Chapter 6

MORE ON CONTINUOUS FUNCTIONS AND SETS

6.1 Introduction

In this chapter[1] we go much more deeply into the analysis of continuous functions. For this we need some new set theoretic ideas and methods.

6.2 Dense Sets

Consider[2] the set $\mathbb{Q}$ of rational numbers and let (a, b) be an open interval in $\mathbb{R}$. How do we show that there is a member of $\mathbb{Q}$ in the interval (a, b), i.e., that $(a, b) \cap \mathbb{Q} \neq \emptyset$?

Suppose first that $0 < a$. Since $b - a > 0$, the Archimedean Property (Theorem 1.9) implies the existence of a positive integer q such that $q(b - a) > 1$, so that $qb > 1 + qa$. The Archimedean Property also implies that $\{m \in \mathbb{N} : m > qa\}$ is non-empty, thus according to the well-ordering principle, there exists $p \in \mathbb{N}$ such that $p - 1 \leq qa < p$. It follows that $qa < p \leq 1 + qa < qb$, which implies $a < \frac{p}{q} < b$. We have shown that, under the assumption $a > 0$, there exists a rational number $r = p/q$ in the interval (a, b).

[1] This chapter may be skipped and the reader can proceed directly to the study of derivatives and integrals in Chapters 7 and 8.

[2] This section reviews material from Section 1.9.

The same is true under the assumption $a < 0$. To see this observe first that if $a < 0 < b$, we can take $r = 0$. If $a < b < 0$, then $0 < -b < -a$, so the argument of the previous paragraph shows that there exists $r \in \mathbb{Q}$ such that $-b < r < -a$. In this case $a < -r < b$.

The preceding discussion proves that every open interval contains a rational number. One often expresses this fact by saying that the set of rational numbets is a *dense* set.

Definition 6.1 A set of real numbers A is said to be *dense* (in $\mathbb{R}$) if for each open interval (a, b) the set $A \cap (a, b)$ is nonempty.

It is important to have a more general concept, that of a set A being dense in a set B.

Definition 6.2 Let A and B be subsets of $\mathbb{R}$. If every open interval that intersects B also intersects A, we say that A *is dense in* B.

Thus Definition 6.1 states the special case of Definition 6.2 that occurs when $B = \mathbb{R}$. We should note that some authors require that $A \subset B$ in their version of Definition 6.2. We find it more convenient not to impose this restriction. Thus, for example, in *our* language $\mathbb{Q}$ is dense in $\mathbb{R} \setminus \mathbb{Q}$.

It is easy to verify that A is dense in B if and only if $\overline{A} \supset B$ (Exercise 6:2.1).

Exercises

6:2.1 Verify that A is dense in B if and only if $\overline{A} \supset B$.

6:2.2 Prove that every set A is dense in its closure $\overline{A}$.

6:2.3 Prove that if A is dense in B and $C \subset B$ then A is dense in C.

6:2.4 Prove that if $A \subset B$ and A is dense in B then $\overline{A} = \overline{B}$. Is the statement correct without the assumption that $A \subset B$?

6:2.5 Is $\mathbb{R} \setminus \mathbb{Q}$ dense in $\mathbb{Q}$?

6:2.6 Below are several pairs (A, B) of sets. In each case determine whether A is dense in B.

(a) $A = \mathbb{N}, B = \mathbb{N}$.

(b) $A = \mathbb{N}, B = \mathbb{Z}$.

(c) $A = \mathbb{N}, B = \mathbb{Q}$.

(d) $A = \left\{x : x = \frac{m}{2^n},\ m \in \mathbb{Z}, n \in \mathbb{N}\right\}, B = \mathbb{Q}$.

6:2.7 Let A and B be subsets of $\mathbb{R}$. Prove that A is dense in B if and only if for every $b \in B$ there exists a sequence $\{a_n\}$ of points from A such that $\lim_{n\to\infty} a_n = b$.

6:2.8 Let B be the set of all irrational numbers. Prove that the set $A = \{q + \sqrt{2} : q \in \mathbb{Q}\}$ is a countable subset of B that is dense in B.

6:2.9 Prove that every subset B of $\mathbb{R}$ has a countable subset A that is dense in B.

6:2.10 Let $f : \mathbb{R} \to \mathbb{R}$ be a strictly increasing continuous function. Does f map dense sets to dense sets, i.e., is it true that $f(E) = \{f(x) : x \in E\}$ is dense if E is dense?

6.3 Nowhere Dense Sets

One might view a set A that is dense in $\mathbb{R}$ as being somehow large—inside every interval, no matter how small, one finds points of A. There is an opposite extreme to this situation: a set is said to be *nowhere dense*, and hence is in some sense very small, if it is not dense in any interval at all. The precise definition of this important concept of smallness follows.

Definition 6.3 The set $A \subset \mathbb{R}$ is said to be *nowhere dense* in $\mathbb{R}$ provided every open interval I contains an open subinterval J such that $A \cap J = \emptyset$.

We can state this another way: A is nowhere dense provided $\overline{A}$ contains no open intervals. (See Exercise 6:3.4.)

Example 6.4 It is easy to construct examples of nowhere dense sets. Each of the sets below is nowhere dense as the reader can verify.

1. Any finite set.
2. $\mathbb{N}$.
3. $\{1/n : n \in \mathbb{N}\}$.

◀

Each of the sets in Example 6.4 is countable and hence also small in the sense of cardinality. It is hard to imagine an uncountable set that is nowhere dense but, as we shall see in Section 6.5, such sets do exist.

We establish a simple result showing that any finite union of nowhere dense sets is again nowhere dense. It is not true that a countable union of nowhere dense sets is again nowhere dense. Indeed countable unions of nowhere dense sets will be very important in our subsequent study.

Theorem 6.5 *Let $A_1, A_2, \dots, A_n$ be nowhere dense in $\mathbb{R}$. Then $A_1 \cup \cdots \cup A_n$ is also nowhere dense in $\mathbb{R}$.*

Proof. Let I be any open interval in $\mathbb{R}$. We seek an open interval $J \subset I$ such that $J \cap A_i = \emptyset$ for $i = 1, 2, \dots, n$.

Since A_1 is nowhere dense, there exists an open interval $I_1 \subset I$ such that $I_1 \cap A_1 = \emptyset$. Now A_2 is also nowhere dense in $\mathbb{R}$, so there exists an open interval $I_2 \subset I_1$ such that $A_2 \cap I_2 = \emptyset$. Proceeding in this way we obtain open intervals $I_1 \supset I_2 \supset I_3 \cdots \supset I_n$ such that for $i = 1, ..., n$, $A_i \cap I_i = \emptyset$. It follows from the fact that $I_n \subset I_i$ for $i = 1, \dots, n$ that $A_i \cap I_n = \emptyset$ for $i = 1, \dots, n$. Thus

$$\left(\bigcup_{i=1}^{n} A_i\right) \cap I_n = \bigcup_{i=1}^{n}(A_i \cap I_n) = \bigcup_{i=1}^{n} \emptyset = \emptyset,$$

as was to be proved. ■

Exercises

6:3.1 Give an example of a sequence of nowhere dense sets whose union is not nowhere dense.

6:3.2 Which of the following statements are true?

(a) Every subset of a nowhere dense set is nowhere dense.

(b) If A is nowhere dense then so too is $A + c = \{t + c : t \in A\}$ for every number c.

(c) If A is nowhere dense then so too is $cA = \{ct : t \in A\}$ for every positive number c.

(d) If A is nowhere dense then so too is A', the set of derived points of A.

(e) A nowhere dense set can have no interior points.

(f) A set that has no no interior points must be nowhere dense.

(g) Every point in a nowhere dense set must be isolated.

(h) If every point in a set is isolated then that set must be nowhere dense.

6:3.3 If A is nowhere dense what can you say about $\mathbb{R} \setminus A$? If A is dense what can you say about $\mathbb{R} \setminus A$?

6:3.4◇ Prove that a set $A \subset \mathbb{R}$ is nowhere dense if and only if $\overline{A}$ contains no intervals; equivalently, the interior of $\overline{A}$ is empty.

6:3.5 What should the statement "A is nowhere dense in the interval I" mean? Give an example of a set that is nowhere dense in $[0, 1]$ but is not nowhere dense in $\mathbb{R}$.

6:3.6◇ Let A and B be subsets of $\mathbb{R}$. What should the statement "A is nowhere dense in the B" mean? Is $\mathbb{N}$ nowhere dense in $[0, 10]$? Is $\mathbb{N}$ nowhere dense in $\mathbb{Z}$? Is $\{4\}$ nowhere dense in $\mathbb{N}$?

6:3.7 Prove that the complement of a dense open subset of $\mathbb{R}$ is nowhere dense in $\mathbb{R}$.

6:3.8 Let $f : \mathbb{R} \to \mathbb{R}$ be a strictly increasing continuous function. Show that f maps nowhere dense sets to nowhere dense sets, i.e., that $f(E) = \{f(x) : x \in E\}$ is nowhere dense if E is nowhere dense.

✂ 6.4 The Baire Category Theorem

In this section we shall establish the Baire Category theorem, that gives a sense in which nowhere dense sets can be viewed as "small": a union of a sequence of nowhere dense sets cannot fill up an interval. This can be considered an important generalization of Cantor's theorem which could be interpreted as asserting that a union of a sequence of finite sets cannot fill up an interval.

