(d) $\sinh(x+iy) = \sinh x \cos y + i \cosh x \sin y$

(e) $\cosh(x+iy) = \cosh x \cos y + i \sinh x \sin y$

17.• Use the equation $\sin z = \sin x \cosh y + i \sinh y \cos x$ where $z = x+iy$ to prove that $|\sinh y| \leq |\sin z| \leq |\cosh y|$.

18. If b is real, prove that $|a^b| = |a|^b$.

19. Is it true that $|a^b| = |a|^{|b|}$ for all $a, b \in \mathbb{C}$?

20. (a) For complex numbers a, b, c, prove that $a^b a^c = a^{b+c}$, using a fixed branch of log.

(b) Show that $(ab)^c = a^c b^c$ if we choose branches so that $\log(ab) = \log a + \log b$ (with no extra $2\pi n i$).

21.• Using polar coordinates, show that $z \mapsto z + 1/z$ maps the circle $|z| = 1$ to the interval $[-2, 2]$ on the x axis.

22. (a) The map $z \mapsto z^3$ maps the first quadrant onto what?

(b) Discuss the geometry of $z \mapsto \sqrt[3]{z}$ as was done in the text for $\sqrt{z}$.

23.• The map $z \mapsto 1/z$ takes the exterior of the unit circle to the interior (excluding zero) and vice versa. To what are lines $\arg z =$ constant mapped?

24. What are the images of vertical and horizontal lines under $z \mapsto \cos z$?

25. Under what conditions does $\log a^b = b \log a$ for complex numbers a, b? (Use the branch of log with $-\pi \leq \theta < \pi$.)

26. (a) Show that under the map $z \mapsto z^2$, lines parallel to the real axis are mapped to parabolas.

(b) Show that under (a branch of) $z \mapsto \sqrt{z}$, lines parallel to the real axis are mapped to hyperbolas.

27. Show that the n nth roots of unity are $1, w, w^2, w^3, \ldots, w^{n-1}$, where $w = e^{2\pi i/n}$.

28. Show that the trigonometric identities can be deduced if $e^{i(x_1+x_2)} = e^{ix_1} \cdot e^{ix_2}$ is assumed.

29.• Show that $\sin z = 0$ iff $z = k\pi, k = 0, \pm 1, \pm 2, \ldots$.

30. Show that the sine and cosine are periodic with minimum period 2π; that is, that

(a) $\sin(z + 2\pi) = \sin z$ for all z.

•(b) $\cos(z + 2\pi) = \cos z$ for all z.

(c) $\sin(z + \omega) = \sin z$ for all z implies $\omega = 2\pi n$ for some integer n.

(d) $\cos(z + \omega) = \cos z$ for all z implies $\omega = 2\pi n$ for some integer n.

31. Find the maximum of $|\cos z|$ on the square

$$0 \leq \operatorname{Re} z \leq 2\pi, 0 \leq \operatorname{Im} z \leq 2\pi.$$

32. Show that $\log z = 0$ iff $z = 1$, using the branch with $-\pi < \arg z \leq \pi$.

33. Compute the following quantities numerically to two significant figures:
(a) $e^{3.2+6.1i}$ (b) $\log(1.2 - 3.0i)$ (c) $\sin(8.1i - 3.2)$

34.• Show that the function $\sin z$ maps the strip $-\pi/2 < \operatorname{Re} z < \pi/2$ onto the set $\mathbb{C}\backslash\{z \mid \operatorname{Im} z = 0 \text{ and } |\operatorname{Re} z| \geq 1\}$.

35.• Discuss the inverse functions $\sin^{-1} z$ and $\cos^{-1} z$. For example, is $\sin z$ one-to-one on the set defined by $0 \leq \operatorname{Re} z < 2\pi$?

1.4 Continuous Functions

In this section and the next, the fundamental notions of continuity and differentiability for complex-valued functions of a complex variable will be analyzed. The results are similar to those learned in the calculus of functions of real variables. These sections will be concerned mostly with the underlying theory, which is applied to the elementary functions in §1.6.

Since $\mathbb{C}$ is $\mathbb{R}^2$ with the extra structure of complex multiplication, many geometric concepts can be translated from $\mathbb{R}^2$ into complex notation. This has already been done for the absolute value, $|z|$, which is the same as the norm, or length, of z regarded as a vector in $\mathbb{R}^2$. Furthermore, we will use calculus for functions of two variables in the study of functions of a complex variable.

Open Sets We will need the notion of an open set. A set $A \subset \mathbb{C} = \mathbb{R}^2$ is called ***open*** when, for each point z_0 in A, there is a real number $\epsilon > 0$ such that $z \in A$ whenever $|z - z_0| < \epsilon$. See Figure 1.4.1. The value of ϵ may depend on z_0; as z_0 gets close to the "edge" of A, ϵ gets smaller. Intuitively, a set is open if it does not contain any of its "boundary" or "edge" points.

For a number $r > 0$, the r ***neighborhood*** or r ***disk*** around a point z_0 in $\mathbb{C}$ is defined to be the set $D(z_0; r) = \{z \in \mathbb{C} \mid |z - z_0| < r\}$. For practice, the student should prove that for each $w_0 \in \mathbb{C}$ and $r > 0$, the disk $A = \{z \in \mathbb{C} \mid |z - w_0| < r\}$ is itself open. A ***deleted*** r ***neighborhood*** is an r neighborhood whose center point has been removed. Thus a deleted r-neighborhood has the form $D(z_0; r)\backslash\{z_0\}$, which stands for the set $D(z_0; r)$ minus the singleton set $\{z_0\}$. See Figure 1.4.2.

A ***neighborhood*** of a point z_0 is, by definition, a set containing some r disk around z_0. Notice that a set A *is open iff for each* z_0 *in* A, *there is an* r *neighborhood of* z_0 *wholly contained in* A.

The basic properties of open sets are collected in the next proposition.

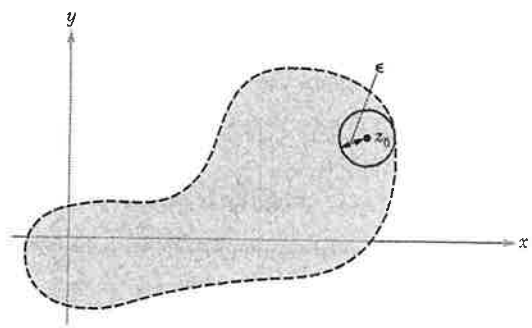

Figure 1.4.1: Open set.

Figure 1.4.2: (a) r-Neighborhood. (b) Deleted r-neighborhood.

