DOCUMENT RESUME

ED 046 775 | 24 | SE 010 736

TITLE	Unified Modern Mathematics, Course 2, Part 2.
INSTITUTION	Secondary School Mathematics Curriculum Improvement Study, New York, N.Y.
SPONS AGENCY	Columbia Univ., New York, N.Y. Teachers College.; Office of Education (DHEW), Washington, D.C. Bureau of Research.
BUREAU NO	BR-7-0711
PUB DATE	68
CONTRACT	OEC-1-7-070711-4420
NOTE	304p.
EDRS PRICE	EDRS Price MF-$0.65 HC-$13.16
DESCRIPTORS	*Curriculum Development, Geometric Concepts, *Instructional Materials, Mathematics, *Modern Mathematics, Probability Theory, *Secondary School Mathematics, Statistics, *Textbooks, Transformations (Mathematics)

ABSTRACT

Topics included in Part 2 of Course II are: real functions; descriptive statistics; transformations in the plane; length, area, and volume; combinatorics; and mass points. The chapter on real functions includes a discussion of properties of functions, composition of functions, inverses of functions and other topics. The chapter on descriptive statistics discusses the graphical representation of sets of data, summation notation, the arithmetic mean, measures of dispersion, and Chebyshev's Inequality. Reflections, translations, rotations, dilations, and similarities are introduced in the section on transformations of the plane. Lengths of line segments, areas of various geometric regions, and volumes of geometric solids are also studied. The Combinatorics chapter considers the counting principle, permutations, and the binomial theorem. The appendix offers a discussion of mass points in the plane and in space. (FL)

Secondary School Mathematics

Curriculum Improvement Study

UNIFIED MODERN MATHEMATICS

COURSE II

PART II

TEACHERS COLLEGE COLUMBIA UNIVERSITY

Secondary School Mathematics

Curriculum Improvement Study

UNIFIED MODERN MATHEMATICS

Course II

Part II

Financial support for the Secondary School Mathematics Curriculum Improvement Study has been provided by the United States Office of Education and Teachers College, Columbia University.

UNIFIED MODERN MATHEMATICS, COURSE II was prepared by the Secondary School Mathematics Curriculum Improvement Study with the cooperation of

Nicholas Branca, Teachers College, Columbia University
Gustave Choquet, Universite de Paris, France
Ray Cleveland, University of Calgary, Canada
John Downes, Emory University
Howard F. Fehr, Teachers College, Columbia University
James Fey, Teachers College, Columbia University
Allan Gewirtz, City University of New York
Abraham Glicksman, Bronx High School of Science, New York
Vincent Haag, Franklin and Marshall College
Thomas Hill, University of Oklahoma
Peter Hilton, Cornell University
Julius Hlavaty, National Council of Teachers of Mathematics
Michael Hoban CFC, Iona College, New York
Meyer Jordan, City University of New York
Burt Kaufman, Southern Illinois University
Howard Kellogg, Teachers College, Columbia University
Erik Kristensen, Aarhus University, Denmark
Howard Levi, City University of New York
Edgar R. Lorch, Columbia University
Lennart Råde, Chalmers Institute of Technology, Sweden
Myron F. Rosskopf, Teachers College, Columbia University
Harry Ruderman, Hunter College High School, New York
Harry Sitomer, C. W. Post College
Hans-Georg Steiner, University of Karlsruhe, Germany
Marshall H. Stone, University of Massachusetts
Stanley Taback, New York University
H. Laverne Thomas, State University College at Oneonta, N. Y.
Albert W. Tucker, Princeton University
Bruce Vogeli, Teachers College, Columbia University
Lucian Wernick, Teachers College, Columbia University

CONTENTS

Chapter X: LENGTH, AREA, VOLUME

Chapter XI: COMBINATORICS

Appendix A: MASS POINTS

Chapter 7

REAL FUNCTIONS

7.1 Mathematical Mappings

The word mapping has a very special and important meaning in mathematics. You recall from Course I that if we are given two sets S and T and a process which assigns to each element in S a unique element in T, we say that this defines a mapping h of S to T. We write this in the form

$$h: \quad S \longrightarrow T.$$

The set S is called the domain of the mapping h, and the set T is called the codomain of h. If $s \in S$ and h assigns $t \in T$ to s, t is called the image of s and s is a pre-image of t. To indicate that h assigns t to s we write

$$s \xrightarrow{h} t.$$

Example 1. If the set S consists of the students in your school and the set T consists of the teachers in your school, the assignment of a homeroom teacher to each student in the school is a mapping h from S to T. Every student is assigned a homeroom teacher, and no student is assigned more than one teacher. There will be teachers who serve as the homeroom teacher for many students--perhaps 25 or 30--and there might be a teacher who does not have responsibility for a homeroom. However, each student is assigned a unique homeroom teacher by the mapping h.

The domain of h is the set of students in your school. The codomain of h is the set of teachers in your school.

Example 2. Another example of a mapping is the assignment of postal zip codes. In this case, set S is the set of all postal addresses in the United States. The set T could be the set of whole numbers. Each element of S is assigned one and only one element of T; that is, one and only one zip code number. There are whole numbers which do not serve as zip codes; for example, 1,267,893. There are many addresses which are assigned the same code number; for example, all homes in Wisconsin Rapids, Wisconsin have zip code 54494. But the important point is that each address is assigned one, but not more than one, zip code number.

Questions. (1) What is the domain of the zip code mapping?

(2) What is the codomain of the mapping?

In the first mapping illustrated, the letters "S," "T," and "h" have a natural relation to the sets and the mapping they symbolize:

S --------------- the set of students

T --------------- the set of teachers

h --------------- the set of homeroom assignments

In this example S and T are chosen for the moment, in this context, as names for particular sets. But the letter "S" is not bound

forever to be the name of a set of students. It is also used in Example 2 to name the set of postal addresses. Here, however, "S" is not suggestive of addresses so a different letter, "A," may be used to be suggestive of the set of addresses. "z" then is a natural candidate to represent the zip coding assignment. Using "W" to represent the set of whole numbers the second mapping can be indicated by

$$z: \quad A \longrightarrow W$$

By contrast to S, T, h, z, and A, we will always use W as a proper name. It represents the set of whole numbers each time that it is used.

Choosing meaningful symbols for the domain, the codomain, and the mapping is a convenient device when we deal with specific sets and mappings. When the domain or codomain is a familiar number system such as Z or Q, these names are proper names and always name the same set. By contrast, "S" and "T" may be used to name different sets in different examples or problems. But in each new situation it must be explained what sets the letters name.

The process which assigns an element in the codomain T to each element in domain S can be one of several types. If S contains only a few elements, a chart or table will give a concise summary of the assignments. For example, the assignment of additive inverses to elements of $(Z_5,+)$ is illustrated in Table 7.1.

$x \in Z_5$	Inverse of x
0	0
1	4
2	3
3	2
4	1

Table 7.1

In other cases there may be a rule which tells how, given an element of S, you can determine its image in T. For the home-room assignment mapping, the rule probably would involve alphabetical and age ranking. The first 30 ninth graders are assigned Mr. Anderson, the next 30 ninth graders are assigned Mr. Charles, and so on.

A third form of assignment process is the arrow diagram. Figure 7.1 (indicating only a few of the correspondences) illustrates that mapping p: $\ell_1 \longrightarrow \ell_2$ assigns A' as the image of A, B' as the image of B, and C' as the image of C by projection along lines parallel to ℓ_3.

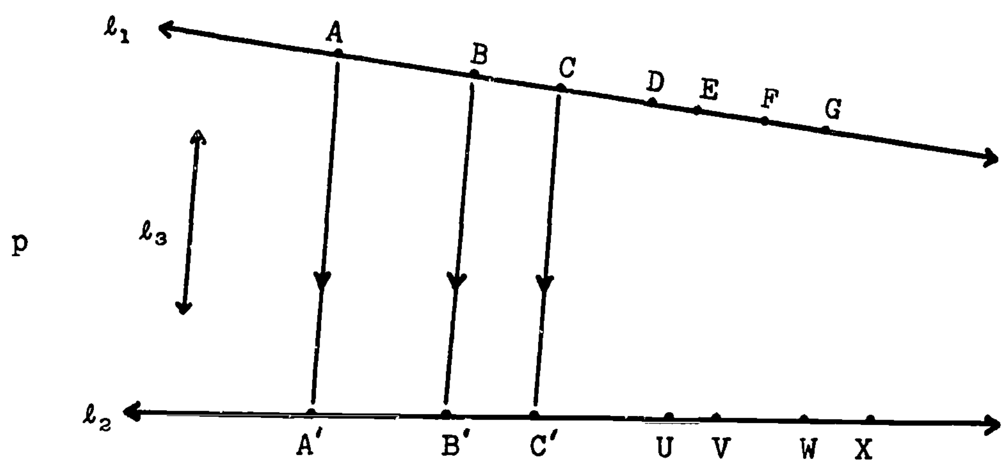


Figure 7.1

<u>Questions</u>. (1) Can you find an image for each point of ℓ_1? If so, how?

(2) Can you find a pre-image for each point of ℓ_2? If so, how?

A fourth common form of assignment process is that given by a formula. For example, if S and T are both the set of rational numbers Q, we can assign to each element in Q its double. Under the mapping d

$$-\frac{7}{2} \xrightarrow{d} -7$$

$$1.35 \xrightarrow{d} 27$$

or in general,

$$q \xrightarrow{d} 2q \text{ for all } q \in Q.$$

The following exercises test your ability to recognize assignments which are mappings and some which are not. You will also be asked to compute images and pre-images using various assignment processes.

7.2 <u>Exercises</u>

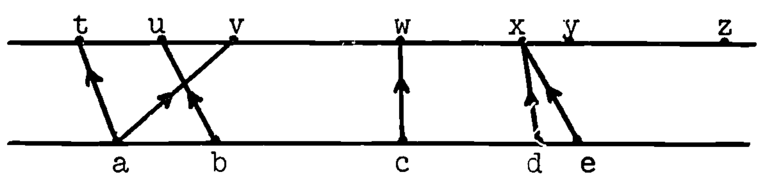
1. For each of the following explain why the given assignment process does or does not define a mapping from set S to set T.

(a)

S	6	2	3	10	7	5
T	1	12	2	9	2	12

(b) Students in mathematics 9X are assigned grades. S is the set of students, T = {A, B, C, D, F}.

(c)



$T = \{t, y, v, w, x, y, z\}$ $S = \{a, b, c, d, e\}$

(d) $S = T =$ the set of rational numbers with assignment process $q \xrightarrow{f} \frac{1}{q}$

(e)

S	T
1	6
12	2
2	3
9	10
2	7
12	5

2. If mapping g: $Q \longrightarrow Q$ has assignment process $x \xrightarrow{g} x + 2$, find the image under g of each of the following numbers.

(a) -12

(b) $-\frac{8}{5}$

(c) 12

(d) $\frac{8}{5}$

(e) 0

3. Using the mapping defined in Exercise 2, find a pre-image (if any exists) for each of the following numbers. (Remember: x is a pre-image of y if $x \xrightarrow{g} y$.)

(a) 6

(b) -2

(c) $-\frac{7}{2}$

(d) 156

(e) 0

4. Repeat the directions of Exercises 2 and 3 using the mapping

f: $Q \longrightarrow Q$ with assignment formula $x \xrightarrow{f} \frac{1}{2}x - 6$.

5. Repeat the directions of Exercises 2 and 3 using the mapping h: $Q \longrightarrow Q$ with assignment formula $x \xrightarrow{h} |x|$

7.3 Properties of Real Functions

As you saw in Section 7.1, the mathematical concept of mapping appears in a variety of settings with a variety of representations. In this section and for the remainder of the chapter, we will focus our attention on a special class of mappings -- those whose domain and codomain are both some subset (frequently all) of the real numbers, R. These mappings are called real functions. The term "function" is synonymous with "mapping," and the adjective "real" is used to indicate that the domain and range set are both subsets of the real numbers. Whenever the domain or codomain of a real function is not specified it is understood to be all of R.

The restriction to real functions may seem like a severe limitation of our study. But the real number system $(R,+,\cdot)$ is a system rich in assignments which are mappings. Since the whole numbers, integers, and rational numbers are subsets of R, many of the mappings you have studied previously are examples of real functions. One of these was the function which assigns to each real number its square,

$$x \xrightarrow{f} x^2.$$

To find the image of any real number x under this mapping, we simply compute $x \cdot x$. Thus

$$3 \xrightarrow{f} 9,$$

$$-7\tfrac{1}{2} \xrightarrow{f} \frac{225}{4},$$

$$60 \xrightarrow{f} 3600,$$

and $$-2 \xrightarrow{f} 4.$$

Instead of writing out the expression "$3 \xrightarrow{f} 9$," you recall from Chapter 2 we use the notation

$f(3) = 9$ (read: "f of 3 equals 9"),

to say that, "the image of 3 under the mapping f is 9." Following this notational convention we have

$$f(-7\tfrac{1}{2}) = \frac{225}{4}$$

$$f(60) = 3600$$

$$f(-2) = 4$$

and so on.

<u>Questions.</u> (1) What is $f(0)$? $f(\sqrt{2})$? $f(-\sqrt{7})$?

(2) Find a replacement of "x" such that $f(x) = \frac{4}{9}$.

If you thought carefully about question (2), you found that "x" could be replaced by $\frac{2}{3}$ or $-\frac{2}{3}$; that is, $\frac{4}{9}$ does not have a unique pre-image under the function f. In Exercise 5 of Section 7.2 the function $x \xrightarrow{h} |x|$ gave rise to a similar situation. The number 6 has two pre-images under the function h; namely, -6 and 6. On the other hand, for the function $x \xrightarrow{g} x + 2$, every real number has only one pre-image. The problem with functions like $x \xrightarrow{f} x^2$ and $x \xrightarrow{h} |x|$ is that they assign the same image to each of two distinct domain elements (see Figure 7.2).

Figure 7.2

The function $x \xrightarrow{g} x + 2$ assigns distinct images to distinct domain elements. For example, $g(2) = 4$, $g(-2) = 0$, $g(\frac{8}{5}) = \frac{18}{5}$, $g(-\frac{8}{5}) = \frac{2}{5}$, $g(0) = 2$, and so on.

This property, which distinguishes the function $x \xrightarrow{g} x + 2$ from the absolute value function and the square function, is an important one. Functions such as g, which assign distinct images to distinct domain elements, are called <u>one-to-one</u> functions. Each element of the domain is assigned its own private image.

<u>Definition</u>. A function $f\colon S \longrightarrow T$ is said to be <u>one-to-one</u> if and only if for all $a, b \in S$, $a \neq b$ implies that $f(a) \neq f(b)$.

If a function f is not one-to-one, this fact can be demonstrated by finding two elements a and b of S such that $a \neq b$ but $f(a) = f(b)$; that is, two distinct elements of the domain which are assigned the same image. For example, $x \xrightarrow{f} x^2$ is not one-to-one because $2 \neq -2$ but $f(2) = f(-2) = 4$. Similarly, $x \xrightarrow{h} |x|$ is not one-to-one because $3 \neq -3$ but $h(3) = h(-3) = 3$.

If you suspect that a function is one-to-one, one way to prove this is the case is to calculate the images of all domain elements and check to see that they are distinct. For example, inspection of Table 7.2 shows quickly that it does define a one-to-one mapping.

S	1	2	3	4	5	6	7
T	3	5	7	9	11	13	15

Table 7.2

If the domain of a function happens to be a large finite set, this procedure will be of little use. If the domain is an infinite set, proof by this approach is impossible. Checking images of several domain elements can give evidence (but not proof) that classification of the function as one-to-one is probably correct.

The function $x \xrightarrow{g} x + 2$ can be distinguished from $x \xrightarrow{f} x^2$ and $x \xrightarrow{h} |x|$ by one other important property. Although all three functions have the same domain and codomain R, the images under f and h are always positive numbers or zero. For example,

$$f(-2) = 4, \qquad h(-2) = 2,$$
$$f(-16) = 256 \qquad h(-16) = 16,$$
$$f(-\sqrt{7}) = 7, \text{ and } \quad h(-\sqrt{7}) = \sqrt{7}.$$

The function g uses every real number at least once as an image for a domain element.

10 has pre-image 8,
77 has pre-image 75
-32 has pre-image -34,
-752,466 has pre-image -752,468,

and so on. For this reason g is called a function from R <u>onto</u> R.

<u>Definition</u> 2. A function k: $S \longrightarrow T$ is a function of S <u>onto</u> T if and only if for each $t \in T$ there is at least one $s \in S$ for which $k(s) = t$.

If a particular function is not onto, this fact can be verified by exhibiting one element of the codomain which is not assigned as an image. For example, f: $R \longrightarrow R$ with rule $x \xrightarrow{f} x^2$ is not a function from R onto R because there is no real number whose square is -2. $((-2)(-2) \neq -2.)$

<u>Question</u>. Can you show that the mapping with rule $x \xrightarrow{h} |x|$ is not a function from R onto R?

As was the case with one-to-one functions, if a function is suspected to be onto, this is usually not easy to prove. If the codomain of the function is finite, it may be possible to check that each element is used as an image. Table 7.3 illustrates a mapping from Z_6 to Z_6, clearly showing it is an onto function.

k:

Z_6	0	1	2	3	4	5
Z_6	2	3	4	5	0	1

Table 7.3

<u>Questions</u>. (1) Is k also one-to-one?

(2) Is there any function from Z_6 <u>onto</u> Z_6 which is not one-to-one?

(3) Is there any function from Z_6 to Z_6 which is one-to-one but not onto?

If the codomain of the given function is an infinite set (or a large finite set) it is impossible to check all codomain

elements. One approach would be to select elements at random from the codomain and check whether or not they serve as images. However, this would only give evidence, not proof, that the mapping is onto.

When a mapping is not onto its codomain, it is often important to specify and name the elements of the codomain which do serve as images.

Definition 3. If k: $S \longrightarrow T$, the range of k is the set of all $t \in T$ for which there is an $s \in S$ such that $k(s) = t$.

According to this definition, the range of $x \xrightarrow{f} x^2$ and of $x \xrightarrow{h} |x|$ is $\{x: x \in R \text{ and } x \geq 0\}$. The range of $x \xrightarrow{g} x + 2$ is R, the same as the codomain of g. In fact, if you put together the definitions of codomain, range, and onto function, you will see that a function is onto if and only if its codomain and range are the same set.

The following exercises are concerned with one-to-one and onto real functions.

7.4 Exercises

1. Let f: $R \longrightarrow R$ have rule of assignment $x \xrightarrow{f} |x| + 2$. For example, $f(-3) = |-3| + 2 = 3 + 2 = 5$.

   (a) Find standard names for:

(i) $f(0)$	(iv) $f(-7\frac{1}{2})$
(ii) $f(5)$	(v) $f(27)$
(iii) $f(-5)$	(vi) $f(-632)$

(b) Find a pre-image (if there is one) for:

(i) 10 (iv) 0

(ii) -2 (v) -7

(iii) 4 (vi) 2

(c) Describe the range of f.

(d) Is f: $R \longrightarrow R$ an onto function? Why or why not?

(e) Does each element of the range have only one pre-image?

(f) Is f: $R \longrightarrow R$ a one-to-one function? Why or why not?

2. Let g: $R \longrightarrow R$ have rule of assignment $x \xrightarrow{g} -x$.

(a) Find standard names for:

(i) $g(0)$ (iv) $g(|-5|)$

(ii) $g(7\frac{1}{3})$ (v) $g(\sqrt{2})$

(iii) $g(-2)$ (vi) $g(-\pi)$

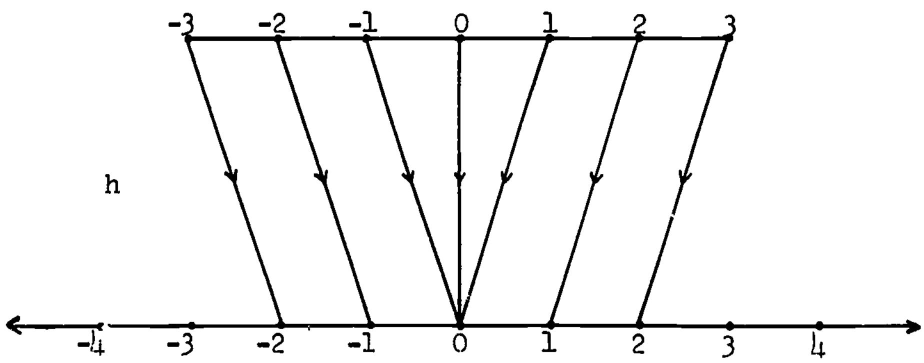
(b) Is g: $R \longrightarrow R$ a one-to-one function? Why or why not?

(c) What is the range of g?

(d) Is g: $R \longrightarrow R$ an onto function? Why or why not?

(e) Are there any real numbers x for which $g(x) = x$?

3. Let h: $\{-3, -2, -1, 0, 1, 2, 3\} \longrightarrow Z$ have the rule of assignment given by the following arrow diagram.



(a) Find standard names for:

(i) $h(-2)$ (iv) $h(0)$

(ii) $h(-1)$ (v) $h(3)$

(iii) $h(1)$

(b) Find a pre-image (if there is one) for:

(i) -1 (ii) 2 (iii) 0

(c) Is h a one-to-one function? Why or why not?

(d) Describe the range of h.

(e) Is h an onto function? Why or why not?

4. Mappings other than real functions can also be classified as one-to-one, onto, or both. Recall the zip code mapping of Section 7.1, $z: A \longrightarrow W$.

(a) Is z a one-to-one mapping? Why or why not?

(b) Is z an onto mapping? If not, describe the range of z.

5. Let $S = \{x: x \in R \text{ and } 1 \leq x \leq 2\}$ and $T = \{x: x \in R \text{ and } 0 \leq x \leq 1\}$. $k: S \longrightarrow T$ is the function with rule of assignment $x \xrightarrow{k} \frac{1}{x}$.

(a) Find the standard names for:

(i) $k(1)$ (v) $k(\frac{9}{8})$

(ii) $k(1\frac{1}{4})$ (vi) $k(\frac{15}{8})$

(iii) $k(1\frac{1}{2})$ (vii) $k(\frac{13}{8})$

(iv) $k(1\frac{3}{4})$ (viii) $k(2)$

(b) Find a pre-image (if there is one) for:

(i) $\frac{3}{4}$ (iv) 1

(ii) $\frac{2}{3}$ (v) $\frac{1}{3}$

(iii) 0 (vi) $\frac{221}{222}$

(c) Is k a one-to-one function? Why or why not?

(d) What is the range of k?

(e) Is k an onto function? Why or why not?

6. Examine the function $x \xrightarrow{f} -\frac{1}{3}x + 2$; that is calculate some images and pre-images to get an idea of the action of the function.

(a) Do you think f: $R \longrightarrow R$ is one-to-one?

(b) Do you think f: $R \longrightarrow R$ is onto?

Be prepared to defend your conclusions.

7. Answer (a) through (d) as true(T) or false(F).

If f: $A \longrightarrow B$ is a real function with range C, then

(a) $C \subset B$ always.

(b) $C \in B$ always.

(c) $C = B$ implies that f is onto.

(d) $B \subset C$ implies that f is one-to-one.

7.5 Representing Real Functions

When mappings of W and Z were discussed in the first course, arrow diagrams were a convenient device for picturing the assignment process. For instance, the mapping (it could also be called a function) d: $Z \longrightarrow Z$ with rule of assignment $x \xrightarrow{d} 2x$ could be partially represented as an incomplete arrow diagram on a single number line as in Figure 7.3(a) or as an arrow diagram between two number lines as in Figure 7.3(b).

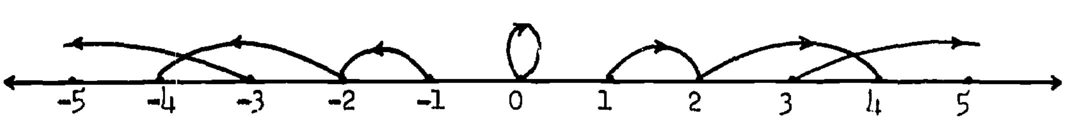


Figure 7.3(a)

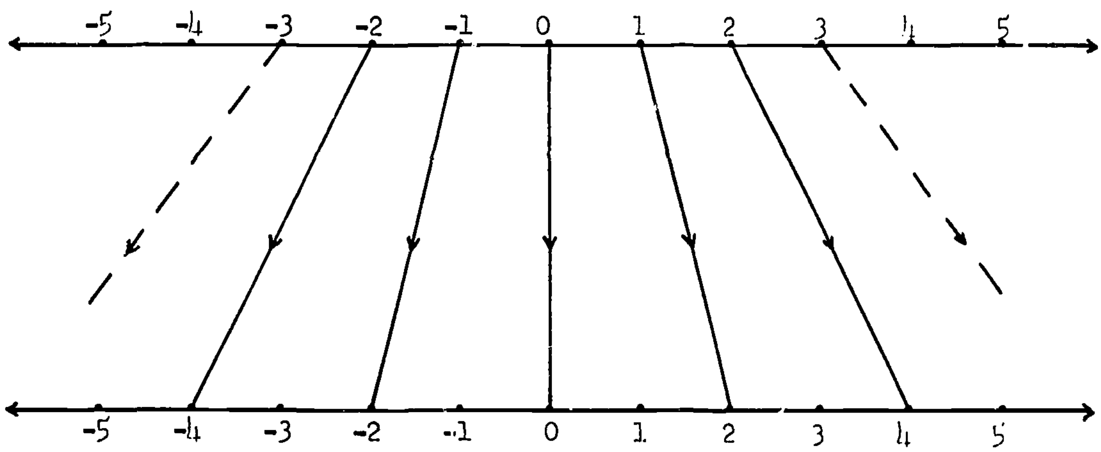


Figure 7.3(b)

However, when functions have R or even Q for domain and co-domain, the arrow diagram is a misleading, or at best incomplete, picture of the function, since only a few assignments can be indicated. Fortunately, a better tool is available.

Now return to the function f: $R \longrightarrow R$ with rule of assignment $x \xrightarrow{f} x^3$. The domain and range of f are both infinite sets. An incomplete arrow diagram is shown in Figure 7.4.

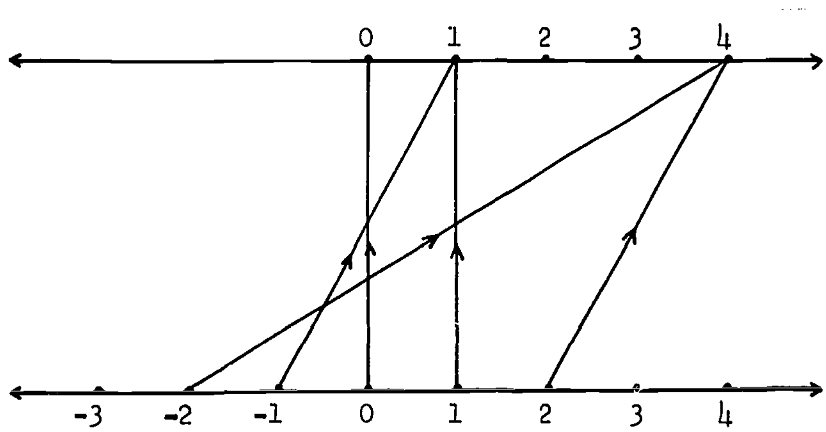


Figure 7.4

But this only shows five assignments. What happens to all the domain numbers between 0 and 1? between 1 and 2? between -1 and -2? greater than 2?

The basic problem is representing an infinite number of assignments with a drawing of limited size. To get a rough picture of the function, let's first make a table of some assignments.

x	-3	$-2\frac{1}{2}$	-2	$-1\frac{1}{2}$	-1	$-\frac{1}{2}$	0	$\frac{1}{2}$	1	$1\frac{1}{2}$	2	$2\frac{1}{2}$	3
$f(x)$	9	$\frac{25}{4}$	4	$\frac{9}{4}$	1	$\frac{1}{4}$	0	$\frac{1}{4}$	1	$\frac{9}{4}$	4	$\frac{25}{4}$	9

Table 7.4

Table 7.4 indicates only a small number of the assignments of f, but it does have a form that suggests different representation procedures.

The function f determines a collection of ordered pairs of numbers -- each real number paired with its square. We can write

$$(3, 9), \qquad (-\tfrac{1}{2}, \tfrac{1}{4}), \qquad (2\tfrac{1}{2}, \tfrac{25}{4}),$$

or, for any $x \in R$, $(x, f(x)) = (x, x^2)$.

From your work in coordinate geometry you know that the set of all ordered pairs of real numbers, $R \times R$, can be represented by the points of a coordinatized plane. Therefore, if we locate on a coordinatized plane those points which represent ordered pairs generated by the function $x \xrightarrow{f} x^2$, we will have a picture or graph of the function. Unless otherwise specified when graphing a function, we will label the domain axis "x," and the codomain axis "y."

As a start, let's locate (see Figure 7.5) the points representing integer pairs: (-3,9), (-2,4), (-1,1), (0,0), (1,1), 4), and (3,9).