We motivate this important theorem by way of a game idea that is due to Stefan Banach (1892–1945) and Stanislaw Mazur (1905–1981). Although the origins of the theorem are due to René Baire after whom the theorem is named, the game approach helps us see why the Baire Category theorem might be true. This Banach-Mazur game is just one of many mathematical games that are used throughout mathematics to develop interesting concepts.

✂ 6.4.1 A Two-player Game

We introduce the Baire Category theorem via a game between two players (A) and (B).

Player (A) is given a subset A of $\mathbb{R}$, and player (B) is given the complementary set $B = \mathbb{R} \setminus A$. Player (A) first selects a closed interval $I_1 \subset \mathbb{R}$, then player (B) chooses a closed interval $I_2 \subset I_1$. The players alternate moves, a move consisting of selecting a closed interval inside the previously chosen interval.

The play of the game thus determines a descending sequence of closed intervals

$$I_1 \supset I_2 \supset I_3 \supset \cdots \supset I_n \supset \ldots$$

where player (A) chooses those with odd index and player (B) those with even index. If

$$A \cap \bigcap_{n=1}^{\infty} I_n \neq \emptyset,$$

then player (A) wins; otherwise player (B) wins.

The goal of player (A) is evidently to make sure that the intersection contains a point of A; the goal of player (B) is to ensure that the intersection is empty or contains only points of B. One expects that player (A) should win if his set A is large while player (B) should win if his set is large. It is not, however, immediately clear what "large" might mean for this game.

Example 6.6 If the set A given to player (A) contains an open interval J, then (A) should choose any interval $I_1 \subset J$. No matter how the game continues, player (A) wins. Another way to say this: if the set given to player (B) is not dense, he loses. ◀

Example 6.7 For a more interesting example let player (A) be dealt the "large" set of all irrational numbers, so that player (B) is dealt the rationals. (Both players have been dealt dense sets now.) Let A consist of the irrational numbers. Player (A) can win by following the strategy we now describe. Let q_1, q_2, q_3, ... be a listing of all of the rational numbers, that is

$$\mathbb{Q} = \{q_1, q_2, q_3, \dots\}.$$

Player (A) chooses the first interval I_1 as any closed interval such that $\{q_1\} \notin I_1$. Inductively, suppose $I_1, I_2, \dots, I_{2n}$ have been chosen according to the rules of the game so that it is now time for player (A) to choose I_{2n+1}. The set $\{q_1, q_2, \dots, q_n\}$ is finite, so there exists a closed interval $I_{2n+1} \subset I_{2n}$ such that

$$I_{2n+1} \cap \{q_1, q_2, \dots, q_n\}$$

is empty. Player (A) chooses such an interval.

Since for each $n \in \mathbb{N}$, $q_n \notin I_{2n+1}$, the set $\bigcap_{n=1}^{\infty} I_n$ contains no rational numbers, but, as a descending sequence of closed intervals, $\bigcap_{n=1}^{\infty} I_n \neq \emptyset$. Thus $A \cap \bigcap_{n=1}^{\infty} I_n \neq \emptyset$, and (A) wins. ◀

In these two examples, using informal language, we can say that player (A) has a *strategy* to win: no matter how player (B) proceeds, player (A) can "answer" each move and win the game.

In both examples player (A) had a clear advantage: the set A was larger than the set B. But in what sense is it larger? It is not the fact that A is uncountable while B is countable that matters here. It is something else—the fact that given an interval I_{2n}, player (A) can choose I_{2n+1} inside I_{2n} in such a way that I_{2n+1} misses the set $\{q_1, q_2, \dots, q_n\}$.

Let us try to see in the second example a general strategy that should work for player (A) in some cases. The set B was the union of the singleton sets $\{q_n\}$. Suppose instead that B is the union of a sequence of "small" sets Q_n. Then the same "strategy" will prevail if given any interval J and given any $n \in \mathbb{N}$, there exists an interval $I \subset J$ such that

$$I \cap (Q_1 \cup Q_2 \cup \cdots \cup Q_n) = \emptyset.$$

The set $\bigcap_{n=1}^{\infty} I_n$ will be nonempty, and will miss the set $\bigcup_{n=1}^{\infty} Q_n$. Thus, if $B = \bigcup_{n=1}^{\infty} Q_n$, player (A) has a winning strategy. It is in this sense that the set B is "small". The set A is "large" because the set B is "small". If we look carefully at the requirement on the sets Q_k, we see it is just that each of these sets is nowhere dense in $\mathbb{R}$.

Thus the key to player (A) winning rests on the concept of a nowhere dense set. But note that it rests on the set B being the union of a sequence of nowhere dense sets.

✂ 6.4.2 The Baire Category Theorem

We can formulate our result from our discussion of the game in several ways:

1. $\mathbb{R}$ cannot be expressed as a countable union of nowhere dense sets.

2. The complement of a countable union of nowhere dense sets is dense.

The second of these provides a sense in which countable unions of nowhere dense sets are "small": no matter which countable collection of nowhere dense sets one chooses, their union leaves a dense set uncovered.

To formulate the Baire Category theorem we need some definitions.

Definition 6.8 Let A be a set of real numbers.

1. A is said to be of the *first category* if it can be expressed as a countable union of nowhere dense sets.

2. A is said to be of the *second category* if it is not of the first category.

3. A is said to be *residual* in $\mathbb{R}$ if the complement $\mathbb{R} \setminus A$ is of the first category.

The following properties of first category sets and their complements, the residual sets, are easily proved and left as exercises.

Lemma 6.9 *A union of any sequence of first category sets is again a first category set.*

Lemma 6.10 *An intersection of any sequence of residual sets is again a residual set.*

Theorem 6.11 (Baire Category Theorem) *A residual subset of $I\!R$ is dense in $I\!R$.*

Proof. The discussion in Section 6.4.1 constitutes a proof. Suppose that player (A) is dealt a set $A = X \cap [a, b]$ where X is residual, i.e.,

$$X = \mathbb{R} \setminus \bigcup_{n=1}^{\infty} Q_n$$

with each Q_n nowhere dense. Then player (A) wins by choosing any interval $I_1 \subset [a, b]$ that avoids Q_1 and continues following the strategy of Section 6.4.1. In particular X must contain a point of the interval $[a, b]$, and hence a point of any interval. ■

Theorem 6.11 provides a sense of largeness of sets that is not shared by dense sets in general. The intersection of two dense sets might be empty but the intersection of two, or even countably many, residual sets must still be dense.

Exercises

6:4.1 Show that the union of any sequence of first category sets is again a first category set.

6:4.2 Show that the intersection of any sequence of residual sets is again a residual set.

6:4.3 Rewrite the proof of Theorem 6.11 without using the games language.

6:4.4 Give an example of two dense sets whose intersection is not dense. Does this contradict Theorem 6.11?

6:4.5 Suppose that $\bigcup_{n=1}^{\infty} A_n$ contains some interval (c, d). Show that there is a set, say A_{n_0}, and a subinterval $(c'd') \subset (c, d)$ so that A_{n_0} is dense in $(c'd')$.

6.4.3 Uniform Boundedness

There are many applications of the Baire Category Theorem in analysis. For now, we present just one application, dealing with the concept of *uniform boundedness.* Suppose we have a collection $\mathcal{F}$ of functions defined on $\mathbb{R}$ with the property that for each $x \in \mathbb{R}$, $\{|f(x)| : f \in \mathcal{F}\}$ is bounded. This means that for each $x \in \mathbb{R}$ there exists a number $M_x \geq 0$ such that $|f(x)| \leq M_x$ for all $f \in \mathcal{F}$. We can describe this situation by saying that $\mathcal{F}$ is *pointwise bounded.* Does this imply that the collection is *uniformly bounded*, i.e., that there is a single number M so that $|f(x)| \leq M$ for all $f \in \mathcal{F}$ and every $x \in \mathbb{R}$?

Example 6.12 Let $q_1, q_2, q_3, \ldots$ be an enumeration of $\mathbb{Q}$. For each $n \in \mathbb{N}$ we define a function f_n by $f_n(q_k) = k$ if $n \leq k$, $f_n(x) = 0$ for all other values x. Let $\mathcal{F} = \{f_n : n \in \mathbb{N}\}$. Then if $x \in \mathbb{R} \setminus \mathbb{Q}$, $f(x) = 0$ for all $f \in \mathcal{F}$, and if $x = q_k$, $|f(x)| \leq k$ for all $f \in \mathcal{F}$. Thus, for each $x \in \mathbb{R}$, the set $\{|f(x)| : f \in \mathcal{F}\}$ is bounded. The bounds can be taken to be 0 if $x \in \mathbb{R} \setminus \mathbb{Q}$ ($M_x = 0$ if $x \in \mathbb{R} \setminus \mathbb{Q}$) and we can take $M_{q_k} = k$. But since $\mathbb{Q}$ is dense in $\mathbb{R}$, none of the functions f_n is bounded on any interval. (Verify this.) Thus a collection of functions may be pointwise bounded but not uniformly bounded on any interval. ◀

The functions f_n in Example 6.12 are everywhere discontinuous. Our next theorem shows that if we had taken a collection $\mathcal{F}$ of *continuous* functions, then not only would each $f \in \mathcal{F}$ be bounded on closed intervals (as Theorem 5.44 guarantees), but there would be an interval I on which the entire collection is *uniformly bounded* , i.e. there exists a constant M such that $|f(x)| \leq M$ for all $f \in \mathcal{F}$ and each $x \in I$.

