained in S. Hence, in the following definition, we will be careful to distinguish between the cases that f is a function of a real variable or is a function of a complex variable.

THE DERIVATIVE OF A FUNCTION

Now begins what is ordinarily thought of as the first main subject of calculus, the derivative.

DEFINITION. Let S be a subset of $\mathbb{R}$, let $f : S \to \mathbb{C}$ be a complex-valued function (of a real variable), and let c be an element of the interior of S. We say that f is *differentiable at c* if

$$\lim_{h\to 0} \frac{f(c+h) - f(c)}{h}$$

exists. (Here, the number h is a real number.)

Analogously, let S be a subset of $\mathbb{C}$, let $f : S \to \mathbb{C}$ be a complex-valued function (of a complex variable), and let c be an element of the interior of S. We say that f is *differentiable at c* if

$$\lim_{h\to 0} \frac{f(c+h) - f(c)}{h}$$

exists. (Here, the number h is a complex number.)

If $f : S \to \mathbb{C}$ is a function either of a real variable or a complex variable, and if S' denotes the subset of S consisting of the points c where f is differentiable, we define a function $f' : S' \to \mathbb{C}$ by

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

The function f' is called the *derivative* of f.

A continuous function $f : [a, b] \to \mathbb{C}$ that is differentiable at each point $x \in (a, b)$, and whose the derivative f' is continuous on (a, b), is called a *smooth* function on

$[a, b]$. If there exists a partition $\{a = x_0 < x_1 < \ldots < x_n = b\}$ of $[a, b]$ such that f is smooth on each subinterval $[x_{i-1}, x_i]$, then f is called *piecewise smooth* on $[a, b]$.

Higher order derivatives are defined inductively. That is, f'' is the derivative of f', and so on. We use the symbol $f^{(n)}$ for the nth derivative of f.

REMARK. In the definition of the derivative of a function f, we are interested in the limit, as h approaches 0, not of f but of the quotient $q(h) = \frac{f(c+h)-f(c)}{h}$. Notice that 0 is not in the domain of the function q, but 0 is a limit point of that domain. This is the reason why we had to make such a big deal above out of the limit of a function. The function q is often called the *differential quotient.*

REMARK. As mentioned in Chapter III, we are often interested in solving for unknowns that are functions. The most common such problem is to solve a differential equation. In such a problem, there is an unknown function for which there is some kind of relationship between it and its derivatives. Differential equations can be extremely complicated, and many are unsolvable. However, we will have to consider certain relatively simple ones in this chapter, e.g., $f' = f$, $f' = -f$, and $f'' = \pm f$.

There are various equivalent ways to formulate the definition of differentiable, and each of these ways has its advantages. The next theorem presents one of those alternative ways.

THEOREM 4.2. *Let c belong to the interior of a set S (either in $\mathbb{R}$ or in $\mathbb{C}$), and let $f : S \to \mathbb{C}$ be a function. Then the following are equivalent.*

(1) *f is differentiable at c. That is,*

$$\lim_{h\to 0} \frac{f(c+h) - f(c)}{h} \text{ exists.}$$

(2)

$$\lim_{x\to c} \frac{f(x) - f(c)}{x - c} \text{ exists.}$$

(3) *There exists a number L and a function θ such that the following two conditions hold:*

$$f(c + h) - f(c) = Lh + \theta(h) \tag{4.1}$$

and

$$\lim_{h\to 0} \frac{\theta(h)}{h} = 0. \tag{4.2}$$

In this case, L is unique and equals $f'(c)$, and the function θ is unique and equals $f(c + h) - f(c) - f'(c)h$.

PROOF. That (1) and (2) are equivalent follows from Exercise 4.5 by writing x as $c + h$.

Suppose next that f is differentiable at c, and define

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

Set

$$\theta(h) = f(c+h) - f(c) - f'(c)h.$$

Then clearly

$$f(c+h) - f(c) = Lh + \theta(h),$$

which is Equation (4.1). Also

$$\begin{aligned} |\frac{\theta(h)}{h}| &= |\frac{f(c+h) - f(c) - f'(c)h}{h}| \\ &= |\frac{f(c+h) - f(c)}{h} - f'(c)|, \end{aligned}$$

which tends to 0 as h approaches 0 because f is differentiable at c. Hence, we have established equations (4.1) and (4.2), showing that (1) implies (3).

Finally, suppose there is a number L and a function θ satisfying Equations (4.1) and (4.2). Then

$$\frac{f(c+h) - f(c)}{h} = L + \frac{\theta(h)}{h},$$

which converges to L as h approaches 0 by Equation (4.2) and part (2) of Theorem 4.1. Hence, $L = f'(c)$, and so $\theta(h) = f(c+h) - f(c) - f'(c)h$. Therefore, (3) implies (1), and the theorem is proved.

REMARK. Though it seems artificial and awkward, Condition (3) of this theorem is very convenient for many proofs. One should remember it.

Exercise 4.6. (a) What is the domain of the function θ of condition (3) in the preceding theorem? Is 0 in this domain? Are there any points in the interior of this domain?

(b) Let L and θ be as in part (3) of the preceding theorem. Prove that, given an $\epsilon > 0$ there exists a $\delta > 0$ such that if $|h| < \delta$ then $|\theta(h)| < \epsilon|h|$.

THEOREM 4.3. *If $f : S \to \mathbb{C}$ is a function, either of a real variable or a complex variable, and if f is differentiable at a point c of S, then f is continuous at c. That is, differentiability implies continuity.*

PROOF. We are assuming that $\lim_{h\to 0}(f(c+h) - f(c))/h = L$. Hence, there exists a positive number δ_0 such that $|\frac{f(c+h)-f(c)}{h} - L| < 1$ if $|h| < \delta_0$, implying that $|f(c+h) - f(c)| < |h|(|L| + 1)$whenever $|h| < \delta_0$. So, if $\epsilon > 0$ is given, let δ be the minimum of δ_0 and $\epsilon/(|L|+1)$. If $y \in S$ and $|y - c| < \delta$, then, thinking of y as being $c + h$,

$$|f(y) - f(c)| = |f(c+h) - f(c)| < |h|(|L| + 1) = |y - c|(|L| + 1) < \epsilon.$$

(Every y can be written as $c + h$ for some h, and $|y - c| = |h|$.)

Exercise 4.7. Define $f(z) = |z|$ for $z \in \mathbb{C}$.

(a) Prove that f is continuous at every point of $\mathbb{C}$.

(b) Show that, if f is differentiable at a point c, then $f'(c) = 0$.

HINT: Using part (b) of Exercise 4.1, evaluate $f'(c)$ in the following two ways.

$$f'(c) = \lim_{n\to\infty} \frac{|c + \frac{1}{n}| - |c|}{\frac{1}{n}}$$

and

$$f'(c) = \lim_{n\to\infty} \frac{|c+\frac{i}{n}| - |c|}{\frac{i}{n}}.$$

Show that the only way these two limits can be equal is for them to be 0.

(c) Conclude that f is not differentiable anywhere. Indeed, if it were, what would the function θ have to be, and why wouldn't it satisfy Equation 4.2?

(d) Suppose $f : \mathbb{R} \to \mathbb{R}$ is the function of a real variable that is defined by $f(x) = |x|$. Show that f is differentiable at every point $x \neq 0$. How does this result not contradict part (c)?

The following theorem generalizes the preceding exercise.

THEOREM 4.4. *Suppose $f : S \to \mathbb{R}$ is a real-valued function of a complex variable, and assume that f is differentiable at a point $c \in S$. Then $f'(c) = 0$. That is, every real-valued, differentiable function f of a complex variable satisfies $f'(c) = 0$ for all c in the domain of f'.*

PROOF. We compute $f'(c)$ in two ways.

$$f'(c) = \lim_n \frac{f(c+\frac{1}{n}) - f(c)}{\frac{1}{n}} \text{ is a real number..}$$

$$f'(c) = \lim_n \frac{f(c+\frac{i}{n}) - f(c)}{\frac{i}{n}} \text{ is a purely imaginary number.}$$

Hence, $f'(c)$ must be 0, as claimed.

REMARK. This theorem may come as a surprise, for it shows that there are very few real-valued differentiable functions of a complex variable. For this reason, whenever $f : S \to \mathbb{R}$ is a real-valued, differentiable function, we will presume that f is a function of a real variable; i.e., that the domain $S \subseteq \mathbb{R}$.

Evaluating $\lim_{h\to 0} q(h)$ in the two different ways, h real, and h pure imaginary, led to the proof of the last theorem. It also leads us to make definitions of what are called "partial derivatives" of real-valued functions whose domains are subsets of $\mathbb{C} \equiv \mathbb{R}^2$. As the next exercise will show, the theory of partial derivatives of real-valued functions is a much richer theory than that of standard derivatives of real-valued functions of a single complex variable.

DEFINITION. Let $f : S \to \mathbb{R}$ be defined on a set $S \subseteq \mathbb{C} \equiv \mathbb{R}^2$, and let $c = (a, b) = + + bi$ be a point in the interior of S. We define the *partial derivative of f with respect to x* at the point $c = (a, b)$ by the formula

$$\frac{\partial f}{\partial x}(a, b) = \lim_{h\to 0} \frac{f(a+h, b) - f(a, b)}{h},$$

and the *partial derivative of f with respect to y* at $c = (a, b)$ by the formula

$$\frac{\partial f}{\partial y}(a, b) = \lim_{h\to 0} \frac{f(a, b+h) - f(a, b)}{h},$$

whenever these limits exist. (In both these limits, the variable h is a real variable.)(

It is clear that the partial derivatives of a function arise when we fix either the real part of the variable or the imaginary part of the variable to be a constant, and then consider the resulting function of the other (real) variable. We will see in Exercise 4.8 that there is a definite difference between a function's being differentiable at a point $c = (a + bi)$ in the complex plane $\mathbb{C}$ versus its having partial derivatives at the point (a, b) in $\mathbb{R}^2$.

Exercise 4.8. (a) Suppose f is a complex-valued function of a complex variable, and assume that both the real and imaginary parts of f are differentiable at a point c. Show that f is differentiable at c and that $f'(c) = 0$.

(b) Let $f = u + iv$ be a complex-valued function of a complex variable that is differentiable at a point c. Prove that both partial derivatives of u and v exist at $c = (a, b)$, and in fact that

$$\frac{\partial u}{\partial x}(c) + i\frac{\partial v}{\partial x}(c) = f'(c)$$

and

$$\frac{\partial u}{\partial y}(c) + i\frac{\partial v}{\partial y}(c) = if'(c).$$

(c) Define a complex-valued function f on $\mathbb{C} \equiv \mathbb{R}^2$ by $f(z) = f(x+iy) = x - iy$. Write $f = u + iv$, and show that both partial derivatives of u and v exist at every point, but that f is not a differentiable function of the complex variable z.

The next theorem is, in part, what we call in calculus the "differentiation formulas."

THEOREM 4.5. *Let f and g be functions (either of a real variable or a complex variable), which are both differentiable at a point c. Let a and b be complex numbers. Then:*

(1) *$af + bg$ is differentiable at c, and $(af + bg)'(c) = af'(c) + bg'(c)$.*
(2) (Product Formula) fg is differentiable at c, and $(fg)'(c) = f'(c)g(c) + f(c)g'(c)$.
(3) (Quotient Formula) f/g is differentiable at c (providing that $g(c) \neq 0$), and

$$(\frac{f}{g})'(c) = \frac{g(c)f'(c) - f(c)g'(c)}{(g(c))^2}.$$

(4) If $f = u+iv$ is a complex-valued function, then f is differentiable at a point c if and only if u and v are differentiable at c, and $f'(c) = u'(c) + iv'(c)$.