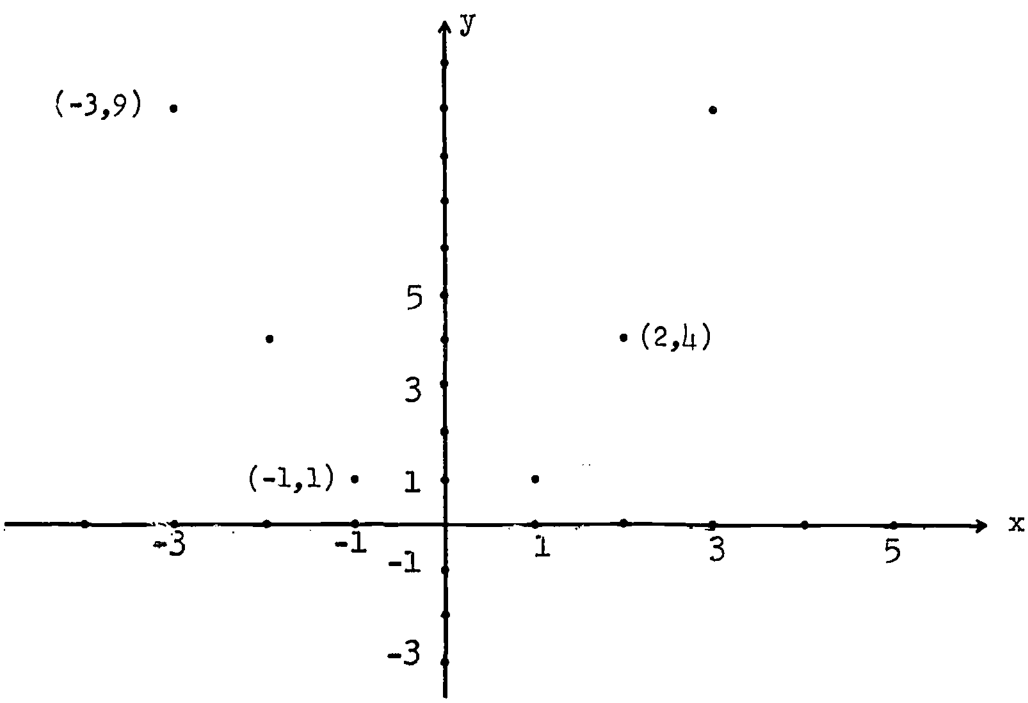


Figure 7.5

(Note: We have chosen perpendicular axes, but for convenience the units on the two axes are not equal. Can you see what would happen if the vertical unit were made as long as the present horizontal unit?)

This is a start, but we are far from finished. The points (4, 16), (-16, 256), (1,000, 1,000,000), and many others generated by f are not yet graphed. In fact, this last point would require a graph so large that we clearly must satisfy ourselves with representing only a limited number of the ordered pairs--perhaps $\{(x, x^2): -3 \leq x \leq 3\}$.

In one respect then, this new method of representing a function

has the same limitation as an arrow diagram. However, for x between -3 and 3 it will do very much better. Let's locate (Figure 7.6) the points corresponding to $(-2\frac{1}{2}, \frac{25}{4})$, $(-1\frac{1}{2}, \frac{9}{4})$, $(-\frac{1}{2}, \frac{1}{4})$, $(\frac{1}{2}, \frac{1}{4})$, $(1\frac{1}{2}, \frac{9}{4})$, $(2\frac{1}{2}, \frac{25}{4})$.

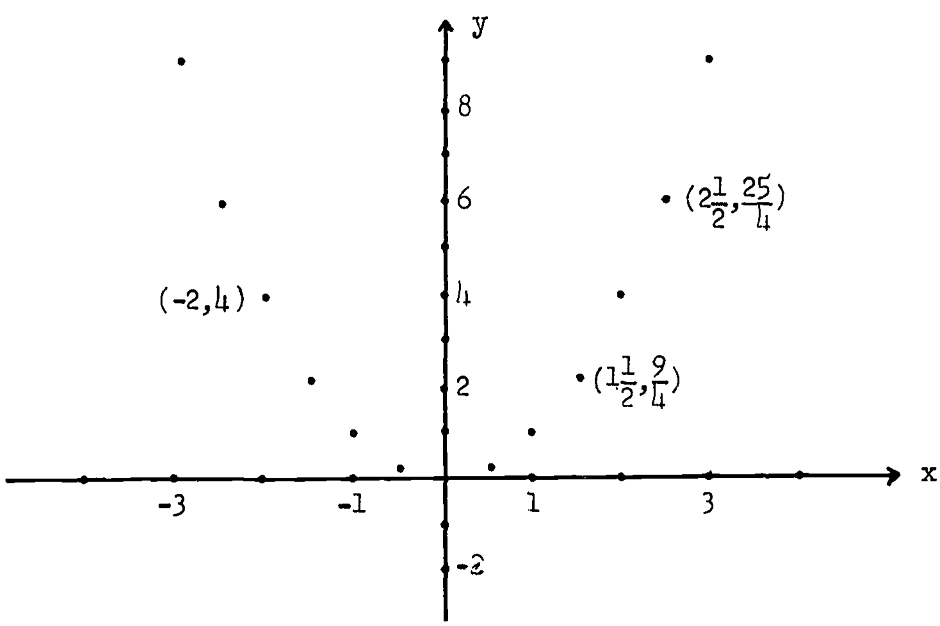


Figure 7.6

We now have a procedure for obtaining still more points in our graph of the function. We could next locate the points obtained when x increases from -3 by steps of $\frac{1}{4}$, then by steps of $\frac{1}{8}$, and so on. However, even if we had the patience to carry out these computations, we could not hope to obtain all points with rational number coordinates, much less the points generated by irrational numbers, such as $(\sqrt{2}, 2)$, $(\frac{\pi}{2}, \frac{\pi^2}{4})$, and so on.

The points already located give a strong indication of the

pattern remaining points will fit. Therefore, the standard procedure is to locate these remaining points without explicitly computing their coordinates. If a and b are both positive or both negative, and if x is between a and b, then x^2 is between a^2 and b^2. Thus it seems reasonable that the graph should look like that in Figure 7.7.

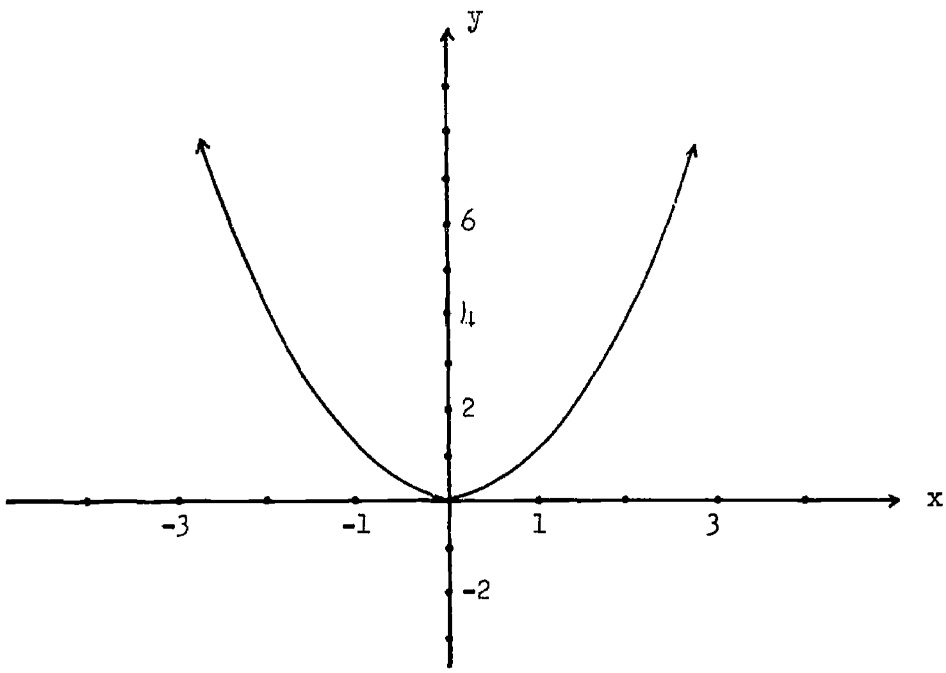


Figure 7.7

This method of representing a function by points in a coordinatized plane is called graphing the function or drawing the graph of the function.

Definition If $f\colon\ S \longrightarrow T$ is a real function, the graph of f is the set of all points in the plane with coordinates $(x, f(x))$ for $x \in S$.

Thus the graph of $x \xrightarrow{f} x^2$ is the set of all points in the plane with coordinates (x, x^2) for $x \in R$.

<u>Question</u>. How does the graph of $x \xrightarrow{f} x^2$ show that as $|x|$ gets larger, $f(x)$ gets larger (whether x is positive or negative)?

In addition to providing a picture of the function $x \xrightarrow{f} x^2$, graphing has an interesting bonus. The graph of f, if constructed carefully, allows us to calculate approximations to certain irrational numbers. For example, we can approximate $\sqrt{3}$ as follows: (See Figure 7.8.)

(1) $\sqrt{3}$ is the real number whose square is 3. Therefore, the point with coordinates $(\sqrt{3}, 3)$ lies on the graph of $x \xrightarrow{f} x^2$.

(2) To locate the point A with coordinates $(\sqrt{3}, 3)$ we move horizontally along a line 3 units above the x-axis until we meet the graph of the function. This line $f(x) = 3$ intersects the graph at two points. However, since $\sqrt{3}$ is by definition a positive number, we choose the point of intersection over the positive x-axis.

(3) To locate the point with x-coordinate $\sqrt{3}$ on the x-axis, we move vertically from $A(\sqrt{3}, 3)$ until we intersect that axis.

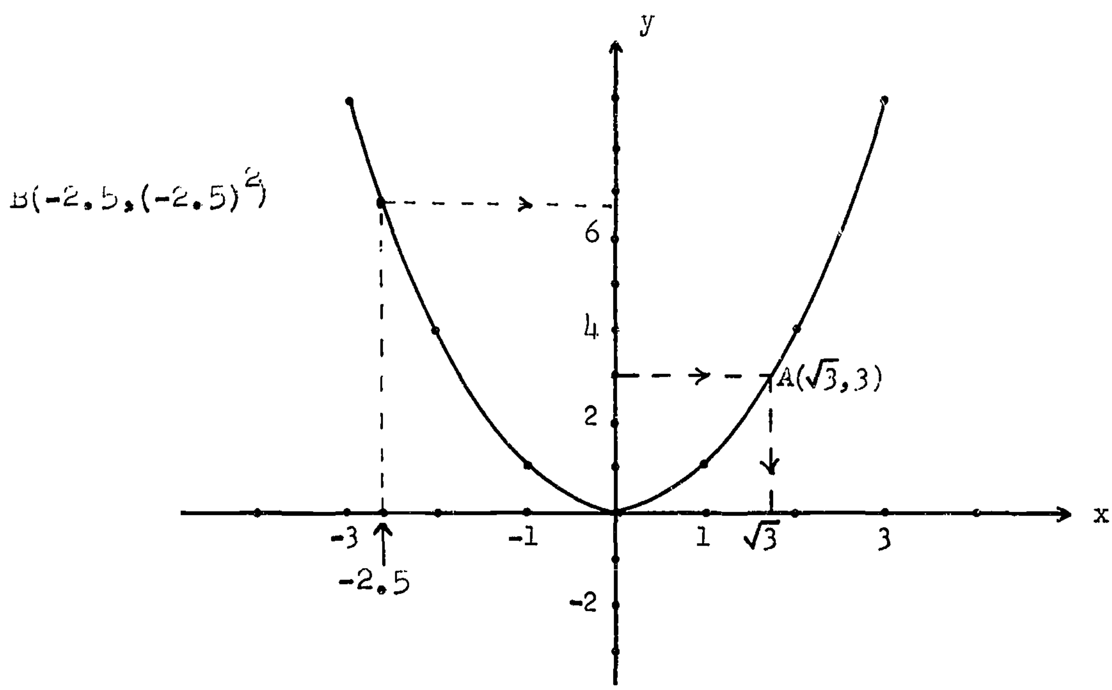


Figure 7.8

On the given graph $\sqrt{3}$ is located at approximately $1\frac{3}{4}$. Checking,

$$(1\tfrac{3}{4})^2 = \frac{49}{16} = 3\frac{1}{16},$$

so $1\frac{3}{4}$ is a reasonable approximation of $\sqrt{3}$.

What we have done is to use the graph of f to locate a positive pre-image of 3. The graph can also be used to locate images of numbers in the domain of f. For example, to find f(-2.5), begin at the point on the x-axis with coordinate -2.5. Then follow the vertical line through that point until it intersects the graph of f. To locate the point with coordinate $(-2.5)^2$ on the y-axis, we move horizontally from B(-2.5, $(-2.5)^2$) until we intersect the y-axis. The y coordinate there is approximately 6.3. Since $(-2.5)^2$ = 6.25, this is a reasonable approximation. Note that in this

process there is no choice as to the point of the graph to be used since a function assigns exactly one image to each element in its domain.

What we have done for the function $x \xrightarrow{f} x^2$ may be done for any real function whatsoever. One note of caution! After we located 13 points, we "filled in the graph" assuming that the pattern already established would continue. In the exercises you will be asked to check that this is indeed the case (at least for a number of other points), with that function. However, you will also encounter several functions that might fool you if you are not careful.

7.6 Exercises

In these exercises, and hereafter, we will write "the point (a, b)" to mean the point with coordinates (a, b).

1. From the following graph of $x \xrightarrow{f} x^2$, determine approximate coordinates of:

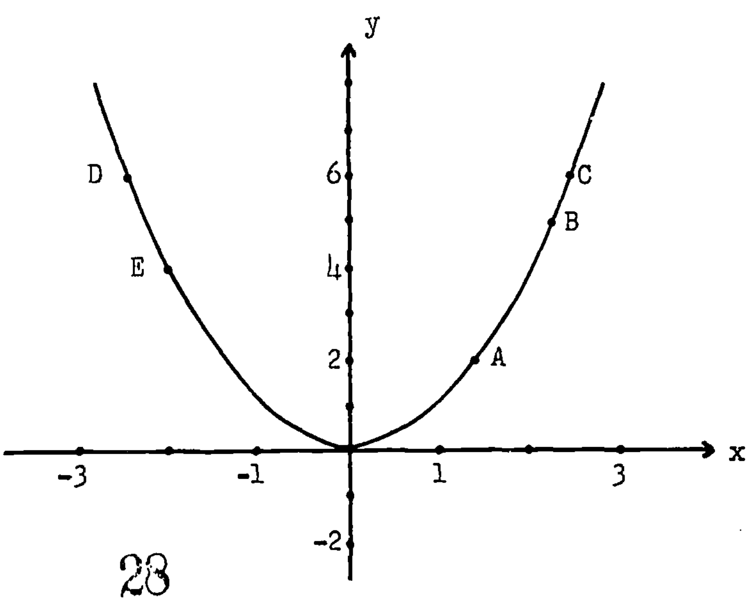
   (a) A

   (b) B

   (c) C

   (d) D

   (e) E



2. Using the graph in Exercise 1, compute approximate values for:

   (a) $\sqrt{2}$ (b) $\sqrt{5}$ (c) $\sqrt{6}$ (d) $\sqrt{7}$

   Then check each approximation by squaring it.

3. For each approximation in Exercise 2, one measure of the error can be calculated as follows: $|(\text{your estimate of } \sqrt{n})^2 - n|$.

   For example $|(1\frac{3}{4})^2 - 3| = \frac{1}{16}$. For each of your estimates, find this measure of error.

4. Let g: $R \longrightarrow R$ have rule $x \xrightarrow{g} 3x + 2$.

   (a) Complete the following table.

x	0		3	-3
g(x)	2	0		

   (b) Construct a pair of equally scaled perpendicular coordinate axes.

   (c) Locate the points $(x, g(x))$ generated in (a) on the set of axes.

   (d) Join the located points in the pattern you feel is likely to continue (i.e. fill in the graph).

   (e) Find standard names for $g(5)$, $g(-5)$, and $g(\frac{1}{3})$.

   (f) Locate $(5, g(5))$, $(-5, g(-5))$, and $(\frac{1}{3}, g(\frac{1}{3}))$ on your graph of g.

5. Let h: $R \longrightarrow R$ have rule $x \xrightarrow{h} |x|$.

   (a) Complete the following table.

x	0	2	4	6
h(x)				

   (b) Locate the points $(0, h(0))$, $(2, h(2))$, and $(6, h(6))$ in a plane rectangular coordinate system; that is, a

plane coordinatized with equally scaled perpendicular axes. (See Chapter 6, Section 6.20.)

(c) Complete the graph in the pattern you feel is likely to continue.

(d) Locate the points $(-1, h(-1))$, $(-3, h(-3))$, and $(-5, h(-5))$.

(e) Is the graph of h the graph you drew for part (c)?

6. Let p_1: $R \longrightarrow R$ be the function which assigns to each real number x the nearest integer greater than or equal to x. For example, $p_1(\frac{1}{2})$ 1, $p_1(1\frac{1}{2}) = 2$, $p_1(-\frac{1}{2}) = 0$, $p_1(7) = 7$, and so on.

(a) Complete the following table

x	$-4\frac{1}{2}$	$-3\frac{1}{2}$	$-2\frac{1}{2}$	$-1\frac{1}{2}$	$-\frac{1}{2}$	$\frac{1}{2}$	$1\frac{1}{2}$	$2\frac{1}{2}$	$3\frac{1}{2}$	$4\frac{1}{2}$
$p_1(x)$					0	1	2			

(b) Locate the points $(x, p_1(x))$ generated in (a) in a plane rectangular coordinate system.

(c) Complete the graph in the pattern you feel is likely to continue.

(d) Find standard names for:

(i) $p_1(1)$

(ii) $p_1(\frac{3}{4})$

(iii) $p_1(\frac{1}{4})$

(iv) $p_1(-\frac{1}{4})$

(v) $p_1(-\frac{3}{4})$

(vi) $p_1(1\frac{1}{4})$

(vii) $p_1(1\frac{3}{4})$

(viii) $p_1(-2\frac{1}{4})$

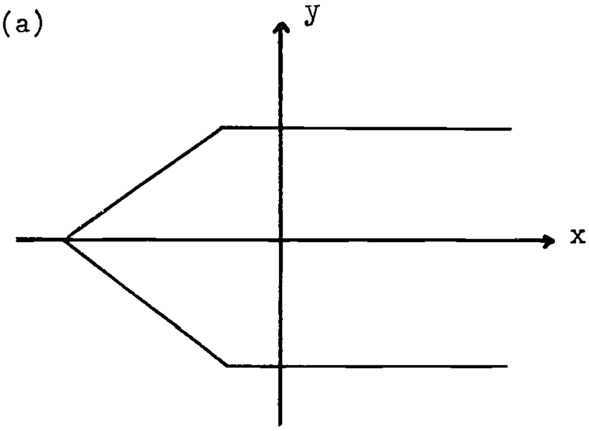
(e) Locate the points $(x, p_1(x))$ calculated in (d) on your graph of part (c).

Graph the function k: $R \longrightarrow R$ with rule $x \xrightarrow{k} -2x - 3$.

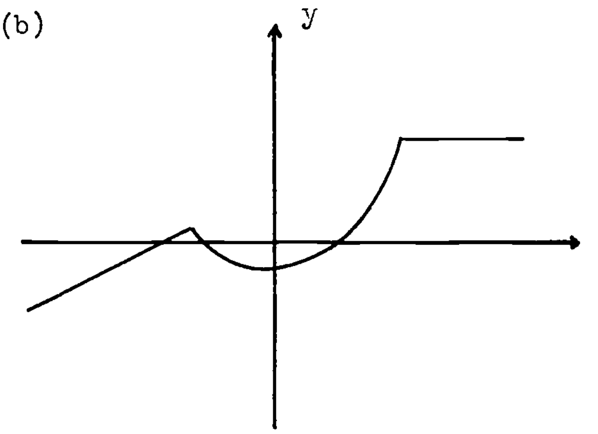
8. Graph the function $m\colon R \longrightarrow R$ with rule $x \xrightarrow{m} x^2 - 2$.

9. Which of the following graphs are graphs of functions?

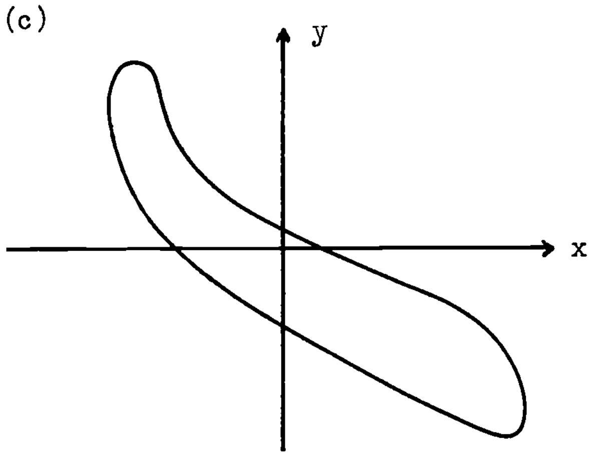
(a)



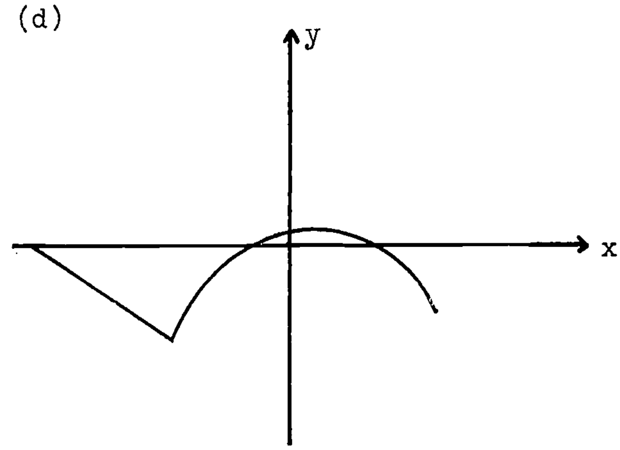
(b)



(c)



(d)



10. Try to formulate a geometric rule for determining whether or not a graph represents a function.

11. Which of the following functions from R to R are one-to-one? (Explain each answer.)

(a) $x \longrightarrow x^2$

(b) $x \longrightarrow 3x + 2$

(c) $x \longrightarrow |x|$

(d) $x \longrightarrow p(x)$ (See Exercise 6.)

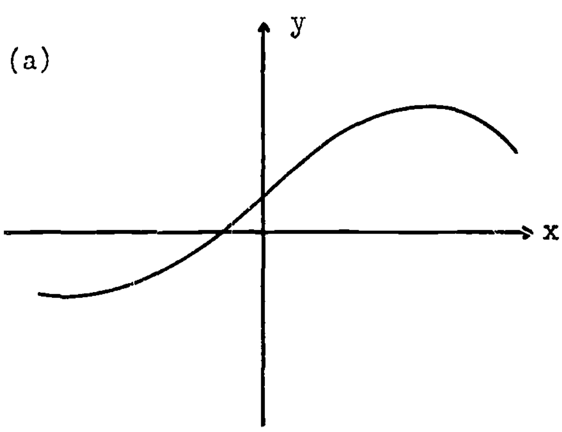
(e) $x \longrightarrow -2x - 3$

(f) $x \longrightarrow x^2 - 2$

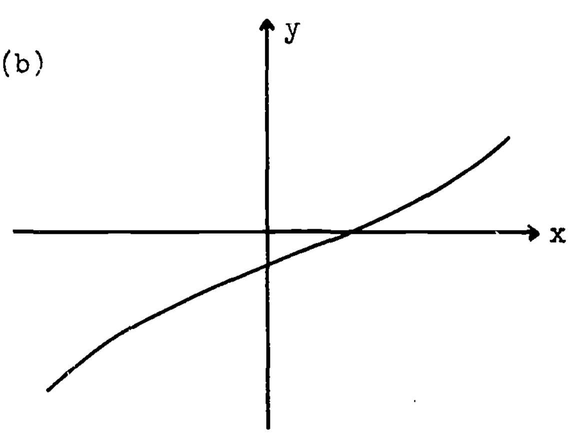
12. Inspect the graphs of the functions listed in Exercise 11. Try to develop a geometric rule for determining whether or not a graph represents a one-to-one function.

13. Which of the following are graphs of one-to-one functions? Why?

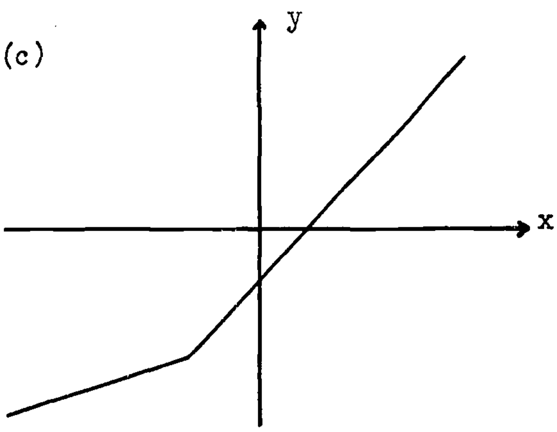
(a)



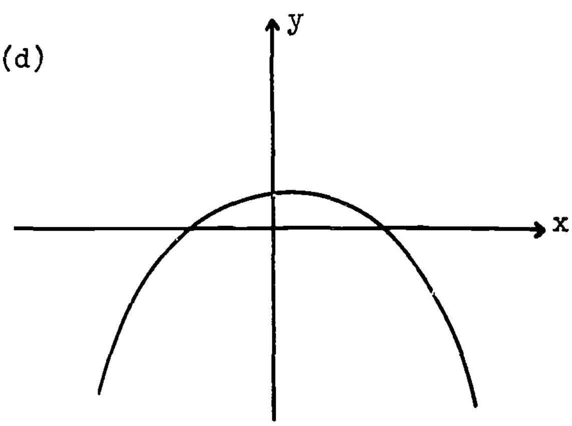
(b)



(c)



(d)



7.7 Composition of Real Functions

The interesting function p_1: $R \longrightarrow R$ which appears in Exercise 6 of Section 7.6 has a familiar interpretation. Called the postal function, it is the assignment process used in calculating postal charges for letters. Since postal rates are figured on the basis of a cost per ounce or partial ounce, it is necessary to round off weight measures in ounces to whole numbers in the following manner:

(1) Weights between 0 and 1 (including 1) round off to 1;

(2) Weights between 1 and 2 (including 2) round off to 2;

(3) Weights between 2 and 3 (including 3) round off to 3;

and so on.

The function p: $R^+ \longrightarrow R$ with $p(x) = p_1(x)$ for all $x \in R^+$ (p is called a restriction of p_1) satisfies the postal weighing requirements. For example,

$$p(\tfrac{3}{4}) = 1, \qquad p(1) = 1, \qquad p(1\tfrac{1}{3}) = 2, \qquad p(10\tfrac{1}{4}) = 11,$$

and so on.

Calculating the weight in whole ounces of a letter is the first step in determining the postage required. This whole number weight must then be multiplied by the rate per ounce, currently 6 cents for first class mail. In other words, the function $n \xrightarrow{r} 6n$ is used. The two step procedure can be summarized as follows:

$$\text{actual weight} \xrightarrow{p} \text{whole ounce weight} \xrightarrow{r} \text{postage charge}$$

In practice these functions produce

$$\tfrac{1}{3} \xrightarrow{p} 1 \xrightarrow{r} 6$$

$$3\tfrac{1}{4} \xrightarrow{p} 4 \xrightarrow{r} 24$$

$$17\tfrac{2}{3} \xrightarrow{p} 18 \xrightarrow{r} 108$$

or in general,

$$x \xrightarrow{p} p(x) \xrightarrow{r} r(p(x)).$$

This application of two functions in sequence should be familiar. It is really composition of functions. You recall that if $f\colon\ A \longrightarrow B$ and $g\colon\ B \longrightarrow C$ are mappings, there is a composite mapping $h\colon\ A \longrightarrow C$ which assigns images as follows: If $f(a) = b$ and $g(b) = c$, then $h(a) = g(f(a))$ or $h(a) = c$. The fact that h is the composite of g with f (or g following f) is indicated

$$h = g \circ f. \quad \text{(See Figure 7.9.)}$$

$$\begin{tikzcd} a \arrow[r, "f"] \arrow[rr, bend right, "h"'] & b \arrow[r, "g"] & c \end{tikzcd}$$

Figure 7.9

h is a function which has the same domain as f, the same codomain as g, and makes the same image assignments as the two step process, g following f.

As another illustration of composition, consider functions $x \xrightarrow{f} x^2$ and $x \xrightarrow{g} 3x + 2$, where f and g have R for domain and codomain.

(1) $f(-2) = 4$ and $g(4) = 14$

so $g \circ f\ (-2) = 14$.

(2) $f(6) = 36$ and $g(36) = 110$

so $g^{o}f\ (6) = 110.$

(3) $f(-\sqrt{2}) = 2$ and $g(2) = 8$

so $g^{o}f\ (-\sqrt{2}) = 8.$

In general,

$f(x) = x^2$ and $g(x^2) = 3x^2 + 2$

so $g^{o}f\ (x) = 3x^2 + 2.$

The composite $g^{o}f$ is a function from R to R with rule of assignment

$$x \xrightarrow{g^{o}f} 3x^2 + 2.$$

Composing two real functions to obtain a third is similar to another familiar mathematical process. If F is the set of all functions with the real numbers for domain and codomain, then composition is a binary operation on F. $(F, ^{o})$ is an operational system because if f: $R \longrightarrow R$ and g: $R \longrightarrow R$, then $g^{o}f$ is a function from R to R.

For instance, if m and n are functions from R to R given by

$$x \xrightarrow{m} x - 2 \text{ and } x \xrightarrow{n} x + 3$$

then n o m is a function from R to R given by

$$x \xrightarrow{n\ o\ m} x + 1.$$

In a similar fashion,

(1) if $x \xrightarrow{m} 3x + 2$ and $x \xrightarrow{n} x^2$, then $x \xrightarrow{n\ o\ m} (3x + 2)^2.$

(2) if $x \xrightarrow{m} 5x$ and $x \xrightarrow{n} -x$, then $x \xrightarrow{n\ o\ m} -5x.$

The composite, n o m, is a function from R to R in each case. You should check these indicated compositions for various values of "x" to see that the rule of n o m makes the same assignments as

the two step process "n following m."

Since $(F,^{o})$ is an operational system, it is natural to ask what properties it has in common with other operational systems. Is ^{o} an associative operation? Is ^{o} a commutative operation? Is there an identity element for $(F,^{o})$? Are there inverses under ^{o} for each element of F? Before reading ahead, make a guess, based on your experience with functions, about the answers to these questions.

The easiest question to answer is that concerning the existence of an identity element in $(F,^{o})$. Consider the function $g\colon\ R \longrightarrow R$ with rule of assignment $x \xrightarrow{g} x$.

$$g(0) = 0,\ g(5) = 5,\ g(-11) = -11$$

and in general, $g(x) = x$. If g is composed with any other real function f, then for all x in R,

$$g \circ f(x) = g(f(x)) = f(x)$$

$$\text{and} \quad f \circ g(x) = f(g(x)) = f(x).$$

Therefore, $\qquad g \circ f = f \circ g = f.$

In a particular case let $x \xrightarrow{m} 3x + 2$. Then we have

$$-7\tfrac{1}{2} \xrightarrow{m} -20\tfrac{1}{2} \xrightarrow{g} -20\tfrac{1}{2}$$

and

$$-7\tfrac{1}{2} \xrightarrow{g} -7\tfrac{1}{2} \xrightarrow{m} -20\tfrac{1}{2}\ .$$

Since identity functions will be important in other situations, we make the following definition.