Proposition 1.4.1 (i) *$\mathbb{C}$ is open.*

(ii) *The empty set $\varnothing$ is open.*

(iii) *The union of any collection of open subsets of $\mathbb{C}$ is open.*

(iv) *The intersection of any finite collection of open subsets of $\mathbb{C}$ is open.*

Proof The first two assertions hold almost by definition; the first because any ϵ will work for any point z_0, and the second because there are no points for which we are required to find such an ϵ. The reader is asked to supply proofs of the last two in Exercises 19 and 20 at the end of this section. ■

Mappings, Limits, and Continuity Let A be a subset of $\mathbb{C}$. Recall that a mapping $f : A \to \mathbb{C}$ is an assignment of a specific point $f(z)$ in $\mathbb{C}$ to each point z in A. The set A is called the ***domain*** of f, and we say f is ***defined on*** A. When the domain and the range (the set of values f assumes) are both subsets of $\mathbb{C}$, as here, we speak of f as a ***complex function of a complex variable.*** Alternatively, we can think of f as a map $f : A \subset \mathbb{R}^2 \to \mathbb{R}^2$; then f is called

a vector-valued function of two real variables. For $f : A \subset \mathbb{C} \to \mathbb{C}$, we can let $z = x + iy = (x, y)$ and define $u(x, y) = \operatorname{Re} f(z)$ and $v(x, y) = \operatorname{Im} f(z)$. Then u and v are the components of f thought of as a vector function. Hence we may uniquely write $f(x + iy) = u(x, y) + iv(x, y)$, where u and v are real-valued functions defined on A.

Next we consider the idea of limit in the setting of complex numbers.

Definition 1.4.2 *Let f be defined on a set containing some deleted r neighborhood of z_0. We say that f has the* ***limit*** *a as $z \to z_0$ and write*

$$\lim_{z \to z_0} f(z) = a,$$

when, for every $\epsilon > 0$, there is a $\delta > 0$ such that for all $z \in D(z_0; r)$ satisfying $z \neq z_0$ and $|z - z_0| < \delta$, we have $|f(z) - a| < \epsilon$.

The expression in this definition has the same intuitive meaning as it has in calculus; namely, $f(z)$ is close to a whenever z is close to z_0. It is not necessary to define f on a *whole* deleted neighborhood to have a valid theory of limits, but deleted neighborhoods are used here for the sake of simplicity and also because such usage will be appropriate later in the text.

Just as with real numbers and real-valued functions, a function can have no more than one limit at a point, and limits behave well with respect to algebraic operations. This is the content of the next two propositions.

Proposition 1.4.3 *Limits are unique if they exist.*

Proof Suppose that $\lim_{z \to z_0} f(z) = a$ and $\lim_{z \to z_0} f(z) = b$ with $a \neq b$. Let $2\epsilon = |a - b|$, so that $\epsilon > 0$. There is a $\delta > 0$ such that $0 < |z - z_0| < \delta$ implies that $|f(z) - a| < \epsilon$ and $|f(z) - b| < \epsilon$. Choose such a point $z \neq z_0$ (because f is defined in a deleted neighborhood of z_0). Then, by the triangle inequality, $|a - b| \leq |a - f(z)| + |f(z) - b| < 2\epsilon$, a contradiction. Thus $a = b$. ■

Proposition 1.4.4 *If $\lim_{z \to z_0} f(z) = a$ and $\lim_{z \to z_0} g(z) = b$, then*

(i) $\lim_{z \to z_0}[f(z) + g(z)] = a + b$.

(ii) $\lim_{z \to z_0}[f(z)g(z)] = ab$.

(iii) $\lim_{z \to z_0}[f(z)/g(z)] = a/b$ *if* $b \neq 0$.

Proof Only assertion (ii) will be proved here. The proof of assertion (i) is easy, and proof of assertion (iii) is slightly more challenging, but the reader can get the necessary clues from the corresponding real-variable case. To prove assertion (ii), we write

$$\begin{aligned} |f(z)g(z) - ab| &\leq |f(z)g(z) - f(z)b| + |f(z)b - ab| \quad \text{(triangle inequality)} \\ &= |f(z)||g(z) - b| + |f(z) - a||b| \quad \text{(factoring).} \end{aligned}$$

To estimate each term, we choose $\delta_1 > 0$ so that $0 < |z - z_0| < \delta_1$ implies that $|f(z) - a| < 1$, and thus $|f(z)| < |a| + 1$, since $|f(z) - a| \geq |f(z)| - |a|$, by Proposition 1.2.5(vi). Given $\epsilon > 0$, select positive numbers δ_2 and δ_3 so that $0 < |z - z_o| < \delta_2$ implies $|f(z) - a| < \epsilon/2(|b| + 1)$ and $0 < |z - z_o| < \delta_3$ implies $|g(z) - b| < \epsilon/2(|a| + 1)$. Let δ be the smallest of δ_1, δ_2, δ_3. If $0 < |z - z_o| < \delta$, we have

$$\begin{aligned}|f(z)g(z) - ab| &\leq |f(z)|\ |g(z) - b| + |f(z) - a|\ |b| \\ &< \frac{\epsilon}{2(|a|+1)}\ |f(z)| + \frac{\epsilon}{2(|b|+1)}\ |b| \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.\end{aligned}$$

Thus $\lim_{z \to z_o} f(z)g(z) = ab$ as claimed. ■

Definition 1.4.5 *Let $A \subset \mathbb{C}$ be an open set and let $f : A \to \mathbb{C}$ be a function. We say f is* ***continuous at*** *$z_0 \in A$ if and only if*

$$\lim_{z \to z_0} f(z) = f(z_0)$$

and that f is ***continuous on*** *A if f is continuous at each point z_0 in A.*

This definition has the same intuitive meaning as it has in elementary calculus: If z is close to z_0, then $f(z)$ is close to $f(z_0)$. From Proposition 1.4.4 we deduce that if f and g are continuous on A, then so are the sum $f + g$ and the product fg, and so is f/g if $g(z_0) \neq 0$ for all points z_0 in A. It is also true that a composition of continuous functions is continuous.

Proposition 1.4.6 (i) *If $\lim_{z \to z_0} f(z) = a$ and h is a function defined on a neighborhood of a and is continuous at a, then $\lim_{z \to z_0} h(f(z)) = h(a)$.*

(ii) *If f is a continuous function on an open set A in $\mathbb{C}$ and h is continuous on $f(A)$, then the composite function $(h \circ f)(z) = h(f(z))$ is continuous on A.*

Proof Given $\epsilon > 0$, there is a $\delta_1 > 0$ such that $|h(w) - h(a)| < \epsilon$ whenever $|w - a| < \delta_1$ and a $\delta > 0$ such that $|f(z) - a| < \delta_1$ whenever $0 < |z - z_0| < \delta$. Therefore we get $|h(f(z)) - h(a)| < \epsilon$ whenever $0 < |z - z_0| < \delta$, which establishes (i). A proof of (ii) follows from (i) and is requested in Exercise 22 at the end of this section. ■

Sequences The concept of convergent sequences of complex numbers is analogous to that for sequences of real numbers studied in calculus. A sequence $z_n, n = 1, 2, 3, \ldots$ of points of $\mathbb{C}$ ***converges to*** z_0 if and only if for every $\epsilon > 0$, there is an integer N such that $n \geq N$ implies $|z_n - z_0| < \epsilon$. The limit of a sequence is expressed as

$$\lim_{n \to \infty} z_n = z_0 \quad \text{or} \quad z_n \to z_0.$$

Limits of sequences have the same properties, obtained by the same proofs, as limits of functions. For example, the limit is unique if it exists; and if $z_n \to z_0$ and $w_n \to w_0$, then

(i) $z_n + w_n \to z_0 + w_0$.