Theorem 6.13 *Let $\mathcal{F}$ be a collection of continuous functions on $\mathbb{R}$ such that for each $x \in \mathbb{R}$ there exists a constant $M_x > 0$ such that $|f(x)| \leq M_x$ for each $f \in \mathcal{F}$. Then there exists an open interval I and a constant $M > 0$ such that $|f(x)| \leq M$ for each $f \in \mathcal{F}$ and $x \in I$.*

Proof. For each $n \in \mathbb{N}$, let $A_n = \{x : |f(x)| \leq n \text{ for all } f \in \mathcal{F}\}$. By hypothesis, $\mathbb{R} = \bigcup_{n=1}^{\infty} A_n$. Also by hypothesis, each $f \in \mathcal{F}$ is continuous and so it is easy to check that each of the sets

$$\{x : |f(x)| \leq n\}$$

must be closed (e.g., Exercise 5:4.31). Thus

$$A_n = \bigcap_{f \in \mathcal{F}} \{x : |f(x)| \leq n\}$$

is an intersection of closed sets and is therefore itself closed. Since $\mathbb{R} = \bigcup_{n=1}^{\infty} A_n$, at least one of the sets, say A_{n_0}, must be dense in some open interval I. (This follows immediately from the Baire Category Theorem.)

Since A_{n_0} is closed and dense in the interval I, A_{n_0} must contain I. This means that $|f(x)| \leq n_0$ for each $f \in \mathcal{F}$ and all $x \in I$. ■

Exercises

6:4.6 Let $\{f_n\}$ be a sequence of continuous functions on an interval $[a, b]$ such that $\lim_{n\to\infty} f_n(x) = f(x)$ exists at every point $x \in [a, b]$. Show that f need not be continuous nor even bounded, but that f must be bounded on some subinterval of $[a, b]$.

6:4.7 Let $\{f_n\}$ be a sequence of continuous functions on $[0, 1]$ such that $\lim_{n\to\infty} f_n(x) = 0$ for all $0 \leq x \leq 1$. Show that there must be an interval $[c, d] \subset [0, 1]$ so that, for all sufficiently large n, $|f_n(x) \leq 1$ for all $x \in [c, d]$.

6:4.8 Give an example of a sequence of functions on $[0, 1]$ with the property that $\lim_{n\to\infty} f_n(x) = 0$ for all $0 \leq x \leq 1$ and yet for every interval $[c, d] \subset [0, 1]$ and every N there is some $x \in [c, d]$ and $n > N$ with $f_n(x) > 1$.

6.5 Cantor Sets ✂

We say that a set is *perfect* if it is a nonempty closed set with no isolated points. The only examples that might come to mind are sets that are finite unions of intervals. It might be difficult to imagine a perfect subset of $\mathbb{R}$ that is also nowhere dense. In this section we obtain such a set, the very important classical Cantor set. We also discuss some of its variants. Such sets have historical significance and are of importance in a number of areas of mathematical analysis.

6.5.1 Construction of the Cantor Ternary Set

We begin with the closed interval [0,1]. From this interval we shall remove a dense open set G. The remaining set $K = [0, 1] \setminus G$ will then be closed and nowhere dense in [0,1]. We construct G is such

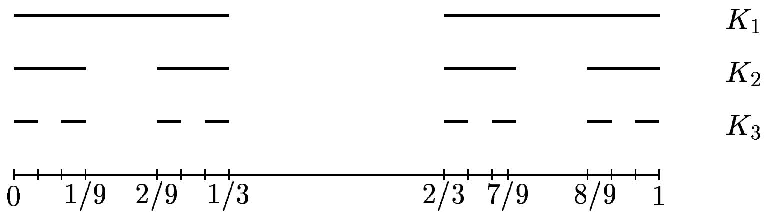

Figure 6.1: Construction of the Cantor Ternary set.

a way that K has no isolated points and is nonempty. Thus K will be a nonempty, nowhere dense perfect subset of [0,1].

It is easiest to understand the set G if we construct it in stages. Let $G_1 = \left(\frac{1}{3}, \frac{2}{3}\right)$, and let $K_1 = [0,1] \setminus G_1$. Thus $K_1 = \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right]$ is what remains when the middle third of the interval [0,1] is removed. This is the first stage of our construction.

We repeat this construction on each of the two component intervals of K_1. Let $G_2 = \left(\frac{1}{9}, \frac{2}{9}\right) \cup \left(\frac{7}{9}, \frac{8}{9}\right)$ and let $K_2 = [0,1] \setminus (G_1 \cup G_2)$. Thus

$$K_2 = \left[0, \frac{1}{9}\right] \cup \left[\frac{2}{9}, \frac{1}{3}\right] \cup \left[\frac{2}{3}, \frac{7}{9}\right] \cup \left[\frac{8}{9}, 1\right].$$

This completes the second stage.

We continue inductively, obtaining two sequences of sets, $\{K_n\}$ and $\{G_n\}$ with the following properties: For each $n \in \mathbb{N}$

1. G_n is a union of 2^{n-1} pairwise disjoint open intervals.

2. K_n is a union of 2^n pairwise disjoint closed intervals.

3. $K_n = [0,1] \setminus (G_1 \cup G_2 \cup \cdots \cup G_n)$.

4. Each component of G_{n+1} is the "middle third" of some component of K_n.

5. The length of each component of K_n is $1/3^n$.

Figure 6.5.1 shows K_1, K_2 and K_3.

Now let

$$G = \bigcup_{n=1}^{\infty} G_n$$

and let

$$K = [0,1] \setminus G = \bigcap_{n=1}^{\infty} K_n.$$

Then G is open and the set K (our Cantor set) is closed.

To see that K is nowhere dense, it is enough, since K is closed, to show that K contains no open intervals (Exercise 6:3.4). Let J be an open interval in [0,1] and let λ be its length. Choose $n \in \mathbb{N}$ such that $1/3^n < \lambda$. By property 5, each component of K_n has length $1/3^n < \lambda$, and by property 2 the components of K_n are pairwise disjoint. Thus K_n cannot contain J, so neither can $K = \bigcap_1^\infty K_n$. We have shown that the closed set K contains no intervals and is therefore nowhere dense.

It remains to show that K has no isolated points. Let $x_0 \in K$. We show that x_0 is a limit point of K. To do this we show that for every $\varepsilon > 0$ there exists $x_1 \in K$ such that $0 < |x_1 - x_0| < \varepsilon$. Choose n such that $1/3^n < \varepsilon$. There is a component L of K_n that contains x_0. This component is a closed interval of length $1/3^n < \varepsilon$. The set $K_{n+1} \cap L$ has two components L_0 and L_1, each of which contains points of K. The point x_0 is in one of the components, say L_0. Let x_1 be any point of $K \cap L_1$. Then $0 < |x_0 - x_1| < \varepsilon$. This verifies that x_0 is a limit point of K. Thus K has no isolated points.

The set K is called the *Cantor set.* Because of its contruction, it is often called the Cantor middle third set. In a moment we shall present a purely arithmetic description of the Cantor set that suggests another common name for K — the "Cantor Ternary set". But first, let's mention a few properties of K and of its complement G that may help the reader visualize these sets.

First note that G is an open dense set in $[0,1]$. Write $G = \bigcup_{k=1}^\infty (a_k, b_k)$. (The component intervals (a_k, b_k) of G can be called the intervals *complementary* to K in $(0,1)$. Each is a 'middle third' of a component interval of some K_n.) Observe that no two of these component intervals can have a common endpoint — if, for example, $b_m = a_n$, then this point would be an isolated point of K, and K has no isolated points.

Next observe that for each $k \in \mathbb{N}$, the points a_k and b_k are points of K. But there are other points of K as well (e.g. $0 \in K$). In fact, we shall see presently that K is uncountable. These other points are all limit points of the endpoints of the complementary intervals. The set of endpoints is countable, but the closure of this set is uncountable as we shall see. Thus, in the sense of cardinality, "most" points of the Cantor set are *not* endpoints of intervals complementary to K.

Each component interval of the set G_n has length $1/3^n$, thus the sum of the lengths of these component intervals is

$$\frac{2^{n-1}}{3^n} = \frac{1}{2}\left(\frac{2}{3}\right)^n.$$

It follows that the lengths of all component intervals of G forms a geometric series with sum

$$\sum_{n=1}^{\infty} \frac{1}{2}\left(\frac{2}{3}\right)^n = 1.$$

(This also gives us a clue as to why K cannot contain an interval—after removing from the unit interval a sequence of pairwise disjoint intervals with length-sum one, no room exists for any intervals in the set K that remains.)

Exercises

6:5.1 Let E be the set of endpoints of intervals complementary to the Cantor set K. Prove $\overline{E} = K$.

6:5.2 Let G be a dense open subset of $\mathbb{R}$ and let $\{(a_k, b_k)\}$ be its set of component intervals. Prove $H = \mathbb{R} \setminus G$ is perfect if and only if no two of these intervals have common endpoints.

✂ **6:5.3** Let K be the Cantor set and let $\{(a_k, b_k)\}$ be the sequence of intervals complementary to K in [0,1]. For each $k \in \mathbb{N}$, let $c_k = (a_k + b_k)/2$ (the midpoint of the interval (a_k, b_k)) and let $N = \{c_k : k \in \mathbb{N}\}$. Prove each of the following:

(a) Every point of N is isolated.

(b) If $c_i \neq c_j$, there exists $k \in \mathbb{N}$ such that c_k is between c_i and c_j (i.e., no point in N has an immediate "neighbor" in N).