PROOF. We prove part (2) and leave parts (1), (3), and (4) for the exercises. We have

$$\begin{aligned}\lim_{h\to 0}\frac{(fg)(c+h)-(fg)(c)}{h} &= \lim_{h\to 0}\frac{f(c+h)g(c+h)-f(c)g(c)}{h}\\ &= \lim_{h\to 0}\frac{f(c+h)g(c+h)-f(c)g(c+h)}{h}\\ &\quad + \lim_{h\to 0}\frac{f(c)g(c+h)-f(c)g(c)}{h}\\ &= \lim_{h\to 0}\frac{f(c+h)-f(c)}{h}\lim_{h\to 0}g(c+h)\\ &\quad + \lim_{h\to 0}f(c)\lim_{h\to 0}\frac{g(c+h)-g(c)}{h}\\ &= f'(c)g(c)+f(c)g'(c),\end{aligned}$$

where we have used Theorems 4.1, 4.2, and 4.3.

Exercise 4.9. (a) Prove parts (1), (3), and (4) of Theorem 4.5.

(b) If f and g are real-valued functions that are differentiable at a point c, what can be said about the differentiability of $\max(f, g)$?

(c) Let f be a constant function $f(z) \equiv k$. Prove that f is differentiable everywhere and that $f'(z) = 0$ for all z.

(d) Define a function f by $f(z) = z$. Prove that f is differentiable everywhere and that $f'(z) = 1$ for all z.

(e) Verify the usual derivative formulas for polynomial functions: If $p(z) = \sum_{k=0}^{n} a_k z^k$, then $p'(z) = \sum_{k=1}^{n} k a_k z^{k-1}$.

What about power series functions? Are they differentiable functions? If so, are their derivatives again power series functions? In fact, everything works as expected.

THEOREM 4.6. *Let f be a power series function $f(z) = \sum_{n=0}^{\infty} a_n z^n$ having radius of convergence $r > 0$. Then f is differentiable at each point z in its open disk $B_r(0)$ of convergence, and*

$$f'(z) = \sum_{n=0}^{\infty} n a_n z^{n-1} = \sum_{n=1}^{\infty} n a_n z^{n-1}.$$

PROOF. The proof will use part (3) of Theorem 4.2. Fix an z with $|z| < r$. Choose r' so that $|z| < r' < r$, and write α for $r' - |z|$, i.e., $|z| + \alpha = r'$. Note first that the infinite series $\sum_{n=0}^{\infty} |a_n| r'^n$ converges to a positive number we will call M. Also, from the Cauchy-Hadamard Formula, we know that the power series function $\sum n a_n w^n$ has the same radius of convergence as does f, and hence the infinite series $\sum n a_n z^{n-1}$ converges to a number we will denote by L. We define a function θ by $\theta(h) = f(z + h) - f(z) - Lh$ from which it follows immediately that

$$f(z + h) - f(z) = Lh + \theta(h),$$

which establishes Equation (4.1). To complete the proof that f is differentiable at z, it will suffice to establish Equation (4.2), i.e., to show that

$$\lim_{h\to 0} \frac{\theta(h)}{h} = 0.$$

That is, given $\epsilon > 0$ we must show that there exists a $\delta > 0$ such that if $0 < |h| < \delta$ then

$$|\theta(h)/h| = |\frac{f(z+h) - f(z)}{h} - L| < \epsilon.$$

Assuming, without loss of generality, that $|h| < \alpha$, we have that

$$\begin{aligned}
|\frac{f(z+h)-f(z)}{h} - L| &= |\frac{\sum_{n=0}^\infty a_n (z+h)^n - \sum_{n=0}^\infty a_n z^n}{h} - L| \\
&= |\frac{\sum_{n=0}^\infty a_n (\sum_{k=0}^n \binom{n}{k} z^{n-k} h^k) - \sum_{n=0}^\infty a_n z^n}{h} - L| \\
&= |\frac{\sum_{n=0}^\infty a_n ((\sum_{k=0}^n \binom{n}{k} z^{n-k} h^k) - z^n)}{h} - L| \\
&= |\frac{\sum_{n=1}^\infty a_n (\sum_{k=1}^n \binom{n}{k} z^{n-k} h^k)}{h} - L| \\
&= |\sum_{n=1}^\infty a_n (\sum_{k=1}^n \binom{n}{k} z^{n-k} h^{k-1}) - \sum_{n=1}^\infty n a_n z^{n-1}| \\
&= |\sum_{n=1}^\infty a_n (\sum_{k=1}^n \binom{n}{k} z^{n-k} h^{k-1}) - \sum_{n=1}^\infty \binom{n}{1} a_n z^{n-1}| \\
&= |\sum_{n=2}^\infty a_n (\sum_{k=2}^n \binom{n}{k} z^{n-k} h^{k-1})| \\
&\leq \sum_{n=2}^\infty \sum_{k=2}^n |a_n| \binom{n}{k} |z|^{n-k} |h|^{k-1} \\
&\leq |h| \sum_{n=2}^\infty |a_n| \sum_{k=2}^n \binom{n}{k} |z|^{n-k} |h|^{k-2} \\
&\leq |h| \sum_{n=2}^\infty |a_n| \sum_{k=2}^n \binom{n}{k} |z|^{n-k} |\alpha|^{k-2} \\
&\leq |h| \frac{1}{\alpha^2} \sum_{n=0}^\infty |a_n| \sum_{k=0}^n \binom{n}{k} |z|^{n-k} \alpha^k \\
&= |h| \frac{1}{\alpha^2} \sum_{n=0}^\infty |a_n| (|z| + \alpha)^n \\
&= |h| \frac{1}{\alpha^2} \sum_{n=0}^\infty |a_n| r'^n \\
&= |h| \frac{M}{\alpha^2},
\end{aligned}$$

so that if $\delta = \epsilon / \frac{M}{\alpha^2}$, then $|\theta(h)/h| < \epsilon$, whenever $|h| < \delta$, as desired.

REMARK. Theorem 4.6 shows that indeed power series functions are differentiable, and in fact their derivatives can be computed, just like polynomials, by differentiating term by term. This is certainly a result we would have hoped was true, but the proof is **not** trivial.

The next theorem, the Chain Rule, is another nontrivial one. It deals with the differentiability of the composition of two differentiable functions. Again, the result is what we would have wanted, the composition of two differentiable functions is itself differentiable, but the argument required to prove it is tricky.

THEOREM 4.7. (Chain Rule) Let $f : S \to \mathbb{C}$ be a function, and assume that f is differentiable at a point c. Suppose $g : T \to \mathbb{C}$ is a function, that $T \subseteq \mathbb{C}$, that the number $f(c) \in T$, and that g is differentiable at $f(c)$. Then the composition $g \circ f$ is differentiable at c and

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

PROOF. Using part (3) of Theorem 4.2, write

$$g(f(c) + k) - g(f(c)) = L_g k + \theta_g(k)$$

and

$$f(c + h) - f(c) = L_f h + \theta_f(h).$$

We know from that theorem that $L_g = g'(f(c))$ and $L_f = f'(c)$. And, we also know that

$$\lim_{k\to 0} \frac{\theta_g(k)}{k} = 0 \text{ and } \lim_{h\to 0} \frac{\theta_f(h)}{h} = 0.$$

Define a function $k(h) = f(c + h) - f(c)$. Then, by Theorem 4.3, we have that $\lim_{h\to 0} k(h) = 0$. We will show that $g \circ f$ is differentiable at c by showing that there exists a number L and a function θ satisfying the two conditions of part (3) of Theorem 4.2. Thus, we have that

$$\begin{aligned} g \circ f(c + h) - g \circ f(c) &= g(f(c + h)) - g(f(c)) \\ &= g(f(c) + k(h)) - g(f(c)) \\ &= L_g k(h) + \theta_g(k(h)) \\ &= L_g(f(c + h) - f(c)) + \theta_g(k(h)) \\ &= L_g(L_f h + \theta_f(h)) + \theta_g(k(h)) \\ &= L_g L_f h + L_g\theta_f(h) + \theta_g(k(h)). \end{aligned}$$

We define $L = L_g l_f = g'(f(c))f'(c)$, and we define the function θ by

$$\theta(h) = L_g\theta_f(h) + \theta_g(k(h)).$$

By our definitions, we have established Equation (4.1)

$$g \circ f(c + h) - g \circ f(c) = Lh + \theta(h),$$

so that it remains to verify Equation (4.2).

We must show that, given $\epsilon > 0$, there exists a $\delta > 0$ such that if $0 < |h| < \delta$ then $|\theta(h)/h| < \epsilon$. First, choose an $\epsilon' > 0$ so that

(4.3). $$|L_g|\epsilon' + |L_f|\epsilon' + {\epsilon'}^2 < \epsilon$$

Next, using part (b) of Exercise 4.6, choose a $\delta' > 0$ such that if $|k| < \delta'$ then $|\theta_g(k)| < \epsilon'|k|$. Finally, choose $\delta > 0$ so that if $0 < |h| < \delta$, then the following two inequalities hold. $|k(h)| < \delta'$ and $|\theta_f(h)| < \epsilon'|h|$. The first can be satisfied because f is continuous at c, and the second is a consequence of part (b) of Exercise 4.6. Then: if $0 < |h| < \delta$,

$$\begin{aligned}
|\theta(h)| &= |L_g\theta_f(h) + \theta_g(k(h))| \\
&\leq |L_g||\theta_f(h)| + |\theta_g(k(h))| \\
&< |L_g|\epsilon'|h| + \epsilon'|k(h)| \\
&= |L_g|\epsilon'|h| + \epsilon'|f(c+h) - f(c)| \\
&= |L_g|\epsilon'|h| + \epsilon'|L_f h + \theta_f(h)| \\
&\leq |L_g|\epsilon'|h| + \epsilon'|L_f||h| + \epsilon'|\theta_f(h)| \\
&< |L_g|\epsilon'|h| + \epsilon'|L_f||h| + \epsilon'\epsilon'|h| \\
&= (|L_g|\epsilon' + |L_f|\epsilon' + {\epsilon'}^2)|h|,
\end{aligned}$$

whence

$$|\theta(h)/h| < (|L_g|\epsilon' + |L_f|\epsilon' + {\epsilon'}^2) < \epsilon,$$

as desired.

Exercise 4.10. (a) Derive the familiar formulas for the derivatives of the elementary transcendental functions:

$$\exp' = \exp,\ \sin' = \cos,\ , sinh' = cosh,\ \cosh' = \sinh \text{ and } \cos' = -\sin.$$

(b) Define a function f as follows. $f(z) = \cos^2(z) + \sin^2(z)$. Use part (a) and the Chain Rule to show that $f'(z) = 0$ for all $z \in \mathbb{C}$. Does this imply that $\cos^2(z) + \sin^2(z) = 1$ for all complex numbers z?

(c) Suppose f is expandable in a Taylor series around the point c : $f(z) = \sum_{n=0}^\infty a_n(z-c)^n$ for all $z \in B_r(c)$. Prove that f is differentiable at each point of the open disk $B_r(c)$, and show that

$$f'(z) = \sum_{n=1}^\infty n a_n (z-c)^{n-1}.$$

HINT: Use Theorem 4.6 and the chain rule.

CONSEQUENCES OF DIFFERENTIABILITY, THE MEAN VALUE THEOREM

DEFINITION. Let $f : S \to \mathbb{R}$ be a real-valued function of a real variable, and let c be an element of the interior of S. Then f is said to *attain a local maximum* at c if there exists a $\delta > 0$ such that $(c-\delta, c+\delta) \subseteq S$ and $f(c) \geq f(x)$ for all $x \in (c-\delta, c+\delta)$.

The function f is said to *attain a local minimum* at c if there exists an interval $(c-\delta, c+\delta) \subseteq S$ such that $f(c) \leq f(x)$ for all $x \in (c-\delta, c+\delta)$.

The next theorem should be a familiar result from calculus.

THEOREM 4.8. (First Derivative Test for Extreme Values) Let $f : S \to \mathbb{R}$ be a real-valued function of a real variable, and let $c \in S$ be a point at which f attains a local maximum or a local minimum. If f is differentiable at c, then $f'(c)$ must be 0.