(ii) $z_n w_n \to z_0 w_0$.

(iii) $z_n / w_n \to z_0 / w_0$ (if w_0 and w_n are not 0).

Also, $z_n \to z_0$ iff $\operatorname{Re} z_n \to \operatorname{Re} z_0$ and $\operatorname{Im} z_n \to \operatorname{Im} z_0$. A proof of this for functions is requested in Exercise 2 at the end of this section.

A sequence z_n is called a ***Cauchy sequence*** if for every $\epsilon > 0$ there is an integer N such that $|z_n - z_m| < \epsilon$ whenever both $n \geq N$ and $m \geq N$. A basic property of real numbers, which we will accept without proof, is that every Cauchy sequence in $\mathbb{R}$ converges. More precisely, if $\{x_n\}_{n=1}^{\infty}$ is a Cauchy sequence of real numbers, then there is a real number x_0 such that $\lim_{n \to \infty} x_n = x_0$. This is equivalent to the ***completeness*** of the real number system.[4] From the fact that $z_n \to z_0$ iff $\operatorname{Re} z_n \to \operatorname{Re} z_0$ and $\operatorname{Im} z_n \to \operatorname{Im} z_0$, we can conclude that *every Cauchy sequence in $\mathbb{C}$ converges*. This is a technical point, but is useful in convergence proofs, as we shall see in Chapter 3.

It should be noted that a link exists between sequences and continuity; namely, $f : A \subset \mathbb{C} \to \mathbb{C}$ is continuous iff for every convergent sequence $z_n \to z_0$ of points in A (that is, $z_n \in A$ and $z_0 \in A$), we have $f(z_n) \to f(z_0)$. The student is requested to prove this in Exercise 18 at the end of this section.

Closed Sets A subset F of $\mathbb{C}$ is said to be ***closed*** if its complement, $\mathbb{C} \backslash F = \{z \in \mathbb{C} \mid z \notin F\}$, is open. By taking complements and using Proposition 1.4.1, one discovers the following properties of closed sets.

Proposition 1.4.7

(i) *The empty set is closed.*

(ii) *$\mathbb{C}$ is closed.*

(iii) *The intersection of any collection of closed subsets of $\mathbb{C}$ is closed.*

(iv) *The union of any finite collection of closed subsets of $\mathbb{C}$ is closed.*

Closed and open sets are important for their relationships to continuous functions and to sequences and for other constructions we will see later.

Proposition 1.4.8 *A set $F \subset \mathbb{C}$ is closed iff whenever $z_1, z_2, z_3, \ldots$ is a sequence of points in F such that $w = \lim_{n \to \infty} z_n$ exists, then $w \in F$.*

---

[4] See, for example, J. Marsden and M. Hoffman, *Elementary Classical Analysis*, Second Edition (New York: W. H. Freeman and Company, 1993).

Proof Suppose F is closed and z_n is a sequence of points in F. If $D(w;r)$ is any disk around w, then by the definition of convergence, z_n is in $D(w;r)$ for large enough n. Thus, $D(w;r)$ cannot be contained in the complement of F. Since that complement is open, w must not be in the complement of F. Therefore, it must be in F.

If F is not closed, then the complement is not open. In other words, there is a point w in $\mathbb{C}\backslash F$ such that no neighborhood of w is contained in $\mathbb{C}\backslash F$. In particular, we may pick points z_n in $F \cap D(w; 1/n)$; this yields a convergent sequence of points of F whose limit is not in F. ∎

Proposition 1.4.9 *If $f : \mathbb{C} \to \mathbb{C}$, the following are equivalent:*

(i) *f is continuous.*

(ii) *The inverse image of every closed set is closed.*

(iii) *The inverse image of every open set is open.*

Proof To show that (i) implies (ii), suppose f is continuous and F is closed. Let $z_1, z_2, z_3, \ldots$ be a sequence of points in $f^{-1}(F)$ and suppose that $z_n \to w$.[5] Since f is continuous, $f(z_n) \to f(w)$. But the points $f(z_n)$ are in the closed set F, and so $f(w)$ is also in F. That is, w is in $f^{-1}(F)$. Proposition 1.4.8 shows that $f^{-1}(F)$ is closed.

To show that (ii) implies (iii), let U be open. Then $F = \mathbb{C}\backslash U$ is closed. If (ii) holds, then $f^{-1}(F)$ is closed. Therefore, $\mathbb{C}\backslash f^{-1}(F) = f^{-1}(\mathbb{C}\backslash F) = f^{-1}(U)$ is open.

To show that (iii) implies (i), fix z_0 and let $\epsilon > 0$. Then z_0 is a member of the open set $f^{-1}(D(f(z_0); \epsilon))$. Hence there is a $\delta > 0$ with

$$D(z_0; \delta) \subset f^{-1}(D(f(z_0); \epsilon)).$$

This says precisely that $|f(z) - f(z_0)| < \epsilon$ whenever $|z - z_0| < \delta$. We thus get exactly the inequality needed to establish continuity. ∎

To handle continuity on a subset of $\mathbb{C}$, it is convenient to introduce the notion of relatively open and closed sets. If $A \subset \mathbb{C}$, a subset B of A is called ***open relative to** A* if $B = A \cap U$ for some open set U. It is said to be ***closed relative to** A* if $B = A \cap F$ for some closed set F. This leads to the following proposition, whose proof is left to the reader.

Proposition 1.4.10 *If $f : A \to \mathbb{C}$, the following are equivalent:*

(i) *f is continuous.*

(ii) *The inverse image of every closed set is closed relative to A.*

(iii) *The inverse image of every open set is open relative to A.*

[5]The set $f^{-1}(F)$, called the ***inverse image of** F* under the map f, is defined to be the collection of all those $z \in \mathbb{C}$ (or more generally, in the domain of f) such that $f(z) \in F$.

Connected Sets This subsection and the next study two important classes of sets which occupy to some extent the place in the theory of complex variables held by intervals and by closed bounded intervals in the theory of functions of a real variable. These are the connected sets and the compact sets.

A connected set should be one that "consists of one piece". This may be approached from a positive point of view—"Any point can be connected to any other"—or from a negative point of view—"The set cannot be split into two parts". This leads to two possible definitions.