<u>Definition</u> 5. A real function $j\colon\ S \longrightarrow S$ given by $x \xrightarrow{j} x$ is called an <u>identity</u> <u>function</u> <u>on</u> <u>S</u>.

This definition actually defines an infinite number of identity functions, one for each choice of subset S of R. The identity function on W (written "j_W" is a different function from the identity function on Z (written "j_Z"). The difference, however, is in the domain and codomain, not in the method of assigning images.

Associativity of composition of functions is also easy to demonstrate. In fact, composition is associative for mappings in general, not just real functions, whenever the compositions are defined. Let's assume that four sets -- A, B, C, and D -- are given with mappings f: $A \longrightarrow B$, g: $B \longrightarrow C$, and h: $C \longrightarrow D$. (Note that the domain of g is the same as the codomain of f and the domain of h is the same as the codomain of g. This is necessary because g must assign images to all range elements of f and h must assign images to all range elements of g.) (See Figure 7.10.)

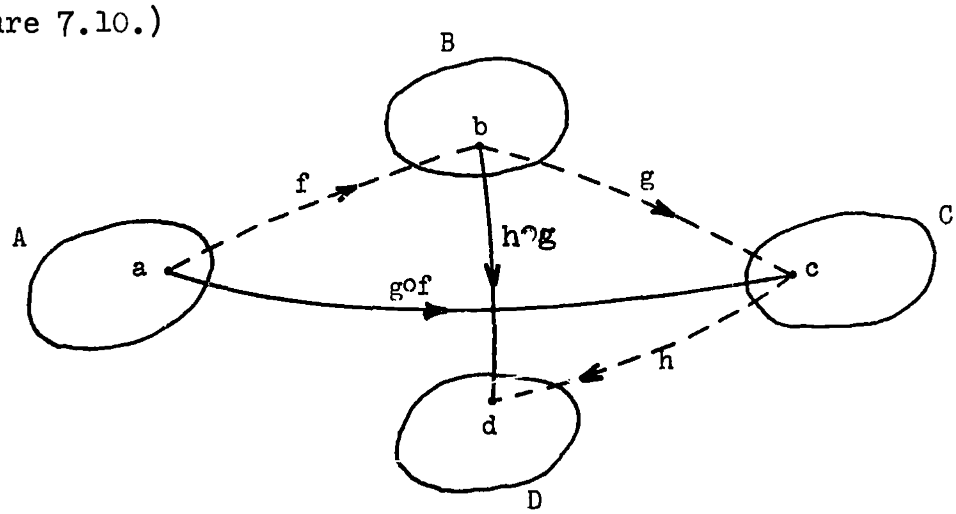


Figure 7.10

To show that $(h \circ g) \circ f = h \circ (g \circ f)$, we must show that the two functions assign images in the same way (clearly they have the same domain A and codomain D.) The picture illustrates the steps

in the proof. $(h \circ g) \circ f$ assigrs the image of a by following f to b and then $(h \circ g)$ directly to d. The function $h \circ (g \circ f)$ assigns the image of a by following $(g \circ f)$ directly to c and then proceeding to d by function h. Regardless of the procedure used -- $(h \circ g) \circ f$ or $h \circ (g \circ f)$ -- the final assignment of the image for a is the same.

Thus $\circ$ is an associative operation on F with identity j. The question of commutativity is considered in the following exercises. Inverses are the subject of Section 7.9.

7.8 Exercises

1. If $x \xrightarrow{f} x + 75$ and $x \xrightarrow{g} \frac{1}{3}x$, find a standard name for:

(a) $f(15)$	(f) $g \circ f(15)$
(b) $f(-30)$	(g) $g \circ f(-30)$
(c) $f(24.9)$	(h) $g \circ f(24.9)$
(d) $f(6)$	(i) $g \circ f(6)$
(e) $f(-25.8)$	(j) $g \circ f(-25.8)$

2. If $x \xrightarrow{f} x^2$ and p is the postal function, find a standard name for:

(a) $f(1\frac{1}{2})$	(g) $p \circ f(1\frac{1}{2})$
(b) $f(\frac{2}{3})$	(h) $p \circ f(\frac{2}{3})$
(c) $f(-\frac{7}{8})$	(i) $p \circ f(-\frac{7}{8})$
(d) $p(1\frac{1}{2})$	(j) $f \circ p(1\frac{1}{2})$
(e) $p(\frac{2}{3})$	(k) $f \circ p(\frac{2}{3})$
(f) $p(\frac{7}{8})$	(l) $f \circ p(\frac{7}{8})$

3. Copy and complete the following table for real functions $x \xrightarrow{h} x - 24.5$, and $x \xrightarrow{k} x + 15.75$.

x	h(x)	k(x)	h∘k(x)	k∘h(x)
0				
19				
-33				
-17.25				
3.14				
-2.7				

Does $h \circ k = k \circ h$?

4. Copy and complete the following table for real functions $x \xrightarrow{m} 18x$, $x \xrightarrow{n} x - 7$.

x	m(x)	n(x)	m∘n(x)	n∘m(x)
0				
43				
-15				
12				

Does $m \circ n = n \circ m$?

5. Is $\circ$ a commutative operation on F? Why or why not?

6. Calculate first class postage for letters weighing:

(a) $\frac{7}{8}$ ounce (b) $\frac{5}{16}$ pound (c) $3\frac{1}{3}$ ounces (d) $\frac{1}{32}$ pound

7. To calculate airmail charges, one needs a function k which counts the weight of a letter in number of half-ounces plus a fractional part (if there is one). For example, $k(\frac{1}{3}) = 1$, $k(\frac{3}{5}) = 2$, $k(1\frac{7}{8}) = 4$, etc.

(a) Find the standard name of:

(i) $k(7)$ (v) $k(1\frac{1}{2})$

(ii) $k(9\frac{5}{8})$ (vi) $k(2\frac{1}{3})$

(iii) $k(3\frac{2}{3})$ (vii) $k(2)$

(iv) $k(\frac{1}{16})$ (viii) $k(3)$

(b) Graph $k: S \longrightarrow R$ where S is the set of real numbers greater than zero and less than 5. (Be careful!)

(c) Airmail letters currently cost 10 cents per half-ounce. At this rate, what is the postage cost of letters weighing:

(i) $5\frac{1}{4}$ ounces (ii) $\frac{7}{8}$ ounce (iii) $3\frac{2}{3}$ ounces

(d) Which of the following (if any) relates the airmail function k to the first class function p? For all $x > 0$:

(i) $k(x) = p(\frac{1}{2}x)$ (ii) $k(x) = p(2x)$ (iii) $k(x) = \frac{1}{2}p(x)$ (iv) $k(x) = 2p(x)$

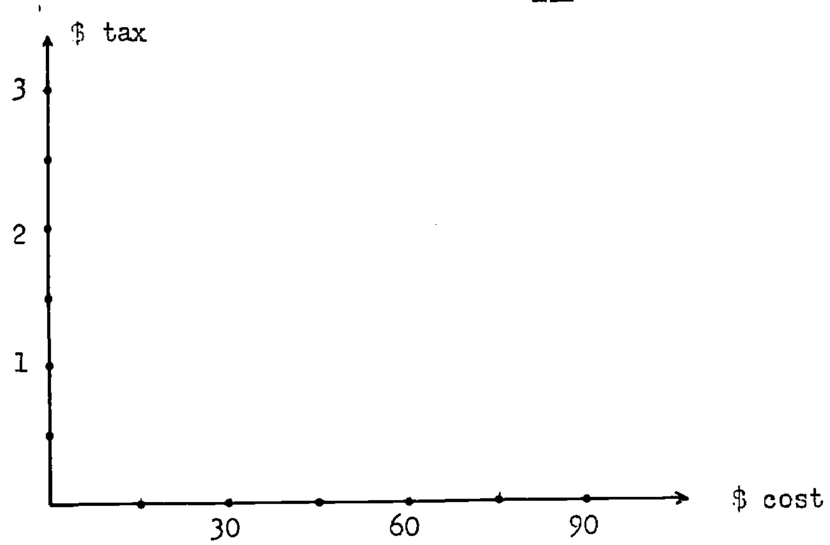
8. In New York, as in most states, there is a 3% state sales tax. The function $x \xrightarrow{t} .03x$ assigns to each purchase price the corresponding tax.

(a) Compute the tax $t(x)$ on items costing

(i) \$ 5.00 (ii) \$ 4.30

(iii) \$ 17.25 (iv) \$ 99.95

(b) Graph the function $t: S \longrightarrow R$ where S is the set of real numbers between 0 and 90 using scales like the following:



9. In practice, the tax calculated as .03x must be rounded off to the nearest penny. If $r\colon R^+ \longrightarrow R$ is the desired rounding function, r(.135) = .14, r(17.133) = 17.13, etc.

(a) Find a standard name of:

(i) r(95.999) (ii) r(32.095) (iii) r(762.012)

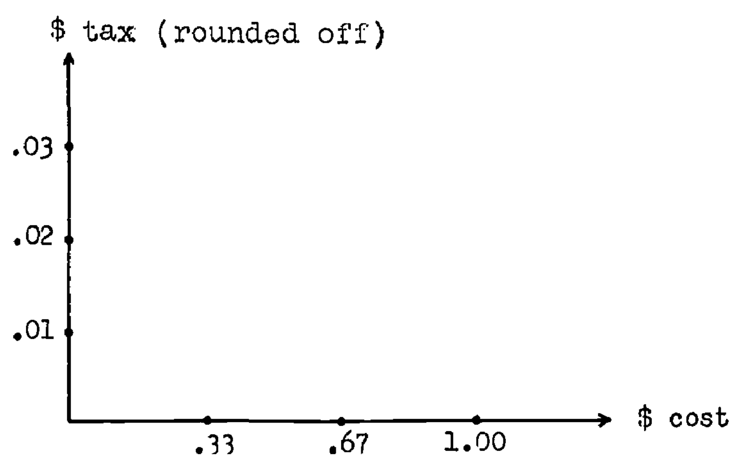
(b) Find a standard name of:

(i) $r \circ t(.10)$ (ii) $r \circ t(.16)$ (iii) $r \circ t(.17)$

(iv) $r \circ t(.49)$ (v) $r \circ t(.50)$ (vi) $r \circ t(.83)$

(vii) $r \circ t(.84)$

(c) Each computation in (b) yields an ordered pair $(x, r \circ t(x))$. Draw the graph of these pairs using a scale like the following.



Does this graph of $r \circ t$ resemble any other familiar graph?

10. Let $x \xrightarrow{f} 3x + 2$ and $x \xrightarrow{g} \frac{1}{3}(x - 2)$ be functions from R to R.

(a) Find standard names for:

(i) $f(0)$	(iv) $g(2)$
(ii) $f(\frac{5}{3})$	(v) $g(7)$
(iii) $f(-2)$	(vi) $g(-4)$

(b) Find standard names for:

(i) $g(f(0))$ (ii) $g(f(\frac{5}{3}))$ (iii) $g(f(-2))$

(c) Find the rule for the composite function $g \circ f$.

(d) Find the rule for $f \circ g$.

11. A mimeographing service advertized the following prices for printing copies of term papers, reports, and similar items:

(a) 50 cents per page for mimeo stencil

(b) 1 cent per printed page for printing

What is the rule which assigns to each whole number n the cost of printing 40 copies of an n page paper?

7.9 Inverses of Real Functions

The set F of all functions from R to R is an operational system under composition. $\circ$ is associative and has an identity j_R. Therefore, if it can be shown that each element of F has an inverse under $\circ$ -- that is, if for each $f \in F$ there is a $g \in F$ satisfying $f \circ g = g \circ f = j_R$ -- $(F, \circ)$ can be called a group. But it this possible?

Let's look at some simple functions in F and try to find their inverses. For example, let $x \xrightarrow{f} x + 5$. This function adds 5

to every real number. Thus a natural choice for the inverse of f is $x \xrightarrow{g} x - 5$ which subtracts 5 from every real number.

$$0 \xrightarrow{f} 5 \xrightarrow{g} 0$$

$$10 \xrightarrow{f} 15 \xrightarrow{g} 10$$

$$-10 \xrightarrow{f} -5 \xrightarrow{g} -10$$

For all real numbers x,

$$g \circ f\ (x) = g\ (x + 5) = (x + 5) - 5 = x$$

and

$$f \circ g\ (x) = f\ (x - 5) = (x - 5) + 5 = x.$$

Therefore g is an inverse for f in $(F,\circ)$ and f is an inverse for g in $(F,\circ)$.

As another example, if $x \xrightarrow{f} \frac{2}{3}x$, then $x \xrightarrow{g} \frac{3}{2}x$ is a likely candidate for its inverse.

$$x \xrightarrow{f} \frac{2}{3}x \xrightarrow{g} \frac{3}{2}\left(\frac{2}{3}x\right) = x$$

and $x \xrightarrow{g} \frac{3}{2}x \longrightarrow \frac{2}{3}\left(\frac{3}{2}x\right) = x.$

Therefore f and g are inverses in $(F,\circ)$.

If you look back at Exercise 10 of Section 8.8, you will see that $x \xrightarrow{f} 3x + 2$ and $x \xrightarrow{g} \frac{1}{3}(x - 2)$ are also inverses of each other in $(F,\circ)$. Is $(F,\circ)$ a group? If you are suspicious that you are being enticed into a false conjecture by carefully chosen examples, your suspicion is justified. $(F,\circ)$ is not a group.

Consider f: $R \longrightarrow R$ with rule $x \xrightarrow{f} x^3$. If f is to have an inverse in $(F,\circ)$, then there must be a function g: $R \longrightarrow R$ with the property $g \circ f = f \circ g = j_R$. But how will this inverse function be defined, for example at 4? We must have

$$-2 \xrightarrow{f} 4 \xrightarrow{g} -2,$$

and

$$2 \xrightarrow{f} 4 \xrightarrow{g} 2.$$

For this to be true, g must somehow assign two images to a single real number 4. Any such assignment is not a function, so f cannot have the desired inverse.

Another function without inverse is the constant function $x \xrightarrow{c} 0$. c assigns 0 as the image of every real number so c is a real function. However, the arrow diagram of Figure 7.11 shows quickly why it is impossible to have an inverse for c.

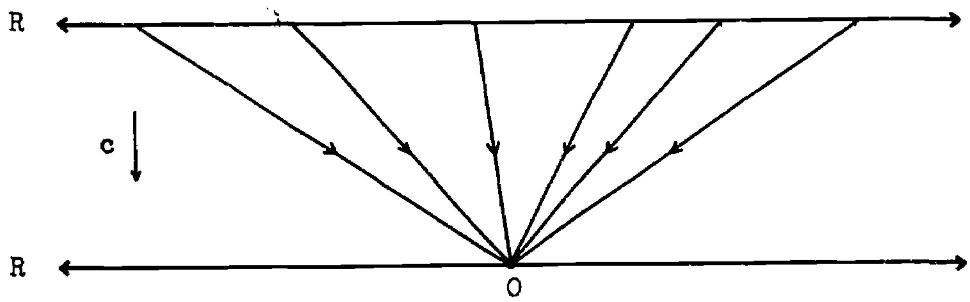


Figure 7.11

The inverse of assignment c, illustrated in Figure 7.12 is clearly not the diagram of a function.

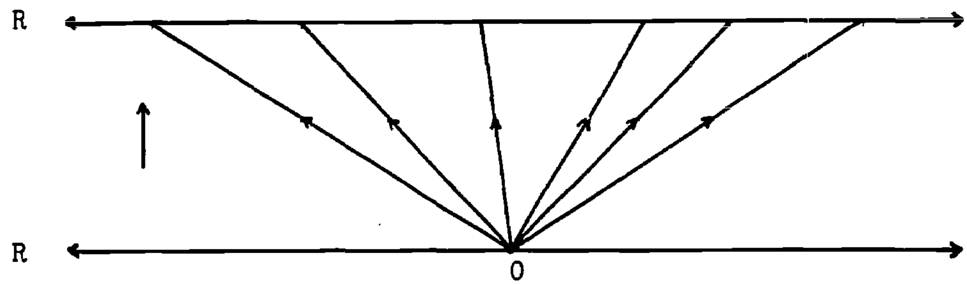


Figure 7.12

If some real functions have inverses and others do not, it would be helpful to have a method of testing functions to see

whether or not they have inverses. Looking carefully at the examples of functions which do not have inverses, you will notice one common difficulty. Function $x \xrightarrow{f} x^2$ has no inverse because $f(-2) = f(2) = 4$, and no function can assign both 2 and -2 as images of 4. Functions $x \xrightarrow{c} 0$ has no inverse because $0 = c(\sqrt{2}) = c(3) = c(-\pi) = \ldots$ and no function can assign two or more images to the real number 0.

This same difficulty will accompany any other function which is not one-to-one. Therefore, we can make the general statement: A function which is not one-to-one has no inverse. Is the following statement true? "A function which <u>is</u> one-to-one always has an inverse"?

Let f: $R \longrightarrow R$ have rule of assignment $x \xrightarrow{f} \frac{x}{|x| + 1}$. According to this rule

$$f(0) - \frac{0}{0 + 1} = 0$$

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

$$f(-3) = \frac{-3}{3 + 1} = -\frac{3}{4}$$

$$f(10) = \frac{10}{10 + 1} = \frac{10}{11}$$

$$f(200) = \frac{-200}{200 + 1} = -\frac{200}{201}$$

The graph of this function is illustrated in Figure 7.13.

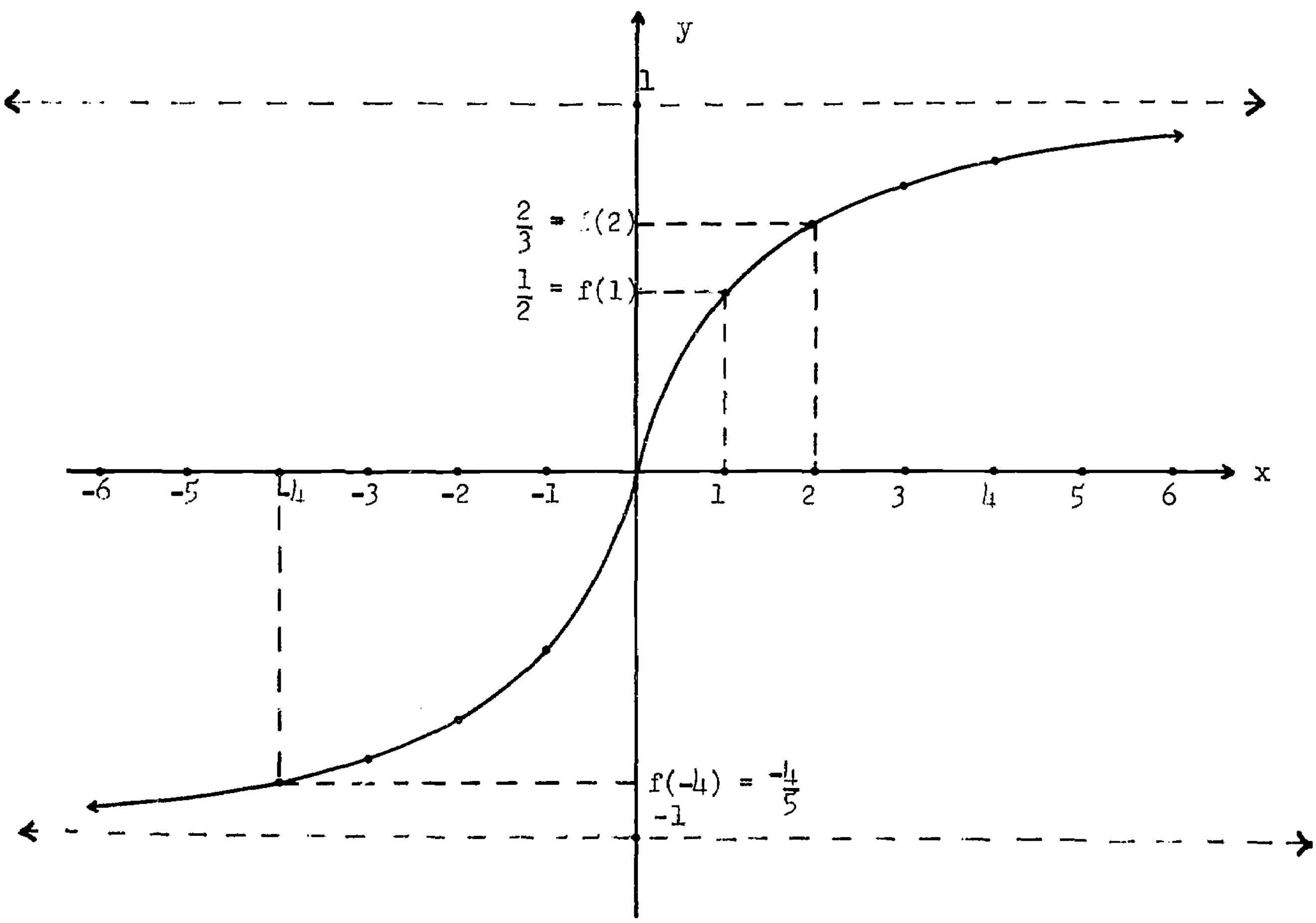


Figure 7.13

If you compute the images of many more real numbers, you will notice an interesting pattern developing. First, the images assigned by the rule $x \xrightarrow{f} \frac{x}{|x| + 1}$ are all numbers between -1 and 1. Second, no number is used as an image more than once and each number between -1 and 1 is the image under f of some real number. Surprising as it may seem, f is a one-to-one function with domain R and range $I = \{x: x \in R \text{ and } -1 < x < 1\}$.

Question. Why is "<" used rather than "≤"?

There are functions from R to R which reverse the assignments of $x \xrightarrow{f} \frac{x}{|x| + 1}$. However, any such function will make

additional assignments unwanted for the inverse of f. If $g: R \longrightarrow R$ reverses the assignments of f,

$$g(0) = 0 \quad \text{since} \quad f(0) = 0$$

$$g(\tfrac{1}{2}) = 1 \quad \text{since} \quad f(1) = \tfrac{1}{2}$$

$$g(-\tfrac{2}{3}) = -2 \quad \text{since} \quad f(-2) = -\tfrac{2}{3}$$

and so on. For all x in R,

$$g \circ f(x) = x.$$

f assigns an image between -1 and 1 and g assigns to each number between -1 and 1 its pre-image under f. But $g: R \longrightarrow R$ must also assign images to numbers outside I, like 10, -23, $\sqrt{7}$, etc. This is where g fails as an inverse of f.

If g assigns a real number m as the image of 10, then $f \circ g(10) = f(m) = \frac{m}{|m| + 1}$. Since the range of f is I, $f \circ g(10)$, or f(m), is in I, and $f \circ g(10) \neq 10$. Therefore $f \circ g \neq j_R$ and f and g are not inverses in $(F, \circ)$.

Although $f: R \longrightarrow R$ does not have an inverse, the function $h: R \longrightarrow I$ with rule of assignment $x \xrightarrow{h} \frac{x}{|x| + 1}$ has the same domain as f, makes the same assignments as f, and has an inverse $k: I \longrightarrow R$ with rule of assignment $x \xrightarrow{k} \frac{x}{1 - |x|}$, such that

$$k \circ h = j_R \text{ and } h \circ k = j_I$$

Even though k and h are not inverses in the operational system $(F, \circ)$ the notion of an inverse of a function is extended to include functions such as k and h as inverses **under** composition.

<u>Definition</u> 6. If $f: A \longrightarrow B$ is a real function, g is called an inverse of f if and only if $g \circ f = j_A$ and $f \circ g = j_B$.

It is clear, by the definition of composition, that g must be a mapping of B to A. This definition, although stated for real functions, describes function inverses in general. The following theorem expresses the fact, which has been illustrated many times, that a function f has an inverse if and only if it is one-to-one and onto.

Theorem A real function f: $A \longrightarrow B$ has an inverse, g, if and only if f is one-to-one and onto.

Since the theorem is in the form of a biconditional, we break the statement into its two component conditions, and prove each one separately.

Proof.

1.

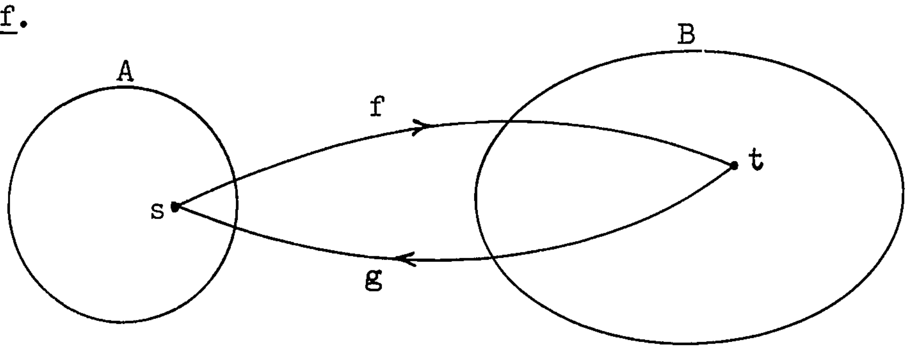


Figure 7.14

Is the assignment constructed by reversing the assignments made by f a mapping of B to A? Since f is onto, for any $t \in B$, t is the image of some element s in A. Since f is one to one, t is the image of exactly one element of A. Hence, the assignment constructed by reversing the assignments is a mapping. Call it g.

2.

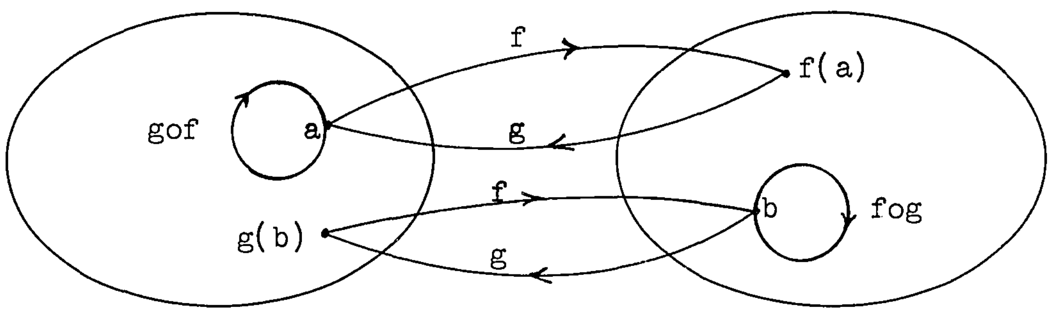


Figure 7.15

Since f and g reverse assignments, for any element $b \in B$, $f \circ g\ (b) = b$ and for any element $a \in A$, $g \circ f\ (a) = a$ (see Figure 7.15). Thus, $g \circ f = j_A$ and $f \circ g = j_B$. By definition, then, g is the inverse of f. Conversely, if a mapping $f\colon\ A \longrightarrow B$ has an inverse g, then f is one to one and onto.

1.

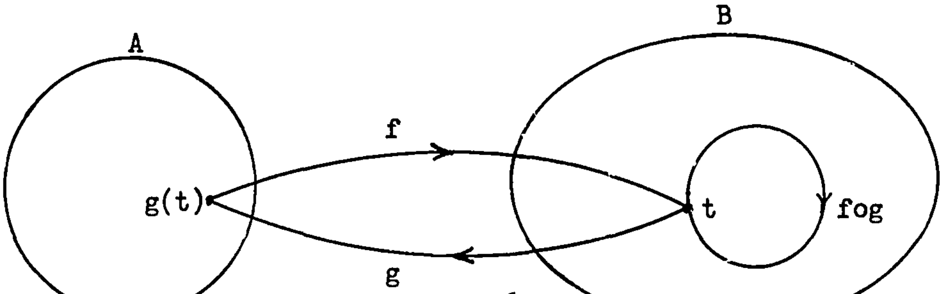


Figure 7.16

If f has an inverse g, we know by definition that $g \circ f = j_A$ and $f \circ g = j_B$. To show that f is onto, take any $t \in B$. Since $f \circ g = j_B$, we have $f \circ g\ (t) = j(t) = t$. But $f \circ g\ (t) = f(g(t))$ where $g(t) = a$ is in A (see Figure 7.16). Hence, for any $t \in B$, $t = f(s)$ for some $a \in A$.

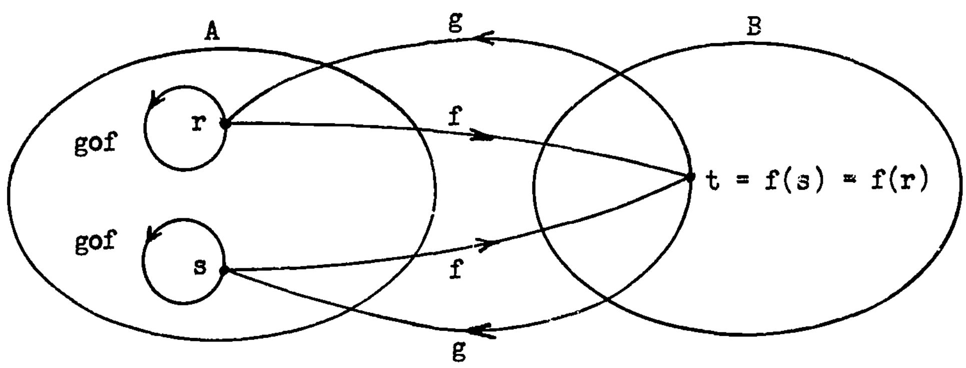


Figure 7.17

To show that f is one to one, we proceed indirectly. Suppose f is not one to one. Then there must be two distinct elements r and s of A that have the same image t in B under f. That is, $f(r) = f(s) = t$. (See Figure 7.17)

Since $g \circ f = j_A$, $r = g \circ f(r) = g(f(r)) = g(t)$. That is, the image of t under g is r. But we also have that

$$s = g \circ f(s) = g(f(s)) = g(t).$$

But $r = g(t)$ and $s = g(t)$ is a contradiction, because g is given to be a mapping. Hence, f is one to one.