(c) Show there is an *order-preserving mapping* $\phi : \mathbb{Q} \cap (0,1) \to N$ (i.e., if $x < y \in \mathbb{Q} \cap (0,1)$, then $\phi(x) < \phi(y) \in N$. This may seem surprising since $\mathbb{Q} \cap (0,1)$ has *no* isolated points while N has *only* isolated points.

6:5.4 It is common now to say that a set E of real numbers is a *Cantor set* if it is nonempty, bounded, perfect, and nowhere dense. Show that the union of a finite number of Cantor sets is also a Cantor set.

6:5.5 Show that every Cantor set is uncountable.

✂ **6:5.6** Let A and B be subsets of $\mathbb{R}$. A function h that maps A onto B, is one-to-one, and with both h and h^{-1} continuous, is called a *homeomorphism* between A and B. The sets A and B are said to be *homeomorphic*. Prove that a set C is a Cantor set if and only if it is homeomorphic to the Cantor ternary set K.

6.5.2 An Arithmetic Construction of K

We turn now to a purely arithmetical construction for the Cantor set. The reader will need some familiarity with ternary (base 3) arithmetic here.

Each $x \in [0,1]$ can be expressed in base 3 as $x = .a_1a_2a_3\ldots$, where $a_i = 0$, 1 or 2, $i = 1,2,3,\ldots$. Certain points have two representations, one ending with a string of zeros, the other in a string of twos. For example, $.1000\cdots = .0222\ldots$ both represent the number 1/3 (base ten). Now, if $x \in (1/3, 2/3)$, $a_1 = 1$, thus each $x \in G_1$ must have '1' in the first position of its ternary expansion. Similarly, if

$$x \in G_2 = \left(\frac{1}{9}, \frac{2}{9}\right) \cup \left(\frac{7}{9}, \frac{8}{9}\right),$$

it must have a '1' in the second position of its ternary expansion — $a_2 = 1$. In general, each point in G_n must have $a_n = 1$. It follows that every point of $G = \bigcup_1^\infty G_n$ must have a '1' someplace in its ternary expansion.

Now endpoints of intervals complementary to K have two representations, one of which involves no '1's. The remaining points of K never fall in the middle third of a component of one of the sets K_n, and so have ternary expansions of the form

$$x = .a_1a_2\ldots \quad a_i = 0 \text{ or } 2.$$

We can therefore describe K arithmetically as the set

$$\{x = .a_1a_2a_3\ldots \text{ (base three) } : a_i = 0 \text{ or } 2 \text{ for each } i \in \mathbb{N}\}.$$

As an immediate result, we see that K is uncountable. In fact, K can be put into 1-1 correspondence with [0,1]: for each

$$x = .a_1a_2a_3\ldots \text{ (base 3)}, \ a_i = 0, 2 \ ,$$

in the set K let there correspond the number

$$y = .b_1b_2b_3\ldots \text{ (base 2)}, \ b_i = a_i/2 \ .$$

This provides a 1-1 correspondence between K (minus endpoints of complementary intervals) and [0,1] (minus the countable set of numbers with two base two representations). By allowing these two countable sets to correspond to each other, we obtain a 1-1 correspondence between K and [0,1].

We end this section by mentioning that variations in the constructions of K can lead to interesting situations. For example, by changing the construction slightly, we can remove intervals in such a way that $G' = \bigcup_{k=1}^\infty (a'_k, b'_k)$ with $\sum_{k=1}^\infty (b'_k - a'_k) = 1/2$ (instead of 1), while still keeping $K' = [0,1] \setminus G'$ nowhere dense and perfect. The resulting set K' created problems for late 19th century mathematicians trying to develop a theory of measure. The "measure" of G' should be 1/2, the "measure" of [0,1] should be 1. Intuition requires that the measure of the nowhere dense set K' should be $1 - \frac{1}{2} = \frac{1}{2}$. How can this be, when K' is so "small"?

Exercises

6:5.7 Find a specific irrational number in the Cantor ternary set.

6:5.8 Show that the Cantor ternary set can be defined as

$$K = \left\{ x \in [0,1] : x = \sum_{n=1}^{\infty} \frac{i_n}{3^n} \text{ for } i_n = 0 \text{ or } 2 \right\}.$$

6:5.9 Let

$$D = \left\{ x \in [0,1] : x = \sum_{n=1}^{\infty} \frac{j_n}{3^n} \text{ for } j_n = 0 \text{ or } 1 \right\}.$$

Show $D + D = \{x + y : x, y \in D\} = [0,1]$. From this deduce, for the Cantor ternary set K, that $K + K = [0,2]$.

6:5.10 Criticize the following "argument" which is far too often seen:

> "If $G = (a,b)$ then $\overline{G} = [a,b]$. Similarly, if $G = \bigcup_{i=1}^{\infty}(a_i, b_i)$ is an open set, then $\overline{G} = \bigcup_{i=1}^{\infty}[a_i, b_i]$. It follows that an open set G and its closure $\overline{G}$ differ by at most a countable set."(?)

✂ 6.5.3 The Cantor Function

The Cantor set allows the construction of a rather bizarre function that is continuous and increasing on the interval $[0,1]$. It has the property that it is constant on every interval complementary to the Cantor set and yet manages to increase from $f(0) = 0$ to $f(1) = 1$ by doing all of its increasing on the Cantor set itself. It has sometimes been called "the devil's staircase".

Define the function f in the following way. On $(1/3, 2/3)$, let $f = 1/2$; on $(1/9, 2/9)$, let $f = 1/4$; on $(7/9, 8/9)$, let $f = 3/4$. Proceed inductively. On the $2^{n-1} - 1$ open intervals appearing at the nth stage, define f to satisfy the following conditions:

(*i*) f is constant on each of these intervals.

(*ii*) f takes the values

$$\frac{1}{2^n}, \frac{3}{2^n}, \dots, \frac{2^n - 1}{2^n}$$

on these intervals.

(*iii*) If x and y are members of different nth-stage intervals with $x < y$, then $f(x) < f(y)$.

This description defines f on $G = [0,1] \setminus K$. Extend f to all of $[0,1]$ by defining $f(0) = 0$ and, for $0 < x \leq 1$, $f(x) = \sup\{f(t) : t \in G, t < x\}$. In order to check that this defines the function that we want, we need to check each of the following.

1. $f(G)$ is dense in $[0,1]$.
2. f is nondecreasing on $[0,1]$. .
3. f is continuous on $[0,1]$.
4. $f(K) = [0,1]$.

These have been left as exercises.

Figure 6.2 illustrates the construction. The function f is called the *Cantor function.* Observe that f "does all its rising" on the set K.

The Cantor function allows a negative answer to many questions that might be asked about functions and derivatives and, hence, has become a popular counterexample. For example let us follow this kind of reasoning. If f is a continuous function on $[0,1]$ and $f'(x) = 0$ for every $x \in (0,1)$ then f is constant. (This is proved in most calculus courses based on the mean value theorem.) Now suppose that we know less, that $f'(x) = 0$ for every $x \in (0,1)$ excepting a "small" set E of points at which we know nothing. If E is finite it is still easy to show that f must be constant. If E is countable it is possible, but a bit more difficult, to show that it is still true that f must be constant. The question then arises, just how small a set E can appear here, i.e., what would we have to know about a set E so that we could say $f'(x) = 0$ for every $x \in (0,1) \setminus E$ implies that f is constant?

The Cantor function is an example of a function constant on every interval complementary to the Cantor set K (and so with a zero derivative at those points) and yet is not constant. The Cantor set, since it is nowhere dense might be viewed as extremely small, but even so it is not insignificant for this problem.

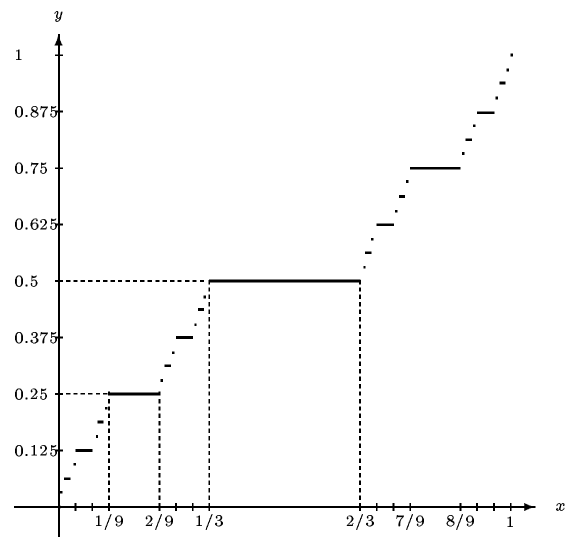

Figure 6.2: The Cantor function.

Exercises

6:5.11 In the construction of the Cantor function complete the verification of details.

(a) Show that $f(G)$ is dense in $[0,1]$.

(b) Show that f is nondecreasing on $[0,1]$. .

(c) Infer from (a) and (b) that f is continuous on $[0,1]$.

(d) Show that $f(K) = [0,1]$ and thus (again) conclude that K is uncountable.

6:5.12 Find the calculus textbook proof for the statement that a continuous function f on an interval $[a,b]$ that has a zero derivative on (a,b) must be constant. Improve the proof to allow a finite set of points on which f is not known to have a zero derivative..