Definition 1.4.11 *A set $C \subset \mathbb{C}$ is **path-connected** if for every pair of points a, b in C there is a continuous map $\gamma : [0,1] \to C$ with $\gamma(0) = a$ and $\gamma(1) = b$. We call γ a **path joining** a and b.*

One can often easily tell if a set is path-connected, as is shown in Figure 1.4.3. The negative point of view suggests a slightly different definition.

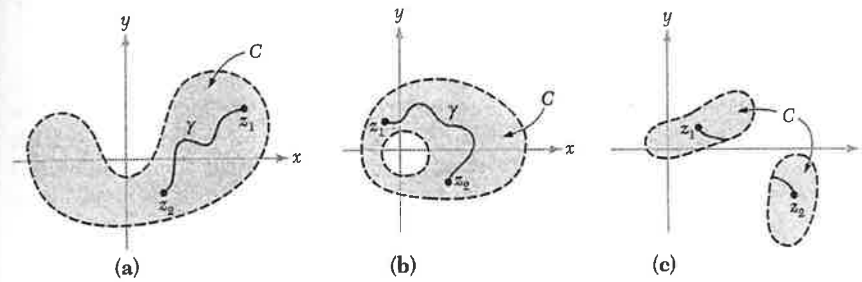

Figure 1.4.3: Regions in (a) and (b) are connected while the region in (c) is not.

Definition 1.4.12 *A set $C \subset \mathbb{C}$ is **not connected** (see Figure 1.4.4) if there are open sets U and V such that*

(i) *$C \subset U \cup V$*

(ii) *$C \cap U \neq \varnothing$ and $C \cap V \neq \varnothing$*

(iii) *$(C \cap U) \cap (C \cap V) = \varnothing$*

*If a set fails to be "not connected", it is called **connected**.*

The notions of relatively open and closed sets allow this to be rephrased in terms of subsets of C. Since the intersection of C with U is the same as its intersection with the complement of V, the set $C \cap U$ is both open and closed relative to C, as is $C \cap V$. This proves the next result.

Proposition 1.4.13 *A set C is connected if and only if the only subsets of C that are both open and closed relative to C are the empty set and C itself.*

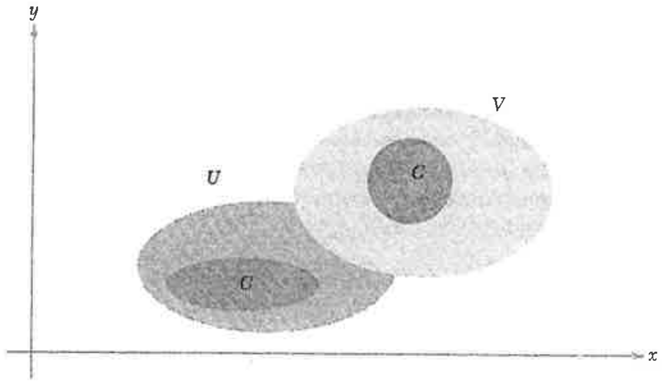

Figure 1.4.4: The set C is not connected.

The next two propositions give the relationship between the two definitions. The two notions are not in general equivalent, but they are for open sets. The proof of this last assertion (given below in Proposition 1.4.15) illustrates a fairly typical way of using the notion of connectivity. One shows that a certain property holds everywhere in C by showing that the set of places where it holds is not empty and is both relatively open and relatively closed.

Proposition 1.4.14 *A path-connected set is connected.*

Proof Suppose C is a path-connected set and D is a nonempty subset of C that is both open and closed relative to C. If $C \neq D$, there is a point z_1 in D and a point z_2 in $C \backslash D$. Let $\gamma : [a, b] \to C$ be a continuous path joining z_1 to z_2. Let $B = \gamma^{-1}(D)$. Then B is a subset of the interval $[a, b]$, since γ is continuous. (See Proposition 1.4.10.) Since a is in B, B is not empty, and $[a, b] \backslash B$ is not empty since it contains b.

This argument shows that it is sufficient to prove the theorem for the case of an interval $[a, b]$. We thus need to establish that intervals on the real line are connected. A proof uses the least upper bound property (or some other characterization of the fact that the system of real numbers is complete). Let $x = \sup B$ (that is, the least upper bound of B). We find that x is in B since B is closed. Since B is open there is a neighborhood of x contained in B (note that $x \neq b$, since b is in $[a, b] \backslash B$). Thus, for some $\epsilon > 0$, the point $x + \epsilon$ is in B. Thus x cannot be the least upper bound. This contradiction shows that such a set B cannot exist. ∎

A connected set need not be path-connected,[6] but if it is open it must be. In fact, more is true.

[6] A standard example is given by letting C be the union of the graph of $y = \sin 1/x$, where $x > 0$, and the line segment $-1 \leq y \leq 1, x = 0$. This set is connected but not path-connected.

Proposition 1.4.15 *If C is an open connected set and a and b are in C, then there is a differentiable path $\gamma : [0, 1] \to C$ with $\gamma(0) = a$ and $\gamma(1) = b$.*[7]

Proof Let a be in C. If z_0 is in C, then since C is open, there is an $\epsilon > 0$ such that the disk $D(z_0; \epsilon)$ is contained in C. By combining a path from a to z_0 with one from z_0 to z that stays in this disk, we see that z_0 can be connected to a by a differentiable path if and only if the same is true for every point z in $D(z_0; \epsilon)$. This shows that both the sets

$$A = \{z \in \mathbb{C} \mid z \text{ can be connected to } a \text{ by a differentiable path}\}$$

and

$$B = \{z \in \mathbb{C} \mid z \text{ cannot be so connected to } a\}$$

are open. Since C is connected, either A or B must be empty. Obviously it must be B. See Figure 1.4.5. ∎

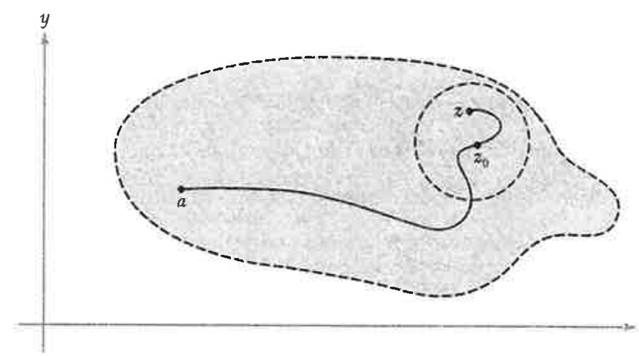

Figure 1.4.5: An open connected set is path-connected.

Because of the importance of open connected sets, they are often designated by a special term. Although the usage is not completely standard in the literature, the words ***region*** and ***domain*** are often used. In this text these terms will be used synonymously to mean an open connected subset of $\mathbb{C}$. The reader should be careful to check the meanings when these words are encountered in other texts.

The notion of connected sets will be of use to us several times. One observation is that a continuous function cannot break apart a connected set.