Definition 7. If $f: A \longrightarrow B$ and $g: A \longrightarrow C$ are functions with the property that $f(a) = g(a)$ for every $a \in A$, then f and g are called equivalent functions.

A pair of equivalent functions are $f: R \longrightarrow R$ and $g: R \longrightarrow R_o^+$ where both f and g have the rule $x \longrightarrow x^2$.

Does every one-to-one function have an inverse? The answer to this question must be "No, a function $f: A \longrightarrow B$ has an inverse if and only if it is one-to-one and onto." If f is one-to-one

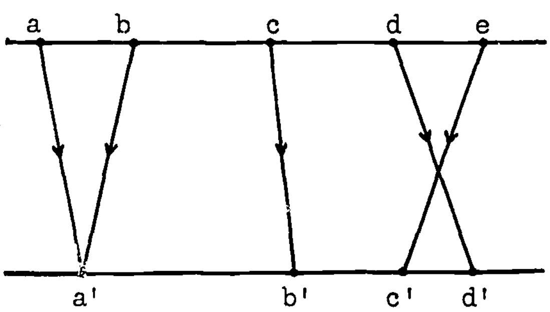
but not onto B, f is equivalent to a function which does have an inverse; namely, the function with the same domain and assignment process as f, but with codomain equal to the range of f.

The interrelationships of one-to-one, onto, range, codomain, and inverse function are illustrated in the following exercises.

7.10 Exercises

1. The following arrow diagrams, graphs, and tables, represent functions from A to B. Explain for each why the given function (i) is or is not one-to-one; (ii) is or is not onto; (iii) has or does not have an inverse.

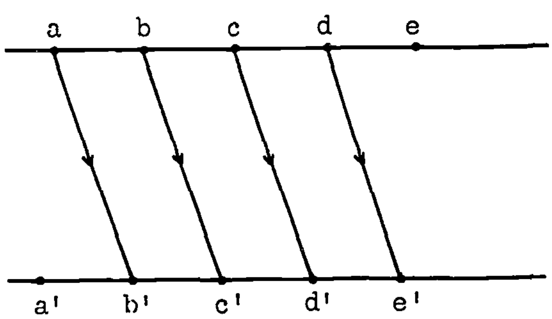
(a)



A = {a, b, c, d, e}

B = {a', b', c', d'}

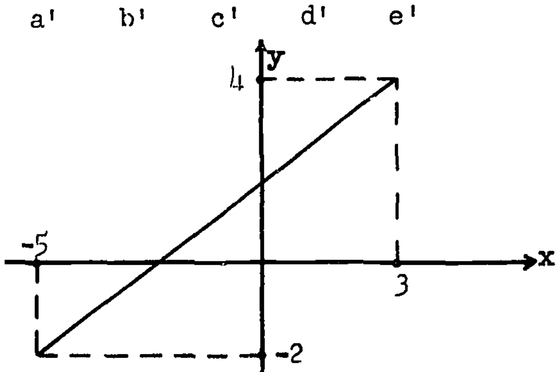
(b)



A = {a, b, c, d}

B = {a', b', c', d', e'}

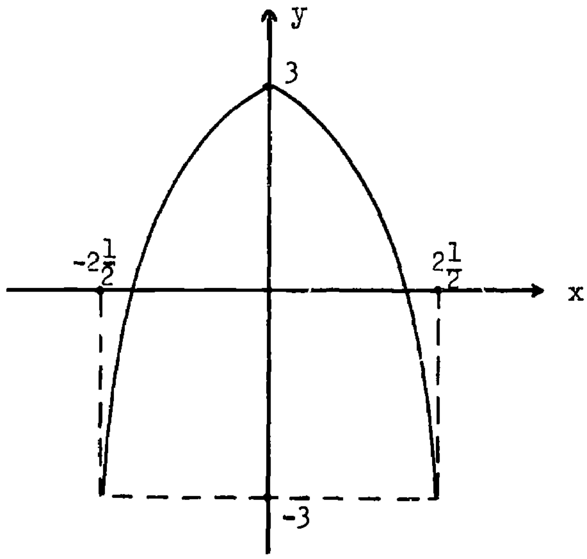
(c)



A = {x: $-5 \leq x \leq 3$}

B = {y: $-2 \leq y \leq 4$}

(d)



$A = \{x: -2\frac{1}{2} \leq x \leq 2\frac{1}{2}\}$

$B = \{y: -3 \leq y \leq 3\}$

(e)

x	1	2	3	4	5	6	7	8
f(x)	10	11	13	16	20	25	31	38

$A = \{1, 2, 3, 4, 5, 6, 7, 8\}$

$B = \{10, 11, 13, 16, 20, 25, 31, 38, 46, \ldots\}$

2. For each function in Exercise 1 that does not have an inverse because it is not onto its codomain, describe the codomain of an equivalent function which does have an inverse.

3. Let g: $R \longrightarrow R$ have rule $x \xrightarrow{g} -x$. Graph g.

   (a) Is g one-to-one? Why?

   (b) Is g onto? Why?

   (c) Does g have an inverse? If not, why? If so, give its rule.

4. Graph $x \xrightarrow{h} 3x$.

   (a) What is the rule for the inverse of h?

   (b) Graph the inverse of h on the same coordinatized plane as h.

(c) Graph $x \xrightarrow{j} x$ with a red pencil on the same coordinatized plane.

5. Repeat the directions of Exercise 3 with $x \xrightarrow{g} \frac{1}{2}x$.

6. Repeat the directions of Exercise 3 with $x \xrightarrow{f} -4x$.

7. Do you see any pattern in the geometry of the graphs of Exercises 3, 4, and 5? Try to state it in the language of reflections or symmetry.

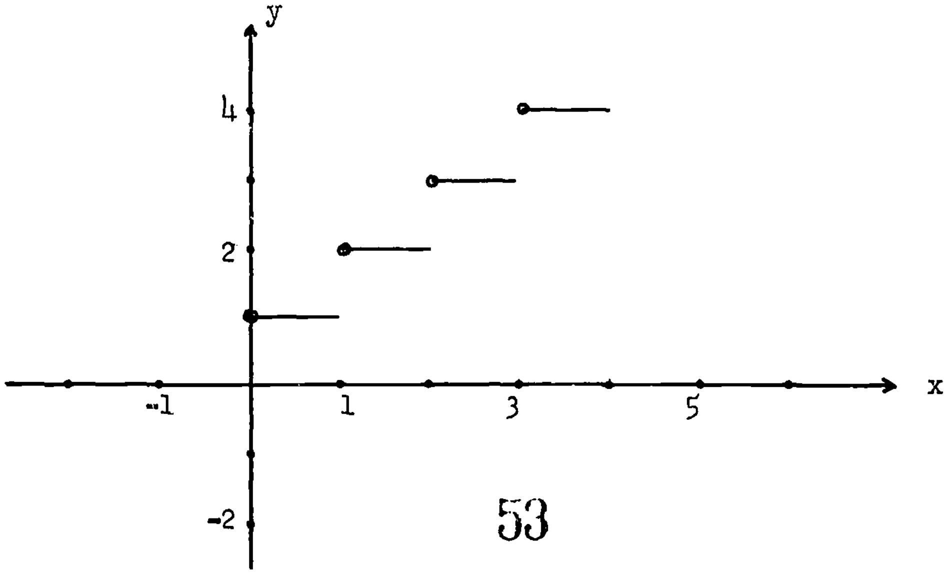
8. Sketch the graph of $x \xrightarrow{h} |x|$.

   (a) Does h have an inverse? Why or why not?

   (b) If your answer to (a) was "yes," make a table showing 10 assignments of the inverse of h.

   (c) If your answer to (a) was "no," can you modify the codomain of h to get a function which does have an inverse? If this is impossible, can you restrict the function by choosing a smaller <u>domain</u>, thus producing a function with the same rule of assignment, but having an inverse?

9. The postal function (see Section 7.7) $p: R^{+} \longrightarrow W$ has the following graph:



(a) Ia p one-to-one? Why or why not?

(b) Is p onto? Why or why not?

(c) Does p have an inverse? Why or why not?

(d) If your answer to (c) was "no," can p be modified to have an inverse by restricting the domain or codomain? (Make the minimal restriction needed.)

10. The function with rule $x \xrightarrow{r} \frac{1}{x}$ has domain and codomain $R \setminus \{0\}$.

(a) Why can 0 not be in the domain of r?

(b) Complete the following table.

x	1	2	3	$\frac{1}{2}$	$1\frac{1}{2}$	$2\frac{1}{2}$	-1	-2	-3	$-\frac{1}{2}$	$-1\frac{1}{2}$	$-2\frac{1}{2}$
$\frac{1}{x}$												

(c) Graph $x \xrightarrow{r} \frac{1}{x}$ using the values computed in (b).

(d) Is r a one-to-one function? An onto function? Why or why not?

(e) Does r have an inverse? If not why not? If so, what is its rule, domain, and codomain?

11. Let h: $R \longrightarrow I$ have rule $x \xrightarrow{h} \frac{x}{|x| + 1}$ and k: $I \longrightarrow R$ have rule $x \xrightarrow{k} \frac{x}{1 - |x|}$. $I = \{x\colon\ x \in R \text{ and } -1 < x < 1\}$.

(a) Find a standard name of:

(i) $h(0)$ (vi) $k(0)$

(ii) $h(1)$ (vii) $k(\frac{1}{2})$

(iii) $h(-1)$ (viii) $k(-\frac{1}{2})$

(iv) $h(2)$ (ix) $k(\frac{2}{3})$

(v) $h(-2)$ (x) $k(-\frac{2}{3})$

(Recall: $\frac{\frac{1}{3}}{\frac{1}{2}} = \frac{1}{3} \div \frac{1}{2}$ etc.)

(b) Find a standard name of:

(i) $k \circ h(0)$ (vi) $h \circ k(0)$

(ii) $k \circ h(1)$ (vii) $h \circ k(\frac{1}{3})$

(iii) $k \circ h(-1)$ (viii) $h \circ k(-\frac{1}{3})$

(iv) $k \circ h(2)$ (ix) $h \circ k(\frac{2}{3})$

(v) $k \circ h(-2)$ (x) $h \circ k(-\frac{2}{3})$

(c) Is $k \circ h(x) = x$ for $x = 0, 1, -1, 2, -2$?

(d) Is $h \circ k(x) = x$ for $x = 0, \frac{1}{3}, -\frac{1}{3}, \frac{2}{3}, -\frac{2}{3}$?

(e) Can you find $x \in R$ for which $k \circ h(x) \neq x$?

(f) Can you find $x \in I$ for which $h \circ k(x) \neq x$?

7.11 <u>[f + g] and [f - g]</u>

One salesman asked to estimate the cost of printing this mathematics book derived the following price formulas:

(1) Cost of printing and binding: \$6000.00 plus an additional \$3.00 per book.

(2) Cost of delivery: \$.10 per book.

These formulas are actually rules of assignment for two functions from W to R given by

$$x \xrightarrow{p} 3x + 6000 \quad \text{(printing and binding cost),}$$

and

$$x \xrightarrow{d} .10x \quad \text{(delivery cost).}$$

To present his estimate to the project director, the salesman had to combine these two price functions into a single estimate function E: $W \longrightarrow R$. You are familiar with one way of combining

two functions to produce a single function -- that is composition. However, composition doesn't seem to be the appropriate operation in this situation.

$$p \circ d(x) = p(.10x) = .3x + 6000$$

and

$$d \circ p(x) = d(3x + 6000) = .3x + 600.$$

Both compositions yield cost functions which assign costs lower than the cost of printing alone!

The natural operation on functions in this case is what we will call addition of functions. The estimated cost of producing and delivering x books is given by the formula

$$x \xrightarrow{E} (3x + 6000) + (.10x)$$

or

$$x \xrightarrow{E} p(x) + d(x).$$

Here are some sample prices calculated with this new sum function.

x	$p(x)$	$d(x)$	$E(x)$ $p(x) + d(x)$
100	6300	10	6310
500	7500	50	7550
1000	9000	100	9100
3000	15000	300	15300

Question. In the formula \$6000.00 plus \$3.00 per copy, what does the \$3.00 per copy represent?

Addition of real functions occurs naturally in many other settings. For instance, if two pumps produce 500 gallons per hour and 1000 gallons per hour respectively and they are run for x hours, then their combined production P is given by the sum of

two functions

$$x \xrightarrow{P_1} 500x \text{ and } x \xrightarrow{P_2} 1000x$$

$$x \xrightarrow{P} (500\ x + 1000x) = 1500x$$

In other cases, it is natural to combine two functions by subtraction. For instance, a fuel dealer buys oil at 15 cents per gallon and sells it at 23 cents per gallon. If he sells x gallons, his gross profit function P: $W \longrightarrow R$ is the difference of his sales function $x \xrightarrow{S} .23x$ and his cost function $x \xrightarrow{C} .15x$.

$$x \xrightarrow{P} (.23x - .15x) = .08x$$

Addition and subtraction of real functions occur often enough to merit more systematic study.

<u>Definition</u> 8. If f: $R \longrightarrow R$ and g: $R \longrightarrow R$, the function h: $R \longrightarrow R$ with rule of assignment $h(x) = f(x) + g(x)$ is called the <u>sum of f and g</u>.

We use "[f + g]" to name the sum of the functions f and g. Therefore, for all real numbers x, $[f + g](x) = f(x) + g(x)$. [f + g] is definitely a function from R to R because for each real number x, f(x) and g(x) are real numbers, and + is an operation on R.

<u>Definition</u> 9. If f: $R \longrightarrow R$ and g: $R \longrightarrow R$, the function h: $R \longrightarrow R$ with rule of assignment $h(x) = f(x) - g(x)$ is called the <u>difference of f and g</u>.

We use "[f - g]" to name the difference of functions f and g. Therefore, for all real numbers x, $[f - g](x) = f(x) - g(x)$. Again we know [f - g] is a real function since subtraction is an operation on (R,+).

Question. Is it possible to define subtraction as an operation on the set of all functions from W to W? From Z to Z?

It is important to keep in mind the double use of the symbols "+" and "-" in expressions such as

(1) $[f + g](x) = f(x) + g(x)$

and

(2) $[f - g](x) = f(x) - g(x)$.

On the left in (1) and (2), "+" and "-" indicate operations on functions; on the right in (1) and (2), "+" and "-" are the familiar arithmetic operations on real numbers. The symbols are used for both operations because addition and subtraction of functions are defined in terms of addition and subtraction of the real number images of the functions. The brackets "[ ]" are used to indicate that [f + g] is a function -- not to be confused with "f(x) + g(x)" which is the sum of two real numbers.

Let's study one more instance of addition and subtraction of functions. Let f and g be function from R to R with rules of assignment $x \xrightarrow{f} \frac{1}{2}x$ and $x \xrightarrow{g} x - 2$.

$$[f + g](0) = f(0) + g(0) = 0 + (-2) = -2$$
$$[f + g](17) = 8\tfrac{1}{2} + 15 = 23\tfrac{1}{2}$$
$$[f + g](-12) = -6 - 14 = -20$$
$$[f - g](0) = f(0) - g(0) = 0 - (-2) = 2$$
$$[f - g](17) = 8\tfrac{1}{2} - 15 = -6\tfrac{1}{2}$$
$$[f - g](-12) = -6 - (-14) = 8.$$

In general

$$x \xrightarrow{[f + g]} \tfrac{3}{2}x - 2; \qquad x \xrightarrow{[f - g]} -\tfrac{1}{2}x + 2.$$

The sum and difference of two functions can be illustrated graphically (see Figure 7.18).

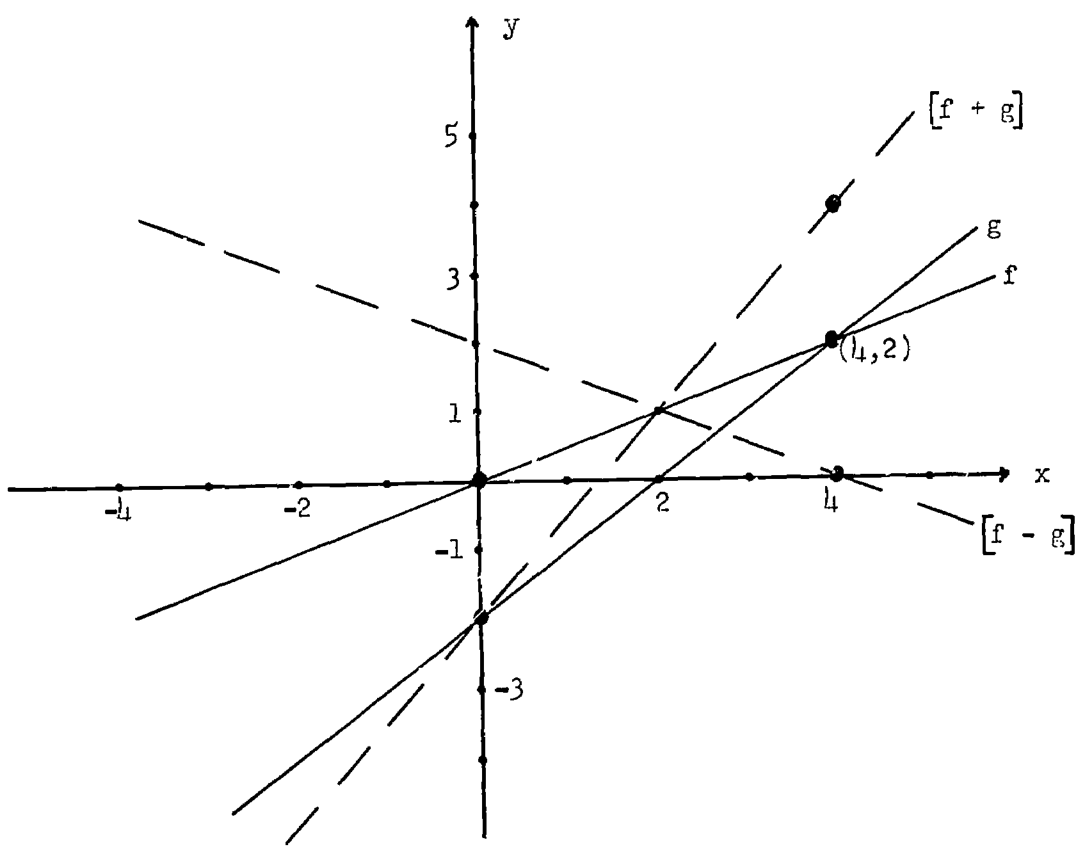


Figure 7.18

The graphs of f and g intersect at the point (4, 2) showing that

$$f(4) = g(4) = 2.$$

This in turn implies that $[f + g](4) = 2 + 2 = 4$. Checking the graph of $[f + g]$ you see that (4, 4) is one of the indicated points. Similarly $[f - g](4) = 2 - 2 = 0$ and (4, 0) is one of the points of the graph of $[f - g]$. Notice that $f(0) = 0$. Thus $[f + g](0) = f(0) + g(0) = -2$. This means that the graphs of $[f + g]$ and g intersect at (0, -2). In the following exercises you will be given practice computing sums and differences of functions, both algebraically and graphically.

7.12 <u>Exercises</u>

1. Let $x \xrightarrow{f} 3$ and $x \xrightarrow{g} |x|$ have domain and codomain R.

(a) Find a standard name for:

(i)	$f(0)$	(vi)	$g(0)$
(ii)	$f(1)$	(vii)	$g(1)$
(iii)	$f(-1)$	(viii)	$g(-1)$
(iv)	$f(16\frac{1}{3})$	(ix)	$g(16\frac{1}{3})$
(v)	$f(-23)$	(x)	$g(-23)$

(b) Find a standard name for:

(i)	$[f + g](0)$	(vi)	$[f - g](0)$
(ii)	$[f + g](1)$	(vii)	$[f - g](1)$
(iii)	$[f + g](-1)$	(viii)	$[f - g](-1)$
(iv)	$[f + g](16\frac{1}{3})$	(ix)	$[f - g](16\frac{1}{3})$
(v)	$[f + g](-23)$	(x)	$[f - g](-23)$

(c) Find a standard name for:

(i)	$f \circ g(0)$	(vi)	$g \circ f(0)$
(ii)	$f \circ g(1)$	(vii)	$g \circ f(1)$
(iii)	$f \circ g(-1)$	(viii)	$g \circ f(-1)$
(iv)	$f \circ g(16\frac{1}{3})$	(ix)	$g \circ f(16\frac{1}{3})$
(v)	$f \circ g(-23)$	(x)	$g \circ f(-23)$

(d) Find a standard name for:

(i)	$[g + f](0)$	(vi)	$[g - f](0)$
(ii)	$[g + f](1)$	(vii)	$[g - f](1)$
(iii)	$[g + f](-1)$	(viii)	$[g - f](-1)$
(iv)	$[g + f](16\frac{1}{3})$	(ix)	$[g - f](16\frac{1}{3})$
(v)	$[g + f](-23)$	(x)	$[g - f](-23)$

(e) What are the rules of assignment for:

(i) $[f + g]$ (iv) $[g - f]$

(ii) $[g + f]$ (v) $g^\circ f$

(iii) $[f - g]$ (vi) $f^\circ g$

(f) Graph f, g, $[f + g]$, and $[f - g]$ on a single coordinatized plane as was done in the text. (You might find it easier to use different colors for the graph of each function.)

2\. Let $x \xrightarrow{h} x^3$ and $x \xrightarrow{k} 3x - 1$ have domain and codomain R.

(a) Complete the following table.

x	h(x)	k(x)	[h + k](x)	[h - k](x)	[k - h](x)
0					
7					
12.5					
-14					
-3					

(b) Find standard names for:

(i) $h^\circ k(0)$ (iv) $k^\circ h(5)$

(ii) $k^\circ h(0)$ (v) $h^\circ k(-2)$

(iii) $h^\circ k(5)$ (vi) $k^\circ h(-2)$

(c) Give the general rule of assignment for:

(i) $[h + k]$ (iii) $[k + h]$

(ii) $[h - k]$ (iv) $[k - h]$

(d) Graph h, j, $[h + k]$ and $[h - k]$ on a single coordinatized plane. (Hint: Make the vertical scale unit smaller than the horizontal scale unit.)

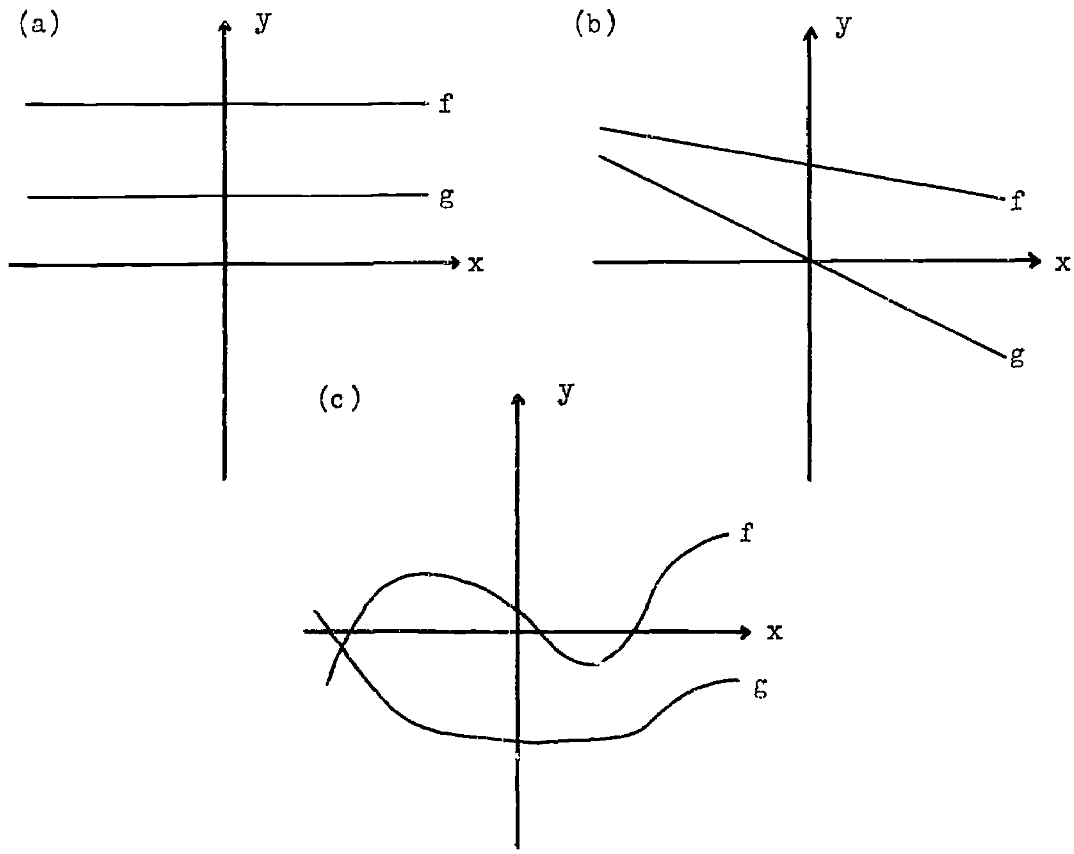
3. Let $x \xrightarrow{f} 2x + 1$ and $x \xrightarrow{k} \frac{1}{2}x - 2$.

   (a) Graph f and g on a single coordinatized plane.

   (b) Graph [f + g] and [f - g] on the same plane.

   Try to do (b) using the graphs of f and g.

4. Copy the graphs below. Then draw the graphs of [f + g] and [f - g] using the same axes.



5. Sketch the graph of the postal function p and the graph of $x \xrightarrow{j} x$ on separate pairs of axes.

   (a) Complete the following table.

x		$\frac{1}{3}$	$\frac{2}{3}$	1	$\frac{4}{3}$	$\frac{5}{3}$	2	$\frac{7}{3}$	$\frac{8}{3}$	3	$\frac{10}{3}$
p(x)		1							3		
p(x) - x		$\frac{2}{3}$							$\frac{1}{3}$		

(b) On a third pair of axes graph $[p - j]$: $R^+ \longrightarrow R$.

(Recall $R^+ = \{x: \ x \in R \text{ and } x > 0\}$.)

6. Sketch the graph of $[p + t]$: $R^+ \longrightarrow R$, where $x \xrightarrow{t} 2$.

7. Graph $x \xrightarrow{f} x^2 + 1$ and $x \xrightarrow{g} -x^2$.

(a) Find standard names for:

(i) $[f + g](0)$

(ii) $[f + g](1)$

(iii) $[f + g](2)$

(iv) $[f + g](-1)$

(v) $[f + g](-2)$

(b) Graph $[f + g]$.

In Exercises 8 - 10, you are asked to explore the following question: "Is (f,+) a group, where, F is the set of all functions from R to R?"

8. Is + associative on F?

(a) If $x \xrightarrow{f} x + 2$, $x \xrightarrow{g} 3x$ and $x \xrightarrow{h} -x^2$,

find the rule for:

(i) $[f + g]$

(ii) $[g + h]$

(iii) $[[f + g] + h]$

(iv) $[f + [g + h]]$

(b) Is $[[f + g] + h]$ equal to $[f + [g + h]]$?

(c) For any functions f, g, and h from R to R, and any $x \in R$,

$$[[f + g] + h](x) = [f + g](x) + h(x) \qquad (1)$$
$$= (f(x) + g(x)) + h(x) \qquad (2)$$

by Definition 8.

Show that $[f + [g + h]](x)$ is equal to (2).

(d) Does (c) help you conclude anything about the associativity of + on F? Explain your answer.

9. Is there a function c: $R \longrightarrow R$ such that

$$[c + f] = [f + c] = f$$

for any function $f \in F$? Do you see that the rule for c is

$x \xrightarrow{c} 0$?

10. Is there an additive inverse for each element of F?

(a) If $x \xrightarrow{f} x^2$ and $x \xrightarrow{g} -x^2$, find a standard name for

(i) $[f + g](0)$

(ii) $[f + g](17)$

(iii) $[f + g](-22.5)$

(iv) $[f + g](-\sqrt{5})$

(b) Complete: $x \xrightarrow{[f + g]} \boxed{}$

(c) For any function h: $R \longrightarrow R$, let $[-h]$ be the function which makes assignments $x \xrightarrow{[-h]} -h(x)$. Is $[-h]$ in F for every h in F? Explain your answer.

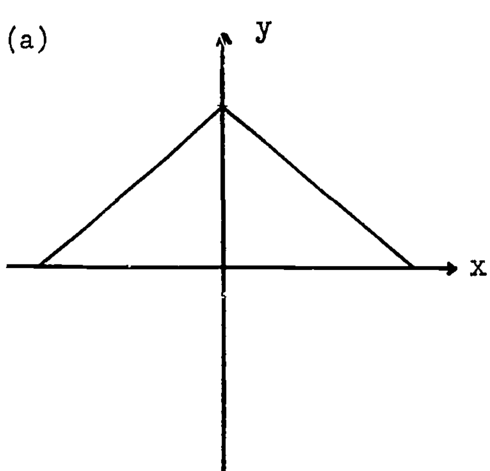
(d) Complete: $x \xrightarrow{h + [-h]} \boxed{}$

(e) Does (d) help answer the question "Is there an additive inverse for each element of F?" If so, how?

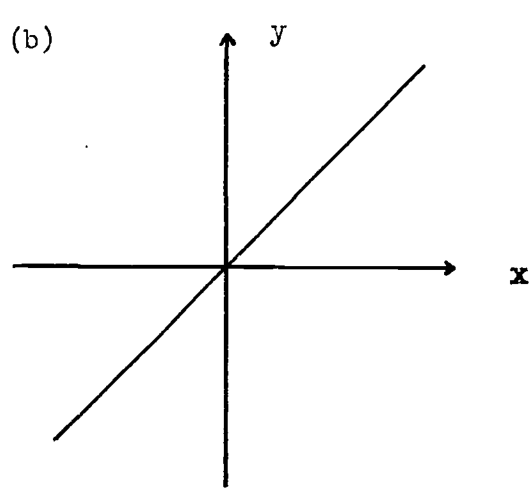
(f) Is (f,+) a group?