PROOF. We prove the theorem when f attains a local maximum at c. The proof for the case when f attains a local minimum is completely analogous.

Thus, let $\delta > 0$ be such that $f(c) \geq f(x)$ for all x such that $|x - c| < \delta$. Note that, if n is sufficiently large, then both $c + \frac{1}{n}$ and $c - \frac{1}{n}$ belong to the interval $(c - \delta, c + \delta)$. We evaluate $f'(c)$ in two ways. First,

$$f'(c) = \lim_n \frac{f(c + \frac{1}{n}) - f(c)}{\frac{1}{n}} \leq 0$$

because the numerator is always nonpositive and the denominator is always positive. On the other hand,

$$f'(c) = \lim_n \frac{f(c - \frac{1}{n}) - f(c)}{\frac{-1}{n}} \geq 0$$

since both numerator and denominator are nonpositive. Therefore, $f'(c)$ must be 0, as desired.

Of course we do not need a result like Theorem 4.8 for functions of a complex variable, since the derivative of every real-valued function of a complex variable necessarily is 0, independent of whether or not the function attains an extreme value.

REMARK. As mentioned earlier, the zeroes of a function are often important numbers. The preceding theorem shows that the zeroes of the derivative f' of a function f are intimately related to finding the extreme values of the function f. The zeroes of f' are often called the *critical points* for f. Part (a) of the next exercise establishes the familiar procedure from calculus for determining the maximum and minimum of a continuous real-valued function on a closed interval.

Exercise 4.11. (a) Let f be a continuous real-valued function on a closed interval $[a, b]$, and assume that f is differentiable at each point x in the open interval (a, b). Let M be the maximum value of f on this interval, and m be its minimum value on this interval. Write S for the set of all $x \in (a, b)$ for which $f'(x) = 0$. Suppose x is a point of $[a, b]$ for which $f(x)$ is either M or m. Prove that x either is an element of the set S, or x is one of the endpoints a or b.

(b) Let f be the function defined on $[0, 1/2)$ by $f(t) = t/(1 - t)$. Show that $f(t) < 1$ for all $t \in [0, 1/2)$.

(c) Let $t \in (-1/2, 1/2)$ be given. Prove that there exists an $r < 1$, depending on t, such that $|t/(1 + y)| < r$ for all y between 0 and t.

(d) Let t be a fixed number for which $0 < t < 1$. Show that, for all $0 \leq s \leq t$, $(t - s)/(1 + s) \leq t$.

Probably the most powerful theorem about differentiation is the next one. It is stated as an equation, but its power is usually as an inequality; i.e., the absolute value of the left hand side is less than or equal to the absolute value of the right hand side.

THEOREM 4.9. (Mean Value Theorem) Let f be a real-valued continuous function on a closed bounded interval $[a,b]$, and assume that f is differentiable at each point x in the open interval (a,b). Then there exists a point $c \in (a,b)$ such that

$$f(b) - f(a) = f'(c)(b-a).$$

PROOF. This proof is tricky. Define a function h on $[a,b]$ by

$$h(x) = x(f(b) - f(a)) - f(x)(b-a).$$

Clearly, h is continuous on $[a,b]$ and is differentiable at each point $x \in (a,b)$. Indeed,

$$h'(x) = f(b) - f(a) - f'(x)(b-a).$$

It follows from this equation that the theorem will be proved if we can show that there exists a point $c \in (a,b)$ for which $h'(c) = 0$. Note also that

$$h(a) = a(f(b) - f(a)) - f(a)(b-a) = af(b) - bf(a)$$

and

$$h(b) = b(f(b) - f(a)) - f(b)(b-a) = af(b) - bf(a),$$

showing that $h(a) = h(b)$.

Let m be the minimum value attained by the continuous function h on the compact interval $[a,b]$ and let M be the maximum value attained by h on $[a,b]$. If $m = M$, then h is a constant on $[a,b]$ and $h'(c) = 0$ for all $c \in (a,b)$. Hence, the theorem is true if $M = m$, and we could use any $c \in (a,b)$. If $m \neq M$, then at least one of these two extreme values is not equal to $h(a)$. Suppose $m \neq h(a)$. Of course, m is also not equal to $h(b)$. Let $c \in [a,b]$ be such that $h(c) = m$. Then, in fact, $c \in (a,b)$. By Theorem 4.8, $h'(c) = 0$.

We have then that in every case there exists a point $c \in (a,b)$ for which $h'(c) = 0$. This completes the proof.

REMARK. The Mean Value Theorem is a theorem about real-valued functions of a real variable, and we will see later that it fails for complex-valued functions of a complex variable. (See part (f) of Exercise 4.16.) In fact, it can fail for a complex-valued function of a real variable. Indeed, if $f(x) = u(x) + iv(x)$ is a continuous complex-valued function on the interval $[a,b]$, and differentiable on the open interval (a,b), then the Mean Value Theorem certainly holds for the two real-valued functions u and v, so that we would have

$$f(b) - f(a) = u(b) - u(a) + i(v(b) - v(a)) = u'(c_1)(b-a) + iv'(c_2)(b-a),$$

which is not $f'(c)(b-a)$ unless we can be sure that the two points c_1 and c_2 can be chosen to be equal. This simply is not always possible. Look at the function $f(x) = x^2 + ix^3$ on the interval $[0,1]$.

On the other hand, if f is a real-valued function of a complex variable (two real variables), then a generalized version of the Mean Value Theorem does hold. See part (c) of Exercise 4.35.

One of the first applications of the Mean Value Theorem is to show that a function whose derivative is identically 0 is necessarily a constant function. This seemingly obvious fact is just **not** obvious. The next exercise shows that this result holds for complex-valued functions of a complex variable, even though the Mean Value Theorem does not.

Exercise 4.12. (a) Suppose f is a continuous real-valued function on (a, b) and that $f'(x) = 0$ for all $x \in (a, b)$. Prove that f is a constant function on (a, b).
HINT: Show that $f(x) = f(a)$ for all $x \in [a, b]$ by using the Mean Value Theorem applied to the interval $[a, x]$.

(b) Let f be a complex-valued function of a real variable. Suppose f is differentiable at each point x in an open interval (a, b), and assume that $f'(x) = 0$ for all $x \in (a, b)$. Prove that f is a constant function.
HINT: Use the real and imaginary parts of f.

(c) Let f be a complex-valued function of a complex variable, and suppose that f is differentiable on a disk $B_r(c) \subseteq \mathbb{C}$, and that $f'(z) = 0$ for all $z \in B_r(c)$. Prove that $f(z)$ is constant on $B_r(c)$.
HINT: Let z be an arbitrary point in $B_r(c)$, and define a function $h : [0, 1] \to \mathbb{C}$ by $h(t) = f((1 - t)c + tz)$. Apply part (b) to h.

The next exercise establishes, at last, two important identities.

Exercise 4.13.) ($\cos^2 + \sin^2 = 1$ and $\exp(i\pi = -1.)$

(a) Prove that $\cos^2(z) + \sin^2(z) = 1$ for all complex numbers z.

(b) Prove that $\cos(\pi) = -1$.
HINT: We know from part (a) that $\cos(\pi) = \pm 1$. Using the Mean Value Theorem for the cosine function on the interval $[0, \pi]$, derive a contradiction from the assumption that $\cos(\pi) = 1$.

(c) Prove that $\exp(i\pi) = -1$.
HINT: Recall that $\exp(iz) = \cos(z) + i \sin(z)$ for all complex z. (Note that this does not yet tell us that $e^{i\pi} = -1$. We do not yet know that $\exp(z) = e^z$.)

(d) Prove that $\cosh^2 z - \sinh^2 z = 1$ for all complex numbers z.

(e) Compute the derivatives of the tangent and hyperbolic tangent functions $\tan = \sin / \cos$ and $\tanh = \sinh / \cosh$. Show in fact that

$$\tan' = \frac{1}{\cos^2} \text{ and } \tanh' = \frac{1}{\cosh^2}.$$

Here are two more elementary consequences of the Mean Value Theorem.

Exercise 4.14. (a) Suppose f and g are two complex-valued functions of a real (or complex) variable, and suppose that $f'(x) = g'(x)$ for all $x \in (a, b)$ (or $x \in B_r(c)$.) Prove that there exists a constant k such that $f(x) = g(x) + k$ for all $x \in (a, b)$ (or $x \in B_r(c)$.)

(b) Suppose $f'(z) = c \exp(az)$ for all z, where c and a are complex constants with $a \neq 0$. Prove that there exists a constant c' such that $f(z) = \frac{c}{a} \exp(az) + c'$. What if $a = 0$?

(c) (A generalization of part (a)) Suppose f and g are continuous real-valued functions on the closed interval $[a, b]$, and suppose there exists a partition $\{x_0 < x_1 < \ldots < x_n\}$ of $[a, b]$ such that both f and g are differentiable on each subinterval

(x_{i-1}, x_i). (That is, we do not assume that f and g are differentiable at the endpoints.) Suppose that $f'(x) = g'(x)$ for every x in each open subinterval (x_{i-1}, x_i). Prove that there exists a constant k such that $f(x) = g(x) + k$ for all $x \in [a, b]$.
HINT: Use part (a) to conclude that $f = g + h$ where h is a step function, and then observe that h must be continuous and hence a constant.

(d) Suppose f is a differentiable real-valued function on (a, b) and assume that $f'(x) \neq 0$ for all $x \in (a, b)$. Prove that f is 1-1 on (a, b).

Exercise 4.15. Let $f : [a, b] \to \mathbb{R}$ be a function that is continuous on its domain $[a, b]$ and differentiable on (a, b). (We do not suppose that f' is continuous on (a, b).)

(a) Prove that f is nondecreasing on $[a, b]$ if and only if $f'(x) \geq 0$ for all $x \in (a, b)$. Show also that f is nonincreasing on $[a, b]$ if and only if $f'(x) \leq 0$ for all $x \in (a, b)$.

(b) Conclude that, if f' takes on both positive and negative values on (a, b), then f is **not** 1-1. (See the proof of Theorem 3.11.)

(c) Show that, if f' takes on both positive and negative values on (a, b), then there must exist a point $c \in (a, b)$ for which $f'(c) = 0$. (If f' were continuous, this would follow from the Intermediate Value Theorem. But, we are not assuming here that f' is continuous.)

(d) Prove the *Intermediate Value Theorem for Derivatives:* Suppose f is continuous on the closed bounded interval $[a, b]$ and differentiable on the open interval (a, b). If f' attains two distinct values $v_1 = f'(x_1) < v_2 = f'(x_2)$, then f' attains every value v between v_1 and v_2.
HINT: Suppose v is a value between v_1 and v_2. Define a function g on $[a, b]$ by $g(x) = f(x) - vx$. Now apply part (c) to g.

Here is another perfectly reasonable and expected theorem, but one whose proof is tough.

THEOREM 4.10. (Inverse Function Theorem) Suppose $f : (a, b) \to \mathbb{R}$ is a function that is continuous and 1-1 from (a, b) onto the interval (a', b'). Assume that f is differentiable at a point $c \in (a, b)$ and that $f'(c) \neq 0$. Then f^{-1} is differentiable at the point $f(c)$, and

$$f^{-1\prime}(f(c)) = \frac{1}{f'(c)}.$$

PROOF. The formula $f^{-1\prime}(f(c)) = 1/f'(c)$ is no surprise. This follows directly from the Chain Rule. For, if $f^{-1}(f(x)) = x$, and f and f^{-1} are both differentiable, then $f^{-1\prime}(f(c))f'(c) = 1$, which gives the formula. The difficulty with this theorem is in proving that the inverse function f^{-1} of f is differentiable at $f(c)$. In fact, the first thing to check is that the point $f(c)$ belongs to the interior of the domain of f^{-1}, for that is essential if f^{-1} is to be differentiable there, and here is where the hypothesis that f is a real-valued function of a real variable is important. According to Exercise 3.12, the 1-1 continuous function f maps $[a, b]$ onto an interval $[a', b']$, and $f(c)$ is in the open interval (a', b'), i.e., is in the interior of the domain of f^{-1}.