Proposition 1.4.16 *If f is a continuous function defined on a connected set C, then the image set $f(C)$ is also connected.*

[7] Differentiability of γ means that each component of γ is differentiable in the usual sense of one-variable calculus.

Proof If U and V are open sets that disconnect $f(C)$, then $f^{-1}(U)$ and $f^{-1}(V)$ are open sets disconnecting C. ∎

Be careful. This proposition works in the opposite direction from the one about open and closed sets. For continuous functions, the inverse images of open sets are open and the inverse images of closed sets are closed. But it is the direct images that are guaranteed to be connected and *not* the inverse images of connected sets. (Can you think of an example?) The same sort of thing will happen with the class of sets studied in the next subsection, the compact sets.

Compact Sets The next special class of sets we introduce is that of the compact sets. These will turn out to be those subsets K of $\mathbb{C}$ that are bounded in the sense that there is a number M such that $|z| \leq M$ for every z in K and that are closed. One of the nice properties of such sets is that every sequence of points in the set must have a subsequence which converges to some point in the set. For example, the sequence $1, \frac{1}{2}, \frac{3}{2}, \frac{1}{3}, \frac{5}{3}, \frac{1}{4}, \frac{7}{4}, \frac{1}{5}, \frac{9}{5}, \ldots$ of points in $]0,2[$ has the subsequence $1, \frac{1}{2}, \frac{1}{3}, \ldots$, which converges to the point 0, which is not in the open interval $]0,2[$ but is in the closed interval $[0,2]$. Note that in the claimed property, the sequence itself is not asserted to converge. All that is claimed is that some subsequence does; the example shows that this is necessary.

As often happens in mathematics, the study consists of three parts:

(i) An easily recognized characterization: closed and bounded

(ii) A property we want: the existence of convergent subsequences

(iii) A technical definition useful in proofs and problems

In the case at hand, the technical definition involves the relationship between compactness and open sets. A collection of open sets U_α for α in some index set $\mathcal{A}$ is called a ***cover*** (or an ***open cover***) of a set K if K is contained in their union: $K \subset \cup_{\alpha \in \mathcal{A}} U_\alpha$. For example, the collection of all open disks of radius 2 is an open cover of $\mathbb{C}$:

$$U_z = D(z;2) \qquad \mathbb{C} \subset \cup_{z \in \mathbb{C}} D(z;2).$$

It may be, as here, that the covering process has been wasteful, using more sets than needed. In that case we may use only some of the sets and talk of a ***subcover***, for example, $\mathbb{C} \subset \cup_{n,m \in \mathbb{Z}} D(n+mi;2)$, where $\mathbb{Z}$ denotes the set of integers.

Definition 1.4.17 *A set K is* ***compact*** *if every open cover of K has a finite subcover.*

That is, if U_α is any collection of open sets whose union contains K, then there is a finite subcollection $U_{\alpha_1}, U_{\alpha_2}, \ldots, U_{\alpha_k}$ such that $K \subset U_{\alpha_1} \cup U_{\alpha_2} \cup \ldots \cup U_{\alpha_k}$.

Proposition 1.4.18 *The following conditions are equivalent for a subset K of $\mathbb{C}$ (or of $\mathbb{R}$):*

(i) *K is closed and bounded.*

(ii) *Every sequence of points in K has a subsequence which converges to some point in K.*

(iii) *K is compact.*

This proposition requires a deeper study of the completeness properties of the real numbers than it is necessary for us to go into here, so the proof is omitted. It may be found in most advanced calculus or analysis texts.[8] It is easy to see why (i) is necessary for (ii) and (iii). If K is not bounded we can select z_1 in K and then successively choose z_2 with $|z_2| > |z_1| + 1$ and, in general, z_n with $|z_n| > |z_{n-1}| + 1$. This gives a sequence with no convergent subsequence. The open disks $D(0;n), n = 1,2,3,\ldots$, would be an open cover with no finite subcover.

If K is a set in $\mathbb{C}$ that is not closed, then there is a point w in $\mathbb{C} \backslash K$ and a sequence $z_1, z_2, \ldots$ of points in K that converges to w. Since the sequence converges, w is the only possible limit of a subsequence, so no subsequence can converge to a point of K. The sets $\{z \text{ such that } |z - w| > 1/n\}$ for $n = 1,2,3,\ldots$ form an open cover of K with no finite subcover.

The utility of the technical Definition 1.4.17 is illustrated in the following results.

Proposition 1.4.19 *If K is a compact set and f is a continuous function defined on K, then the image set $f(K)$ is also compact.*

Proof If U_α is an open cover of $f(K)$, then the sets $f^{-1}(U_\alpha)$ form an open cover of K. Selection of a finite subcover gives

$$K \subset f^{-1}(U_{\alpha_1}) \cup \ldots \cup f^{-1}(U_{\alpha_k})$$

so that $f(K) \subset U_{\alpha_1} \cup \ldots \cup U_{\alpha_k}$. ∎

Theorem 1.4.20 (Extreme Value Theorem) *If K is a compact set and $f : K \to \mathbb{R}$ is continuous, then f attains finite maximum and minimum values.*

Proof The image $f(K)$ is compact, hence closed and bounded. Since it is bounded, the numbers $M = \sup\{f(z) \mid z \in K\}$ and $m = \inf\{f(z) \mid z \in K\}$ are finite. Since $f(K)$ is closed, m and M are included in $f(K)$. ∎

Another illustration of the use of compactness is given by the following lemma, which asserts that the distance from a compact set to a closed set is positive. That is, there must be a definite gap between the two sets.

Lemma 1.4.21 (Distance Lemma) *Suppose K is compact, C is closed, and $K \cap C = \varnothing$. Then the distance $d(K,C)$ from K to C is greater than 0. That is, there is a number $\rho > 0$ such that $|z - w| > \rho$ whenever z is in K and w is in C.*

[8] See, for example, J. Marsden and M. Hoffman, *Elementary Classical Analysis*, Second Edition (New York: W. H. Freeman and Company, 1993).

Proof The complement of C, namely the set $U = \mathbb{C}\backslash C$, is an open set and $K \subset U$, so that each point z in K is the center of some disk $D(z;\rho(z)) \subset U$. The collection of smaller disks $D(z;\rho(z)/2)$ also covers K, and by compactness there is a finite number of disks, which we denote by $D_k = D(z_k;\rho(z_k)/2), k = 1,2,3,\ldots,N$ that cover K. (See Figure 1.4.6.) Let $\rho_k = \rho(z_k)/2$ and $\rho = \min(\rho_1,\rho_2,\ldots,\rho_N)$. If z is in K and w is in C, then z is in D_k for some k, and so $|z - z_k| < \rho_k$. But $|w - z_k| > \rho(z_k) = 2\rho_k$. Thus, $|z - w| > \rho_k \geq \rho$. ■

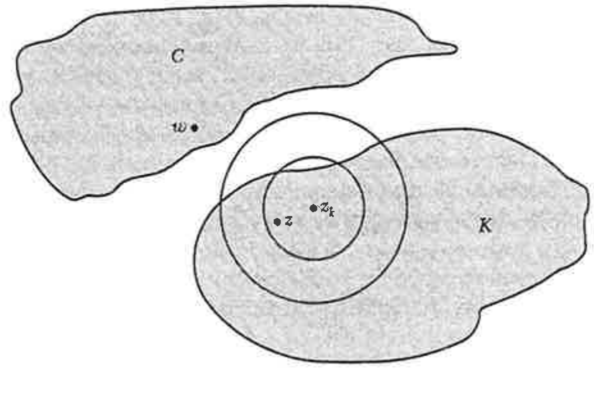

Figure 1.4.6: The distance between a closed set C and a compact set K is greater than zero.