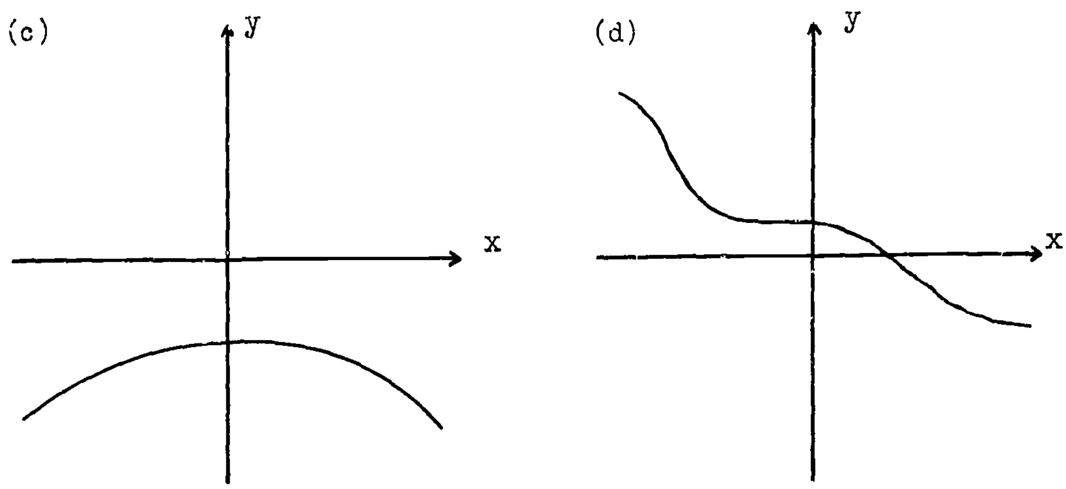
11. Copy each of the following graphs of functions and sketch the graph of the corresponding additive inverse in F on the same axes.

(a)



(b)





12. Is addition of real functions commutative? Explain your answer. Is subtraction of real functions commutative? Explain your answer.

7.13 $[f \cdot g]$ and $[\frac{f}{g}]$

In Section 7.11, addition and subtraction of functions were defined so that $[f + g](x) = f(x) + g(x)$ and $[f - g](x) = f(x) - g(x)$ for all real numbers x. These definitions depend upon the fact that all images under f and g are real numbers. Addition and subtraction of functions can therefore be defined in terms of addition and subtraction of the images of the functions.

Following the pattern set in defining $[f + g]$, there is a natural definition of the product of two functions.

If $x \xrightarrow{f} x^3$ and $x \xrightarrow{g} 3x - 1$, then $f(2) = 8$ and $g(2) = 5$. Therefore if k is to be the product of f and g, k(2) should equal

$f(2)\cdot g(2)$ $(=8\cdot 5 = 40)$. Similarly

$k(-5) = f(-5)\cdot g(-5) = (-125)(-16) = 2{,}000$; $k(10) = 1000\cdot 29 = 29{,}000$;

and in general,

for all $x \in R$, $k(x) = f(x)\cdot g(x) = x^3(3x - 1)$.

<u>Definition</u> 10. If $f\colon R \longrightarrow R$ and $g\colon R \longrightarrow R$, the function $k\colon R \longrightarrow R$ with rule of assignment $x \xrightarrow{k} f(x)\cdot g(x)$ is called the <u>product of f and g</u>.

We use "$[f\cdot g]$" to name the product of functions f and g. Therefore, $[f\cdot g](x) = f(x)\cdot g(x)$ for all real numbers x.

Addition, subtraction, and multiplication are all operations on the set F of functions from R to R. Can division be defined in a natural way as an operation on F? If $x \xrightarrow{f} x^3$ and $x \xrightarrow{g} 3x - 1$, then it is reasonable to expect that the quotient of f and g will be defined so that

$$[\tfrac{f}{g}](2) = \frac{f(2)}{g(2)} = \frac{8}{5},$$

$$[\tfrac{f}{g}](-5) = \frac{f(-5)}{g(-5)} = \frac{-125}{-16},$$

and in general

$$[\tfrac{f}{g}](x) = \frac{f(x)}{g(x)} = \frac{x^3}{3x - 1}$$

<u>Question</u>. Evaluate $f(\frac{1}{3})$, $g(\frac{1}{3})$, and $[\frac{f}{g}](\frac{1}{3})$.

What difficulty does this present to defining a new function $[\frac{f}{g}]$?

Although $x \xrightarrow{[\frac{f}{g}]} \frac{f(x)}{g(x)}$ assigns images to most real numbers, it does not define a function from R to R because there is no real

number assigned as the quotient of $f(\frac{1}{3}) = \frac{1}{27}$ and $g(\frac{1}{3}) = 0$. In another case, when $x \xrightarrow{c} 0$, $x \xrightarrow{\frac{f}{c}} \frac{f(x)}{c(x)}$ again fails to define a function from R to R because no number is assigned as the quotient of $f(2) = 8$ and $c(2) = 0$ or $f(-3) = -27$ and $c(-3) = 0$ or any other pair $f(x) = x^3$ and $c(x) = 0$.

These examples show that division of functions is <u>not</u> an operation on F.

However we can make the following definition.

<u>Definition</u> 11. If $f\colon R \longrightarrow R$ and $g\colon R \longrightarrow R$ then the function $[\frac{f}{g}]\colon A \longrightarrow R$ with rule of assignment $x \xrightarrow{[\frac{f}{g}]} \frac{f(x)}{g(x)}$ is called the <u>quotient of f and g</u>. $(A = \{x\colon x \in R \text{ and } g(x) \neq 0\})$

We write "$[\frac{f}{g}]$" with the understanding that the domain of this function is always restricted to those real numbers x for which $g(x) \neq 0$.

The restrictions on dividing functions suggest that $(F,\circ)$ is not a group because some functions in F don't have multiplicative inverses. The function $c\colon R \longrightarrow R$ with rule $x \xrightarrow{c} 0$ is a certain troublemaker, but since this is the identity for $(F,+)$, we should investigate $(F\setminus\{c\},\cdot)$.

Before we can look for inverses in $(F\setminus\{c\},\cdot)$ we must know if there is an identity. You recall that j_R was the identity for $(F,^\circ)$, and $x \xrightarrow{c} 0$ the identity for $(F,+)$. (See Exercise 9 of Section 7.12.) Let $i\colon R \longrightarrow R$ have the rule $x \xrightarrow{i} 1$. Then for any function f in $f\setminus\{c\}$

$$[f\cdot i](x) = f(x)\cdot i(x) = f(x)\cdot 1 = f(x)$$

and

$$[i\cdot f](x) = i(x)\cdot f(x) = 1\cdot f(x) = f(x)$$

for all real numbers x. Therefore $x \xrightarrow{i} 1$ is the required identity of $(F,\cdot)$.

Let $x \xrightarrow{k} 2x$ be a function from R to R. If h is a function proposed as the multiplicative inverse of k, then we must have for all real numbers x

$$[h\cdot k](x) = h(x)k(x) = 1.$$

But $k(0) = 0$ and

$$[h\cdot k](0) = h(0)\cdot k(0) = h(0)\cdot 0 = 0$$

and h is not the required inverse. Thus no function in F will do the job. Thus $(F\setminus\{c\},\cdot)$ is not a group. Since there are elements in $F\setminus\{c\}$ without inverses, $(F,+,\cdot)$ is not a field. (See Chapter 4, Section 4.1.) However, $(F,+)$ is a group (see Exercises 8-10 Section 7.12), and it is natural to ask just what structure the two-fold system $(F,+,\cdot)$ has. This question is the theme of the following set of exercises.

7.14 Exercises

1. Let $x \xrightarrow{h} 2x$ and $x \xrightarrow{k} \frac{1}{3}x$ be functions from R to R.

   (a) Find standard names for:

(i)	$h(0)$	(vi)	$k(0)$
(ii)	$h(6)$	(vii)	$k(6)$
(iii)	$h(8.4)$	(viii)	$k(8.4)$
(iv)	$h(-12.6)$	(ix)	$k(-12.6)$
(v)	$h(-27)$	(x)	$k(-27)$

(b) Find standard names for:

(i) $[h \cdot k](0)$
(ii) $[h \cdot k](6)$
(iii) $[h \cdot k](8.4)$
(iv) $[h \cdot k](-12.6)$
(v) $[h \cdot k](-27)$
(vi) $[k \cdot h](0)$
(vii) $[k \cdot h](6)$
(viii) $[k \cdot h](8.4)$
(ix) $[k \cdot h](-12.6)$
(x) $[k \cdot h](-27)$

(c) Find standard names for:

(i) $[\frac{h}{k}](0)$
(ii) $[\frac{h}{k}](6)$
(iii) $[\frac{h}{k}](8.4)$
(iv) $[\frac{k}{h}](0)$
(v) $[\frac{k}{h}](6)$
(vi) $[\frac{k}{h}](8.4)$

(d) Find standard names for:

(i) $h \circ k(0)$
(ii) $h \circ k(6)$
(iii) $k \circ h(8.4)$
(iv) $k \circ h(-27)$

2. Graph the functions h and k of Exercise 1 on a single coordinatized plane. Then graph $[h \cdot k]$ on the same axes. (Note: $x \xrightarrow{[h \cdot k]} \frac{2}{3}x^2$ but $x \xrightarrow{h \circ k} \frac{2}{3}x$!)

3. Complete the following table.

x	f(x)	g(x)	[f + g](x)	[f·g](x)
2	13	$-\frac{1}{2}$		
$-7\frac{1}{4}$	-24	$3\frac{1}{4}$		
5	20	-17		
π	0	1		

4. Let $x \xrightarrow{h} |x| + 1$ and $x \xrightarrow{g} \frac{1}{|x| + 1}$ be assignment process for functions from R to R.

(a) What is the range of h? of g? of [h • g]?

(b) Complete the following table.

x	0	1	-1	2	-2	3	-3	$\frac{1}{2}$	$-\frac{1}{2}$
$\lvert x\rvert+1$			2						
$\frac{1}{\lvert x\rvert+1}$					$\frac{1}{3}$				

(c) Graph h and g on the same coordinatized plane.

(d) Graph $x \xrightarrow{i} 1$ (see Section 7.13) on the same plane (perhaps in red).

5. Let $x \xrightarrow{f} 2x^2 + 3$ and $x \xrightarrow{i} 1$ be functions from R to R.

(a) Find standard names for:

(i) f(-5) (iv) i(-5)

(ii) f(14) (v) i(14)

(iii) f(-6.5) (vi) i(-5)

(b) Find standard names for:

(i) [f • i](-5) (iv) [i • f](-5)

(ii) [f • i](14) (v) [i • f](14)

(iii) [f • i](-6.5) (vi) [i • f](-6.5)

6. Graph $x \xrightarrow{f} x^2 + 1$ on a large coordinatized plane.

(a) Compute $\frac{1}{x^2 + 1}$ for x equal to:

(i) 0 (v) -1

(ii) $\frac{1}{2}$ (vi) $1\frac{1}{2}$

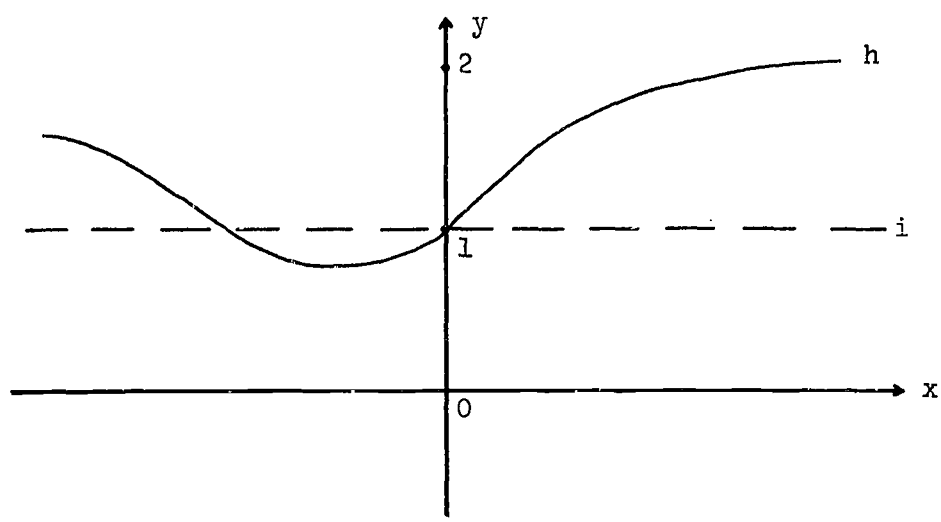
(iii) $-\frac{1}{2}$ (vii) $-1\frac{1}{2}$

(iv) 1 (viii) 2

(ix) -2

(b) Graph $x \xrightarrow{g} \frac{1}{x^2 + 1}$ using the same set of axes as used for f.

7. Below is a graph of a function h: $R \longrightarrow R$. Copy the given graph, and using the line i: $R \longrightarrow R$, $x \xrightarrow{i} 1$, sketch the graph of $x \longrightarrow \frac{1}{h(x)}$ on the same axes.



8. If f, g and h are functions from R to R, then for every real number x, f(x), g(x) and h(x) are real numbers. It is also true that for each $x \in R$,

$$[[f \cdot g] \cdot h](x) = ([f \cdot g](x)) \cdot h(x) \qquad (1)$$

$$= (f(x) \cdot g(x)) \cdot h(x) \qquad (2)$$

(a) Show that $[f \cdot [g \cdot h]](x)$ is equal to (2).

(b) What, if anything, does (a) suggest about a property of $\cdot$ on F?

9. If $x \xrightarrow{f} x^3$, $x \xrightarrow{g} 3x - 1$, and $x \xrightarrow{h} 2x$ are functions from R to R, copy the following table, and complete it.

x	f(x)	g(x)	h(x)	$[f\cdot g](x)$	$[f\cdot h](x)$	$[f\cdot g](x)+[f\cdot h](x)$
2						
0						
-2						

10. Extend the table in Exercise 9 to include:

$[g + h](x)$	$[f\cdot[g + h]](x)$

11. What property of $(F,+,\cdot)$ is suggested by Exercises 9 and 10?

12. Re-examine Sections 7.11, 7.12, 7.13, and the preceding exercises in this section. Try to list all the significant properties of $(F,+,\cdot)$ and also those familiar properties of two-fold operational systems that are not true in $(F,+,\cdot)$.

13. If $f\colon\ R \longrightarrow R$, we can define a new function $[3f]\colon\ R \longrightarrow R$ as follows: $x \xrightarrow{3f} 3\cdot f(x)$. In general, for any real number a, $[af]\colon\ R \longrightarrow R$ has rule of assignment $x \xrightarrow{[af]} a\cdot f(x)$. This new function $[af]$ is called the scalar product of f by the real number a. Let $x \xrightarrow{f} 2x + 3$ be a function from R to R.

(a) Find standard names for:

(i) $[3f](0)$ (ii) $[3f](5)$ (iii) $[3f](-7)$

(b) What is the rule for $[3f]$?

Let $x \xrightarrow{g} x^2$ be a function from R to R.

(c) Find standard names for:

(i) $[3g](0)$ (ii) $[3g](5)$ (iii) $[3g](-7)$

(d) What is the rule for $[3g]$?

(e) Find standard names for:

(i) $[3[f + g]](0)$ (ii) $[3[f + g]](5)$

(iii) $[3[f + g]](-7)$ (iv) $[3f](0) + [3g](0)$

(v) $[3f](5) + [3g](5)$ (vi) $[3f](-7) + [3g](-7)$

7.15 The Square Root and Cube Root Functions

The function $x \xrightarrow{f} x^2$ was studied in Sections 7.3 and 7.5. This squaring function does not have an inverse for composition because it is not one-to-one. However, if the domain is restricted to the positive real numbers, the restricted function does have an inverse. You recall that an accurate graph of $x \xrightarrow{f} x^2$ was used to calculate rational approximations of $\sqrt{2}$ and $\sqrt{3}$ (see Figure 7.19).

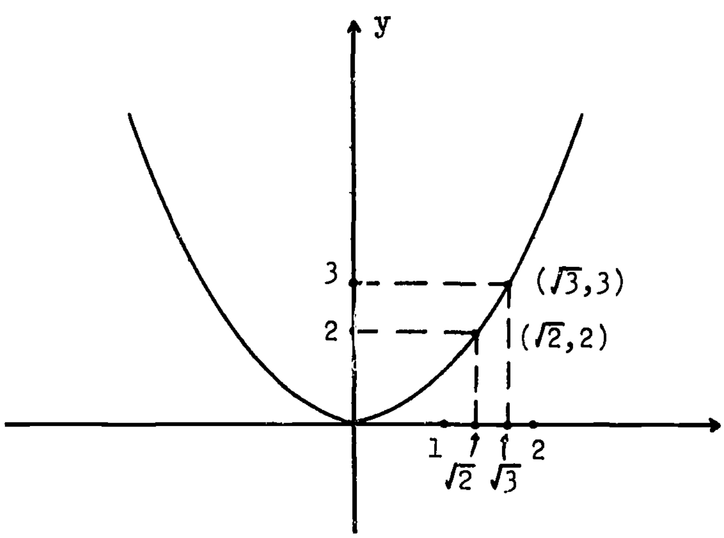


Figure 7.19

Following this procedure we could theoretically obtain an approximation to the square root of any non-negative real number. Given a positive real number, we can locate its square root as the x-coordinate of a point on the x-axis.

Therefore, we can define a new function $\sqrt{\ }$ with domain and codomain R_0^+. This function has the property that $b \xrightarrow{\sqrt{\ }} a$ if and only if $a \geq 0$ and $b = a^2$. Instead of the usual function notation, with which we would write "$\sqrt{\ }(b) = a$," we will write "$\sqrt{b} = a$" to indicate that the square root function assigns a as the image of b or that the square root of b is a. Some square roots are easy to compute. For example, $\sqrt{0} = 0$, $\sqrt{1} = 1$, $\sqrt{4} = 2$, $\sqrt{\frac{9}{4}} = \frac{3}{2}$, $\sqrt{\frac{25}{4}} = \frac{5}{2}$, $\sqrt{9} = 3$, $\sqrt{\frac{1}{4}} = \frac{1}{2}$.

An easy way to obtain the graph of the $\sqrt{\ }$ function is to use the fact that for $a \geq 0$ and $b \geq 0$, $a \xrightarrow{\sqrt{\ }} b$ if and only if $b \xrightarrow{f} a$, when $x \xrightarrow{f} x^2$. That is, (a, b) is in the graph of $\sqrt{\ }$ if and only if (b, a) is in the graph of f restricted to R_0^+. For example, we found $\sqrt{2}$ and $\sqrt{3}$ from the graph of f by finding the points $(\sqrt{2}, 2)$ and $(\sqrt{3}, 3)$ in the graph of f. Thus, $(2, \sqrt{2})$ and $(3, \sqrt{3})$ are in the graph of $\sqrt{\ }$. Hence, we graph the reversed ordered pairs of f restricted to R_0^+ to obtain the graph of $\sqrt{\ }$.

But, even better, there is a nice geometric relationship between the graph of f and the graph of $\sqrt{\ }$. This is illustrated in Figure 7.20, where g and h are any two functions such that $A(x, y)$ is in the graph of h if and only if $A'(y, x)$ is in the graph of g.

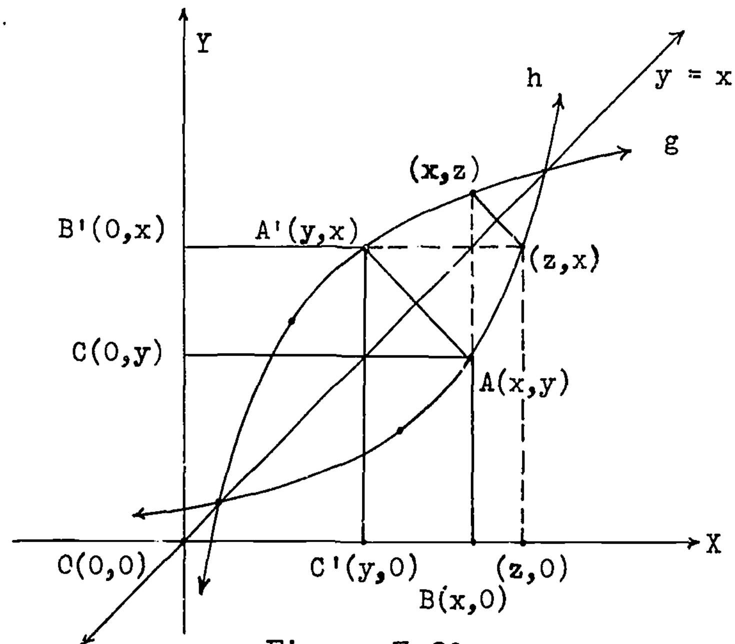


Figure 7.20

This can occur if and only if A' is the image of A under the reflection of the plane in the line $y = x$. To see this, observe that rectangle O B' A' C' must be the image of O B A C under the reflection in the line $y = x$. Since the coordinates of A are (x, y), the coordinates of A' must be (y, x). Likewise, A is the image of A' under the same reflection so that the figure formed by the graphs of f and g is its own image under this reflection and hence symmetric with respect to the line $y = x$.

<u>Question</u>. Are g and h inverse functions if their domain and codomain are both R_o^+ (assuming the graphs continue symmetrically)?

The graph of $\sqrt{\ }$ is now constructed from the graph of f restricted to R_o^+ using the reflection in the line $y = x$. See if you can figure out a neat way to construct this graph yourself using paper folding or tracing paper.

We saw how a rational approximation of $\sqrt{3}$ and $\sqrt{2}$ can be found using a graph of $x \xrightarrow{f} x^2$. A similar procedure can be used for finding a rational approximation to $\sqrt[3]{2}$, the real number (and there is only one) which when cubed yields 2. The graph of $x \xrightarrow{g} x^3$ is shown in Figure 7.21.

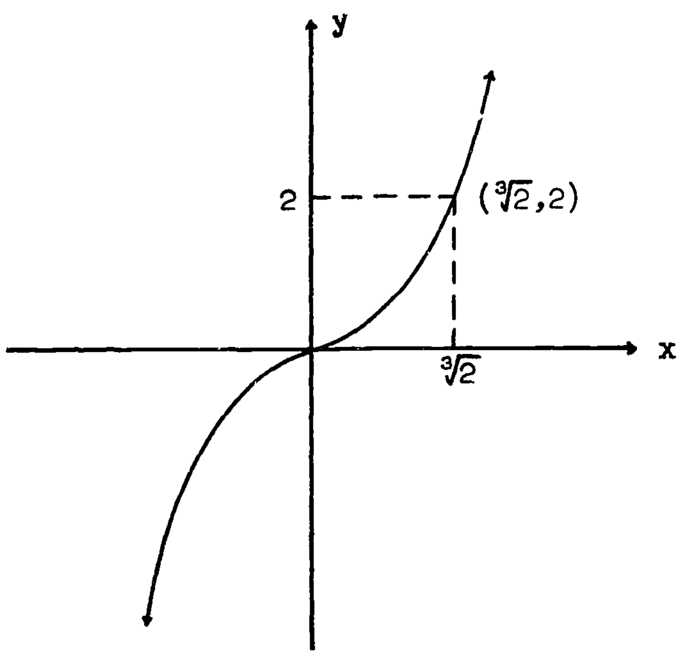


Figure 7.21

There are two important differences between $x \xrightarrow{f} x^2$ and $x \xrightarrow{g} x^3$ that are evident from a comparison of the graphs of the two functions. (See Figures 7.19 and 7.21.) The squaring function is neither one-to-one nor onto R. The cubing function is one-to-one and onto R and thus has an inverse $\sqrt[3]{}$: $R \longrightarrow R$ which assigns to each real number its cube root; that is $b = a^3$ if and only if $\sqrt[3]{b} = a$.

Some cube roots are easy to compute: $\sqrt[3]{0} = 0$, $\sqrt[3]{1} = 1$, $\sqrt[3]{-1} = -1$,

$\sqrt[3]{\frac{27}{8}} = \frac{3}{2}$, $\sqrt[3]{\frac{-27}{8}} = -\frac{3}{2}$, and so on. But, as with $\sqrt{\ }$, the graph of $\sqrt[3]{\ }$ can best be constructed by reflecting the graph of g in the line y = x(see Figure 7.22).

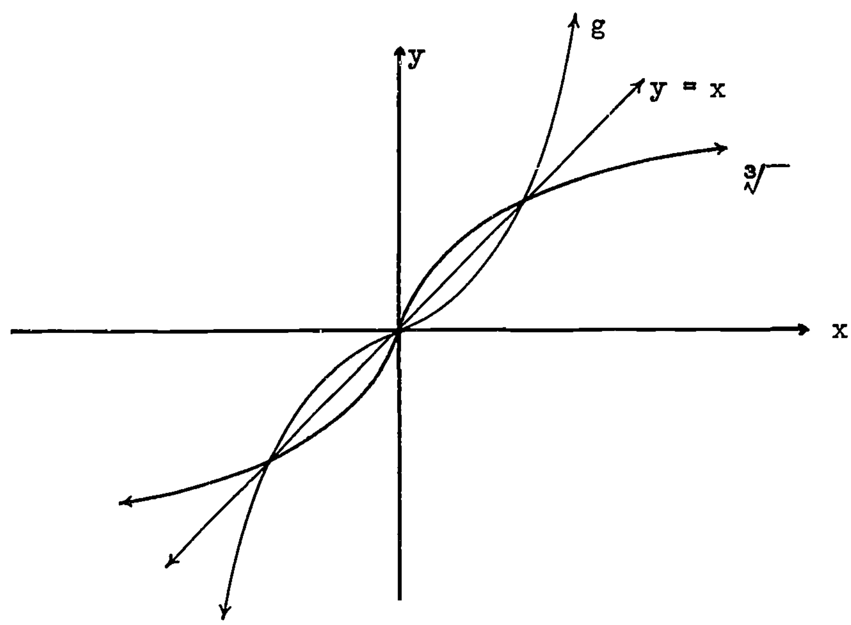


Figure 7.22

The graphs of $\sqrt{\ }$ and $\sqrt[3]{\ }$ are used in the following exercises to find rational approximations of the cube and square roots of certain real numbers. Both the functions $\sqrt{\ }$ and $\sqrt[3]{\ }$ have a special property, called the multiplicative property. A real function h has the multiplicative property if and only if for every a and b in the domain of h, $h(ab) = h(a)\ h(b)$. The function f: $R \longrightarrow R$ given by $x \xrightarrow{f} x^2$ has this property.

$f(ab) = (ab)^2 = (ab)(ab) = (aa)(bb) = a^2b^2 = f(a)\ f(b)$. This argument follows from the rule of f and the associative and commutative laws of $(R,\cdot)$.

The function $\sqrt{\ }$ "inherits" this multiplicative property from its inverse, f restricted to R_0^+ in the following way.

For $a \geq 0$ and $b \geq 0$ $\sqrt{ab} = \sqrt{a} \cdot \sqrt{b}$ if and only if $(\sqrt{ab})^2 = (\sqrt{a} \cdot \sqrt{b})^2$ by definition of $\sqrt{\ }$. But,

$$(\sqrt{ab})^2 = ab \text{ by definition of } \sqrt{\ },$$

and,

$$\begin{aligned}(\sqrt{a} \cdot \sqrt{b})^2 &= (\sqrt{a})^2 \cdot (\sqrt{b})^2 \text{ (f has the multiplicative property)} \\ &= a \cdot b \quad \text{(definition of } \sqrt{\ } \text{ and } \quad)\end{aligned}$$

"$\sqrt{12}$" and "$\sqrt[3]{12}$" (called <u>radicals</u>) are names for range elements of $\sqrt{\ }$ and $\sqrt[3]{\ }$.

The domain element on which $\sqrt{\ }$ or $\sqrt[3]{\ }$ acts is called the <u>radicand</u>. "$\sqrt[3]{12}$" is said to be a radical of <u>index</u> 3 with radicand 12. "$\sqrt{12}$" is a radical of index 2, even though the 2 is not written.

The multiplicative property of $\sqrt{\ }$ and $\sqrt[3]{\ }$ allow many useful transformations of radical names for real numbers.

<u>Example</u> 1. $\sqrt{12} = \sqrt{4 \cdot 3} = \sqrt{4} \cdot \sqrt{3} = 2\sqrt{3}$

<u>Example</u> 2. $\frac{\sqrt{3}}{\sqrt{4}} = \sqrt{\frac{1}{4} \cdot 3} = \sqrt{\frac{1}{4}} \cdot \sqrt{3} = \frac{1}{2}\sqrt{3}$

<u>Example</u> 3. $\sqrt{\frac{2}{3}} = \sqrt{\frac{6}{9}} = \sqrt{\frac{1}{9}} \cdot \sqrt{6} = \frac{1}{3}\sqrt{6}$

<u>Example</u> 4. $\frac{\sqrt{2}}{\sqrt{5}} = \frac{\sqrt{2}}{\sqrt{5}} \cdot \frac{\sqrt{2}}{\sqrt{2}} = \frac{\sqrt{2} \cdot \sqrt{2}}{\sqrt{5} \cdot \sqrt{2}} = \frac{2}{\sqrt{10}}$

<u>Example</u> 5. $\sqrt[3]{54} = \sqrt[3]{27} \cdot \sqrt[3]{2} = 3\sqrt[3]{2}$

<u>Example</u> 6. $\sqrt[3]{\frac{3}{2}} = \sqrt[3]{\frac{12}{8}} = \sqrt[3]{\frac{1}{8}} \cdot \sqrt[3]{12} = \frac{1}{2}\sqrt[3]{12}$.

Some real functions have an analogous <u>additive property</u> e. g. for all a and b in the domain of the real function h, $h(a + b) = h(a) + h(b)$. For example, f: $R \longrightarrow R$ given by $x \xrightarrow{f} 7x$ has this rty since $f(a + b) = 7(a + b) = 7a + 7b = f(a) + f(b)$.

$\sqrt{}$, however, does not have the additive property. One counter-example suffices.

Take a = 4 and b = 9. Then,

$\sqrt{4+9} = \sqrt{13}$. But $\sqrt{4} + \sqrt{9} = 2 + 3 = 5$. Clearly, $\sqrt{13} \neq 5$.