✂ 6.6 Borel Sets

In our study of continuous functions we have seen that the classes of open sets and closed sets play a significant role. But the class of sets that are of importance in analysis goes beyond merely the open and closed sets. It was recognized by E. Borel (1871–1956) that for many operations of analysis one needed to form countable intersections and countable unions of classes of sets. The collection of Borel sets was introduced exactly to allow these operations. We recall that a countable union of closed sets may not be closed (or open) and that a countable intersection of open sets, also, may not be open (or closed).

In this section we introduce two additional types of sets of importance in analysis — sets of type $\mathcal{G}_\delta$ and sets of type $\mathcal{F}_\sigma$. These classes form just the beginning of the large class of Borel sets. We shall find that they are precisely the right classes of sets to solve some fundamental questions about real functions.

6.6.1 Sets of Type G_δ

Recall that the union of a collection of open sets is open (regardless of how many sets are in the collection), but the intersection of a collection of open sets need not be open if the collection has infinitely many sets. For example,

$$\bigcap_{n=1}^{\infty} \left(-\frac{1}{n}, \frac{1}{n}\right) = \{0\}.$$

Similarly, if $q_1, q_2, q_3, \dots$ is an enumeration of $\mathbb{Q}$, then

$$\bigcap_{k=1}^{\infty} (\mathbb{R} \setminus \{q_k\}) = \mathbb{R} \setminus \mathbb{Q},$$

the set of irrational numbers. The set $\{0\}$ is closed (not open), and $\mathbb{R} \setminus \mathbb{Q}$ is neither open nor closed. The set $\mathbb{R} \setminus \mathbb{Q}$ is a countable intersection of open sets. Such sets are of sufficient importance to give them a name.

Definition 6.14 A subset H of $\mathbb{R}$ is said to be of *type* $\mathcal{G}_\delta$ (or a $\mathcal{G}_\delta$ *set*) if it can be expressed as a countable intersection of open sets, that is, if there exist open sets $G_1, G_2, G_3, \dots$ such that $H = \bigcap_{k=1}^{\infty} G_k$.

Example 6.15 A closed interval $[a, b]$ or a half-open interval $(a, b]$ is of type $\mathcal{G}_\delta$ since

$$[a, b] = \bigcap_{n=1}^{\infty} \left(a - \frac{1}{n}, b + \frac{1}{n}\right)$$

and

$$(a, b] = \bigcap_{n=1}^{\infty} \left(a, b + \frac{1}{n}\right).$$

◀

Theorem 6.16 *Every open set and every closed set in $I\!R$ is of type $\mathcal{G}_\delta$.*

Proof. Let G be an open set in $\mathbb{R}$. It is clear that G is of type $\mathcal{G}_\delta$. We also show that G can be expressed as a countable union of closed sets. Express G in the form $G = \bigcup_{k=1}^{\infty}(a_k, b_k)$ where the intervals (a_k, b_k) are pairwise disjoint. Now for each $k \in \mathbb{N}$, there exist sequences $\{c_{k_j}\}$ and $\{d_{k_j}\}$ such that the sequence $\{c_{k_j}\}$ decreases to a_k, the sequence $\{b_{k_j}\}$ increases to b_k and $c_{k_j} < d_{k_j}$ for each $j \in \mathbb{N}$. Thus $(a_k, b_k) = \bigcup_{j=1}^{\infty}[c_{k_j}, d_{k_j}]$. We have expressed each component interval of G as a countable union of closed sets. It follows that

$$G = \bigcup_{k=1}^{\infty}\bigcup_{j=1}^{\infty}[c_{k_j}, d_{k_j}] = \bigcup_{j,k=1}^{\infty}[c_{k_j}, d_{k_j}]$$

is also a countable union of closed sets. Now take complements. This shows that $\mathbb{R} \setminus G$ can be expressed as a countable intersection of open sets (by using the de Morgan Laws). Since every closed set F can be written $F = \mathbb{R} \setminus G$ for some open set G we have shown that any closed set is of type $\mathcal{G}_\delta$. ■

We observed in Section 6.4 that a dense set can be small in the sense of category. For example, $\mathbb{Q}$ is a first category set. Our next result shows that a dense set *of type $\mathcal{G}_\delta$* must be large in the sense of category.

Theorem 6.17 *Let H be of type $\mathcal{G}_\delta$ and be dense in $I\!R$. Then H is residual.*

Proof. Write $H = \bigcap_{k=1}^{\infty} G_k$ with each of the sets G_k open. Since H is dense by hypothesis and $H \subset G_k$ for each $k \in \mathbb{N}$, each of the open sets G_k is also dense. Thus $\mathbb{R} \setminus G_k$ is nowhere dense for every $k \in \mathbb{N}$, and G is residual. The result now follows from Theorem 6.11 part (2). ■

Exercises

6:6.1 Which of the sets below are of type $\mathcal{G}_\delta$?

(a) $\mathbb{N}$

(b) $\left\{\frac{1}{n} : n \in \mathbb{N}\right\}$.

(c) The set $\{C_n : n \in \mathbb{N}\}$ of midpoints of intervals complementary to the Cantor set.

(d) A finite union of intervals (that need not be open or closed).

6:6.2 Prove Theorem 6.17 for the interval $[a, b]$ in place of $\mathbb{R}$.

6:6.3 Prove that a set E of type $\mathcal{G}_\delta$ in $\mathbb{R}$ is either residual or else there is an interval containing no points of E.

6.6.2 Sets of Type F_σ

Just as the countable intersections of open sets form a larger class of sets, the $\mathcal{G}_\delta$ sets, so also the countable unions of closed sets form a larger class of sets.

The complements of open sets are closed. By dealing with complements of $\mathcal{G}_\delta$ sets we arrive at the dual notion of a set of type $\mathcal{F}_\sigma$.

Definition 6.18 A subset E of $\mathbb{R}$ is said to be of *type* $\mathcal{F}_\sigma$ (or an $\mathcal{F}_\sigma$ set) if it can be expressed as a countable union of closed sets — that is, if there exist closed sets $F_1, F_2, F_3, \dots$ such that $E = \bigcup_{k=1}^{\infty} F_k$.

Example 6.19 The set of rational numbers, $\mathbb{Q}$ is a set of type $\mathcal{F}_\sigma$. This is clear since it can be expressed as

$$\mathbb{Q} = \bigcup_{n=1}^{\infty} \{r_n\}$$

where $\{r_n\}$ is any enumeration of the rationals. The singleton sets $\{r_n\}$ are clearly closed. But note that $\mathbb{Q}$ is *not* of type $\mathcal{G}_\delta$ also. It follows from Theorem 6.17 that a dense set of type $\mathcal{G}_\delta$ must be uncountable (because a countable set is first category). In particular, $\mathbb{Q}$ is not of type $\mathcal{G}_\delta$, (and therefore $\mathbb{R} \setminus \mathbb{Q}$ is not of type $\mathcal{F}_\sigma$.) ◀

Using the de Morgan laws, one verifies easily that the complement of a $\mathcal{G}_\delta$ set is an $\mathcal{F}_\sigma$ and vice-versa (Exercise 6:6.4). This is closely related to the fact that a set is open if and only if its complement is closed.

Theorem 6.20 *A set is of type $\mathcal{G}_\delta$ if and only if its complement is of type $\mathcal{F}_\sigma$.*

Example 6.21 A half-open interval $(a, b]$ is both of type $\mathcal{G}_\delta$ and of type $\mathcal{F}_\sigma$:

$$(a, b] = \bigcap_{n=1}^{\infty} \left(a, b + \frac{1}{n}\right) = \bigcup_{n=1}^{\infty} \left[a + \frac{b-a}{n}, b\right].$$

◀

Remark. The only subsets of $\mathbb{R}$ that are both open and closed are the empty set and $\mathbb{R}$ itself. There are however many sets that are of type $\mathcal{G}_\delta$ and also of type $\mathcal{F}_\sigma$. See Exercise 6:6.1.

We can now enlarge on Theorem 6.16. There we showed that all open sets and all closed sets are in the class $\mathcal{G}_\delta$. We now show they are also in the class $\mathcal{F}_\sigma$.

Theorem 6.22 *Every open set and every closed set in $\mathbb{R}$ is both of type $\mathcal{F}_\sigma$ and $\mathcal{G}_\delta$.*

Proof. In the proof of Theorem 6.16 we showed explicitly how to express any open set as an $\mathcal{F}_\sigma$. Thus open sets are of type $\mathcal{F}_\sigma$ as well as of type $\mathcal{G}_\delta$ (the latter being trivial). The part pertaining to closed sets now follows by considering complements — the complement of a closed set is open and the complement of an $\mathcal{F}_\sigma$ set is a $\mathcal{G}_\delta$ set. ■

Exercises

6:6.4 Verify that a subset A of $\mathbb{R}$ is an $\mathcal{F}_\sigma$ ($\mathcal{G}_\delta$) if and only if $\mathbb{R} \setminus A$ is a $\mathcal{G}_\delta$ ($\mathcal{F}_\sigma$).

6:6.5 Which of the sets below are of type $\mathcal{F}_\sigma$?

(a) $\mathbb{N}$

(b) $\left\{\frac{1}{n} : n \in \mathbb{N}\right\}$.

(c) The set $\{C_n : n \in \mathbb{N}\}$ of midpoints of intervals complementary to the Cantor set.

(d) A finite union of intervals (that need not be open or closed).

✂ **6:6.6** Prove that a set of type $\mathcal{F}_\sigma$ in $\mathbb{R}$ is either first category or contains an open interval.