Uniform Continuity Remember that a function is said to be continuous on a set K if it is continuous at each point of K. This is called a ***local property*** since it is defined in terms of the behavior of the function at or near each point and can be determined for each point by looking only near the point and not at the whole set at once. This is in contrast to ***global properties*** of a function, which depend on its behavior on the whole set.

An example of a global property is boundedness. Saying that a function f is bounded by some number M on a set K is an assertion that depends on the whole set at once. If the function is continuous it is certainly bounded near each point, but that would not automatically say that it is bounded on the whole set. For example, the function $f(x) = 1/x$ is continuous on the open interval $]0,1[$ but is certainly not bounded there. We have seen that if a function f is continuous on a compact set K, then it is bounded on K and in fact the bounds are attained. Thus, compactness of K allowed us to carry the local boundedness near each point given by continuity over to the whole set. Compactness often can be used to make such a shift from a local property to a global one. The following is a global version of the notion of continuity.

Definition 1.4.22 *A function $f : A \to \mathbb{C}$ (or $\mathbb{R}$) is **uniformly continuous** on A*

if for every choice of $\epsilon > 0$ there is a $\delta > 0$ such that $|f(s) - f(t)| < \epsilon$ whenever s and t are in A and $|s - t| < \delta$.

Notice that the difference between this and the definition of ordinary continuity is that now the choice of δ can be made so that the same δ will work everywhere in the set A. Obviously, uniformly continuous functions are continuous. On a compact set the opposite is true as well.

Proposition 1.4.23 *A continuous function on a compact set is uniformly continuous.*

Proof Suppose f is a continuous function on a compact set K, and let $\epsilon > 0$. For each point t in K, there is a number $\delta(t)$ such that $|f(s) - f(t)| < \epsilon/2$ whenever $|s - t| < \delta(t)$. The open sets $D(t;\delta(t)/2)$ cover K, so by compactness there are a finite number of points $t_1, t_2, \ldots, t_N$ such that the sets $D_k = D(t_k;\delta(t_k)/2)$ cover K. Let $\delta_k = \delta(t_k)/2$ and set δ equal to the minimum of $\delta_1, \delta_2, \ldots, \delta_N$. If $|s-t| < \delta$, then t is in D_k for some k, and so $|t - t_k| < \delta_k$. Thus $|f(t) - f(t_k)| < \epsilon/2$. But also,

$$|s - t_k| = |s - t + t - t_k| \leq |s - t| + |t - t_k| \leq \delta + \delta_k \leq \delta(t_k)$$

and so $|f(s) - f(t_k)| \leq \epsilon/2$. Thus

$$\begin{aligned} |f(s) - f(t)| &= |f(s) - f(t_k) + f(t_k) - f(t)| \\ &\leq |f(s) - f(t_k)| + |f(t_k) - f(t)| < \epsilon/2 + \epsilon/2 = \epsilon. \end{aligned}$$

We have produced a single δ that works everywhere in K, and so f is uniformly continuous. ■

Path-Covering Lemma The notion of uniform continuity is a very powerful one that will be useful to us several times. We use it first in conjunction with the Distance Lemma and some of the properties of compact sets to establish a useful geometric lemma about curves in open subsets of the complex plane. This lemma will be useful later in the text, particularly for studying integrals along such curves. It says that the curve can be covered by a finite number of disks centered along the curve in such a way that each disk is contained in the open set and each contains the centers of both the preceding and the succeeding disks along the curve. (See Figure 1.4.7.)

Lemma 1.4.24 (Path-Covering Lemma) *Suppose $\gamma : [a,b] \to G$ is a continuous path from the interval $[a,b]$ into an open subset G of $\mathbb{C}$. Then there are a number $\rho > 0$ and a subdivision of the interval $a = t_0 < t_1 < t_2 < \ldots < t_n = b$ such that*

(i) $D(\gamma(t_k);\rho) \subset G$ *for all k*

(ii) $\gamma(t) \in D(\gamma(t_0);\rho)$ *for $t_0 \leq t \leq t_1$*

(iii) $\gamma(t) \in D(\gamma(t_k);\rho)$ *for $t_{k-1} \leq t \leq t_{k+1}$*

(iv) $\gamma(t) \in D(\gamma(t_n);\rho)$ *for $t_{n-1} \leq t \leq t_n$*

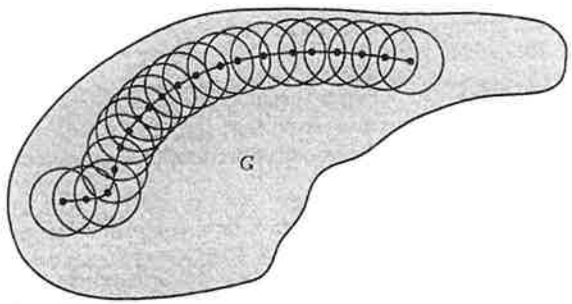

Figure 1.4.7: A continuous path in an open set can be covered by a finite number of well-overlapping disks.

Proof Since γ is continuous and the closed interval $[a, b]$ is compact, the image curve $K = \gamma([a, b])$ is compact. By the Distance Lemma 1.4.21 there is a number ρ such that each point on the curve is a distance at least ρ from the complement of G. Therefore, $D(\gamma(t); \rho) \subset G$ for every t in $[a, b]$. Also, since γ is continuous on the compact set $[a, b]$, it is uniformly continuous, and there is a number $\delta > 0$ such that $|\gamma(t) - \gamma(s)| < \rho$ whenever $|s - t| < \delta$. Thus if the subdivision is chosen fine enough so that $t_k - t_{k-1} < \delta$ for all $k = 1, 2, 3, \ldots, N$, then the conclusions of the theorem hold. ■

Riemann Sphere and Point at Infinity For some purposes it is convenient to introduce a point ∞ in addition to the points $z \in \mathbb{C}$. One must be careful in doing so, since it can lead to confusion and abuse of the symbol ∞. But with care it can be useful, and we certainly want to be able to talk intelligently about infinite limits and limits at infinity.