According to part (2) of Theorem 4.2, we can prove that f^{-1} is differentiable at $f(c)$ by showing that

$$\lim_{x \to f(c)} \frac{f^{-1}(x) - f^{-1}(f(c))}{x - f(c)} = \frac{1}{f'(c)}.$$

That is, we need to show that, given an $\epsilon > 0$, there exists a $\delta > 0$ such that if $0 < |x - f(c)| < \delta$ then

$$|\frac{f^{-1}(x) - f^{-1}(f(c))}{x - f(c)} - \frac{1}{f'(c)}| < \epsilon.$$

First of all, because the function $1/q$ is continuous at the point $f'(c)$, there exists an $\epsilon' > 0$ such that if $|q - f'(c)| < \epsilon'$, then

$$(4.4). \qquad |\frac{1}{q} - \frac{1}{f'(c)}| < \epsilon$$

Next, because f is differentiable at c, there exists a $\delta' > 0$ such that if $0 < |y-c| < \delta'$ then

$$(4.5). \qquad |\frac{f(y) - f(c)}{y - c} - f'(c)| < \epsilon'$$

Now, by Theorem 3.10, f^{-1} is continuous at the point $f(c)$, and therefore there exists a $\delta > 0$ such that if $|x - f(c)| < \delta$ then

$$(4.6). \qquad |f^{-1}(x) - f^{-1}(f(c)| < \delta'$$

So, if $|x - f(c)| < \delta$, then

$$|f^{-1}(x) - c| = |f^{-1}(x) - f^{-1}(f(c))| < \delta'.$$

But then, by Inequality 4.5,

$$|\frac{f(f^{-1}(x)) - f(c)}{f^{-1}(x) - c} - f'(c)| < \epsilon',$$

from which it follows, using Inequality 4.4, that

$$|\frac{f^{-1}(x) - f^{-1}(f(c))}{x - f(c)} - \frac{1}{f'(c)}| < \epsilon,$$

as desired.

REMARK. A result very like Theorem 4.10 is actually true for complex-valued functions of a complex variable. We will have to show that if c is in the interior of the domain S of a one-to-one, continuously differentiable, complex-valued function f of a complex variable, then $f(c)$ is in the interior of the domain $f(S)$ of f^{-1}. But, in the complex variable case, this requires a somewhat more difficult argument. Once that fact is established, the proof that f^{-1} is differentiable at $f(c)$ will be the same for complex-valued functions of complex variables as it is here for real-valued functions of a real variable. Though the proof of Theorem 4.10 is reasonably complicated for real-valued functions of a real variable, the corresponding result for complex functions is much more deep, and that proof will have to be postponed to a later chapter. See Theorem 7.10.

THE EXPONENTIAL AND LOGARITHM FUNCTIONS

We derive next the elementary properties of the exponential and logarithmic functions. Of course, by "exponential function," we mean the power series function exp. And, as yet, we have not even defined a logarithm function.

Exercise 4.16. (a) Define a complex-valued function $f : \mathbb{C} \to \mathbb{C}$ by $f(z) = \exp(z)\exp(-z)$. Prove that $f(z) = 1$ for all $z \in \mathbb{C}$.

(b) Conclude from part (a) that the exponential function is never 0, and that $\exp(-z) = 1/\exp(z)$.

(c) Show that the exponential function is always positive on $\mathbb{R}$, and that $\lim_{x\to-\infty}\exp(x) = 0$.

(d) Prove that exp is continuous and 1-1 from $(-\infty,\infty)$ onto $(0,\infty)$.

(e) Show that the exponential function is **not** 1-1 on $\mathbb{C}$.

(f) Use parts b and e to show that the Mean Value Theorem is not in any way valid for complex-valued functions of a complex variable.

Using part (d) of the preceding exercise, we make the following important definition.

DEFINITION. We call the inverse $\exp^{-1}$ of the restriction of the exponential function to $\mathbb{R}$ the (natural) *logarithm function,* and we denote this function by $\ln$.

The properties of the exponential and logarithm functions are strongly tied to the simplest kinds of differential equations. The connection is suggested by the fact, we have already observed, that $\exp' = \exp$. The next theorem, corollary, and exercises make these remarks more precise.

THEOREM 4.11. *Suppose $f : \mathbb{C} \to \mathbb{C}$ is differentiable everywhere and satisfies the differential equation $f' = af$, where a is a complex number. Then $f(z) = c\exp(az)$, where $c = f(0)$.*

PROOF. Consider the function $h(z) = f(z)/\exp(az)$. Using the Quotient Formula, we have that

$$h'(z) = \frac{\exp(az)f'(z) - a\exp(az)f(z)}{[\exp(az)]^2} = \frac{\exp(az)(f'(z) - af(z))}{[\exp(z)]^2} = 0.$$

Hence, there exists a complex number c such that $h(z) = c$ for all z. Therefore, $f(z) = c\exp(az)$ for all z. Setting $z = 0$ gives $f(0) = c$, as desired.

COROLLARY. (Law of Exponents) For all complex numbers z and w, $\exp(z + w) = \exp(z)\exp(w)$.

PROOF OF THE COROLLARY. Fix w, define $f(z) = \exp(z + w)$, and apply the preceding theorem. We have $f'(z) = \exp(z + w) = f(z)$, so we get

$$\exp(z + w) = f(z) = f(0)\exp(z) = \exp(w)\exp(z).$$

Exercise 4.17. (a) If n is a positive integer and z is any complex number, show that $\exp(nz) = (\exp(z))^n$.

(b) If r is a rational number and x is any real number, show that $\exp(rx) = (\exp(x))^r$.

Exercise 4.18. (a) Show that ln is continuous and 1-1 from $(0,\infty)$ onto $\mathbb{R}$.

(b) Prove that the logarithm function ln is differentiable at each point $y \in (0,\infty)$ and that $\ln'(y) = 1/y$.

HINT: Write $y = \exp(c)$ and use Theorem 4.10.

(c) Derive the first law of logarithms: $\ln(xy) = \ln(x) + \ln(y)$.

(d) Derive the second law of logarithms: That is, if r is a rational number and x is a positive real number, show that $\ln(x^r) = r\ln(x)$.

We are about to make the connection between the number e and the exponential function. The next theorem is the first step.

THEOREM 4.12. $\ln(1) = 0$ *and* $\ln(e) = 1$.

PROOF. If we write $1 = \exp(t)$, then $t = \ln(1)$. But $\exp(0) = 1$, so that $\ln(1) = 0$, which establishes the first assertion.

Recall that

$$e = \lim_n (1 + \frac{1}{n})^n.$$

Therefore,

$$\begin{aligned}
\ln(e) &= \ln(\lim_n (1 + \frac{1}{n})^n) \\
&= \lim_n \ln((1 + \frac{1}{n})^n) \\
&= \lim_n n\ln(1 + \frac{1}{n}) \\
&= \lim_n \frac{\ln(1 + \frac{1}{n})}{\frac{1}{n}} \\
&= \lim_n \frac{\ln(1 + \frac{1}{n}) - \ln(1)}{\frac{1}{n}} \\
&= \ln'(1) \\
&= 1/1 \\
&= 1.
\end{aligned}$$

This establishes the second assertion of the theorem.

Exercise 4.19. (a) Prove that

$$e = \sum_{n=0}^{\infty} \frac{1}{n!}.$$

HINT: Use the fact that the logarithm function is 1-1.

(b) For r a rational number, show that $\exp(r) = e^r$.

(c) If a is a positive number and $r = p/q$ is a rational number, show that

$$a^r = \exp(r\ln(a)).$$

(d) Prove that e is irrational.

HINT: Let p_n/q_n be the nth partial sum of the series in part (a). Show that $q_n \leq n!$, and that $\lim q_n(e - p_n/q_n) = 0$. Then use Theorem 2.19.

We have finally reached a point in our development where we can make sense of raising any positive number to an arbitrary complex exponent. Of course this includes raising positive numbers to irrational powers. We make our general definition based on part (c) of the preceding exercise.

DEFINITION. For a a positive real number and z an arbitrary complex number, define a^z by

$$a^z = \exp(z\ln(a)).$$

REMARK. The point is that our old understanding of what a^r means, where $a > 0$ and r is a rational number, coincides with the function $\exp(r\ln(a))$. So, this new definition of a^z coincides and is consistent with our old definition. And, it now allows us to raies a positive number a to an arbitrary complex exponent.

REMARK. Let the bugles sound!! Now, having made all the appropriate definitions and derived all the relevant theorems, we can finally prove that $e^{i\pi} = -1$. From the definition above, we see that if $a = e$, then we have $e^z = \exp(z)$. Then, from part (c) of Exercise 4.13, we have what we want:

$$e^{i\pi} = -1.$$

Exercise 4.20. (a) Prove that, for all complex numbers z and w, $e^{z+w} = e^z e^w$.

(b) If x is a real number and z is any complex number, show that

$$(e^x)^z = e^{xz}.$$

(c) Let a be a fixed positive number, and define a function $f : \mathbb{C} \to \mathbb{C}$ by $f(z) = a^z$. Show that f is differentiable at every $z \in \mathbb{C}$ and that $f'(z) = \ln(a)a^z$.

(d) Prove the general laws of exponents: If a and b are positive real numbers and z and w are complex numbers,

$$a^{z+w} = a^z a^w,$$

$$a^z b^z = (ab)^z,$$

and, if x is real,

$$a^{xw} = (a^x)^w.$$

(e) If y is a real number, show that $|e^{iy}| = 1$. If $z = x + iy$ is a complex number, show that $|e^z| = e^x$.

(f) Let $\alpha = a + bi$ be a complex number, and define a function $f : (0, \infty) \to \mathbb{C}$ by $f(x) = x^\alpha = e^{\alpha\ln(x)}$. Prove that f is differentiable at each point x of $(0, \infty)$ and that $f'(x) = \alpha x^{\alpha-1}$.

(g) Let $\alpha = a + bi$ be as in part (f). For $x > 0$, show that $|x^\alpha| = x^a$.

THE TRIGONOMETRIC AND HYPERBOLIC FUNCTIONS

The laws of exponents and the algebraic connections between the exponential function and the trigonometric and hyperbolic functions, give the following "addition formulas:"

THEOREM 4.13. *The following identities hold for all complex numbers z and w.*

$$\sin(z + w) = \sin(z)\cos(w) + \cos(z)\sin(w).$$

$$\cos(z + w) = \cos(z)\cos(w) - \sin(z)\sin(w).$$

$$\sinh(z + w) = \sinh(z)\cosh(w) + \cosh(z)\sinh(w).$$

$$\cosh(z+w) = \cosh(z)\cosh(w) + \sinh(z)\sinh(w).$$

PROOF. We derive the first formula and leave the others to an exercise.

First, for any two real numbers x and y, we have

$$\begin{aligned}\cos(x+y) + i\sin(x+y) &= e^{i(x+y)} \\ &= e^{ix}e^{iy} \\ &= (\cos x + i\sin x) \times (\cos y + i\sin y) \\ &= \cos x\cos y - \sin x\sin y + i(\cos x\sin y + \sin x\cos y),\end{aligned}$$

which, equating real and imaginary parts, gives that

$$\cos(x+y) = \cos x\cos y - \sin x\sin y$$

and

$$\sin(x+y) = \sin x\cos y + \cos x\sin y.$$

The second of these equations is exactly what we want, but this calculation only shows that it holds for real numbers x and y. We can use the Identity Theorem to show that in fact this formula holds for all complex numbers z and w. Thus, fix a real number y. Let $f(z) = \sin z\cos y + \cos z\sin y$, and let

$$g(z) = \sin(z+y) = \frac{1}{2i}(e^{i(z+y)} - e^{-i(z+y)} = \frac{1}{2i}(e^{iz}e^{iy} - e^{-iz}e^{-iy}).$$

Then both f and g are power series functions of the variable z. Furthermore, by the previous calculation, $f(1/k) = g(1/k)$ for all positive integers k. Hence, by the Identity Theorem, $f(z) = g(z)$ for all complex z. Hence we have the formula we want for all complex numbers z and all real numbers y.