✂ **6:6.7** Let $\{f_n\}$ be a sequence of real functions defined on $\mathbb{R}$ and suppose that $f_n(x) \to f(x)$ at every point x. Show that

$$\{x : f(x) > \alpha\} = \bigcup_{m=1}^{\infty} \bigcup_{r=1}^{\infty} \bigcap_{n=r}^{\infty} \{x : f_n(x) \geq \alpha + 1/m\}.$$

If each function f_n is continuous what can you assert about the set $\{x : f(x) > \alpha\}$?

6.7 Oscillation and Continuity

In this section we return to a problem that we began investigating in Section 5.9 about the nature of the set of discontinuity points of a function. To discuss this set we shall need the notions of $\mathcal{F}_\sigma$ and $\mathcal{G}_\delta$ sets and we need to introduce a new tool—the oscillation of a function.

We begin with an example of a function f that is discontinuous at every rational number and continuous at every irrational number.

Example 6.23 Let $q_1, q_2, q_3, \dots$ be an enumeration of $\mathbb{Q}$. Define a function f by

$$f(x) = \begin{cases} \frac{1}{k}, & \text{if } x = q_k \\ 0, & \text{if } x \in \mathbb{R} \setminus \mathbb{Q}. \end{cases}$$

Since $\mathbb{R} \setminus \mathbb{Q}$ is dense in $\mathbb{R}$, f can be continuous at a point x only if $f(x) = 0$ — i.e. only if $x \in \mathbb{R} \setminus \mathbb{Q}$. Thus f is discontinuous at every $x \in \mathbb{Q}$. To check that f is continuous at each point of $\mathbb{R} \setminus \mathbb{Q}$, let $x_0 \in \mathbb{R} \setminus \mathbb{Q}$ and let $\varepsilon > 0$. Choose $k \in \mathbb{N}$ such that $1/k < \varepsilon$. Since the set $q_1, q_2, \dots, q_k$ is a finite set not containing x_0, there exists $\delta > 0$ such that $|q_i - x_0| \geq \delta$ for each $i = 1, \dots, k$. Thus if $x \in \mathbb{R}$ and $|x - x_0| < \delta$, then either $x \in \mathbb{R} \setminus \mathbb{Q}$ or $x = q_j$ for some $j > k$. In either case $|f(x) - f(x_0)| \leq \frac{1}{k} < \varepsilon$. This verifies the continuity of f at x_0. Since x_0 was an arbitrary irrational point we see that f is continuous at every irrational. ◀

Our example shows that it is possible for a function to be continuous at every irrational number and discontinuous at every rational number. Is it possible for the opposite to occur? Does there exist a function f continuous on $\mathbb{Q}$ and discontinuous on $\mathbb{R} \setminus \mathbb{Q}$? More generally, what sets can be the set of points of continuity of some function f defined on an interval.

We answer this question in this section. The principle tool is that of *oscillation* of a function at a point.

6.7.1 Oscillation of a Function

In order to describe a point of discontinuity we need a way of measuring that discontinuity. For monotonic functions the jump was used previously for such a measure. For general, nonmonotonic, functions a different tool is used.

Definition 6.24 Let f be defined on a non-degenerate interval I.

We define the *oscillation of f on I* as the quantity

$$\omega f(I) = \sup_{x,y \in I} |f(x) - f(y)|.$$

Let's see how oscillation relates to continuity. Suppose f is defined in a neighborhood of x_0, and f is continuous at x_0. Then

$$\inf_{\delta>0} \omega f((x_0 - \delta, x_0 + \delta)) = 0, \tag{1}$$

To see this, let $\varepsilon > 0$. Since f is continuous at x_0, there exists $\delta_0 > 0$ such that $|f(x) - f(x_0)| < \varepsilon/2$ if $|x - x_0| < \delta_0$. If

$$x_0 - \delta_0 < x_1 \le x_2 < x_0 + \delta_0,$$

then

$$|f(x_1) - f(x_2)| \le |f(x_1) - f(x_0)| + |f(x_0) - f(x_2)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \tag{2}$$

Since (2) is valid for all $x_1, x_2 \in (x_0 - \delta_0, x_0 + \delta_0)$, we have

$$\sup\{|f(x_1) - f(x_2)| : x_0 - \delta_0 < x_1 \le x_2 < x_0 + \delta_0\} \le \varepsilon. \tag{3}$$

But (3) implies that if $0 < \delta < \delta_0$, then $\omega f([x_0 - \delta, x_0 + \delta]) \le \varepsilon$. Since ε was arbitrary, the result follows.

The converse is also valid. Suppose (1) holds. Let $\varepsilon > 0$. Choose $\delta > 0$ such that $\omega f(x_0 - \delta, x_0 + \delta) < \varepsilon$. Then

$$\sup\{|f(x) - f(x_0)| : x \in (x_0 - \delta, x_0 + \delta)\} < \varepsilon,$$

so $|f(x) - f(x_0)| < \varepsilon$ whenever $|x - x_0| < \delta$. This implies continuity of f at x_0.

We summarize the preceding as a theorem.

Theorem 6.25 *Let f be defined on an interval I and let $x_0 \in I$. Then f is continuous at x_0 if and only if*

$$\inf_{\delta>0} \omega f((x_0 - \delta, x_0 + \delta)) = 0.$$

The quantity in the statement of the theorem is sufficiently important to have a name.

Definition 6.26 Let f be defined in a neighborhood of x_0. The quantity

$$\omega_f(x_0) = \inf_{\delta>0} \omega f((x_0 - \delta, x_0 + \delta))$$

is called the *oscillation of f at x_0*.

Theorem 6.25 thus states that a function f is continuous at a point x_0 if and only if $\omega_f(x_0) = 0$. Returning to the function that introduced this section, we see that

$$\omega_f(x) = \begin{cases} 1/k, & \text{if } x = q_k \\ 0, & \text{if } x \in \mathbb{R} \setminus \mathbb{Q}. \end{cases}$$

Let's now see how the concept of oscillation relates to the set of points of continuity of a function.

Theorem 6.27 *Let f be defined on a closed interval I (which may be all of $\mathbb{R}$). Let $\gamma > 0$. Then the set*

$$\{x : \omega_f(x) < \gamma\}$$

is open and the set

$$\{x : \omega_f(x) \geq \gamma\}$$

is closed.

Proof. Let $A = \{x : \omega_f(x) < \gamma\}$ and let $x_0 \in A$. We wish to find a neighborhood U of x_0 such that $U \subset A$, i.e., such that $\omega_f(x) < \gamma$ for all $x \in U$.

Let $\omega_f(x_0) = \alpha < \gamma$ and let $\beta \in (\alpha, \gamma)$. From Definition 6.26 we infer the existence of a number $\delta > 0$ such that $|f(u) - f(v)| \leq \beta$ for $u, v \in (x_0 - \delta, x_0 + \delta)$. Let $U = (x_0 - \delta, x_0 + \delta)$ and let $x \in U$. Since U is open, there exists $\delta_1 < \delta$ such that $(x - \delta_1, x + \delta_1) \subset U$. Then

$$\begin{aligned} \omega_f(x_0) &\leq \sup\{|f(t) - f(s)| : t, s \in (x - \delta_1, x + \delta_1)\} \\ &\leq \sup\{|f(u) - f(v)| : u, v \in U\} \leq \beta < \gamma, \end{aligned}$$

so $x \in A$. This proves A is open. It follows then that the complement of A in I, the set

$$\{x : \omega_f(x) \geq \gamma\}$$

must be closed. ■

We use the oscillation in the next subsection to answer a question about the nature of the set of points of continuity of a function. We shall encounter the oscillation concept again in Chapter 8 when we study the integrability of functions.

Exercises

6:7.1 Suppose that f is bounded on an interval I. Prove that

$$\omega f(I) = \sup_{x \in I} f(x) - \inf_{x \in I} f(x).$$

6:7.2 The statement below is *false*.

$$\omega_f(x_0) = \limsup_{x \to x_0} f(x) - \liminf_{x \to x_0} f(x).$$

How can it fail, even for bounded functions?

6:7.3 Prove that

$$\omega f(x) = \lim_{\delta \to 0+} \omega f((x_0 - \delta, x_0 + \delta)).$$

6:7.4 Calculate $\omega f(0)$ for each of the following functions.

(a) $f(x) = \begin{cases} x, & \text{if } x \neq 0 \\ 4, & \text{if } x = 0. \end{cases}$

(b) $f(x) = \begin{cases} 0, & \text{if } x \in \mathbb{Q} \\ 1, & \text{if } x \notin \mathbb{Q}. \end{cases}$

(c) $f(x) = \begin{cases} n, & \text{if } x = \frac{1}{n} \\ 0, & \text{otherwise.} \end{cases}$

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

(e) $f(x) = \begin{cases} \sin \frac{1}{x}, & \text{if } x \neq 0 \\ 7, & \text{if } x = 0. \end{cases}$

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

6:7.5 In the proof of Theorem 6.27 we let $\omega f(x_0) = \alpha < \gamma$ and let $\beta \in (\alpha, \gamma)$. Why was the β introduced? Would the proof have worked if we had used $\beta = \gamma$?

6.7.2 The Set of Points Where a Function is Continuous

Given an arbitrary function how can we describe the nature of the set of points where f is continuous? Can it be any set? Given a set E how can we know whether there is a function which is continuous at every point of E and discontinuous at every point not in E?