In contrast to the real line, to which $+\infty$ and $-\infty$ can be added, we have only one ∞ for $\mathbb{C}$. The reason is that $\mathbb{C}$ has no natural ordering as $\mathbb{R}$ does. Formally we add a symbol ∞ to $\mathbb{C}$ to obtain the *extended complex plane*, $\mathbb{C}$, and define operations with ∞ by the rules

$$\begin{aligned} z + \infty &= \infty \\ z \cdot \infty &= \infty \qquad \text{provided } z \neq 0 \\ \infty + \infty &= \infty \\ \infty \cdot \infty &= \infty \\ \frac{z}{\infty} &= 0 \end{aligned}$$

for $z \in \mathbb{C}$. Notice that some things are not defined: $\infty/\infty, 0 \cdot \infty, \infty - \infty$, and so forth are ***indeterminate forms*** for essentially the same reasons that they are in the calculus of real numbers. We also define appropriate limit concepts:

$\lim_{z\to\infty} f(z) = z_0$ *means: For any* $\epsilon > 0$, *there is an* $R > 0$ *such that* $|f(z) - z_0| < \epsilon$ *whenever* $|z| \geq R$.

$\lim_{z\to z_0} f(z) = \infty$ *means: For any* $R > 0$, *there is a* $\delta > 0$ *such that* $|f(z)| > R$ *whenever* $|z - z_0| < \delta$.

For sequences:

$\lim_{z\to\infty} z_n = \infty$ *means: For any* $R > 0$, *there is an* $N > 0$ *such that* $|z_n| > R$ *whenever* $n > N$.

Thus a point $z \in \mathbb{C}$ is "close to ∞" when it lies outside a large circle. This type of closeness can be pictured geometrically by means of the *Riemann sphere* shown in Figure 1.4.8. By the method of *stereographic projection* illustrated in this figure, a point z' on the sphere is associated with each point z in $\mathbb{C}$. Exactly one point on the sphere $\mathbb{S}$ has been omitted—the "north" pole. We assign ∞ in $\bar{\mathbb{C}}$ to the north pole of $\mathbb{S}$. We see geometrically that z is close to ∞ if and only if the corresponding points are close on the Riemann sphere in the usual sense of closeness in $\mathbb{R}^3$. Proof of this is requested in Exercise 24.

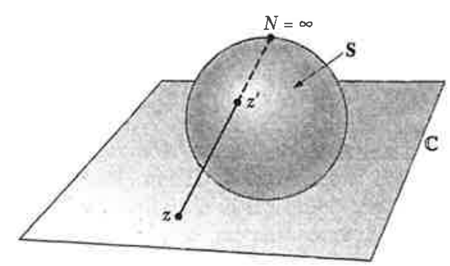

Figure 1.4.8: Riemann sphere.

The Riemann sphere $\mathbb{S}$ represents a convenient geometric picture of the extended plane $\bar{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$. The sphere does point up one fact about the extended plane that is sometimes useful in further theory. Since $\mathbb{S}$ is a closed bounded subset of $\mathbb{R}^3$, it is compact. Therefore every sequence in it has a convergent subsequence. Since stereographic projection makes convergence on $\mathbb{S}$ coincide with convergence of the sequence of corresponding points in $\bar{\mathbb{C}}$, the same is true there. That is, $\bar{\mathbb{C}}$ is compact. Every sequence of points in $\bar{\mathbb{C}}$ must have a subsequence convergent in $\bar{\mathbb{C}}$. *Caution*: Since the convergence is in the extended plane, the limit might be ∞, in that case we would normally say that the limit does not exist. Basically we have thrown in the point at infinity as another available limit so that sequences that did not formerly have a limit now have one. The sphere can be used both to help visualize and to make precise some notions about the behavior of functions "at infinity" that we will meet in future chapters.

Worked Examples

Example 1.4.25 *Where is the function*

$$f(z) = \frac{z^3 + 2z + 1}{z^3 + 1}$$

continuous?

Solution Since sums, products, and quotients of continuous functions are continuous except where the denominator is 0, this function is continuous on the whole plane except at the cube roots of -1. In other words, this function is continuous on the set $\mathbb{C}\backslash\{e^{\pi i/3}, e^{5\pi i/3}, -1\}$.

Example 1.4.26 *Show that the set* $\{z \mid \operatorname{Re} z > 0\}$ *is open.*

Solution A proof can be based on the following properties of complex numbers (see Exercise 1): *If* $w \in \mathbb{C}$, *then*

(i) $|\operatorname{Re} w| \leq |w|$

(ii) $|\operatorname{Im} w| \leq |w|$

(iii) $|w| \leq |\operatorname{Re} w| + |\operatorname{Im} w|$

Let $U = \{z \mid \operatorname{Re} z > 0\}$ and let z_0 be in U. We claim that the disk $D(z_0; \operatorname{Re} z_0)$ lies in U. To see this, let z be in this disk. Then $|\operatorname{Re} z - \operatorname{Re} z_0| = |\operatorname{Re}(z - z_0)| \leq |z - z_0| < \operatorname{Re} z_0$, and so $\operatorname{Re} z > 0$ and z is in U. Thus, $D(z_0; \operatorname{Re} z_0)$ is a neighborhood of z_0 that is contained in U. Since this can be done for any point z_0 which is in U, the set U is open. See Figure 1.4.9.

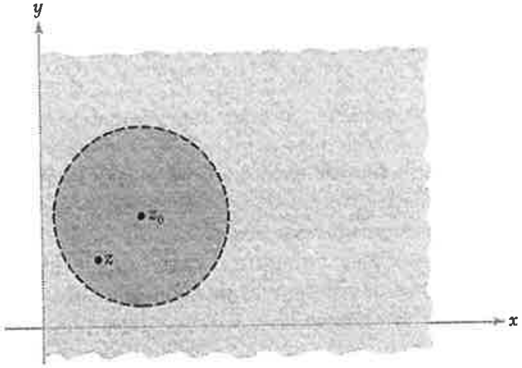

Figure 1.4.9: Open right half-plane.

Example 1.4.27 *Prove the following statement: Let* $A \subset \mathbb{C}$ *be an open set and* $z_0 \in A$, *and suppose that* $D_r = \{z \text{ such that } |z - z_0| \leq r\} \subset A$. *Then there is a number* $\rho > r$ *such that* $D(z_0; \rho) \subset A$.

Solution We know from the Extreme Value Theorem 1.4.20 that a continuous real-valued function on a closed bounded set in $\mathbb{C}$ attains its maximum and minimum at some point of the set. For z in D_r let $f(z) = \inf\{|z - w| \text{ such that } w \in \mathbb{C}\backslash A\}$. (Here "inf" means the greatest lower bound.) In other words, $f(z)$ is the distance from z to the complement of A. Since A is open, $f(z) > 0$ for each z in D_r. We can also verify that f is continuous. Thus f assumes its minimum at some point z_1 in D_r. Let $\rho = f(z_1) + r$, and check that this ρ has the desired properties. See Figure 1.4.10.