To finish the proof, we do the same trick one more time. Fix a complex number z. Let $f(w) = \sin z\cos w + \cos z\sin w$, and let

$$g(w) = \sin(z+w) = \frac{1}{2i}(e^{i(z+w)} - e^{-i(z+w)} = \frac{1}{2i}(e^{iz}e^{iw} - e^{-iz}e^{-iw}).$$

Again, both f and g are power series functions of the variable w, and they agree on the sequence $\{1/k\}$. Hence they agree everywhere, and this completes the proof of the first addition formula.

Exercise 4.21. (a) Derive the remaining three addition formulas of the preceding theorem.

(b) From the addition formulas, derive the two "half angle" formulas for the trigonometric functions:

$$\sin^2(z) = \frac{1-\cos(2z)}{2},$$

and

$$\cos^2(z) = \frac{1+\cos(2z)}{2}.$$

THEOREM 4.14. *The trigonometric functions* sin *and* cos *are periodic with period* 2π; *i.e.,* $\sin(z+2\pi) = \sin(z)$ *and* $\cos(z+2\pi) = \cos(z)$ *for all complex numbers* z.

PROOF. We have from the preceding exercise that $\sin(z+2\pi) = \sin(z)\cos(2\pi) + \cos(z)\sin(2\pi)$, so that the periodicity assertion for the sine function will follow if we show that $\cos(2\pi) = 1$ and $\sin(2\pi) = 0$. From part (b) of the preceding exercise, we have that

$$0 = \sin^2(\pi) = \frac{1-\cos(2\pi)}{2}$$

which shows that $\cos(2\pi) = 1$. Since $\cos^2 + \sin^2 = 1$, it then follows that $\sin(2\pi) = 0$.

The periodicity of the cosine function is proved similarly.

Exercise 4.22. (a) Prove that the hyperbolic functions sinh and cosh are periodic. What is the period?

(b) Prove that the hyperbolic cosine $\cosh(x)$ is never 0 for x a real number, that the hyperbolic tangent $\tanh(x) = \sinh(x)/\cosh(x)$ is bounded and increasing from $\mathbb{R}$ onto $(-1,1)$, and that the inverse hyperbolic tangent has derivative given by ${\tanh^{-1}}'(y) = 1/(1-y^2)$.

(c) Verify that for all $y \in (-1,1)$

$$\tanh^{-1}(y) = \ln\Big(\sqrt{\frac{1+y}{1-y}}\Big).$$

Exercise 4.23. (Polar coordinates) Let z be a nonzero complex number. Prove that there exists a unique real number $0 \le \theta < 2\pi$ such that $z = re^{i\theta}$, where $r = |z|$. HINT: If $z = a+bi$, then $z = r(\frac{a}{r} + \frac{b}{r}i$. Observe that $-1 \le \frac{a}{r} \le 1$, $-1 \le \frac{b}{r} \le 1$, and $(\frac{a}{r})^2 + (\frac{b}{r})^2 = 1$. Show that there exists a unique $0 \le \theta < 2\pi$ such that $\frac{a}{r} = \cos\theta$ and $\frac{b}{r} = \sin\theta$.

L'Hopital's Rule

Many limits of certain combinations of functions are difficult to evaluate because they lead to what's known as "indeterminate forms." These are expressions of the form $0/0$, ∞/∞, 0^0, $\infty - \infty$, 1^∞, and the like. They are precisely combinations of functions that are not covered by our limit theorems. See Theorem 4.1. The very definition of the derivative itself is such a case: $\lim_{h\to0}(f(c+h) - f(c)) = 0$, $\lim_{h\to0} h = 0$, and we are interested in the limit of the quotient of these two functions, which would lead us to the indeterminate form $0/0$. The definition of the number e is another example: $\lim(1+1/n) = 1$, $\lim n = \infty$, and we are interested in the limit of $(1+1/n)^n$, which leads to the indeterminate form 1^∞. L'Hopital's Rule, Theorem 4.16 below, is our strongest tool for handling such indeterminate forms.

To begin with, here is a useful generalization of the Mean Value Theorem.

THEOREM 4.15. (Cauchy Mean Value Theorem) Let f and g be continuous real-valued functions on a closed interval $[a,b]$, suppose $g(a) \neq g(b)$, and assume that both f and g are differentiable on the open interval (a,b). Then there exists a point $c \in (a,b)$ such that

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

Exercise 4.24. Prove the preceding theorem.
HINT: Define an auxiliary function h as was done in the proof of the original Mean Value Theorem.

The following theorem and exercise comprise what is called L'Hopital's Rule.

THEOREM 4.16. *Suppose f and g are differentiable real-valued functions on the bounded open interval (a,b) and assume that*

$$\lim_{x\to a+0} \frac{f'(x)}{g'(x)} = L,$$

where L is a real number. (Implicit in this hypothesis is that $g'(x) \neq 0$ for x in some interval $(a, a+\alpha)$.) Suppose further that either

$$\lim_{x\to a+0} f(x) = \lim_{x\to a+0} g(x) = 0$$

or

$$\lim_{x\to a+0} f(x) = \lim_{x\to a+0} g(x) = \infty.$$

then

$$\lim_{x\to a+0} \frac{f(x)}{g(x)} = L.$$

PROOF. Suppose first that

$$\lim_{x\to a+0} f(x) = \lim_{x\to a+0} g(x) = 0.$$

Observe first that, because $g'(x) \neq 0$ for all x in some interval $(a, a+\alpha)$, $g'(x)$ is either always positive or always negative on that interval. (This follows from part (d) of Exercise 4.15.) Therefore the function g must be strictly monotonic on the interval $(a, a+\alpha)$. Hence, since $\lim_{x\to a+0} g(x) = 0$, we must have that $g(x) \neq 0$ on the interval $(a, a+\alpha)$.

Now, given an $\epsilon > 0$, choose a positive $\delta < \alpha$ such that if $a < c < a+\delta$ then $|\frac{f'(c)}{g'(c)} - L| < \epsilon$. Then, for every natural number n for which $1/n < \delta$, and every $a < x < a+\delta$, we have by the Cauchy Mean Value Theorem that there exists a point c between $a+1/n$ and x such that

$$|\frac{f(x)-f(a+1/n)}{g(x)-g(a+1/n)} - L| = |\frac{f'(c)}{g'(c)} - L| < \epsilon.$$

Therefore, taking the limit as n approaches ∞, we obtain

$$|\frac{f(x)}{g(x)} - L| = \lim_{n\to\infty} |\frac{f(x)-f(a+1/n)}{g(x)-g(a+1/n)} - L| \leq \epsilon$$

for all x for which $a < x < a + \delta$. This proves the theorem in this first case.

Next, suppose that

$$\lim_{x \to a+0} f(x) = \lim_{x \to a+0} g(x) = \infty.$$

This part of the theorem is a bit more complicated to prove. First, choose a positive α so that $f(x)$ and $g(x)$ are both positive on the interval $(a, a+\alpha)$. This is possible because both functions are tending to infinity as x approaches a. Now, given an $\epsilon > 0$, choose a positive number $\beta < \alpha$ such that

$$|\frac{f'(c)}{g'(c)} - L| < \frac{\epsilon}{2}$$

for all $a < c < a + \beta$. We express this absolute value inequality as the following pair of ordinary inequalities:

$$L - \frac{\epsilon}{2} < \frac{f'(c)}{g'(c)} < L + \frac{\epsilon}{2}.$$

Set $y = a + \beta$. Using the Cauchy Mean Value Theorem, and the preceding inequalities, we have that for all $a < x < y$

$$L - \frac{\epsilon}{2} < \frac{f(x) - f(y)}{g(x) - g(y)} < L + \frac{\epsilon}{2},$$

implying that

$$(L - \frac{\epsilon}{2})(g(x) - g(y)) + f(y) < f(x) < (L + \frac{\epsilon}{2})(g(x) - g(y)) + f(y).$$

Dividing through by $g(x)$ and simplifying we obtain

$$L - \frac{\epsilon}{2} - \frac{(L - \frac{\epsilon}{2})g(y)}{g(x)} + \frac{f(y)}{g(x)} < \frac{f(x)}{g(x)} < L + \frac{\epsilon}{2} - \frac{(L + \frac{\epsilon}{2})g(y)}{g(x)} + \frac{f(y)}{g(x)}.$$

Finally, using the hypothesis that $\lim_{x \to a+0} g(x) = \infty$, and the fact that $L, \epsilon, g(y)$, and $f(y)$ are all constants, choose a $\delta > 0$, with $\delta < \beta$, such that if $a < x < a + \delta$, then

$$|-\frac{(L - \frac{\epsilon}{2})g(y)}{g(x)} + \frac{f(y)}{g(x)}| < \frac{\epsilon}{2}$$

and

$$|-\frac{(L + \frac{\epsilon}{2})g(y)}{g(x)} + \frac{f(y)}{g(x)}| < \frac{\epsilon}{2}.$$

Then, for all $a < x < a + \delta$, we would have

$$L - \epsilon < \frac{f(x)}{g(x)} < L + \epsilon,$$

implying that

$$|\frac{f(x)}{g(x)} - L| < \epsilon,$$

and the theorem is proved.

Exercise 4.25. (a) Show that the conclusions of the preceding theorem also hold if we assume that

$$\lim_{x \to a+0} \frac{f'(x)}{g'(x)} = \infty.$$

HINT: Replace ϵ by a large real number B and show that $f(x)/g(x) > B$ if $0 < x - a < \delta$.

(b) Show that the preceding theorem, as well as part (a) of this exercise, also holds if we replace the (finite) endpoint a by $-\infty$.
HINT: Replace the δ's by negative numbers B.

(c) Show that the preceding theorem, as well as parts a and b of this exercise, hold if the limit as x approaches a from the right is replaced by the limit as x approaches b from the left.
HINT: Replace $f(x)$ by $f(-x)$ and $g(x)$ by $g(-x)$.

(d) Give an example to show that the converse of L'Hopital's Rule need not hold; i.e., find functions f and g for which $\lim_{x\to a+0} f(x) = \lim_{x \to a+0} g(x) = 0$,

$$\lim_{x \to a+0} \frac{f(x)}{g(x)} \text{ exists, but } \lim_{x \to a+0} \frac{f'(x)}{g'(x)} \text{ does not exist.}$$

(e) Deduce from the proof given above that if $\lim_{x\to a+0} f'(x)/g'(x) = L$ and $\lim_{x\to a+0} g(x) = \infty$, then $\lim_{x\to a+0} f(x)/g(x) = L$ independent of the behavior of f.

(f) Evaluate $\lim_{x\to\infty} x^{1/x}$, and $\lim_{x\to 0}(1-x)^{1/x}$.
HINT: Take logarithms.

HIGHER ORDER DERIVATIVES

DEFINITION. Let S be a subset of $\mathbb{R}$ (or $\mathbb{C}$), and Let $f : S \to \mathbb{C}$ be a function of a real (or complex) variable. We say that f is *continuously differentiable* on S^0 if f is differentiable at each point x of S^0 and the function f' is continuous on S^0. We say that $f \in C^1(S)$ if f is continuous on S and continuously differentiable on S^0. We say that f is *2-times continuously differentiable* on S^0 if the first derivative f' is itself continuously differentiable on S^0. And, inductively, we say that f is *k-times continuously differentiable* on S^0 if the $k-1$st derivative of f is itself continuously differentiable on S^0. We write $f^{(k)}$ for the kth derivative of f, and we write $f \in C^k(S)$ if f is continuous on S and is k times continuously differentiable on S^0. Of course, if $f \in C^k(S)$, then all the derivatives $f^{(j)}$, for $j \le k$, exist nd are continuous on S^0. (Why?)