The distributive law aids in simplifying some expressions involving radicals.

Example 7. $3\sqrt{5} - \sqrt{5} + 4\sqrt{5} = (3 - 1 + 4)\sqrt{5} = 6\sqrt{5}$.

Example 8. $(\sqrt{2} - 3)(\sqrt{2} + 3) = (\sqrt{2} - 3)\sqrt{2} + (\sqrt{2} - 3)3$

$$= \sqrt{2} \cdot \sqrt{2} - 3\sqrt{2} + 3\sqrt{2} - 9$$

$$= 2 - 9 = -7.$$

7.16 Exercises

1. Use the graph of $\sqrt{} : R_0^+ \longrightarrow R$ given by $x \xrightarrow{\sqrt{}} \sqrt{x}$ to find rational approximations for:

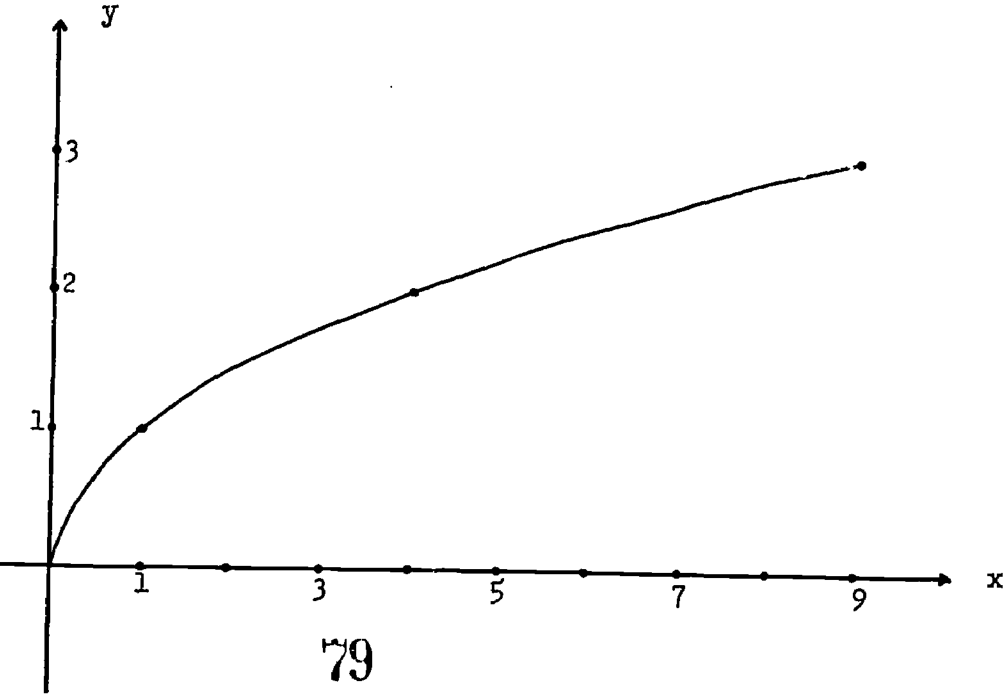
   (a) $\sqrt{5}$

   (b) $\sqrt{6}$

   (c) $\sqrt{7}$

   (d) $\sqrt{2\frac{1}{4}}$

   (e) $\sqrt{4\frac{1}{4}}$



2. Check each approximation in Exercise 1 by squaring. For example, if you found $\sqrt{5} \approx 2\frac{1}{4}$ then $(2\frac{1}{4})(2\frac{1}{4}) = \frac{81}{16} = 5\frac{1}{16}$.

3. Use the graph of $x \xrightarrow{\sqrt[3]{\ }} \sqrt[3]{x}$ to find rational approximations for:

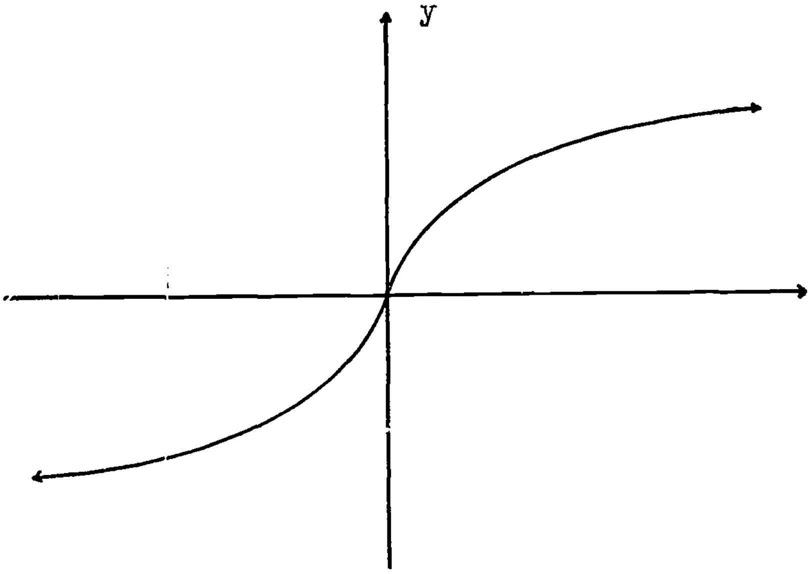
(a) $\sqrt[3]{8}$

(b) $\sqrt[3]{6}$

(c) $\sqrt[3]{7}$

(d) $\sqrt[3]{4}$

(e) $\sqrt[3]{2}$



4. Check each approximation in Exercise 3 by cubing. For example, if you found $\sqrt[3]{8} \approx 2\frac{1}{2}$, $(2\frac{1}{2})^3 = \frac{125}{8} = 15\frac{5}{8}$.

5. Answer the following questions about the square, cube, square root, and cube root functions. For which real numbers x is:

(a) $x^2 > x$	(j) $x^3 > x$
(b) $x^2 < x$	(k) $x^3 < x$
(c) $x^2 = x$	(l) $x^3 = x$
(d) $\sqrt{x} = x$ (and $x \geq 0$)	(m) $\sqrt[3]{x} = x$
(e) $\sqrt{x} < x$ (and $x \geq 0$)	(n) $\sqrt[3]{x} < x$
(f) $\sqrt{x} > x$ (and $x \geq 0$)	(o) $\sqrt[3]{x} > x$
(g) $\sqrt{x} > x^2$ (and $x \geq 0$)	(p) $\sqrt[3]{x} > x^3$
(h) $\sqrt{x} < x^2$ (and $x \geq 0$)	(q) $\sqrt[3]{x} < x^3$
(i) $\sqrt{x} = x^2$ (and $x \geq 0$)	(r) $\sqrt[3]{x} = x^3$

(Hint: It might help to graph the functions on a single coordinatized plane, and also to draw the line $y = x$.)

6. Transform each of the following so that the radicand is the smallest possible positive integer and so that the radical does not appear in the denominator of any fraction (as done in Examples 1--3, 5, 6).

(a) $\sqrt{121}$	(e) $\sqrt{44}$	(i) $\sqrt{\frac{9}{2}}$	(m) $\sqrt[3]{-\frac{3}{4}}$
(b) $\sqrt[3]{1000}$	(f) $\sqrt{147}$	(j) $\sqrt[3]{\frac{1}{8}}$	(n) $\frac{5}{\sqrt{2}}$
(c) $\sqrt{64}$	(g) $\sqrt{500}$	(k) $\sqrt{\frac{3}{5}}$	(o) $\frac{\sqrt{3}}{\sqrt{5}}$
(d) $\sqrt[3]{64}$	(h) $\sqrt[3]{-108}$	(l) $\sqrt{\frac{1}{8}}$	(p) $\frac{\sqrt[3]{2}}{\sqrt[3]{-9}}$

7. Perform the following operations. Use the properties of $\sqrt{\ }$ and $\sqrt[3]{\ }$ and the real numbers to write your final answer compactly.

(a) $\sqrt{7} \cdot \sqrt{14}$	(e) $5\sqrt{2} \cdot 3\sqrt{18}$	(i) $(\sqrt{2} - \sqrt{3})(\sqrt{2} + \sqrt{3})$
(b) $2\sqrt{3} - 3\sqrt{3}$	(f) $2\sqrt{44} \div \sqrt{11}$	(j) $(8 - \sqrt{12})(4 + \sqrt{3})$
(c) $\sqrt{\frac{8}{27}} - \sqrt{\frac{2}{3}}$	(g) $(2 + \sqrt{3})(5 - \sqrt{2})$	
(d) $\frac{5}{\sqrt{2}} + \sqrt{50}$	(h) $(\sqrt{6} - 5)(\sqrt{10} + 4)$	

8. Prove that $g\colon R \longrightarrow R$ with rule $x \longrightarrow x^3$ is multiplicative. Use this to show that its inverse, $\sqrt[3]{\ }\colon R \longrightarrow R$ is multiplicative.

7.17 Summary

In this chapter we have studied some properties of real functions -- those with domain and codomain contained in R. Since real

functions are a special class of mappings, the properties considered were often familiar: one-to-one, onto, inverse, identity, and composition. To illustrate these properties a number of real functions were studied in some detail: $x \xrightarrow{h} |x|$, the postal function p, $x \xrightarrow{j} x$, $x \xrightarrow{c} 0$, $x \xrightarrow{f} \frac{x}{|x| + 1}$, $x \xrightarrow{i} 1$. The fact that all real functions assign images that are real numbers permits definition of many new operations on functions: $[f + g]$, $[f - g]$, $[f \cdot g]$, and $[af]$. The operational systems $(F, \circ)$, $(F, +)$, $(F, \cdot)$, and $(F, +, \cdot)$ have familiar structures.

Some important points of the chapter are the following:

(1) A function $f\colon R \longrightarrow R$ has an inverse if and only if it is one-to-one and onto. However, if f is one-to-one and not onto, it is equivalent to a function which has an inverse.

(2) F, the set of all functions from R to R, is an operational system under composition. $(F, \circ)$ is not a group.

(3) Any real function f can be represented partially by a graph -- the set of points on a coordinatized plane with coordinates of the form "(x,f(x))."

(4) The functions $x \xrightarrow{f} x^2$ and $x \xrightarrow{g} x^3$ and their graphs can be used to obtain rational approximations of $\sqrt{x}$ for any non-negative real number x and of $\sqrt[3]{x}$ for any real number x.

(5) $\sqrt{}\colon R_0^+ \longrightarrow R_0^+$ is the inverse of $f\colon R_0^+ \longrightarrow R_0^+$ where $x \xrightarrow{f} x^2$. And $\sqrt[3]{}\colon R \longrightarrow R$ is the inverse of $g\colon R \longrightarrow R$ where $x \xrightarrow{g} x^3$.

(6) $(F, +, \cdot)$ is not a field.

7.18 Review Exercises

1. Let $x \xrightarrow{h} x^2$ and $x \xrightarrow{g} 2x - 1$ be functions from R to R.

   Copy and complete the following table.

x	$h(x)$	$g(x)$	$g \circ h(x)$	$[h + g](x)$	$[h \cdot g](x)$
0					
1					
-1					
$2\frac{1}{2}$					
$-6\frac{1}{3}$					
12					
-19					
$\sqrt{2}$					
$\|-3\|$					

2. Graph the function $x \xrightarrow{h} 2|x|$ for $-3 \leq x \leq 3$.

3. Graph $x \xrightarrow{f} 3x - 1.5$ for $-4 \leq x \leq 4$.

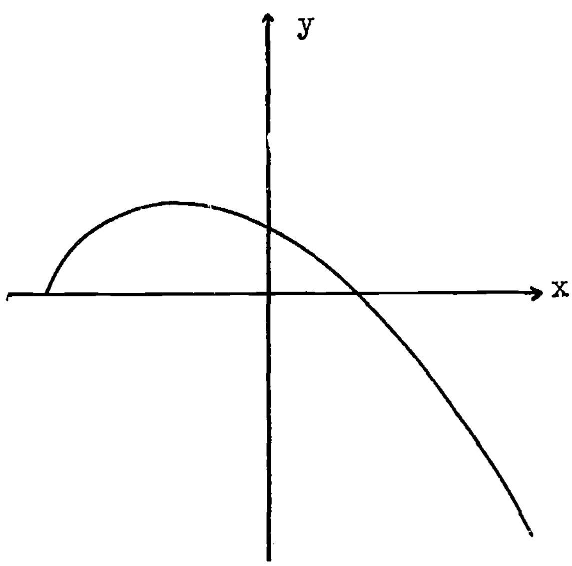
4. For each of the graphs given below:

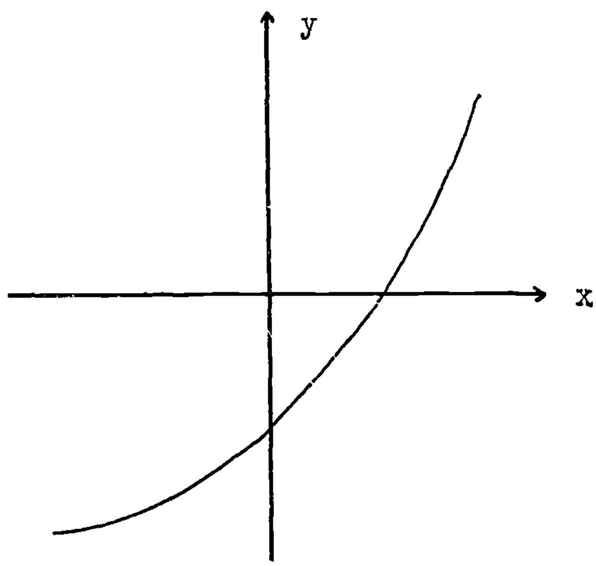
   (a) is the function one-to-one?

   (b) what is the image of 2?

   (c) what is the pre-image (if any) of -2?

   (d) copy the given graph and sketch the graph of -f.





5. For each of the following functions from R to R determine:

   (a) Is it one-to-one? If not, show why by counterexample.

   (b) Is it onto? If not, show why by counterexample.

   (c) Give a rule for the inverse if it exists.

   (1) $x \xrightarrow{f} |x|$ (4) $x \xrightarrow{k} x + \sqrt{2}$

   (2) $x \xrightarrow{g} x^3$ (5) $x \xrightarrow{m} 3.14$

   (3) $x \xrightarrow{h} 17x - 289$ (6) $x \xrightarrow{n} 5x$

6. For the functions given in Exercise 5 find a standard name for:

(a)	$f(-7.51)$	(i)	$h^\circ g(2)$
(b)	$g(-13)$	(j)	$h^\circ g(-2)$
(c)	$h(17)$	(k)	$[f + g](15)$
(d)	$k(-\sqrt{2})$	(l)	$[f + [-g]](15)$
(e)	$m(752)$	(m)	$[n \cdot g](15)$
(f)	$n(.25)$	(n)	$[g \cdot n](10)$
(g)	$m^\circ n(.25)$	(o)	$[4h](12)$
(h)	$n^\circ m(.25)$	(p)	$[f \cdot g](-3)$

7. For the functions given in Exercise 5, determine the rule of assignment for:

(a) $[h + n]$

(b) $[h \cdot n]$

(c) $n^{\circ}m$

(d) $m^{\circ}n$

(e) $[n \cdot g]$

(f) $g^{\circ}n$

(g) $[5g]$

(h) $g^{\circ}k$

CHAPTER 8

DESCRIPTIVE STATISTICS

8.1 Introduction

You have often heard, on radio and television, the warning: Don't be a "statistic"! What does that mean? What are statistics? What are statistical data? The World Almanac and daily newspapers are of statistical data and of statistics. You are a walking bundle of statistical data. Your age, your height, the number of members in your family, your street address, are all examples. In all of these cases we are using the words statistical data (in brief, data) to stand for numbers used to describe observations.

A statistic (descriptive statistic) is a number computed from statistical data. Thus an "average" of two bits of statistical data is a statistic.

In this chapter we will study the gathering of statistical data, its presentation in the form of tables and graphs, and the calculation of certain statistics such as the average and the standard deviation (a measure of the spread of statistical data).

8.2 Examples of Sets of Data and their Graphical Presentation

A mathematics teacher gave the same test on Probability to two different classes each of which had 35 students. There were ten question each of which was worth five points for a correct answer and zero points otherwise. The set of possible test

grades was:

$$\{0,\ 5,\ 10,\ \ldots,\ 45,\ 50\}$$

The results for Class I and Class II are recorded in Table 8.1 and Table 8.2 respectively:

20	30	10	30	0
5	45	50	0	25
20	35	40	20	35
25	40	35	30	30
15	25	45	15	40
40	40	35	35	35
25	35	25	40	20

Table 8.1

TEST GRADES FOR CLASS I

20	15	15	10	45
15	10	20	25	20
25	25	35	20	15
35	20	25	20	15
15	15	30	10	5
20	20	30	30	20
20	30	20	20	15

Table 8.2

TEST GRADES FOR CLASS II

There were only eleven possible grades for the test on probability. It would not make sense for this type of measure to assign other real numbers between two consecutive multiples of five as test grades. You would be quite puzzled, if on such a test your teacher assigned you a grade of the square root of twenty-nine. The test data is an example of one type of discrete data. Later you will work with data gathered from heights, weights, time and other measures where we will think in terms of subsets of the real numbers which include all real numbers within some interval.

This type of data is said to be continuous data.

The data in Table 8.1 and Table 8.2 was taken directly from the teacher's record book in alphabetical order of the students names. For this reason it is difficult to get much of the information that a teacher might desire about the test results. The Frequency Table, Table 8.3, presents the data in such a way that much information is readily obtained from the table.

	CLASS I		CLASS II	
Grades	Frequency	Cumulative Frequency	Frequency	Cumulative Frequency
0	2	2	0	0
5	1	3	1	1
10	1	4	3	4
15	2	6	8	12
20	4	10	12	24
25	5	15	4	28
30	4	19	4	32

35	7	26	2	34
40	6	32	0	34
45	2	34	1	35
50	1	35	0	35

Table 8.3

FREQUENCY TABLE FOR TEST GRADES OF

CLASS I AND CLASS II

<u>Question</u>. Use Table 8.3 to find the difference between the greatest and the least grade received in Class I. Do the same for Class II. Are the differences the same for Class I and Class II?

<u>Definition</u> 1. In any set of data the difference between the greatest and the least measure is called the <u>range</u> of the set of data.

<u>Question</u>. Use the cumulative frequency column for Class I in Table 8.3 to find the middle grade for Class I. Do the same for Class II. For which class was the middle grade greater?

<u>Definition</u> 2. If a discrete set of data has an odd number of measures, the middle measure is called the <u>median</u> of the set of data. If the number of measures is even, the "average" of the two middle measures is the <u>median</u>.

<u>Question</u>. Again use Table 8.3 to find the grade with the greatest frequency in Class I. Do the same for Class II. Are these numbers the same for both classes?

<u>Definition</u> 3. In a set of data a measure that occurs with a frequency at least as large as the frequency of any other measure in the set is called a <u>mode</u> of the set of data.

<u>Example</u>. In the set of data {3,4,4,4,5,5,6,6,6,7,8} both 4 and 6 are modes. Note the three 4's represent <u>different</u> measurements and therefore are different elements of the set of data. The frequency diagrams in Figures 8.1 and 8.2 represent graphically the same set of data that was represented in tabular form in Table 8.3. Some aspects of the data that are obscure in the table are more apparent in the figures.

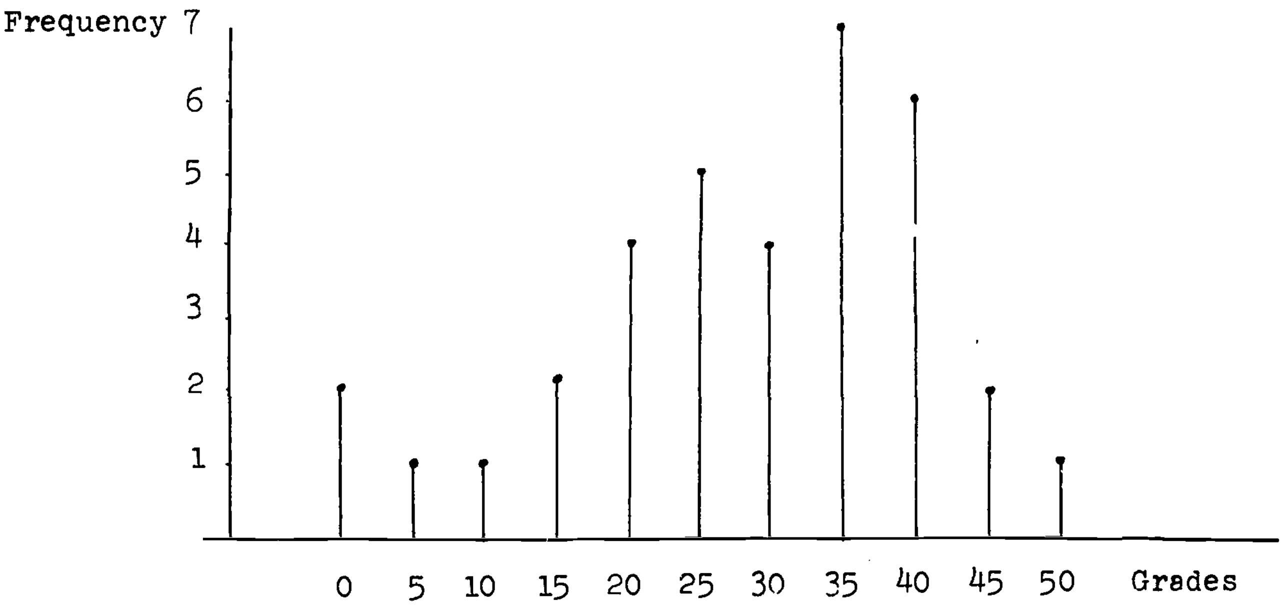


Figure 8.1

Frequency Diagram For Data in Table 8.3

Grades For Class I

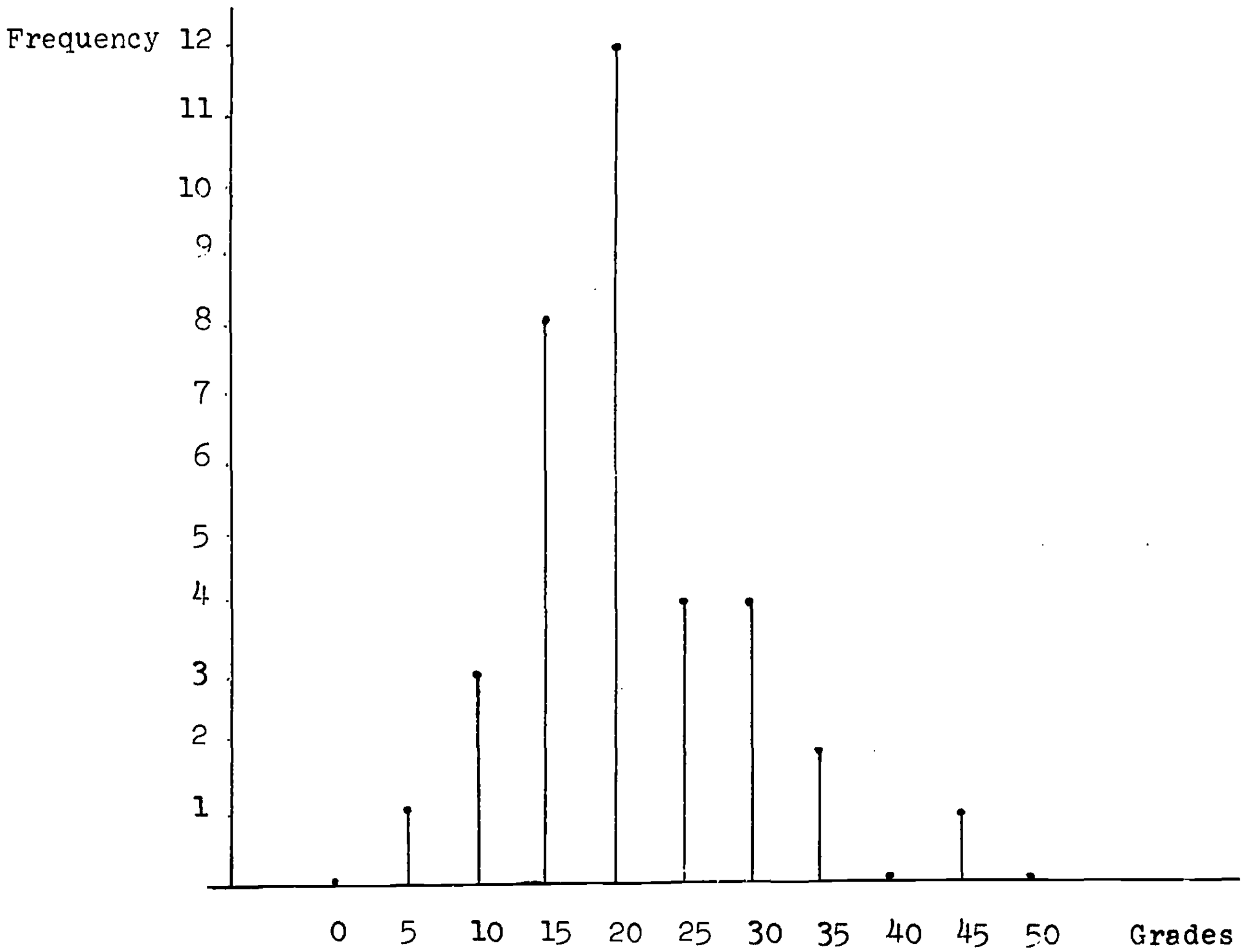


Figure 8.2

Frequency Diagram For Data in Table 8.3

Grades For Class II

Notice that in the graphs, or frequency diagrams, in Figures 8.1 and 8.2 the lengths of segments represent the frequencies of the various grades. Notice also that the range and mode are very easy to determine from the graph. But in addition to this, it is also easy to get some idea of how the grades are scattered or spread by examining the diagram. The two diagrams are placed on the page

in such a way that it is easy to compare the sets of data that are represented.

When you take a test, you may be interested in how well you did on the test with respect to the rest of the class. One way is to find out how many students in the class received a grade that was less than or equal to the grade that you received. The cumulative frequency diagrams in Figures 8.3 and 8.4 provide this information for Class I and Class II. For example, suppose you were in Class I and received a grade of 25. You would look along the horizontal axis until you come to the grade 25. Then you would look at the point on the y-axis which is the same distance from the x-axis as the length of the segment above the grade 25. In Figure 8.3 this shows that there are 15 grades which are which are less than or equal to the grade of 25 for this test. Another way to think of your standing is that you did at least as well as about 43 percent of the class. Looking at it another way, about 57% of the class did better than you.