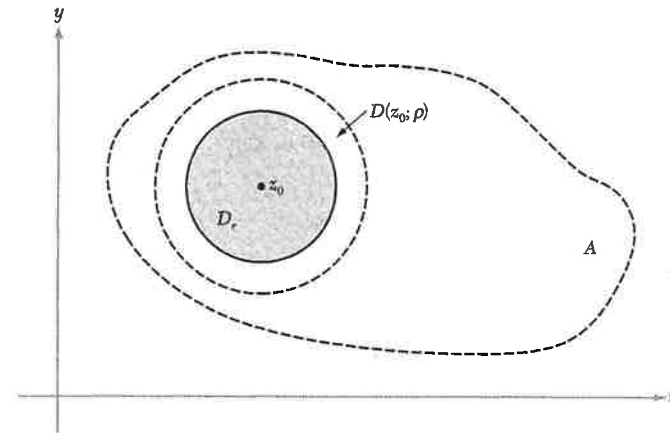

Figure 1.4.10: A closed disk in an open set may be enlarged.

Example 1.4.28 *Find* $\lim_{z\to\infty} \frac{3z^4 + 2z^2 - z + 1}{z^4 + 1}$.

Solution

$$\lim_{z\to\infty} \frac{3z^4 + 2z^2 - z + 1}{z^4 + 1} = \lim_{z\to\infty} \frac{3 + 2z^{-2} - z^{-3} + z^{-4}}{1 + z^{-4}} = 3$$

using $\lim_{z\to\infty} z^{-1} = 0$ and the basic properties of limits.

Exercises

(1) Show that if $w \in \mathbb{C}$, then

(a) $|\operatorname{Re} w| \le |w|$
(b) $|\operatorname{Im} w| \le |w|$
(c) $|w| \le |\operatorname{Re} w| + |\operatorname{Im} w|$

2.• (a) Show that

$$|\operatorname{Re} z_1 - \operatorname{Re} z_2| \le |z_1 - z_2| \le |\operatorname{Re} z_1 - \operatorname{Re} z_2| + |\operatorname{Im} z_1 - \operatorname{Im} z_2|$$

for any two complex numbers z_1 and z_2.
(b) If $f(z) = u(x,y) + iv(x,y)$, show that

$$\lim_{z\to z_0} f(z) = \lim_{\substack{x\to x_0\\ y\to y_0}} u(x,y) + i \lim_{\substack{x\to x_0\\ y\to y_0}} v(x,y)$$

exists if both limits on the right of the equation exist. Conversely, if the limit on the left exists, show that both limits on the right exist as well and equality holds. Show that $f(z)$ is continuous iff u and v are.

3. Prove: If f is continuous and $f(z_0) \ne 0$, there is a neighborhood of z_0 on which f is $\ne 0$.

4. If $z_0 \in \mathbb{C}$, show that the set $\{z_0\}$ is closed.

5. Prove: The complement of a finite number of points is an open set.

6. Use the fact that a function is continuous if and only if the inverse image of every open set is open to show that a composition of two continuous functions is continuous.

7. Show that $f(z) = \bar{z}$ is continuous.

8. Show that $f(z) = |z|$ is continuous.

9. What is the largest set on which the function $f(z) = 1/(1 - e^z)$ is continuous?

10. Prove or find a counterexample if false: If $\lim_{z\to z_0} f(z) = a$, h is defined at the points $f(z)$, and $\lim_{w\to a} h(w) = c$, then $\lim_{z\to z_0} h(f(z)) = c$. [*Hint:* Could we have $h(a) \ne c$?]

11. For what z does the sequence $z_n = nz^n$ converge?

12.• Define $f : \mathbb{C} \to \mathbb{C}$ by setting $f(0) = 0$ and by setting $f(r[\cos\theta + i\sin\theta]) = \sin\theta$ if $r > 0$. Show that f is discontinuous at 0 but is continuous everywhere else.

13. For each of the following sets, state (i) whether or not it is open and (ii) whether or not it is closed.

(a) $\{z \text{ such that } |z| < 1\}$
(b) $\{z \mid 0 < |z| \le 1\}$

(c) $\{z \mid 1 \le \operatorname{Re} z \le 2\}$

14. For each of the following sets, state (i) whether or not it is open and (ii) whether or not it is closed.

(a) $\{z \mid \operatorname{Im} z > 2\}$
(b) $\{z \mid 1 \le |z| \le 2\}$
(c) $\{z \mid -1 < \operatorname{Re} z \le 2\}$

15. For each of the following sets, state (i) whether or not it is connected and (ii) whether or not it is compact.

(a) $\{z \mid 1 \le |z| \le 2\}$
(b) $\{z \text{ such that } |z| \le 3 \text{ and } |\operatorname{Re} z| \ge 1\}$
(c) $\{z \text{ such that } |\operatorname{Re} z| \le 1\}$
(d) $\{z \text{ such that } |\operatorname{Re} z| \ge 1\}$

16. For each of the following sets, state (i) whether or not it is connected and (ii) whether or not it is compact.

(a) $\{z \mid 1 < \operatorname{Re} z \le 2\}$
(b) $\{z \mid 2 \le |z| \le 3\}$
(c) $\{z \text{ such that } |z| \le 5 \text{ and } |\operatorname{Im} z| \ge 1\}$

17. If $A \subset \mathbb{C}$ and $f : \mathbb{C} \to \mathbb{C}$, show that $\mathbb{C}\backslash f^{-1}(A) = f^{-1}(\mathbb{C}\backslash A)$.

18. Show that $f : A \subset \mathbb{C} \to \mathbb{C}$ is continuous if and only if $z_n \to z_0$ in A implies that $f(z_n) \to f(z_0)$.

19. Show that the union of any collection of open subsets of $\mathbb{C}$ is open.

20. Show that the intersection of any finite collection of open subsets of $\mathbb{C}$ is open.

21. Give an example to show that the statement in Exercise 20 is false if the word "finite" is omitted.

22. Prove part (ii) of Proposition 1.4.6 by using part (i).

23. Show that if $|z| > 1$, then $\lim_{n\to\infty}(z^n/n) = \infty$.

24. Introduce the ***chordal metric*** ρ on $\bar{\mathbb{C}}$ by setting $\rho(z_1, z_2) = d(z_1', z_2')$ where z_1' and z_2' are the corresponding points on the Riemann sphere and d is the usual distance between points in $\mathbb{R}^3$.

(a) Show that $z_n \to z$ in $\mathbb{C}$ if and only if $\rho(z_n, z) \to 0$.
(b) Show that $z_n \to \infty$ if and only if $\rho(z_n, \infty) \to 0$.
(c) If $f(z) = (az + b)/(cz + d)$ and $ad - bc \ne 0$, show that f is continuous at ∞.