For completeness, we define $f^{(0)}$ to be f itself, and we say that $f \in C^\infty(S)$ if f is continuous on S and has infinitely many continuous derivatives on S^0; i.e., all of its derivatives exist and are continuous on S^0.

As in Chapter III, we say that f is *real-analytic* (or *complex-analytic*) on S if it is expandable in a Taylor series around each point $c \in S^0$

REMARK. Keep in mind that the definition above, as applied to functions whose domain S is a nontrivial subset of $\mathbb{C}$, has to do with functions of a complex variable that are continuously differentiable on the set S^0. We have seen that this is quite different from a function having continuous partial derivatives on S^0. We will return to partial derivatives at the end of this chapter.

THEOREM 4.17. *Let S be an open subset of $\mathbb{R}$ (or $\mathbb{C}$).*

(1) *Suppose WS is a subset of $\mathbb{R}$. Then, for each $k \geq 1$, there exists a function in $C^k(S)$ that is not in $C^{k+1}(S)$. That is, $C^{k+1}(S)$ is a proper subset of $C^k(S)$.*
(2) *If f is real-analytic (or complex-analytic) on S, then $f \in C^\infty(S)$.*
(3) *There exists a function in $C^\infty(\mathbb{R})$ that is not real-analytic on $\mathbb{R}$. That is, the set of real-analytic functions on $\mathbb{R}$ is a proper subset of the set $C^\infty(\mathbb{R})$.*

REMARK. Suppose S is an open subset of $\mathbb{C}$. It is a famous result from the Theory of Complex Variables that if f is in $C^1(S)$, then f is necessarily complex analytic on S. We will prove this amazing result in Theorem 7.5. Part (3) of the theorem shows that the situation is quite different for real-valued functions of a real variable.

PROOF. For part (1), see the exercise below. Part (2) is immediate from part (c) of Exercise 4.10. Before finishing the proof of part (3), we present the following lemma:

LEMMA. *Let f be the function defined on all of $\mathbb{R}$ as follows.*

$$f(x) = \begin{cases} 0 & x \leq 0 \\ \frac{p(x)e^{-1/x}}{x^n} & x > 0 \end{cases}$$

where $p(x)$ is a fixed polynomial function and n is a fixed nonnegative integer. Then f is continuous at each point x of $\mathbb{R}$.

PROOF OF THE LEMMA. The assertion of the lemma is clear if $x \neq 0$. To see that f is continuous at 0, it will suffice to prove that

$$\lim_{x \to 0+0} \frac{p(x)e^{-1/x}}{x^n} = 0.$$

(Why?) But, for $x > 0$, we know from part (b) of Exercise 3.22 that $e^{1/x} > 1/(x^{n+1}(n+1)!)$, implying that $e^{-1/x} < x^{n+1}(n+1)!$. Hence, for $x > 0$,

$$|f(x)| = \frac{|p(x)|e^{-1/x}}{x^n} < (n+1)!x|p(x)|,$$

and this tends to 0 as x approaches 0 from the right, as desired.

. Returning to the proof of Theorem 4.17, we verify part (3) by observing that if f is as in the preceding lemma then f is actually differentiable, and its derivative f' is a function of the same sort. (Why?) It follows that any such function belongs to $C^\infty(\mathbb{R})$. On the other hand, a nontrivial such f cannot be expandable in a Taylor series around 0 because of the Identity Theorem. (Take $x_k = -1/k$.) This completes the proof.

Exercise 4.26. (a) Prove part (1) of Theorem 4.17. Use functions of the form $x^n \sin(1/x)$.

(b) Prove that any function of the form of the f in the lemma above is everywhere differentiable on $\mathbb{R}$, and its derivative has the same form. Conclude that any such function belongs to $C^\infty(\mathbb{R})$.

(c) For each positive integer n, define a function f_n on the interval $(-1,1$ by $f_n(x) = |x|^{1+1/n}$. Prove that each f_n is differentiable at every point in $(-1,1)$, including 0. Prove also that the sequence $\{f_n\}$ converges uniformly to the function $f(x) = |x|$. (See part (h) of Exercise 3.28.) Conclude that the uniform limit of differentiable functions of a real variable need not be differentiable. (Again, for functions of a complex variable, the situation is very different. In that case, the uniform limit of differentiable functions **is** differentiable. See Theorem 7.11.)

Exercise 4.27. (A smooth approximation to a step function.) Suppose $a < b < c < d$ are real numbers. Show that there exists a function χ in $C^\infty(\mathbb{R})$ such that $0 \leq \chi(x) \leq 1$ for all x, $\chi(x) \equiv 1$ for $x \in [b,c]$, and $\chi(x) \equiv 0$ for $x \notin (a,d)$. (If a is close to b and c is close to d, then this function is a C^∞ approximation to the step function that is 1 on the interval $[b,c]$ and 0 elsewhere.)

(a) Let f be a function like the one in the lemma. Think about the graphs of the functions $f(x-c)$ and $f(b-x)$. Construct a C^∞ function g that is 0 between b and c and positive everywhere else.

(b) Construct a C^∞ function h that is positive between a and d and 0 everywhere else.

(c) Let g and h be as in parts (a) and (b). If $j = g + h$, show that j is never 0, and write k for the C^∞ function $k = 1/j$.

(d) Examine the function hk, and show that it is the desired function χ.

THEOREM 4.18. (Formula for the coefficients of a Taylor Series function) Let f be expandable in a Taylor series around a point c :

$$f(x) = \sum a_n (x-c)^n.$$

Then for each n, $a_n = f^{(n)}(c)/n!$.

PROOF. Because each derivative of a Taylor series function is again a Taylor series function, and because the value of a Taylor series function at the point c is equal to its constant term a_0, we have that $a_1 = f'(c)$. Computing the derivative of the derivative, we see that $2a_2 = f''(c) = f^{(2)}(c)$. Continuing this, i.e., arguing by induction, we find that $n!a_n = f^{(n)}(c)$, which proves the theorem.

TAYLOR POLYNOMIALS AND TAYLOR'S REMAINDER THEOREM

DEFINITION. Let f be in $C^n(B_r(c))$ for c a fixed complex number, $r > 0$, and n a positive integer. Define the *Taylor polynomial* of degree n for f at c to be the polynomial $T^n \equiv T^n_{(f,c)}$ given by the formula:

$$(T^n_{(f,c)})(z) = \sum_{j=0}^{n} a_j (z-c)^j,$$

where $a_j = f^{(j)}(c)/j!$.

REMARK. If f is expandable in a Taylor series on $B_r(c)$, then the Taylor polynomial for f of degree n is nothing but the nth partial sum of the Taylor series for f on $B_r(c)$. However, any function that is n times differentiable at a point c has a Taylor polynomial of order n. Functions that are infinitely differentiable have

Taylor polynomials of all orders, and we might suspect that these polynomials are some kind of good approximation to the function itself.

Exercise 4.28. Prove that f is expandable in a Taylor series function around a point c (with radius of convergence $r > 0$) if and only if the sequence $\{T^n_{(f,c)}\}$ of Taylor polynomials converges pointwise to f; i.e.,

$$f(z) = \lim(T^n_{(f,c)})(z)$$

for all z in $B_r(c)$.

Exercise 4.29. Let $f \in C^n(B_r(c))$. Prove that $f' \in C^{n-1}(B_r(c))$. Prove also that $(T^n_{(f,c)})' = T^{n-1}_{(f',c)}$.

The next theorem is, in many ways, the fundamental theorem of numerical analysis. It clearly has to do with approximating a general function by polynomials. It is a generalization of the Mean Value Theorem, and as in that case this theorem holds only for real-valued functions of a real variable.

THEOREM 4.19. (Taylor's Remainder Theorem) Let f be a real-valued function on an interval $(c-r, c+r)$, and assume that $f \in C^n((c-r, c+r))$, and that $f^{(n)}$ is differentiable on $(c-r, c+r)$. Then, for each x in $(c-r, c+r)$ there exists a y between c and x such that

$$f(x) - (T^n_{(f,c)})(x) = \frac{f^{(n+1)}(y)}{(n+1)!}(x-c)^{n+1}. \tag{4.7}$$

REMARK. If we write $f(x) = T^n_{f,c})(x) + R_{n+1}(x)$, where $R_{n+1}(x)$ is the error or remainder term, then this theorem gives a formula, and hence an estimate, for that remainder term. This is the evident connection with Numerical Analysis.

PROOF. We prove this theorem by induction on n. For $n = 0$, this is precisely the Mean Value Theorem. Thus,

$$f(x) - T^0_{f,c}(x) = f(x) - f(c) = f'(y)(x - c.$$

Now, assuming the theorem is true for all functions in $C^{n-1}((c-r, c+r))$, let us show it is true for the given function $f \in C^n((c-r, c+r))$. Set $g(x) = f(x) - (T^n_{(f,c)})(x)$ and let $h(x) = (x-c)^{n+1}$. Observe that both $g(c) = 0$ and $h(c) = 0$. Also, if $x \neq c$, then $h(x) \neq 0$. So, by the Cauchy Mean Value Theorem, we have that

$$\frac{g(x)}{h(x)} = \frac{g(x) - g(c)}{h(x) - h(c)} = \frac{g'(w)}{h'(w)}$$

for some w between c and x. Now

$$g'(w) = f'(w) - (t^n_{(f,c)})'(w) = f'(w) - (T^{n-1}_{(f',c)})(w)$$

(See the preceding exercise.), and $h'(w) = (n+1)(w-c)^n$. Therefore,

$$\begin{aligned}\frac{f(x) - (T^n_{(f,c)})(x)}{(x-c)^{n+1}} &= \frac{g(x)}{h(x)} \\ &= \frac{g'(w)}{h'(w)} \\ &= \frac{f'(w) - (T^{n-1}_{(f',c)})(w)}{(n+1)(w-c)^n}.\end{aligned}$$

We apply the inductive hypotheses to the function f' (which is in $C^{n-1}((c-r, c+r))$) and obtain

$$\begin{aligned}\frac{f(x) - (T^n_{(f,c)})(x)}{(x-c)^{n+1}} &= \frac{f'(w) - (T^{n-1}_{(f',c)})(w)}{(n+1)(w-c)^n} \\ &= \frac{\frac{f'^{(n)}(y)}{n!}(w-c)^n}{(n+1)(w-c)^n} \\ &= \frac{f'^{(n)}(y)}{(n+1)!} \\ &= \frac{f^{(n+1)}(y)}{(n+1)!}\end{aligned}$$

for some y between c and w. But this implies that

$$f(x) - (T^n_{(f,c)})(x) = \frac{f^{(n+1)}(y)(x-c)^{n+1}}{(n+1)!},$$

for some y between c and x, which finishes the proof of the theorem.

Exercise 4.30. Define $f(x) = 0$ for $x \leq 0$ and $f(x) = e^{-1/x}$ for $x > 0$. Verify that $f \in C^\infty(\mathbb{R})$, that $f^{(n)}(0) = 0$ for all n, and yet f is not expandable in a Taylor series around 0. Interpret Taylor's Remainder Theorem for this function. That is, describe the remainder $R_{n+1}(x)$.

As a first application of Taylor's Remainder Theorem we give the following result, which should be familiar from calculus. It is the generalized version of what's ordinarily called the "second derivative test."