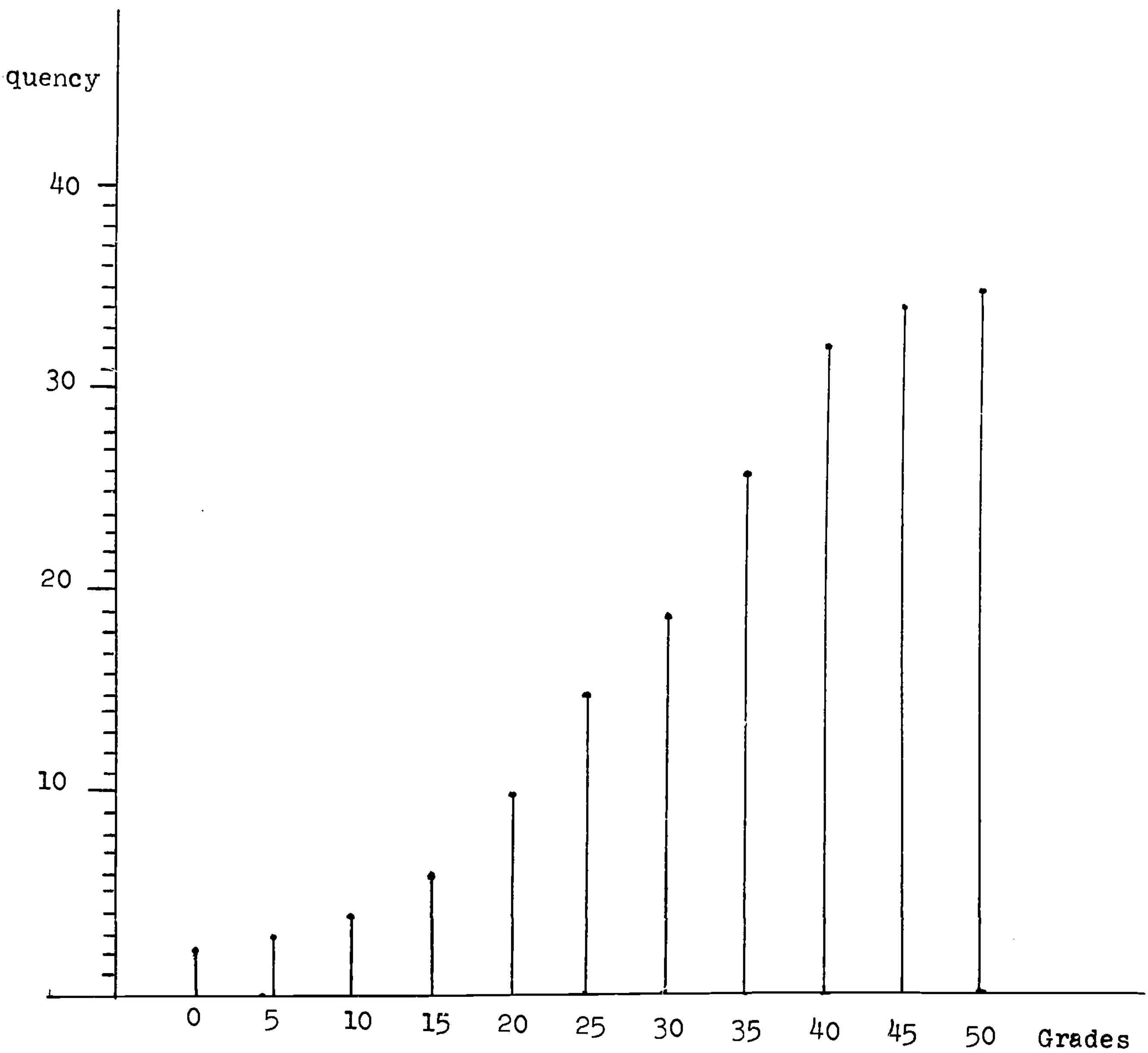


Figure 8.3

CUMULATIVE FREQUENCY DIAGRAM FOR GRADES

OF CLASS I

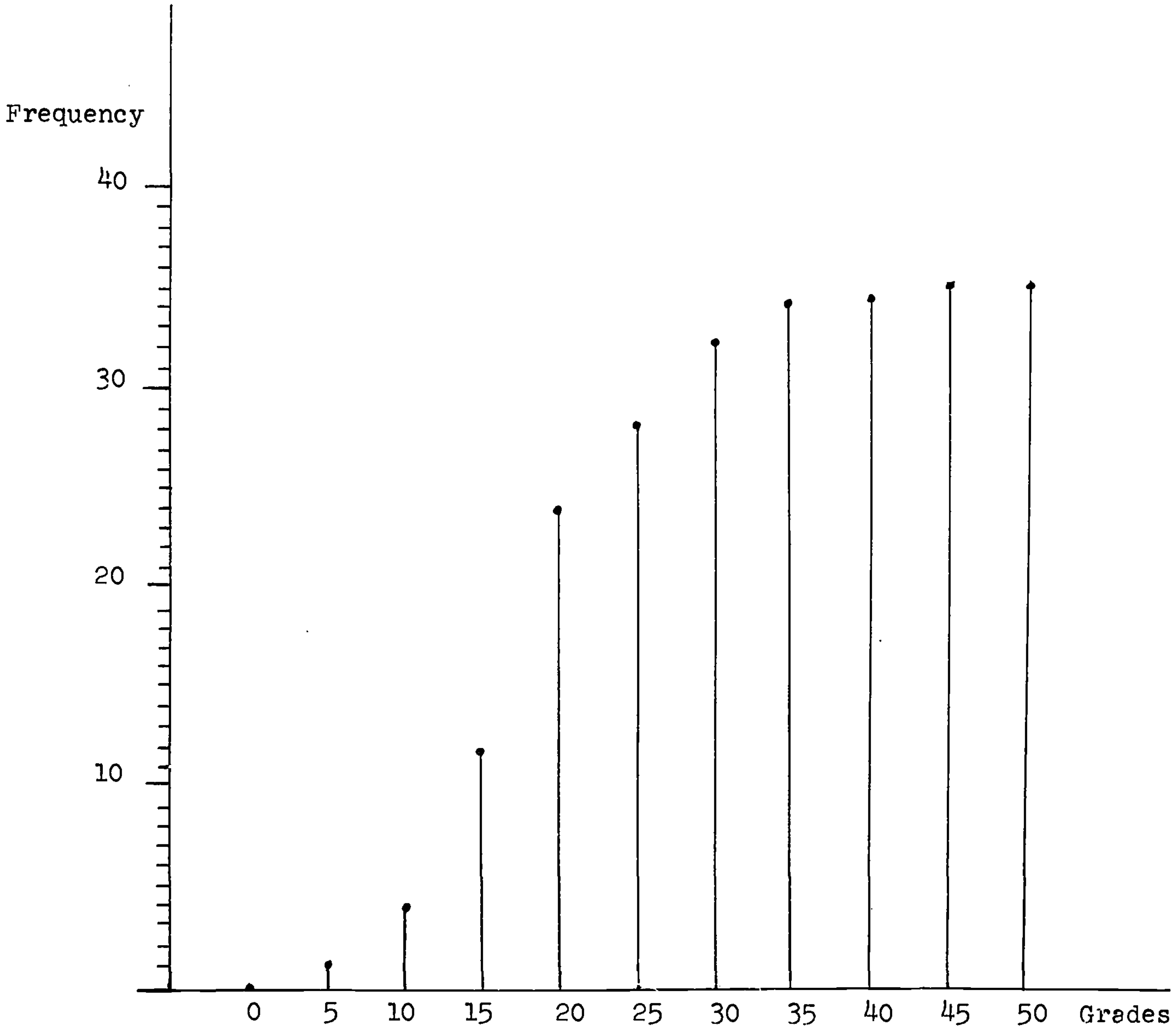


Figure 8.4

CUMULATIVE FREQUENCY DIAGRAM FOR GRADES

OF CLASS II

Next we give an example of a set of data which is based on continuous measure of time. We will illustrate appropriate types of graphical representation for such data.

The quality control engineer in a light bulb manufacturing plant took 50 bulbs from a large lot of 60-watt bulbs manufactured during a day (making sure to select bulbs in such a way that the likelihood of getting a representative sample would be high), and determined the length of life for each of the 50 bulbs by burning them until they expired. To simplify the presentation each of the 50 measurements was rounded off to the nearest multiple of ten. (See Table 8.4.)

910	1110	1010	1070	1290
1000	990	880	780	1150
1150	1030	1030	1270	1310
1050	1170	1180	1380	1080
1230	1060	1130	860	960
1220	930	1050	1080	940
1230	1030	1010	1200	1060
1220	1320	1290	1110	1100
1020	1120	1110	1070	1210
1130	960	1170	950	1070

Table 8.4

50 BULB-LIVES IN HOURS TO NEAREST MULTIPLE OF TEN

Here we have 50 observations. How can we picture and study them? We notice that the shortest life is 780 and the longest 1380. Since the range is fairly large and there are not too many repetitions, we simplify the presentation by grouping the data in intervals of 50 hours with the agreement that the right end-point of each interval belongs to that interval. That is, if a bulb should "expire" at the end of one interval and the beginning of the next, it belongs to the lower interval. If we group the data into intervals of 50 hours starting with 775--825, 825--875, 875--925 etc. that have mid-points 800, 850 etc. we get 13 intervals.

Table 8.5 shows a frequency table for the data on the lives of light bulbs grouped into intervals of 50 hours.

Interval	Midpoints	Frequency	Cumulative Frequency
775--825	800	1	1
825--875	850	1	2
875--925	900	2	4
925--975	950	5	9
975--1025	1000	5	14
1025--1075	1050	10	24
1075--1125	1100	7	31
1125--1175	1150	6	37
1175--1225	1200	5	42
1225--1275	1250	3	45
1275--1325	1300	4	49
1325--1375	1350	0	49
1375--1425	1400	1	50

Table 8.5

FREQUENCY TABLE FOR THE GROUPED DATA ON LIGHT BULBS

In getting information from this table note that the sixth entry in the cumulative frequency column, 24, tells us that there were 24 bulbs that had lives of 1075 hours or less. The last entry in this column shows us that there were 50 light bulbs that had lives of 1425 hours or less.

We can represent the data of Table 8.5 graphically and illustrate certain aspects of the data more clearly by use of a frequency polygon, frequency histogram and a cumulative frequency polygon.

Figure 8.5 shows both a frequency histogram and a frequency polygon.

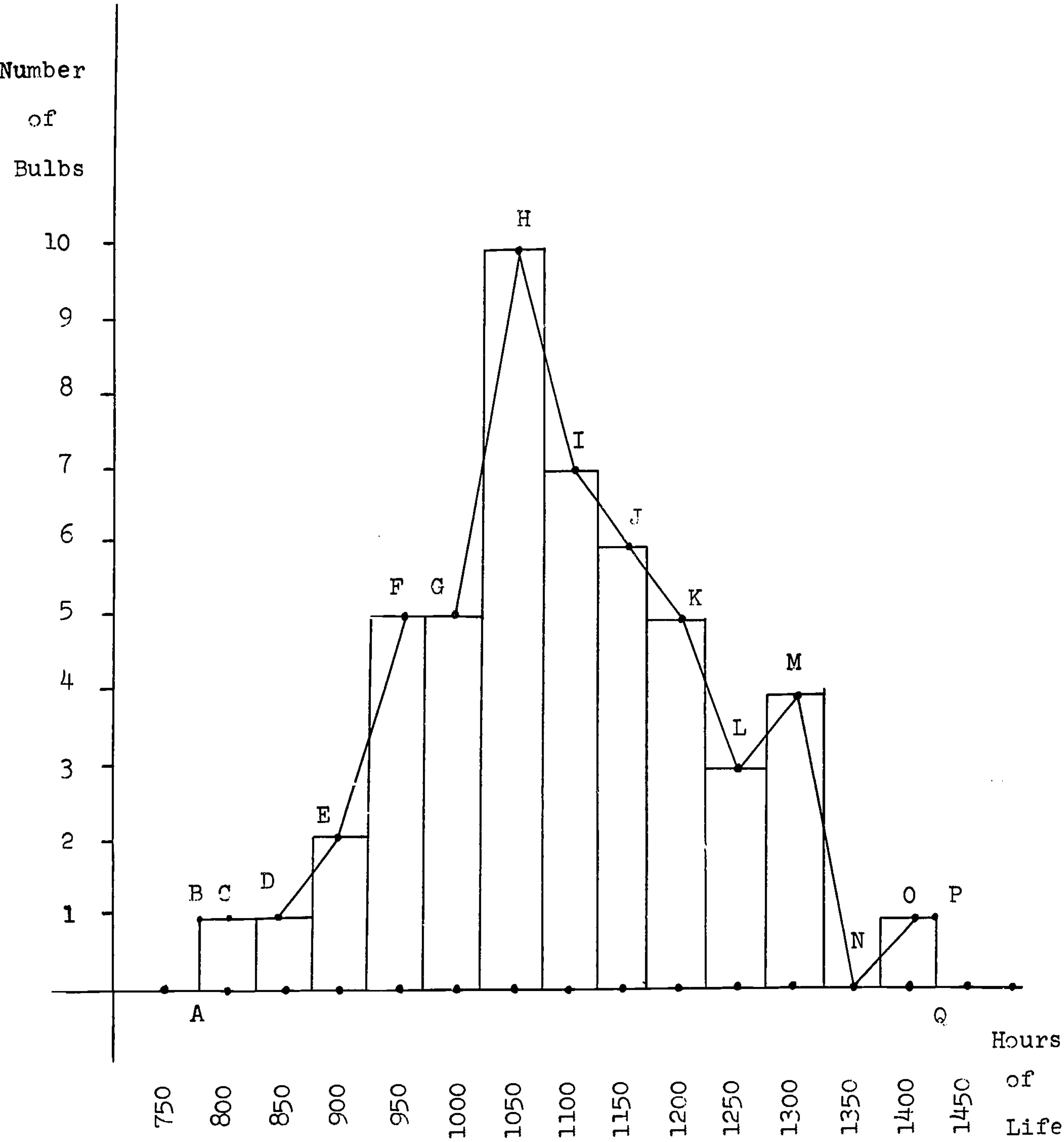


Figure 8.5

FREQUENCY HISTOGRAM AND FREQUENCY POLYGON
FOR THE DATA ON THE LIVES OF THE 50 BULBS

The frequency histogram consists of a number of rectangles, one for each interval whose frequency is not zero in our table. The base of a rectangle is the width of the interval, and the height of the rectangle is the number of bulbs with burning time within the interval.

In Figure 8.5 certain points have been assigned letters to call atention to the Frequency Polygon ABCDEFGHIJKLMNOPQ. Each such labeled point except the first two (A and B) and the last two (P and Q) is the midpoint of either the upper base of a rectangle of the frequency histogram, or of an interval of the histogram (when there is no rectangle over that interval). The points A and B are the lower left-hand vertex and upper left-hand vertex, respectively, of the first rectangle; the points P and Q are the upper right-hand vertex and lower right-hand vertex, respectively, of the last rectangle of the histogram.

Question. Can you give an informal argument to show that the area of the frequency polygon is the same number as the sum of the areas of the rectangles in the histogram?

Question. Can you describe what will happen to the frequency polygon if you keep increasing the number of data and decreasing the size of the intervals?

It is often the case that one wishes to know the number of measures in a set of data that are less than or equal to some particular number. For example in the data on light bulbs, one may be interested in the number of bulbs that had a burning time

less than or equal to 1175 hours. It is also very possible that a quality control engineer might be interested in knowing what percent of the lot of 50 bulbs had a life which was less than or equal to 1150 hours. The cumulative frequency polygon in Figure 8.6 can provide this information. For example the number of bulbs associated with 1175 hours by the polygon is 37.

CUMULATIVE FREQUENCE POLYGON FOR THE DATA IN Table 8.5

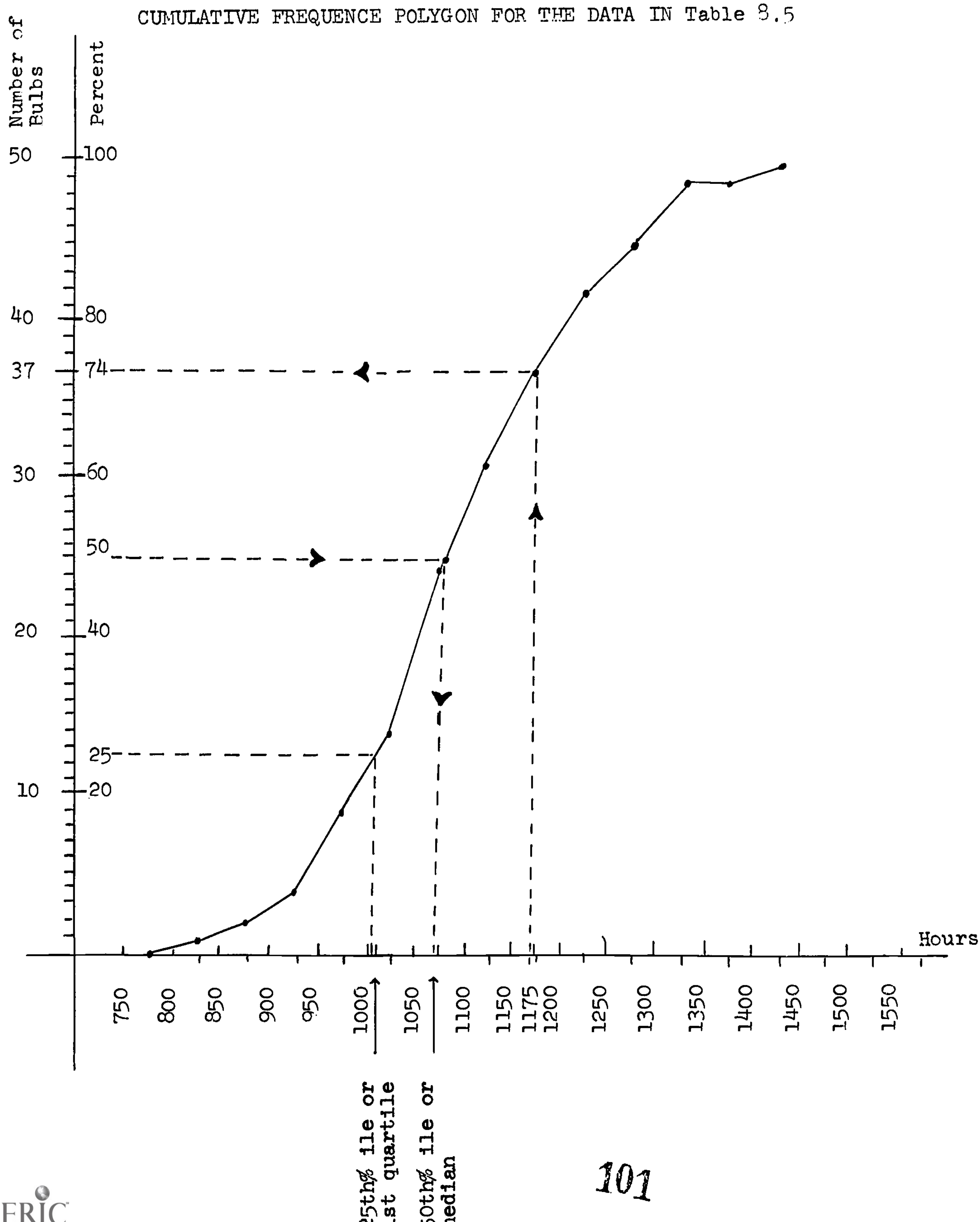


Figure 8.6

From the cumulative frequency polygon in Figure 8.6 it is possible to get many types of information:

(1) You can find the number of bulbs that had a burning time less than or equal to some particular number of hours. For example if you follow the dotted line in Figure 8.6 from 1175 on the horizontal axis up to the graph and then horizontally to the vertical axis you will see that there were 37 bulbs that had a life which was less than or equal to 1175 hours.

(2) You can find the percent of the bulbs that had a life which was less than or equal to some particular number of hours. For example you can see also from the graph (polygonal) that if we label the vertical axis in percent, as we did in Figure 8.6, with 50 bulbs on the original scale corresponding to 100%, we can read off the median and the quartiles. For example, to get the median, we start at the 50% coordinate on the vertical scale, move horizontally until we meet the graph; then move straight down until we hit a point on the horizontal axis. The hours-coordinate of this point is the <u>median</u> for the grouped data. (This means that 50% of the bulbs had lives of 1075 hours or less.)

<u>Definition</u> 4. The 25th percentile is called the <u>first quartile</u> (or lower quartile). The 50th percentile is called the <u>second quartile</u>

(or the median). The 75th percentile is called the third quartile (or upper quartile).

Question (1). From the cumulative frequency polygon in Figure 8.6 what is the 8th percentile of the set of data represented on the horizontal axis?

(2). Between which two numbers represented on the horizontal axis does the median lie?

8.3 Exercises

After you do the following exercises, save the results. They will be used in later exercises.

1. The length (number of words) of 40 sentences taken from a certain portion of Toynbee's A Study of History were as follows:

24	39	46	22	51	20	48	58	39	60
28	19	44	54	80	35	36	23	15	21
43	18	12	19	26	38	25	7	17	22
17	70	42	12	15	65	39	73	26	42

(a) Make a frequency diagram and a frequency table of these numbers.

(b) Determine the range and median of the sentence lenghts.

2. In a deck of cards let the Ace, 2, ..., 10 be assigned measurements 1 through 10 respectively, and let Jack, Queen, King be assigned 11, 12, 13, respectively. Draw 25 samples

of 3 cards each and record the sum of the measurements on the three cards. Shuffle before each draw. Before beginning write down your guess for the median of these data.

(a) Make a frequency diagram and a frequency table of your data.

(b) Determine the range and median of the data.

(c) Compare your guess with the actual result.

3. Throw 20 pennies 25 times and record the total number of heads obtained on each throw. Before beginning write down your guess for the median of your data.

(a) Make a frequency diagram and a frequency table of these data.

(b) Indicate the median from the frequency table and compare this with your guess.

4. Throw 3 dice 20 times and count the total number of dots which turn up on each throw. Before beginning write down your guess for the median of your data.

(a) Make a frequency diagram and a cumulative frequency table for the 20"measurements" thus obtained.

(b) Indicate the median from the cumulative frequency table and compare it with your guess.

5. Open a telephone directory at random and select one column on the chosen page.

(a) Make a listing of the last digit of each telephone number in the chosen column.

(b) Before you list the digits, can you make a prediction of what you'll find?

(c) Summarize your observations in 3 different ways.

6. The heights of the 14 year - old boys (in inches, to the nearest inch) in a junior high school were recorded as follows:

58	53	56	53	57	51	60	55
61	54	65	58	54	54	56	57
54	55	59	54	56	57	55	54
57	53	54	55	62	59	58	58
53	59	56	52	55	55	55	55
55	54	57	57	53	56	56	50
63	52	61	55	55	53	52	56
56	57	56	60	60	58	57	59

(a) Group these measurements into intervals of the same length and construct a frequency table as shown in Table 8.5.

(b) Make a frequency histogram and a frequency polygon for the data in part (a) similar to the one in Figure 8.5.

(c) Construct a cumulative frequency polygon for the data, and on it find the median and the quartiles.

7. Check the construction of the frequency histogram (Figure 8.5) and the cumulative frequency polygon (Figure 8.6) for the

illustrative example in Section 8.2.

8. From the cumulative frequency polygon of Exercise 7 estimate the median, the upper quartile, and the lower quartile.

9. Suppose the Scholastic Aptitude Test scores of 180 seniors of a certain school range from 330 to 788.
What interval mid-points and interval boundaries would you use for a frequency table? Make up the first two columns for such a table.

10. Each of 50 measurements is given to four decimal places, the smallest measurement being 0.9967 inches and the largest being 1.0048 inches. Determine equal intervals for grouping the measurements, and suggest interval boundaries and interval midpoints for a frequency table that you might construct.

11. Tabulate the set of weight-measures for students in your class (include yourself).
    (a) Decide on intervals and midpoints to condense the observations.
    (b) Construct a frequency table.
    (c) Construct a frequency histogram and a cumulative frequency polygon.
    (d) Determine the median and the 1st and 3rd quartiles.

8.4 The symbol Σ and summation.

So far in this chapter you have been working with finite sets

of data and representing them in various ways. You will frequently deal with sums of such sets of data and represent the sums in various ways. Some of these ways will be more convenient than others depending on your purpose.

Consider the data 5, -3, 11, 8. This can be represented more generally by the symbols x_1, x_2, x_3, x_4 which is read "x-sub-one, x-sub-two, x-sub-three, x-sub-four". This is a representation for any set of data with four elements and each of the sub-scripts, 1, 2, 3, and 4 serves as an index here for the purpose of distinguishing among the four elements. Moreover we can relate the specific and general representations by writing, $x_1=5$, $x_2=(-3)$, $x_3=11$, $x_4=8$.

Question. How can you interpret the above representations as a mapping of a set of data onto a subset of the natural numbers?

Similarly we now have two representations for the sum of the above data:

$$x_1 + x_2 + x_3 + x_4 = 5 + -3 + 11 + 8.$$

Now, if we let i be a variable for the set of natural numbers we can use the symbol x_i, read "x-sub-i", to stand for any definite number in a set of data that we do not wish to specify (and do not really need to).

You will agree that if we had a very large set of data, the above representation would be quite awkward. Suppose we had a set containing 99 numbers. We could improve our notation some by the following:

$x_1 + x_2 + x_3 + \ldots + x_i + \ldots + x_{99}$. In this case the "$x_i$" as used here, is sometimes called a typical element of the summation. The subscript i is called the index of summation.

Mathematicians take the notation one more step to this shorthand form of representing a sum:

$$\sum_{i=1}^{99} x_i$$

This symbol represents the same sum of 99 numbers represented above. Actually both symbols can represent any 99 element sum. This is read:

"The summation of x-sub-i as i goes from 1 through 99".

The symbol $\sum$ is the capital Greek letter, Sigma, which corresponds to the English capital letter S.

The "1" below the sigma and the "99" above represent the limits of summation.

Now let us use our new notation on the four element set that we mentioned at the beginning of this section, {5, -3, 11, 8}. We write:

$$\sum_{i=1}^{4} x_i = x_1 + x_2 + x_3 + x_4 = 5 + -3 + 11 + 8 = 21$$

Example 1. Suppose $x_1 = 1$, $x_2 = 2$ and in general $x_i = i$ up to the upper limit of summation. It then is sensible to replace x_i by i in the summation notation:

$$\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} i = 1 + 2 + \ldots + n$$

Example 2. Suppose the typical element is some function of x_i such as $x_i - a$. Then the summation will be as follows:

$$\sum_{i=1}^{n} (x_i - a) = (x_1 - a) + (x_2 - a) + \ldots + (x_n - a)$$

Example 3. In general:

$$\sum_{i=1}^{n} f(x_i) = f(x_1) + f(x_2) + \ldots + f(x_n).$$

8.5 Exercises.

1. If $x_1 = 5$, $x_2 = 6$, $x_3 = 3$, $x_4 = 1$ find:

(a) $\sum_{i=1}^{4} x_i$ (b) $\sum_{i=1}^{3} x_i^2$ (c) $\sum_{i=1}^{4} 5x_i$

(d) $\sum_{i=1}^{4} (x_i + 5)$ (e) $\sum_{i=1}^{4} (x_i + k)$

2. From illustrative Example 1, $\sum_{i=1}^{5} i$ means $1 + 2 + 3 + 4 + 5$.

Using this idea, find:

(a) $\sum_{i=1}^{7} i$ (b) $\sum_{i=1}^{5} i(i+1)$ (c) $\sum_{i=0}^{4} i^2$

3. Show that $\sum_{i=1}^{n} k x_i = k\sum_{i=1}^{n} x_i$, when k is a constant.

(See Exercise 1 (c).)

4. Show that $\sum_{i=1}^{n} (a_i + b_i) = \sum_{i=1}^{n} a_i + \sum_{i=1}^{n} b_i$

(a) First try this using the value n = 3.

(b) Can you state the equation in words?

5. Show that $\sum_{i=1}^{n} (a_i - b_i) = \sum_{i=1}^{n} a_i - \sum_{i=1}^{n} b_i$.

6. Show that $\sum_{i=1}^{n} k = kn$, where k is a constant. (Compare with a part of Exercise 1 (d).)

7. Show that

$$\sum_{i=1}^{n} (x_i - m)^2 = \sum_{i=1}^{n} x_i^2 - 2m\sum_{i=1}^{n} x_i + nm^2$$

8.6 The Arithmetic Mean, its Computation and Properties.

An important descriptive statistic of a set of numbers is its arithmetic mean. Suppose that the members of a club pooled all the dollars that each of them possessed and then distributed these dollars equally among the members. (A very unlikely event!) The number of dollars that they would then each have would be the

arithmetic mean of the set of numbers of dollars that they originally had individually. A more likely event is that your teacher would assign you the arithmetic mean of your monthly grades as a final grade for the year.

<u>Definition</u> 5. The <u>arithmetic</u> <u>mean</u> (or just <u>mean</u>) of a set of n numbers, $x_1, x_2, \ldots, x_n$, is:

$$\frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{n}(x_1 + x_2 + x_3 + \ldots + x_n)$$

Using $\overline{x}$, read "x-bar," as a symbol to express the arithmetic mean of a set of n numbers, the definition can be expressed as follows:

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

From the formula above we immediately get a new formula.

$$n\,\overline{x} = \sum_{i=1}^{n} x_i$$

which states that the sum of the measures in an experiment is equal to the product of the mean and the number of measures.

<u>Example</u> 1. Given the numbers

$$x_1 = 7,\ x_2 = 10,\ x_3 = 15$$

$$\sum_{i=1}^{3} x_i = 7 + 10 + 15 = 32$$

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{3} x_i = \frac{1}{3} \cdot 32 = 10\tfrac{2}{3}$$

Then $n\,\bar{x} = (3)\ (10\frac{2}{3}) = 32 = \sum_{i=1}^{3} x_i.$

The problems we have done and the symbolism we have introduced give us shortcuts for computation.

Example 2. Suppose we find a set of 15 scores which are distributed as follows:

Score	Frequency
60	5
55	7
50	3

How do we find the mean? The definition of **mean is** the sum of the scores divided by the number of scores.
It is easy to see from this data that:

$$\bar{x} = \frac{60(5) + 55(7) + 50(3)}{5 + 7 + 3}$$

This suggests a more general procedure outlined as follows:

Scores	Frequency
x_1	f_1
x_2	f_2
x_3	f_3

It follows from the definition of mean that we multiply each score by its frequency, add the products, and divide by the total of the frequencies, that is

$$\bar{x} = \frac{x_1 f_1 + x_2 f_2 + x_3 f_3}{f_1 + f_2 + f_3} = \frac{\sum_{i=1}^{3} x_i f_i}{\sum_{i=1}^{3} f_i}$$

For our problem, the computation can be presented by introducing another column:

x_i	f_i	$x_i f_i$
60	5	300
55	7	385
50	3	150
	$\sum_{i=1}^{3} f_i = 15$	$835 = \sum_{i=1}^{3} x_i f_i$

Then the mean would be:

$$\bar{x} = \frac{\sum_{i=1}^{3} x_i f_i}{\sum_{x=1}^{3} f_i} = \frac{835}{15} = 55\frac{2}{3}\ .$$

We can use the technique illustrated above to advantage with the data of Table 8.5 as illustrated in Table 8.6.

x_i (midpoints)	f_i	$y_i = x_i - 1100$	$z_i = \frac{y_i}{50}$	$f_i z_i$	$f_i z_i^2$
800	1	-300	-6	-6	36
850	1	-250	-5	-5	25
900	2	-200	-4	-8	32
950	5	-150	-3	-15	45

1000	5	-100	-2	-10	20
1050	10	- 50	-1	-10	10
1100	7	0	0	0	0
1150	6	50	1	6	6
1200	5	100	2	10	20
1250	3	150	3	9	27
1300	4	200	4	16	64
1350	0	250	5	0	0
1400	1	300	6	6	36
	$\sum_{i=1}^{13} f_i = 50$			$\sum_{i=1}^{13} f_i z_i = -7$	$\sum_{i=1}^{13} f_i z_i^2 = 321$

Table 8.6

When you look at the x_i column and the f_i column in the table it may seem that the computation of the mean would be formidable but we can use what we learned about summation to obtain two ideas about the mean that will simplify the process.

Since $\sum_{i=1}^{n}(x_i + k) = \sum_{i=1}^{n} x_i + nk$ (see Section 8.5 Exercises 1(d) and 1(e)),

it follows that $\frac{1}{n}\sum_{i=1}^{n}(x_i + k) = \frac{1}{n}(\sum_{i=1}^{n} x_i + nk) = \frac{1}{n}\sum_{i=1}^{n} x_i + k.$

In other words, adding the same number to every element in a set of data increases the mean of the set by that number.

Also since $\sum_{i=1}^{n} kx_i = k\sum_{i=1}^{n} x_i$ (See Section 8.5 Exercise 3.)

it follows that $\frac{1}{n}\sum_{i=1}^{n} kx_i = k \cdot \frac{1}{n}\sum_{i=1}^{n} x_i.$

In other words, multiplying every element in a set of data by the ne number multiplies the mean of the set by the same number. The above two generalizations can also be shown to apply to subtraction and divi-

sion since they are defined in terms of addition and multiplication.

Notice in Table 8.6 we subtracted 1100 from each of the x_i to get the y_i. This means that the mean of the y_i is 1100 less than the mean of the x_i. We then divided each of the y_i by 50 to get the z_i. This means that the mean of the z_i's is one-fiftieth of the mean of the y_i's.