We saw in Example 6.23 that a function exists whose set of continuity points is exactly the irrationals. Can a function exist whose set of continuity points is exactly the rationals? By characterizing the set of such points we can answer this and other questions about the structure of functions.

We now prove the main result of this section using primarily the notion of oscillation introduced in Section 6.7.1.

Theorem 6.28 *Let f be defined on a closed interval I (which may be all of $\mathbb{R}$). Then the set C_f of points of continuity of f is of type $\mathcal{G}_\delta$, and the set D_f of points of discontinuity of f is of type $\mathcal{F}_\sigma$. Conversely, if H is a set of type $\mathcal{G}_\delta$, then there exists a function f defined on $\mathbb{R}$ such that $C_f = H$.*

Proof. To prove the first part, let $f : I \to \mathbb{R}$. We show that $\{x : \omega f(x) = 0\}$ is of type $\mathcal{G}_\delta$. For each $k \in \mathbb{N}$, let $B_k = \left\{x : \omega f(x) \geq \frac{1}{k}\right\}$. By Theorem 6.27, each of the sets B_k is closed. Thus $B = \bigcup_{k=1}^{\infty} B_k$ is of type $\mathcal{F}_\sigma$. By Theorem 6.25, $D_f = B$. Therefore $C_f = I \setminus B$. Since the complement of an $\mathcal{F}_\sigma$ is a $\mathcal{G}_\delta$, C_f is a $\mathcal{G}_\delta$.

To prove the converse, let H be any subset of $\mathbb{R}$ of type $\mathcal{G}_\delta$. Then H can be expressed in the form $H = \bigcap_{k=1}^{\infty} G_k$ with each of the sets G_k being open. We may assume without loss of generality that $G_1 = \mathbb{R}$ and that $G_i \supset G_{i+1}$ for each $i \in \mathbb{N}$. (Verify this.)

Let $\{\alpha_k\}$ and $\{\beta_k\}$ be sequences of positive numbers, each converging to zero, with $\alpha_k > \beta_k > \alpha_{k+1}$, for all $k \in \mathbb{N}$. Define a function $f : \mathbb{R} \to \mathbb{R}$ by

$$f(x) = \begin{cases} 0 & \text{if } x \in H \\ \alpha_k & \text{if } x \in (G_k \setminus G_{k+1}) \cap \mathbb{Q} \\ \beta_k & \text{if } x \in (G_k \setminus G_{k+1}) \cap (\mathbb{R} \setminus \mathbb{Q}). \end{cases}$$

We show that f is continuous at each point of H and discontinuous at each point of $\mathbb{R} \setminus H$.

Let $x_0 \in H$ and let $\varepsilon > 0$. Choose n such that $\alpha_n < \varepsilon$. Since $x_0 \in H = \bigcap_{k=1}^{\infty} G_k$, $x_0 \in G_n$. The set G_n is open, so there exists $\delta > 0$ such that $(x_0 - \delta, x_0 + \delta) \subset G_n$. From the definition of G_n, we see that $0 \leq f(x) \leq \alpha_n < \varepsilon$ for all $x \in (x_0 - \delta, x_0 + \delta)$. Thus

$$|f(x) - f(x_0)| = |f(x) - 0| = |f(x)| < \varepsilon$$

if $|x - x_0| < \delta$, so f is continuous at x_0.

Now let $x_0 \in \mathbb{R} \setminus H$. Then there exists $k \in \mathbb{N}$ such that $x_0 \in G_k \setminus G_{k+1}$, so $f(x_0) = \alpha_k$ or $f(x_0) = \beta_k$, say $f(x_0) = \alpha_k$.

If x_0 is an interior point of $G_k \setminus G_{k+1}$, then x_0 is a limit point of

$$\{x : x \in (G_k \setminus G_{k+1}) \cap (\mathbb{R} \setminus \mathbb{Q})\} = \{x : f(x) = \beta_k\},$$

so f is discontinuous ar x_0.

The argument is similar if x_0 is a boundary point of $G_k \setminus G_{k+1}$. Again, assume $f(x_0) = \alpha_k$. Arbitrarily close to x_0 there are points of $\mathbb{R} \setminus (G_k \setminus G_{k+1})$. At these points, f takes on values in the set

$$S = \{0\} \cup \bigcup_{i \neq k} \alpha_i \cup \bigcup_{j \neq k} \beta_j.$$

The only limit point of this set is zero and so S is closed. In particular, α_k is *not* a limit point of this set and does not belong to the set. Let ε be half the distance from the point α_k to the closed set S, i.e., in symbols let $\varepsilon = \frac{1}{2} d(\alpha_k, S)$. Arbitrarily close to x_0 there are points x such that $f(x) \in S$. For such a point,

$$|f(x) - f(x_0)| = |f(x) - \alpha_k| > \varepsilon,$$

so f is discontinuous at x_0. ■

Observe that Theorem 6.28 answers a question we asked earlier: is there a function f continuous on $\mathbb{Q}$ and discontinuous at every point of $\mathbb{R} \setminus \mathbb{Q}$? The answer is negative, since $\mathbb{Q}$ is not of type $\mathcal{G}_\delta$.

Exercises

6:7.6 In the second part of the proof of Theorem 6.28 we provided a construction for a function f with $C_f = H$, where H is an arbitrary set of type $\mathcal{G}_\delta$. Exhibit explicitly sets G_k that will give rise to a function f such that $C_f = \mathbb{R} \setminus \mathbb{Q}$. Can you do this in such a way that the resulting function is the one we obtained at the beginning of this section?

6:7.7 In the proof of the 'converse part' of Theorem 6.28 we took $\varepsilon = \frac{1}{2}d(\alpha_k, S)$. Show that this number equals

$$\frac{1}{2} \min_{i \neq k} \{\min\{|\alpha_i - \alpha_k|, |\beta_i - \beta_k|\}\}.$$

✂ 6.8 Sets of Measure Zero

In analysis there are a number of ways in which a set might be considered as "small". For example the Cantor set is not small in the sense of counting: it is uncountable. It is small in another different sense: it is nowhere dense, that is there is no interval at all in which it is dense. Now we turn to another way in which the Cantor set can be considered small: it has "zero length".

Example 6.29 Suppose we wish to measure the "length" of the Cantor set. Since the Cantor set is rather bizarre we might look instead at the sequence of intervals that has been removed. There is no difficulty in assigning a meaning of length to an interval; the length of (a, b) is $b - a$. What is the total length of the intervals removed in the construction of the Cantor set? From the interval $[0, 1]$ we remove first a middle third interval of length $1/3$, then two middle third intervals of length $1/9$, and so on so that at the nth stage we remove 2^{n-1} intervals each of length 3^{-n}. The sum of the lengths of all intervals so removed is

$$1/3 + 2(1/9) + 4(1/27) + \cdots =$$
$$1/3\,(1 + 2/3 + (2/3)^2 + (2/3)^3 + \dots) = 1.$$

From the interval $[0, 1]$ we appear to have removed all of the length. What is left over, the Cantor set, must have length zero.

This method of computing lengths has some merit but it is not the one we wish to adopt here. Another approach to "measuring" the length of the Cantor set is to consider the length that *remains* at each stage. At the first stage the Cantor set is contained inside the union

$$[0, 1/3] \cup [2/3, 1]$$

which has length $2(1/3)$. At the next stage it is contained inside a union of four intervals, with total length $4(1/9)$. Similarly at the nth stage the Cantor set is contained inside the union of 2^n intervals each of length 3^{-n}. The sum of the lengths of all these intervals is $(2/3)^n$ and this tends to zero as n gets large. Thus, as before, it seems we should assign zero length to the Cantor set. ◀

We convert the second method of the example into a definition of what it means for a set to be of measure zero. "Measure" is the technical term used to describe the "length" of sets that need not be intervals. In the example we used closed intervals while in our definition we have employed open intervals. There is no difference (see Exercise 6:8.13). In the example we covered the Cantor set with a finite sequence of intervals while in our defininition we have employed an infinite sequence. For the Cantor set there is no difference but for other sets (sets that are not bounded or are not closed) there is a difference.

Definition 6.30 Let E be a set of real numbers. Then E is said to have *measure zero* if for every $\varepsilon > 0$ there is a finite or infinite sequence

$$(a_1, b_1), (a_2, b_2), (a_3, b_3), (a_4, b_4), \dots$$

of open intervals covering the set E so that

$$\sum_{k=1}^{\infty} (b_k - a_k) \leq \varepsilon.$$

Note. In the definition of measure zero sets is there a change if we insist on an *infinite* sequence of intervals, disallowing finite sequences? Suppose that the sequence

$$(a_1, b_1), (a_2, b_2), (a_3, b_3), (a_4, b_4), \dots (a_N, b_N)$$

of open intervals covers the set E so that

$$\sum_{k=1}^{N} (b_k - a_k) < \varepsilon/2.$$

Then to satisfy the definition we could add in some further intervals that do not amount in length to more than $\varepsilon/2$. For example take $(a_{N+p}, b_{N+p}) = (0, \varepsilon/2^{p+1})$ for $p = 1, 2, 3, \ldots$. Then

$$\sum_{k=1}^{\infty}(b_k - a_k) = \sum_{k=1}^{N}(b_k - a_k) + \sum_{p=1}^{\infty}\varepsilon/2^{p+1} < \varepsilon.$$

Thus the definition would not be changed if we had required infinite coverings.

Here are some examples of sets of measure zero.