THEOREM 4.20. (Test for Local Maxima and Minima) Let f be a real-valued function in $C^n(c-r, c+r)$, suppose that the $n+1$st derivative $f^{(n+1)}$ of f exists everywhere on $(c-r, c+r)$ and is continuous at c, and suppose that $f^{(k)}(c) = 0$ for all $1 \leq k \leq n$ and that $f^{(n+1)}(c) \neq 0$. Then:

(1) If n is even, f attains neither a local maximum nor a local minimum at c. In this case, c is called an *inflection point*.
(2) If n is odd and $f^{(n+1)}(c) < 0$, then f attains a local maximum at c.
(3) If n is odd and $f^{(n+1)}(c) > 0$, then f attains a local minimum at c.

PROOF. Since $f^{(n+1)}$ is continuous at c, there exists a $\delta > 0$ such that $f^{(n+1)}(y)$ has the same sign as $f^{(n+1)}(c)$ for all $y \in (c-\delta, c+\delta)$. We have by Taylor's Theorem that if $x \in (c-\delta, c+\delta)$ then there exists a y between x and c such that

$$f(x) = (T^n_{(f,c)})(x) + \frac{f^{(n+1)}(y)}{(n+1)!}(x-c)^{n+1},$$

from which it follows that

$$\begin{aligned}f(x) - f(c) &= \sum_{k=1}^{n} f^{(k)}(c)k!(x-c)^k + \frac{f^{(n+1)}(y)}{(n+1)!}(x-c)^{n+1} \\ &= \frac{f^{(n+1)}(y)}{(n+1)!}(x-c)^{n+1}.\end{aligned}$$

Suppose n is even. It follows then that if $x < c$, the sign of $(x-c)^{n+1}$ is negative, so that the sign of $f(x) - f(c)$ is the opposite of the sign of $f^{(n+1)}(c)$. On the other hand, if $x > c$, then $(x - c)^{n+1} > 0$, so that the sign of $f(x) - f(c)$ is the same as the sign of $f^{(n+1)}(c)$. So, $f(x) > f(c)$ for all nearby x on one side of c, while $f(x) < f(c)$ for all nearby x on the other side of c. Therefore, f attains neither a local maximum nor a local minimum at c. This proves part (1).

Now, if n is odd, the sign of $f(x) - f(c)$ is the same as the sign of $f^{(n+1)}(y)$, which is the same as the sign of $f^{(n+1)}(c)$, for all $x \in (c - \delta, c + \delta)$. Hence, if $f^{(n+1)}(c) < 0$, then $f(x) - f(c) < 0$ for all $x \in (c - \delta, c + \delta)$, showing that f attains a local maximum at c. And, if $f^{(n+1)}(c) > 0$, then the sign of $f(x) - f(c)$ is positive for all $x \in (c - \delta, c + \delta)$, showing that f attains a local minimum at c. This proves parts (2) and (3).

The General Binomial Theorem

We use Taylor's Remainder Theorem to derive a generalization of the Binomial Theorem to nonintegral exponents. First we must generalize the definition of binomial coefficient.

DEFINITION. Let α be a complex number, and let k be a nonnegative integer. We define the general *binomial coefficient* $\binom{\alpha}{k}$ by

$$\binom{\alpha}{k} = \frac{\alpha(\alpha-1)\dots(\alpha-k+1)}{k!}.$$

If α is itself a positive integer and $k \leq \alpha$, then $\binom{\alpha}{k}$ agrees with the earlier definition of the binomial coefficient, and $\binom{\alpha}{k} = 0$ when $k > \alpha$. However, if α is not an integer, but just an arbitrary complex number, then every $\binom{\alpha}{k} \neq 0$.

Exercise 4.31. Estimates for the size of binomial coefficients. Let α be a fixed complex number.

(a) Show that

$$|\binom{\alpha}{k}| \leq \prod_{j=1}^{k}(1 + \frac{|\alpha|}{j})$$

for all nonnegative integers k.
HINT: Note that

$$|\binom{\alpha}{k}| \leq \frac{|\alpha|(|alpha|+1)(|alpha|+2)\dots(|\alpha|+k-1)}{k!}.$$

(b) Use part (a) to prove that there exists a constant C such that

$$|\binom{\alpha}{k}| \leq C2^k$$

for all nonnegative integers k.
HINT: Note that $(1 + |\alpha|/j) < 2$ for all $j > |\alpha|$.

(c) Show in fact that for each $\epsilon > 0$ there exists a constant C_ϵ such that

$$|\binom{\alpha}{k}| \leq C_\epsilon (1+\epsilon)^k$$

for all nonnegative integers k.

(d) Let $h(t)$ be the power series function given by $h(t) = \sum_{k=0}^\infty \binom{\alpha}{k} t^k$. Use the ratio test to show that the radius of convergence for h equals 1.

REMARK. The general Binomial Theorem, if there is one, should be something like the following:

$$(x+y)^\alpha = \sum_{k=0}^\infty \binom{\alpha}{k} x^{\alpha-k} y^k.$$

The problem is to determine when this infinite series converges, i.e., for what values of the three variables x, y, and α does it converge. It certainly is correct if $x = 0$, so we may as well assume that $x \neq 0$, in which case we are considering the validity of the formula

$$(x+y)^\alpha = x^\alpha (1+t)^\alpha = x^\alpha \sum_{k=0}^\infty \binom{\alpha}{k} t^k,$$

where $t = y/x$. Therefore, it will suffice to determine for what values of t and α does the infinite series

$$\sum_{k=0}^\infty \binom{\alpha}{k} t^k$$

equal

$$(1+t)^\alpha.$$

The answer is that, for n arbitrary complex number α, this series converges to the correct value for all $t \in (-1, 1)$. (Of course, t must be larger than -1 for the expression $(1+t)^\alpha$ even to be defined.) However, the next theorem only establishes this equality for t's in the subinterval $(-1/2, 1/2)$. As mentioned earlier, its proof is based on Taylor's Remainder Theorem. We must postpone the complete proof to the next chapter, where we will have a better version of Taylor's Theorem.

THEOREM 4.21. *Let $\alpha = a + bi$ be a fixed complex number. Then*

$$(1+t)^\alpha = \sum_{k=0}^\infty \binom{\alpha}{k} t^k$$

for all $t \in (-1/2, 1/2)$.

PROOF. Of course, this theorem is true if α is a nonnegative integer, for it is then just the original Binomial Theorem, and in fact in that case it holds for every complex number t. For a general complex number α, we have only defined x^α for positive x's, so that $(1+t)^\alpha$ is not even defined for $t < -1$.

Now, for a general $\alpha = a + bi$, consider the function $g : (-1/2, 1/2) \to \mathbb{C}$ defined by $g(t) = (1+t)^\alpha$. Observe that the nth derivative of g is given by

$$g^{(n)}(t) = \frac{\alpha(\alpha-1)\ldots(\alpha-n+1)}{(1+t)^{n-\alpha}}.$$

Then $g \in C^\infty((-1/2, 1/2))$. (Of course, g is actually in $C^\infty(-1,1)$, but the present theorem is only concerned with t's in $(-1/2, 1/2)$.)

For each nonnegative integer k define

$$a_k = g^{(k)}(0)/k! = \frac{\alpha(\alpha-1)\dots(\alpha-k+1)}{k!} = \binom{\alpha}{k},$$

and set h equal to the power series function given by $h(t) = \sum_{k=0}^\infty a_k t^k$. According to part (d) of the preceding exercise, the radius of convergence for the power series$\sum a_k t^k$ is 1. The aim of this theorem is to show that $g(t) = h(t)$ for all $-1/2 < t < 1/2$. In other words, we wish to show that g agrees with this power series function at least on the interval $(-1/2, 1/2)$. It will suffice to show that the sequence $\{S_n\}$ of partial sums of the power series function h converges to the function g, at least on $(-1/2, 1/2)$. We note also that the nth partial sum of this power series is just the nth Taylor polynomial T_g^n for g.

$$S_n(t) = \sum_{k=0}^n \binom{\alpha}{k} t^k = \sum_{k=0}^n \frac{g^{(k)}(0)}{k!} t^k.$$

Now, fix a t strictly between $-1/2$ and $1/2$, and let $r < 1$ be as in part (c) of Exercise 4.11. That is, $|t/(1+y)| < r$ for every y between 0 and t. (This is an important inequality for our proof, and this is one place where the hypothesis that $t \in (-1/2, 1/2)$ is necessary.) Note also that, for any $y \in (-1/2, 1/2)$, we have $|(1+y)^\alpha| = (1+y)^a$, and this is trapped between $(1/2)^a$ and $(3/2)^a$. Hence, there exists a number M such that $|(1+y)^\alpha| \le M$ for all $y \in (-1/2, 1/2)$.

Next, choose an $\epsilon > 0$ for which $\beta = (1+\epsilon)r < 1$. We let C_ϵ be a constant satisfying the inequality in Part (c) of Exercise 4.31. So, using Taylor's Remainder Theorem, we have that there exists a y between 0 and t for which

$$\begin{aligned}
|g(t) - \sum_{k=0}^n a_k t^k| &= |g(t) - (T_{(g,0)}^n(t)| \\
&= |\frac{g^{(n+1)}(y)}{(n+1)!} t^{n+1}| \\
&= |\frac{\alpha(\alpha-1)\dots(\alpha-n)}{(n+1)!(1+y)^{n+1-\alpha}} t^{n+1}| \\
&\le |\binom{\alpha}{n+1}||(1+y)^\alpha|| \frac{t}{1+y}|^{n+1} \\
&\le C_\epsilon (1+\epsilon)^{n+1} M | \frac{t}{1+y}|^{n+1} \\
&\le C_\epsilon (1+\epsilon)^{n+1} M r^{n+1} \\
&\le C_\epsilon M \beta^{n+1},.
\end{aligned}$$

Taking the limit as n tends to ∞, and recalling that $\beta < 1$, shows that $g(t) = h(t)$ for all $-1/2 < t < 1/2$, which completes the proof.

MORE ON PARTIAL DERIVATIVES

We close the chapter with a little more concerning partial derivatives. Thus far, we have discussed functions of a single variable, either real or complex. However, it is difficult not to think of a function of one complex variable $z = x + iy$ as equally well being a function of the two real variables x and y. We will write (a, b) and $a+bi$ to mean the same point in $\mathbb{C} \equiv \mathbb{R}^2$, and we will write $|(a, b)|$ and $|a+bi|$ to indicate the same quantity, i.e., the absolute value of the complex number $a + bi \equiv (a, b)$. We have seen in Theorem 4.4 that the only real-valued, differentiable functions of a complex variable are the constant functions. However, this is far from the case if we consider real-valued functions of two real variables, as is indicated in Exercise 4.8. Consequently, we make the following definition of differentiability of a real-valued function of two real variables. Note that it is clearly different from the definition of differentiability of a function of a single complex variable, and though the various notations for these two kinds of differentiability are clearly ambiguous, we will leave it to the context to indicate which kind we are using.

DEFINITION. Let $f : S \to \mathbb{R}$ be a function whose domain is a subset S of $\mathbb{R}^2$, and let $c = (a, b)$ be a point in the interior S^0 of S. We say that f is *differentiable, as a function of two real variables,* at the point (a, b) if there exists a pair of real numbers L_1 and L_2 and a function θ such that

$$f(a + h_1, b + h_2) - f(a, b) = L_1 h_1 + L_2 h_2 + \theta(h_1, h_2) \tag{4.8}$$

and

$$\lim_{|(h_1,h_2)|\to 0} \frac{\theta(h_1, h_2)}{|(h_1, h_2)|} = 0. \tag{4.9}$$

One should compare this definition with part (3) of Theorem 4.2.

Each partial derivative of a function f is again a real-valued function of two real variables, and so it can have partial derivatives of its own. We use simplifying notation like f_{xyxx} and $f_{yyyxyy\ldots}$ to indicate "higher order" mixed partial derivatives. For instance, f_{xxyx} denotes the fourth partial derivative of f, first with respect to x, second with respect to x again, third with respect to y, and finally fourth with respect to x. These higher order partial derivatives are called *mixed partial derivatives.*

DEFINITION. Suppose S is a subset of $\mathbb{R}^2$, and that f is a continuous real-valued function on S. If both partial derivatives of f exist at each point of the interior S^0 of S, and both are continuous on S^0, then f is said to belong to $C^1(S)$. If all kth order mixed partial derivatives exist at each point of S^0, and all of them are continuous on S^0, then f is said to belong to $C^k(S)$. Finally, if all mixed partial derivatives, of arbitrary orders, exist and are continuous on S^0, then f is said to belong to $C^\infty(S)$.

Exercise 4.32. (a) Suppose f is a real-valued function of two real variables and that it is differentiable, as a function of two real variables, at the point (a, b). Show that the numbers L_1 and L_2 in the definition are exactly the partial derivatives of f at (a, b). That is,

$$L_1 = \frac{\partial f}{\partial x}(a, b) = \lim_{h\to 0} \frac{f(a + h, b) - f(a, b)}{h}$$

and

$$L_2 = \frac{\partial f}{\partial y}(a,b) = \lim_{h \to 0} \frac{f(a,b+h) - f(a,b)}{h}.$$

(b) Define f on $\mathbb{R}^2$ as follows: $f(0,0) = 0$, and if $(x,y) \neq (0,0)$, then $f(x,y) = xy/(x^2+y^2)$. Show that both partial derivatives of f at $(0,0)$ exist and are 0. Show also that f is **not**, as a function of two real variables, differentiable at $(0,0)$.
HINT: Let h and k run through the numbers $1/n$.

(c) What do parts (a) and (b) tell about the relationship between a function of two real variables being differentiable at a point (a,b) and its having both partial derivatives exist at (a,b)?

(d) Suppose $f = u + iv$ is a complex-valued function of a complex variable, and assume that f is differentiable, as a function of a complex variable, at a point $c = a + bi \equiv (a,b)$. Prove that the real and imaginary parts u and v of f are differentiable, as functions of two real variables. Relate the five quantities

$$\frac{\partial u}{\partial x}(a,b), \frac{\partial u}{\partial y}(a,b), \frac{\partial v}{\partial x}(a,b), \frac{\partial v}{\partial y}(a,b), \text{ and } f'(c).$$

Perhaps the most interesting theorem about partial derivatives is the "mixed partials are equal" theorem. That is, $f_{xy} = f_{yx}$. The point is that this is**not** always the case. An extra hypothesis is necessary.

THEOREM 4.22. (Theorem on mixed partials) Let $f : S \to \mathbb{R}$ be such that both second order partials derivatives f_{xy} and f_{yx} exist at a point (a,b) of the interior of S, and assume in addition that one of these second order partials exists at every point in a disk $B_r(a,b)$ around (a,b) and that it is continuous at the point (a,b). Then $f_{xy}(a,b) = f_{yx}(a,b)$.

PROOF. Suppose that it is f_{yx} that is continuous at (a,b). Let $\epsilon > 0$ be given, and let $\delta_1 > 0$ be such that if $|(c,d) - (a,b)| < \delta_1$ then $|f_{yx}(c,d) - f_{yx}(a,b)| < \epsilon$. Next, choose a δ_2 such that if $0 < |k| < \delta_2$, then

$$|f_{xy}(a,b) - \frac{f_x(a,b+k) - f_x(a,b)}{k}| < \epsilon,$$

and fix such a k. We may also assume that $|k| < \delta_1/2$. Finally, choose a $\delta_3 > 0$ such that if $0 < |h| < \delta_3$, then

$$|f_x(a,b+k) - \frac{f(a+h,b+k) - f(a,b+k)}{h}| < |k|\epsilon,$$

and

$$|f_x(a,b) - \frac{f(a+h,b) - f(a,b)}{h}| < |k|\epsilon,$$

and fix such an h. Again, we may also assume that $|h| < \delta_1/2$.

In the following calculation we will use the Mean Value Theorem twice.

$$
\begin{aligned}
0 &\leq |f_{xy}(a,b) - f_{yx}(a,b)| \\
&\leq |f_{xy}(a,b) - \frac{f_x(a,b+k) - f_x(a,b)}{k}| \\
&\quad + |\frac{f_x(a,b+k) - f_x(a,b)}{k} - f_{yx}(a,b)| \\
&\leq \epsilon + |\frac{f_x(a,b+k) - \frac{f(a+h,b+k)-f(a,b+k)}{h}}{k}| \\
&\quad + |\frac{\frac{f(a+h,b)-f(a,b)}{h} - f_x(a,b)}{k}| \\
&\quad + |\frac{f(a+h,b+k) - f(a,b+k) + (f(a+h,b) - f(a,b))}{hk} - f_{yx}(a,b)| \\
&< 3\epsilon + |\frac{f(a+h,b+k) - f(a,b+k) + (f(a+h,b) - f(a,b))}{hk} - f_{yx}(a,b)| \\
&= 3\epsilon + |\frac{f_y(a+h,b') - f_y(a,b')}{h} - f_{yx}(a,b)| \\
&= 3\epsilon + |f_{yx}(a',b') - f_{yx}(a,b)| \\
&< 4\epsilon,
\end{aligned}
$$

because b' is between b and $b+k$, and a' is between a and $a+h$, so that $|(a',b') - (a,b)| < \delta_1/\sqrt{2} < \delta_1$. Hence, $|f_{xy}(a,b) - f_{yx}(a,b) < 4\epsilon$, for an arbitrary ϵ, and so the theorem is proved.

Exercise 4.33. Let f be defined on $\mathbb{R}^2$ by $f(0,0) = 0$ and, for $(x,y) \neq (0,0)$, $f(x,y) = x^3y/(x^2+y^2)$.

(a) Prove that both partial derivatives f_x and f_y exist at each point in the plane.

(b) Show that $f_{yx}(0,0) = 1$ and $f_{xy}(0,0) = 0$.

(c) Show that f_{xy} exists at each point in the plane, but that it is not continuous at $(0,0)$.

The following exercise is an obvious generalization of the First Derivative Test for Extreme Values, Theorem 4.8, to real-valued functions of two real variables.

Exercise 4.34. Let $f : S \to \mathbb{R}$ be a real-valued function of two real variables, and let $c = (a,b) \in S^0$ be a point at which f attains a local maximum or a local minimum. Show that if either of the partial derivatives $\partial f/\partial x$ or $\partial f/\partial y$ exists at c, then it must be equal to 0.
HINT: Just consider real-valued functions of a real variable like $x \to f(x,b)$ or $y \to f(a,y)$, and use Theorem 4.8.

Whenever we make a new definition about functions, the question arises of how the definition fits with algebraic combinations of functions and how it fits with the operation of composition. In that light, the next theorem is an expected one.

THEOREM 4.23. (Chain Rule again) Suppose S is a subset of $\mathbb{R}^2$, that (a,b) is a point in the interior of S, and that $f : S \to \mathbb{R}$ is a real-valued function that is differentiable, as a function of two real variables, at the point (a,b). Suppose that T is a subset of $\mathbb{R}$, that c belongs to the interior of T, and that $\phi : T \to \mathbb{R}^2$ is differentiable at the point c and $\phi(c) = (a, b)$. Write $\phi(t) = (x(t), y(t))$. Then the composition $f \circ \phi$ is differentiable at c and

$$f \circ \phi'(c) = \frac{\partial f}{\partial x}(a,b)x'(c) + \frac{\partial f}{\partial y}(a,b)y'(c) = \frac{\partial f}{\partial x}(\phi(c))x'(c) + \frac{\partial f}{\partial y}(\phi(c))y'(c).$$

PROOF. From the definition of differentiability of a real-valued function of two real variables, write

$$f(a+h_1, b+h_2) - f(a,b) = L_1h_1 + L_2h_2 + \theta_f(H_1, h_2).$$

and from part (3) of Theorem 4.2, write

$$\phi(c+h) - \phi(c) = \phi'(c)h + \theta_\phi(h),$$

or, in component form,

$$x(c+h) - x(c) = x(c+h) - a = x'(c)h + \theta_x(h)$$

and

$$y(c+h) - y(c) = y(c+h) - b = y'(c)h + \theta_y(h).$$

We also have that

$$\lim_{|(h_1,h_2)|\to 0} \frac{\theta_f((h_1,h_2))}{|(h_1,h_2)|} = 0,$$

$$\lim_{h\to 0} \frac{\theta_x(h)}{h} = 0,$$

and

$$\lim_{h\to 0} \frac{\theta_y(h)}{h} = 0.$$

We will show that $f \circ \phi$ is differentiable at c by showing that there exists a number L and a function θ satisfying the two conditions of part (3) of Theorem 4.2.

Define

$$k_1(h), k_2(h)) = \phi(c+h) - \phi(c) = (x(c+h) - x(c), y(c+h) - y(c)).$$

Thus, we have that

$$\begin{aligned} f \circ \phi(c+h) - f \circ \phi(c) &= f(\phi(c+h)) - f(\phi(c)) \\ &= f(x(c+h), y(c+h)) - f(x(c), y(c)) \\ &= f(a + k_1(h), b + k_2(h)) - f(a,b) \\ &= L_1k_1(h) + L_2k_2(h) + \theta_f(k_1(h), k_2(h)) \\ &= l_1(x(c+h) - x(c)) + L_2(y(c+h) - y(c)) \\ &\qquad + \theta_f(k_1(h), k_2(h)) \\ &= L_1(x'(c)h + \theta_x(h)) + L_2(y'(c)h + \theta_y(h)) \\ &\qquad + \theta_f(k_1(h), k_2(h)) \\ &= (L_1x'(c) + L_2y'(c))h \\ &\qquad + L_1\theta_x(h) + L_2\theta_y(h) + \theta_f(k_1(h), k_2(h)). \end{aligned}$$

We define $L = (L_1x'(c) + L_2y'(c))$ and $\theta(h) = l_1\theta_x(h) + L_2\theta_y(h) + \theta_f(k_1(h), k_2(h))$. By these definitions and the calculation above we have Equation (4.1)

$$f \circ \phi(c + h) - f \circ \phi(c) = Lh + \theta(h),$$

so that it only remains to verify Equation (4.2) for the function θ. We have seen above that the first two parts of θ satisfy the desired limit condition, so that it is just the third part of θ that requires some proof. The required argument is analogous to the last part of the proof of the Chain Rule (Theorem 4.7), and we leave it as an exercise.

Exercise 4.35. (a) Finish the proof to the preceding theorem by showing that

$$\lim_{h \to 0} \frac{\theta_f(k_1(h), k_2(h))}{h} = 0.$$

HINT: Review the corresponding part of the proof to Theorem 4.7.

(b) Suppose $f : S \to \mathbb{R}$ is as in the preceding theorem and that ϕ is a real-valued function of a real variable. Suppose f is differentiable, as a function of two real variables, at the point (a, b), and that ϕ is differentiable at the point $c = f(a, b)$. Let $g = \phi \circ f$. Find a formula for the partial derivatives of the real-valued function g of two real variables.

(c) (A generalized Mean Value Theorem) Suppose u is a real-valued function of two real variables, both of whose partial derivatives exist at each point in a disk $B_r(a, b)$. Show that, for any two points (x, y) and (x', y') in $B_r(a, b)$, there exists a point $(\hat{x}, \hat{y})$ on the line segment joining (x, y) to (x', y') such that

$$u(x, y) - u(x', y') = \frac{\partial u}{\partial x}(\hat{x}, \hat{y})(x - x') + \frac{\partial u}{\partial y}(\hat{x}, \hat{y})(y - y').$$

HINT: Let $\phi : [0, 1] \to \mathbb{R}^2$ be defined by $\phi(t) = (1 - t)(x', y') + t(x, y)$. Now use the preceding theorem.

(d) Verify that the assignment $f \to \partial f/\partial x$ is linear; i.e., that

$$\frac{\partial(f + g)}{\partial x} = \frac{\partial f}{\partial x} + \frac{\partial g}{\partial x}.$$

Check that the same is true for partial derivatives with respect to y.