Example 6.31 Every finite set has measure zero. The empty set is easily handled. If

$$E = \{x_1, x_2, \ldots x_N\}$$

and $\varepsilon > 0$ then the sequence of intervals

$$\left(x_i - \frac{\varepsilon}{2N}, x_i + \frac{\varepsilon}{2N}\right) \quad i = 1, 2, 3, \ldots N$$

covers the set E and the sum of all the lengths is ε. ◀

Example 6.32 Every infinite, countable set has measure zero. If

$$E = \{x_1, x_2, \ldots\}$$

and $\varepsilon > 0$ then the sequence of intervals

$$\left(x_i - \frac{\varepsilon}{2^{i+1}}, x_i + \frac{\varepsilon}{2^{i+1}}\right) \quad i = 1, 2, 3,$$

covers the set E and the sum of all the lengths is

$$\sum_{k=1}^{\infty} 2\left(\frac{\varepsilon}{2^{k+1}}\right) = \sum_{k=1} \varepsilon 2^{-k} = \varepsilon.$$

◀

Example 6.33 The Cantor set has measure zero. Let $\varepsilon > 0$. Choose n so that $(2/3)^n < \varepsilon$. Then the nth stage intervals in the construction of the Cantor set give us 2^n closed intervals each of length $(1/3)^n$. This covers the Cantor set with 2^n closed intervals of total length $(2/3)^n$ which is less than ε. If the closed intervals trouble you (the definition requires open intervals) see Exercise 6:8.13 or argue as follows. Since $(2/3)^n < \varepsilon$ there is a positive number δ so that

$$(2/3)^n + \delta < \varepsilon.$$

Enlarge each of the closed intervals to form a slightly larger open interval, but change the length of each only enough so that the sum of the lengths of all the 2^n closed intervals does not increase by more than δ. The resulting collection of open intervals also covers the

Cantor set and the sum of the length of these intervals is less than ε.
◀

One of the most fundamental of the properties of sets having measure zero is how sequences of such sets combine. We recall that the union of any sequence of countable sets is also countable. We now prove that the union of any sequence of measure zero sets is also a measure zero set.

Theorem 6.34 *Let E_1, E_2, E_3, ... be a sequence of sets of measure zero. Then the set E formed by taking the union of all the sets in the sequence is also of measure zero.*

Proof. Let $\varepsilon > 0$. We shall construct a cover of E consisting of a sequence of open intervals of total length less than ε. Since E_1 has measure zero there is a sequence of open intervals

$$(a_{11}, b_{11}), (a_{12}, b_{12}), (a_{13}, b_{13}), (a_{14}, b_{14}), \dots$$

covering the set E_1 and so that the sum of the lengths of these intervals is smaller than $\varepsilon/2$. Since E_2 has measure zero there is a sequence of open intervals

$$(a_{21}, b_{21}), (a_{22}, b_{22}), (a_{23}, b_{23}), (a_{24}, b_{24}), \dots$$

covering the set E_2 and so that the sum of the lengths of these intervals is smaller than $\varepsilon/4$. In general for each $k = 1, 2, 3, \dots$ there is a sequence of open intervals

$$(a_{k1}, b_{k1}), (a_{k2}, b_{k2}), (a_{k3}, b_{k3}), (a_{k4}, b_{k4}), \dots$$

covering the set E_k and so that the sum of the lengths of these intervals is smaller than $\varepsilon/2^k$. The totality of all these intervals can be arranged into a single sequence of open intervals that covers every point in the union of the sequence $\{E_k\}$. The sum of the lengths of all the intervals in the large sequence is smaller than

$$\varepsilon/2 + \varepsilon/4 + \varepsilon/8 + \cdots = \varepsilon.$$

■

Let us return to the situation for the Cantor set once again. For each $\varepsilon > 0$ we were able to choose a finite cover of open intervals with total length less than ε. This is not the case for all sets of measure zero. For example the set of all rational numbers on the real line is countable and hence also of measure zero. Any finite collection of intervals must fail to cover that set, in fact cannot come close to covering all rational numbers. For what sets is it possible to select finite coverings of small length? The answer is that this is possible for compact sets of measure zero.

Theorem 6.35 *Let E be a compact set of measure zero. Then for every $\varepsilon > 0$ there is a finite collection of open intervals*

$$(a_1, b_1), (a_2, b_2), (a_3, b_3), (a_4, b_4), \ldots (a_N, b_N)$$

that covers the set E and so that

$$\sum_{k=1}^{N} (b_k - a_k) < \varepsilon.$$

Proof. Since E has measure zero it is certainly possible to select a sequence of open intervals

$$(a_1, b_1), (a_2, b_2), (a_3, b_3), (a_4, b_4), \ldots$$

that covers the set E and so that

$$\sum_{k=1}^{\infty} (b_k - a_k) < \varepsilon.$$

But how can we reduce this collection to a finite one that also covers the set E? The reader who is already familiar the Heine-Borel theorem (Theorem 4.33) knows precisely how.

For readers who have skipped over that section we shall present here a proof that uses the Bolzano-Weierstrass theorem instead. We claim that we can find an integer N so that all points of E are in one of the intervals

$$(a_1, b_1), (a_2, b_2), (a_3, b_3), (a_4, b_4), \ldots (a_N, b_N).$$

This will prove the theorem.

We prove this by contradiction. If this is not so then for each integer $k = 1, 2, 3, \ldots$ we must be able to find a point $x_k \in E$ but x_k is not in any of the intervals

$$(a_1, b_1), (a_2, b_2), (a_3, b_3), (a_4, b_4), \ldots (a_k, b_k).$$

The sequence $\{x_k\}$ is bounded because E is bounded. By the Bolzano-Weierstrass theorem the sequence has a convergent subsequence $\{x_{n_j}\}$. Let z be the limit of the convergent subsequence. Since E is closed z is in E. The original sequence of intervals covers all of E and so there must be an interval (a_M, b_M) that contains z. For large values of j the points x_{n_j} also belong to (a_M, b_M). But this is impossible since x_{n_j} cannot belong to the interval (a_M, b_M) for $n_j \geq M$. Since this is a contradiction the proof is done. ■

Exercises

6:8.1 Show that every subset of a set of measure zero also has measure zero.

6:8.2 If E has measure zero show that the translated set

$$E + \alpha = \{x + \alpha : x \in E\}$$

also has measure zero.

6:8.3 If E has measure zero show that the expanded set

$$cE = \{cx : x \in E\}$$

also has measure zero for any $c > 0$.

6:8.4 If E has measure zero show that the reflected set

$$-E = \{-x : x \in E\}$$

also has measure zero.

6:8.5 Without referring to the proof of Theorem 6.34 show that the union of any two sets of measure zero also has measure zero.

6:8.6 If $E_1 \subset E_2$ and E_1 has measure zero but E_2 has not, what can you say about the set $E_2 \setminus E_1$?

6:8.7 Show that any interval (a, b) or $[a, b]$ is not of measure zero.

6:8.8 Give an example of a set that is not of measure zero and does not contain any interval $[a, b]$.

6:8.9 If a set E has measure zero is it true that the closure $\overline{E}$ must also have measure zero?

6:8.10 If a set E has measure zero what can you say about interior points of that set?

6:8.11 Suppose that a set E has the property that $E \cap [a, b]$ has measure zero for every compact interval $[a, b]$. Must E also have measure zero?

6:8.12 Show that the set of real numbers in the interval $[0, 1]$ that do not have a 7 in their infinite decimal expansion is of measure zero.

6:8.13 In Definition 6.30 show that closed intervals may be used without changing the definition.

6:8.14 Describe completely the class of sets E with the following property: for every $\varepsilon > 0$ there is a *finite* collection of open intervals

$$(a_1, b_1), (a_2, b_2), (a_3, b_3), (a_4, b_4), \ldots (a_N, b_N)$$

that covers the set E and so that

$$\sum_{k=1}^{N} (b_k - a_k) < \varepsilon.$$

(These sets are said to have *zero content.*)

6:8.15 Show that a set E has measure zero if and only if there is a sequence of intervals

$$(a_1, b_1), (a_2, b_2), (a_3, b_3), (a_4, b_4), \dots$$

so that every point in E belongs to infinitely many of the intervals and $\sum_{k=1}^{\infty}(b_k - a_k)$ converges.

6:8.16 By altering the construction of the Cantor set construct a nowhere dense closed subset of $[0, 1]$ so that the sum of the lengths of the intervals removed is not equal to 1. Will this set have measure zero?

6.9 Additional Problems for Chapter 6

6:9.1 Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous function. Assume that for every positive number ε the sequence $\{f(n\varepsilon)\}$ converges to zero as $n \to \infty$. Prove that

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

6:9.2 Let f_n be a sequence of continuous functions defined on an interval $[a, b]$ such that $\lim_{n\to\infty} f_n(x) = 0$ for each $x \in [a, b]$. Show that for any $\varepsilon > 0$ there is an interval $[c, d] \subset [a, b]$ and an integer N so that

$$|f_n(x)| < \varepsilon$$

for every $n \geq N$ and every $x \in [c, d]$. Show that this need not be true for $[c, d] = [a, b]$.

6:9.3 Let f_n be a sequence of continuous functions defined on an interval $[a, b]$ such that $\lim_{n\to\infty} f_n(x) = \infty$ for each $x \in [a, b]$. Show that for any $M > 0$ there is an interval $[c, d] \subset [a, b]$ and an integer N so that

$$f_n(x) > M$$

for every $n \geq N$ and every $x \in [c, d]$. Show that this need not be true for $[c, d] = [a, b]$.