> Note. To better understand the transformations that we have been using on the sets of data above, see Section 6.4, Relating Two Coordinate Systems on a Line in Chapter 6, Coordinate Geometry in this text.

Now we have a set of small integers, the z_i, whose mean we can find by the formula:

$$\overline{z} = \frac{\sum_{i=1}^{13} z_i f_i}{\sum_{i=1}^{13} f_i}$$

In our example $\sum_{i=1}^{13} z_i f_i = -7$, $\sum_{i=1}^{13} f_i = 50$

Then $\overline{z} = -\frac{7}{50} = -.14$

We previously learned the following facts:

Multiplying each of a set of numbers by c, multiplies their mean by c.

Adding h to each of a set of numbers increases their mean by h.

In our example:

$$\overline{y} = 50\ \overline{z} = 50(-.14) = -7$$

$$\overline{x} = \overline{y} + 1100 = -7 + 1100 = 1093$$

We have found the **mean** of the set of bulb-lives from the grouped data. We assumed that all the measurements in an interval were concentrated at the midpoint.

We close this section with an important theorem about the sum of the deviations of the numbers in a set from the mean of the set. That is, you subtract the mean of the set from each of the members in the set, and then add the differences. (Note that some of the differences will be negative.)

Theorem 1. The sum of the deviations from the mean is equal to zero.

Given the n measurements $x_1, x_2, \ldots, x_n$ and their mean $\overline{x}$, to prove

$$\sum_{i=1}^{n} (x_i - \overline{x}) = 0$$

Proof. $\sum_{i=1}^{n} (x_i - \overline{x}) = \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} \overline{x}$. (Exercise 5 in Section 8.5.)

$$\sum_{i=1}^{n} x_i = n\,\overline{x}. \quad \text{(Why?)}$$

$$\sum_{i=1}^{n} \overline{x} = n\,\overline{x}$$

(Since $\overline{x}$ is a constant; see Exercise 6 of Section 8.5).

Thus:

$$\sum_{i=1}^{n} (x_i - \overline{x}) = \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} \overline{x} = n\,\overline{x} - n\,\overline{x} = 0$$

8.7 Exercises

1. Find the mean of the measurements 1,6, 8. If 7 is added to each of these measurements, what is the mean of the new set of measurements?

2. Given the two sets of measurements 10,9,2 and 3,5,1 find the mean of each set. Now form a new set of measurement by adding corresponding measurements in the two sets (the first number in the new set is 10 + 3 = 13.) Find the mean of the new set of three measurements, and relate it to the means of the first two sets.

3. Given the measurements 1,6,8 form a new set of measurements by multiplying each of the original ones by 7. Find the means of the two sets and compare them.

4. Given the measurements 4,5,9 find the mean. Also find the mean of a new set of measurements obtained by multiplying each measurement by 7 and then adding 5 to the result. Relate the new mean to the old mean.

5. In changing from the Centigrade temperature scale to the Fahrenheit Scale, the relation is $F = \frac{9}{5} C + 32$ where C is the measurement on the Centigrade scale and F the corresponding measurement on the Fahrenheit scale. In a chemical

experiment the following 10 measurements (in degrees centigrade) were taken:

40, 35, 38, 36, 42, 41, 37, 42, 37, 35.

(a) Find the mean of these measurements.

(b) What is the mean of these measurements in the Fahrenheit scale?

6. The number of students in five different algebra classes ~~are 32~~, 36, 30, 41, and 38. Find the mean number of students per class, and compare it with the median number of students.

*7. Suppose that you have a mapping, $x_i \longrightarrow (cx_i + h)$ where i=1, ..., n and where c and h are real numbers. How will the mean of the domain be related to the mean of the range? Try to prove your conjecture!

8. Find the average (arithmetic mean) of the life-lengths of the bulbs in the illustrative example of Section 8.2.

(a) Now group these 50 numbers into groups of 5, and compute each group average; then average the 10 averages. What do you find?

(b) Can you think of some other short cut for finding the average of the 50 measurements?

9. Find the mean of the numbers given in Exercise 1 of Section 8.3.

10. Find the mean of the observations in your experiment in Section 8.3 (Exercise 2,3 or 4). Use any short cut you can think of.

11. Find the mean of the heights of 14-year old boys in Exercise 6 of Section 8.3.

    (a) Before you do your calculations try to guess the mean.

    (b) Can you find a shortcut by using your <u>frequency table</u>?

    (c) Compare the mean of these measurements with the median and the mode.

12. Find the mean of the observations in Exercise 11 of Section 8.3.

13. Find the mean and the median for the numbers: 15, 18, 18, 21, 45, 63, 69, 78, 45, 45, 27, 36, 60.

14. Find the mean and the median for the numbers: 5.7, 4.6, 8.2, 5.7, 3.6, 2.8, 4.9, 5.7, 6.2, 9.1.

15. In a class of 30 pupils, on a certain examination, 5 got grades of 65, 10 got 70, 12 got 80 and 3 got 90. Calculate the mean of these marks by using the formula

$$\frac{1}{n}\sum_{i=1}^{n} x_i f_i$$

16. In a certain plant employing 100 workers the average salary is \$90 per week. In another plant, employing 50 workers, the average salary is \$110 per week.

    (a) What is the total weekly payroll for the two plants?

    (b) What is the average salary for all the workers in the two plants?

(c) Do you get the same result by averaging \$90 and \$110? Why?

17. The table below gives the frequencies of pupils of ages 12 - 17 in a certain school.

Age	Frequency
12	10
13	40
14	175
15	105
16	55
17	10

(a) What interval limits are indicated by this frequency table?

(b) Construct a histogram and a cumulative frequency polygon for these data.

18. A chicken farmer found that his hens averaged 350 eggs per day in a certain week. His records for six days of that week show the following counts: 347, 351, 358, 345, 350, and 353, but he lost the record for the seventh day. What must it have been?

19. Prove: $$\sum_{i=1}^{n} (x_i - a)^2 = \sum_{i=1}^{n} (x_i - \overline{x})^2 + n(\overline{x} - a)^2$$

Hint: Expand $(x_i - a)^2$ and $(x_i - \overline{x})^2$. Then apply results of the preceding section to simplify.

8.8 Measures of Dispersion

We have studied certain descriptive statistics of sets of data that measure central tendency such as the mode, median and the arithmetic mean. So far the only measure of dispersion or scatter of a set of data that has been mentioned is the range which does not seem to be too significant since the computation of the range is completely dependent on only two numbers of the set.

Before looking for a good measure of dispersion let us examine the frequency diagrams (a) -- (e) in Figure 8.7. If you compute the mean from the diagram for each of the sets of data, you will discover that each set of data has the same mean. See if you can find a line of symmetry for each of the diagrams! It is apparent from the diagrams that what we should like to call the dispersion of these sets differs for any pair.

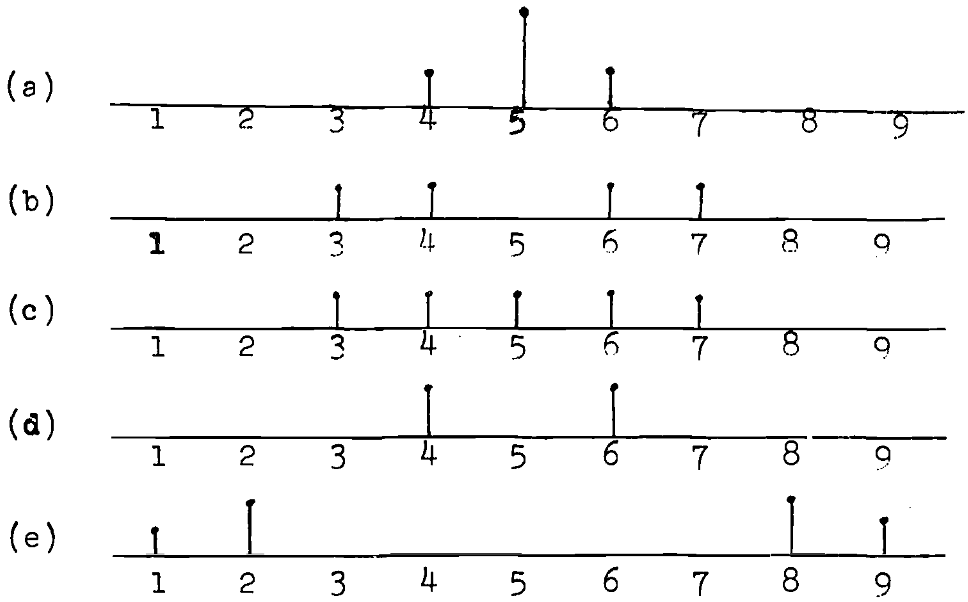


Figure 8.7

Frequency Diagram ($\overline{x}=5$)

In Figure 8.8 (f) and (g) have the same mean but are dispersed differently and likewise for (h) and (i).

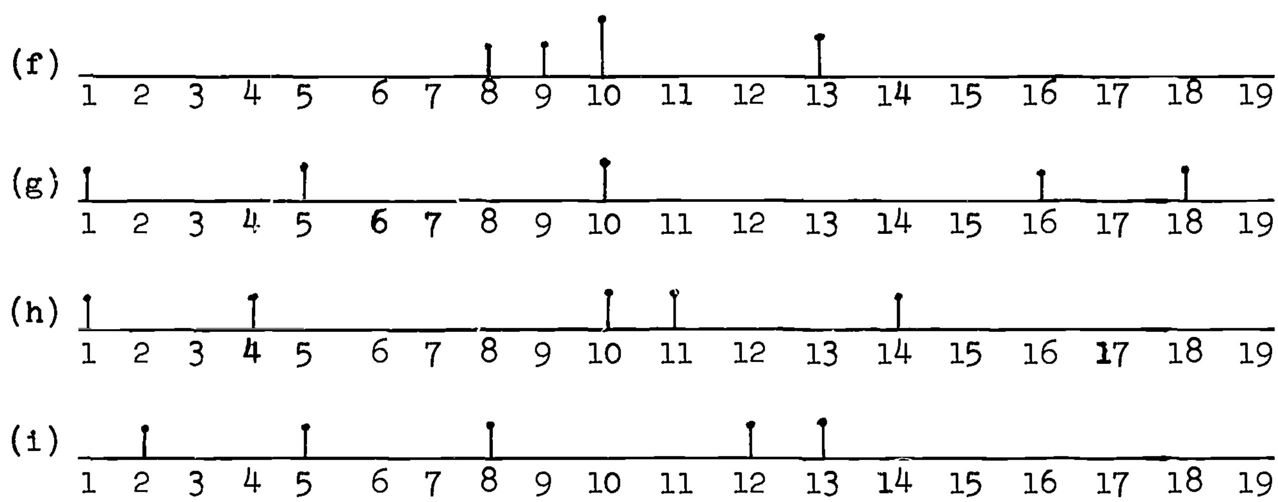


Figure 8.8

Another Frequency Diagram

The examples above should be enough to convince us that any measure of dispersion must be independent of the arithmetic mean. We would like to find a measure that is small when the numbers seem to cluster and large when they are scattered.

An obvious suggestion for measuring the dispersion of a set of numbers is to find how their measures differ, on the average, from the mean. However, we already know that this is not useful. We know that the sum of the deviations from the mean is zero. (See Theorem 1 in Section 8.6)

One way around this difficulty is to disregard the signs of the deviations. You will be asked to work out the averages of the absolute deviations for the sets represented in Figures 8.7 and 8.8 in the exercises of Section 8.9. The measure found in this

way is called the <u>mean absolute deviation</u>.

Although the mean absolute deviation has the advantage of always being non-negative, it has disadvantages. One is that it is not feasible to compute the mean absolute deviation of large sets of data from the mean absolute deviations of their proper subsets.

Another measure of dispersion, one that is widely used, is called the <u>variance</u> of the set of data.

<u>Example</u>.

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
4	-1	1
5	0	0
5	0	0
5	0	0
6	+1	1
$\Sigma x_i = 25$; $\bar{x} = 5$		$2 = \Sigma(x_i - \bar{x})^2$

Variance $= \frac{1}{5}\sum_{i=1}^{5}(x_i - \bar{x})^2 = \frac{2}{5}$

Table 8.7

<u>Definition</u> 6. The <u>variance</u> of a set of numbers x_1, ---, x_n, denoted s^2, is defined:

$$s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

There is one objection to the variance as a measure of dispersion. This objection is that, since the computation of the variance involved squaring the deviations from the mean, we have

in a sense changed the dimension of the original. That is, the variance is not expressed in the same units of measure as the numbers in the original set of data. Fortunately, this is easily remedied. Simply find the square root of the variance and you have a measure of dispersion that is in the same units as the original data. This new measure is called the standard deviation of a set of data.

Definition 7. The standard deviation s is the square root of the variance s^2.

Sometimes the variance is referred to as the "average of the squared deviations". Similarly, the standard deviation is sometimes expressed as the "square root of the average of the square deviations".

8.9 Exercises

1. Complete the Table below for the observations given in Figures 8.7 and 8.8.

	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	
Measures	4,5, 5,5, 6	3,4, 6,7								(1)
n	5	4	5	4	6	5	5	5	5	(2)
$\bar{x}$	5	5	5	5	5	10	10	8	8	(3)
Range	2	4	4	2	8	5	17	13	11	(4)
$x_i - \bar{x}$	-1,0, 0,0, +1	-2,-1 +1,+2								(5)
$\Sigma\lvert x_i - \bar{x}\rvert$	2	6								(6)
$\frac{\Sigma\lvert x_i - \bar{x}\rvert}{n}$	.4	1.5								(7)
$(x_i - \bar{x})^2$	1,0 0,0 +1	4,1 1,4								(8)
$\Sigma(x_i - \bar{x})^2$	2	10								(9)
variance s^2	.4	2.5								(10)
Standard Deviations	.63 app.	1.6 app.								(11)

(a) Lines 7, 10 and 11 give us several different measures of dispersion. Compare them. Do they rank the dispersions of the sets as you would intuitively? Are they easy to compute?

(b) Can you find the standard deviation from the variance by intelligent guessing?

2. The numbers of students in 5 different mathematic classes are 22, 26, 20, 31, and 26.

(a) Find the mean number of students per class and compare it with the median.

(b) Find the mean absolute deviation.

(c) Find the variance.

(d) Find the standard deviation.

3. Find the standard deviation of the seven measurements in Exercise 18 Section 8.7.

4. Given the measurements 8, 10, 24 compute the mean, the variance, and the standard deviation.

(a) Now subtract 3 from each of the measurements and compute the mean, the variance, and the standard deviation.

(b) What observation do you make?

5. Compute the mean and the variance of the three measurements 1,6, and 8. If 9 is added to each of these measurements what is the variance of the new set?

6. Show that the variance of a set of measurements is unchanged when the same constant is added to each measurement in the set.

7. Compute the mean and the variance of the measurements 8999, 8997, 9001. Choose a convenient number, say 9000, and subtract it from each of the measurements, getting -1, -3, +1. Compute the mean and the variance of the new triple. Can you obtain the mean and the variance of the original measurements from your results?
8. If each measurement in a set is multiplied by the same constant k, show that the variance is multiplied by k^2, and that the standard deviation is multiplied by $|k|$.

8.10 Simplified Computation of the Variance and the Standard Deviation

While the definitions of the variance and the standard deviation are simple enough, the computations for any sizeable number of data are formidable. For example, using the data in Table 8.6 to calculate the variance and standard deviation would seem to require the following computations:

1. Add the fifty measurements and divide by fifty to find their mean.
2. Calculate the differences of the fifty individual scores from the mean.
3. Square each of the fifty differences thus obtained.
4. Sum these squares, and divide the sum by fifty to find the variance.
5. Take the square root of the variance to get the standard deviation.

Are there any shortcuts we can use?

The problems in the preceding section and the shortcuts we used in calculating the mean in Section 8.6 may give us some ideas.

Our definition for the variance is, using subscripts and the sigma notation;

$$s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

We can expand this in the following way: (Review Sections 8.4 and 8.5.)

$$s^2 = \frac{1}{n}\sum(x_i^2 - 2x_i\bar{x} + \bar{x}^2)$$

$$= \frac{1}{n}\left[\sum x_i^2 - 2\bar{x}\sum x_i + \sum\bar{x}^2\right]$$

$$= \frac{1}{n}\sum x_i^2 - (\frac{1}{n} \cdot 2 \cdot \bar{x} \cdot n\bar{x}) + \bar{x}^2$$

$$= \frac{1}{n}\sum x_i^2 - 2\bar{x}^2 + \bar{x}^2$$

$$= \frac{1}{n}\sum x_i^2 - \bar{x}^2$$

You may have observed that we replaced $\sum_{i=1}^{n}$ by $\sum$. When the content of the problem leaves no room for doubt we will often drop the limits of the summation.

Question: Can you show:

$$s^2 = \frac{n\sum x_i^2 - (\sum x_i)^2}{n^2}?$$

For data grouped into k classes the formula for variance is:

$$s^2 = \frac{1}{n}\sum_{i=1}^{k}(f_i x_i - \overline{x})^2 \text{ where } \overline{x} = \frac{1}{k}\sum_{i=1}^{k} f_i x_i \text{ and } \sum_{i=1}^{k} f_i = n.$$

For computation we use $s^2 = \frac{1}{n}\sum_{i=1}^{k} f_i x_i^2 - \overline{x}^2$.

Using these results on the data in Table 8.6 we find that:

$$\sum f_i z_i^2 = 321$$

Then the variance for the z scores is:

$$s^2 = \frac{1}{n}\sum f_i z_i^2 - \overline{z}^2 = 6.42 - (-.14)^2 = 6.42 - .0196 \approx 6.42 - .02 = 6.40$$

Now using the results of Exercises 6 and 8 in Section 8.9, we get the variance of the original x-scores by multiplying the variance for the z-scores by $(50)^2$. Thus the variance for the x-scores is approximately 16 000.

The standard deviation of the z-scores is approximately equal to $\sqrt{6.40} \approx 2.5$

The standard deviation of the x-scores, then, is

$$50\ (2.5) = 125.$$

(Check that $\sqrt{16000}$ gives the same result.)

8.11 Exercises

Use shortcuts where possible in your computations.

1. Check the computations for the variance and standard deviations of the data in Table 8.6.
2. Find the variance and the standard deviations of the data in Exercise 11 of Section 8.3.

3. Find the variance and the standard deviation of the observations in any experiment that you performed in Section 8.3, for example Exercises 3,4 or 5.

*8.12 The Chebyshev Inequality

In this section we will prove an inequality, that illustrates the fact that the standard deviation, s, is a measure of scatter (or dispersion) in a set of observations.

Suppose that the observations are $x_1, x_2, \ldots, x_n$ and that the mean is $\bar{x}$ and the standard deviation is s. Thus we have:

$$(1)\quad \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \quad \text{and} \quad s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

Note. You will notice that previously we have represented the variance as the sum of the squared deviations multiplied by 1/n. Now we are multiplying by 1/(n-1). The reason is beyond the scope of this chapter. For large n the numerical difference is slight.

Let k be a real number greater than 1, i.e. $k > 1$.

We define the following two sets A and B:

$$A = \{x_i : |x_i - \bar{x}| < ks\} \; ; \quad B = \{x_i : |x_i - \bar{x}| \geq ks\}$$

In Figure 8.9 we have illustrated the two sets A and B:

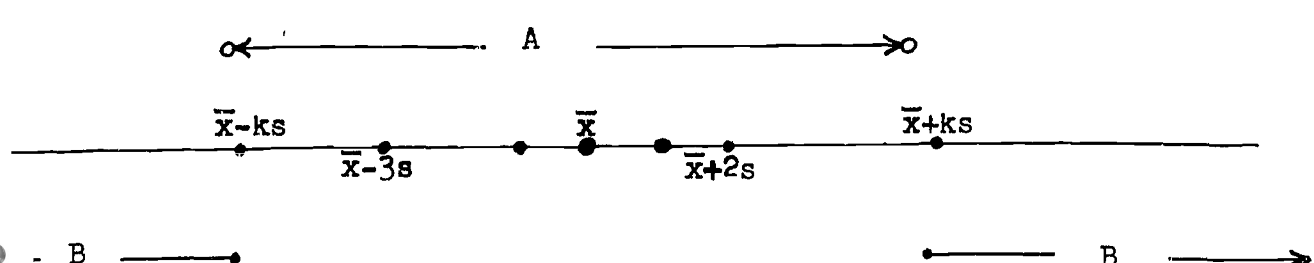


Figure 8.9

We see that A is the set of observations with a distance from $\overline{x}$ less than ks and B that set of observations with a distance at least ks from $\overline{x}$.

Let m(B) be the number of observations in B.

The Chebyshev inequality now says that:

$$(2) \quad \frac{m(B)}{n} < \frac{1}{k^2}$$

This means that the relative frequency, $\frac{m(B)}{n}$, of observations in B is less than $1/k^2$. To prove this we split the sum in s^2 in two parts as follows:

$$(3) \quad s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2 = \frac{1}{n-1}\sum_{x_i \in A}(x_i - \overline{x})^2 + \frac{1}{n-1}\sum_{x_i \in B}(x_i - \overline{x})^2$$

In the first part we sum over the observations in A and in the second part we sum over the observations in B. We now drop the first sum, (i.e. the mean of the squared deviations of all observations between $\overline{x}$ - ks and $\overline{x}$ + ks, see Figure 8.9) which certainly is non-negative (can it be equal to zero?). We then get the inequality:

$$(4) \quad s^2 \geq \frac{1}{n-1}\sum_{x_i \in B}(x_i - \overline{x})^2$$

But for $x_i \in B$ we have $|x_i - \overline{x}| \geq ks$ or $(x_i - \overline{x})^2 \geq k^2 s^2$.

(see Figure 8.9)

Thus we get:

(5) $$s^2 \geq \frac{1}{n-1} \sum_{x_i \in B} k^2 s^2 = \frac{1}{n-1} \cdot m(B) k^2 s^2$$

Note: $k^2 s^2$ is a constant. Also note the theorem in the section on summation about the summation of a constant. Here $m(B)$ plays the role of the n in the theorem.

If we multiply both sides of this inequality by $\frac{n-1}{ns^2k^2}$ we get:

(6) $$\frac{n-1}{nk^2} \geq \frac{m(B)}{n}$$

which gives:

(7) $$\frac{m(B)}{n} \leq \frac{n-1}{n} \cdot \frac{1}{k^2} < \frac{1}{k^2}$$

Which in turn gives:

(8) $$\frac{m(B)}{n} < \frac{1}{k^2} .$$

By this time you may be somewhat mystified about the meaning of this theorem. Perhaps if we describe the final result in English, it will help. First of all notice that set B is the set of all observations which are at least k standard deviations from the mean. This is quite apparent from Figure 8.9. Then notice that $\frac{m(B)}{n}$ is the relative frequency of observations in set B. Putting these two ideas together we see that the final result says, in effect, that the relative frequency with which an observation is at least k standard deviations from the mean is less than $1/k^2$.

Example: Let $k = 2$

The theorem then states that the relative frequency with which an observation is at least 2 standard deviations from the mean is less than $\frac{1}{4}$.

Question: What does the theorem state if $k = 3$?

We summarize in

Theorem 2. Given observations $x_1, \ldots, x_n$ with mean $\bar{x} = \frac{1}{n}\Sigma x_i$ and variance $s^2 = \frac{1}{n-1}\Sigma(x_i - \bar{x})^2$; let $k > 1$ and let $B = \{x_i : |x_i - \bar{x}| \geq ks\}$. If $m(B)$ is the number of x_i in B then $\frac{m(B)}{n} < \frac{1}{k^2}$.

8.13 Summary

In this chapter we reviewed and extended ideas about the collection and presentation of data, the definitions and computations of several descriptive statistics of sets of data and mathematics that you needed in connection with this work such as summation.

1. Statistical data can be represented in the form of frequency tables, cumulative frequency tables, frequency diagrams, frequency histograms, cumulative frequency polygons, and in other forms depending on the nature of the data and the purpose of the particular presentation.
2. The mean, median and mode of a set of data are statistics and examples of measures of central tendency of the set.
3. The range, mean absolute deviation, variance and standard deviation are statistics and examples of measures of the dispersion or spread of a set of a data.
4. Summation is used so extensively in the presentation of statistical ideas and statistical computation that the theorems about summation and the symbolism of summation are necessary parts of the study of statistics.

*5. The Chebyshev inequality is an example of a statistic that gives information both about central tendency and dispersion of data.

8.14 Review Exercises

1. A sample of 25 thermostatic switches was taken at random from a lot just manufactured and the "trigger" temperature was determined for each with the following results (in degrees Fahrenheit):

55	56	56	56	54
49	56	54	52	51
55	57	50	52	54
50	53	56	55	56
52	56	57	54	53

(a) Construct a frequency diagram.

(b) Construct a cumulative frequency polygon. Mark the vertical axis in frequencies and percentages.

(c) Determine the median and the two quartiles. Show these on your graph.

(d) What is the mode of these measurements?

(e) Find the range.

(f) Construct a frequency table and from it calculate the mean.

(g) Compute the variance and the standard deviation by using a shortcut with the original data.

2. A sample of 30 aluminum castings, when tested, yielded the following tensile strengths in pounds per square inch -- to the nearest 100 pounds:

29,300	37,700	25,800
34,900	34,900	23,700
36,800	26,700	28,700
30,100	34,800	32,400
34,000	38,000	28,200
30,800	25,700	34,000
35,400	25,800	34,500
31,300	26,500	29,200
32,200	28,000	28,700
33,400	24,600	29,800

(a) What is the range of the given data?

(b) Group the data using convenient intervals and midpoints and construct a frequency table. (Suggestion: use <u>either</u> intervals 23,000-- 24,000 with midpoint 23,500 etc., <u>or</u> intervals 22,000-- 24,000 with midpoint 23,000 etc.)

(c) What is the mode of the distribution?

(d) Construct a frequency histogram.

(e) Construct a cumulative frequency polygon.

(f) Estimate the median from your cumulative frequency polygon. <u>Calculate</u> the median from your frequency table and compare your results.

(g) Compute the variance and the standard deviation by first transforming the measurements by an equation

$$y_i = ax_i + b$$

(choose convenient constants a and b), and then using the formula

$$y = \frac{\sum f_i y_i^2}{\sum f_i} - \bar{y}^2$$

CHAPTER 9

TRANSFORMATIONS IN THE PLANE: ISOMETRIES

9.1 What is a Transformation?

A transformation is a special kind of mapping. Under a plane transformation each point of a plane is mapped onto one point of the plane. Because we learn much about properties of geometric figures through the study of plane transformations, this chapter is a continuation of our study of geometry.

To clarify the nature of transformations, let us look at two examples that illustrate transformations and one that does not.

Example 1. Let O be a given point of a plane. For any point A in the plane there is exactly one point A' such that O is the midpoint of $\overline{AA'}$. (See Figure 9.1.) A' is the image of A.

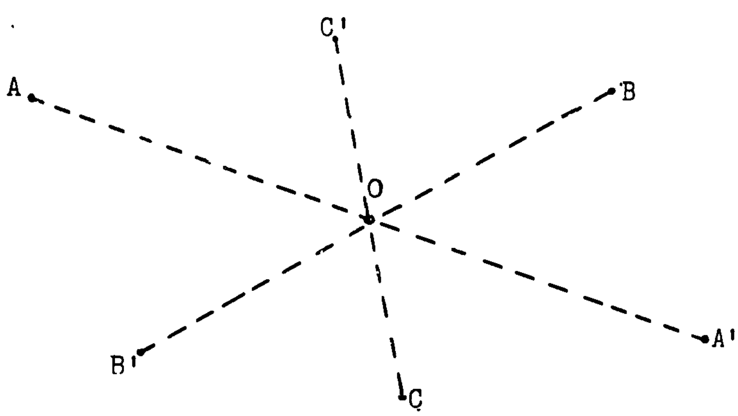


Figure 9.1

By this method of assigning images to points, B is mapped onto B' and C is mapped onto C'. It happens that by this mapping the image of A' is A, that is, the image of the image of A is A. But this is not a feature of all transformations. What makes this a transformation are the following characteristics:

(1) Every point of the plane is assigned exactly one point of the plane.

(2) Every point of the plane is the image of exactly one point in the plane.

We can summarize these two characteristics by saying that the plane is mapped onto itself by a one-to-one mapping*.

<u>Example</u> 2. In this example we use a plane rectangular coordinate system. Let P have coordinates (1,2). We assign P' to P, if the x-coordinate of P' is twice that of P and its y-coordinate is 1 more than that of P, that is, if P' has coordinates (2,3). The rule of assignment is

$$x \longrightarrow 2x, \quad y \longrightarrow y + 1$$

$$\text{or } (x,y) \longrightarrow (2x, y + 1)$$

A rule, such as this one, is called a <u>coordinate</u> rule. Study Figure 9.2 to see how this rule assigns images to

* Some use the term transformation to include many-to-one mappings.

A, B, C:

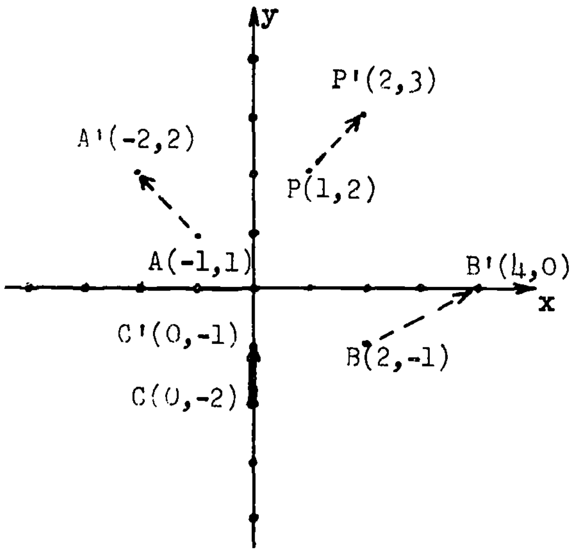


Figure 9.2

Does every point of the plane have exactly one image in the plane by this method of assignment? Is every point in the plane the image of exactly one point in the plane? If both answers are yes (as they are), then this is indeed a transformation.

Example 3. Again we use a rectangular plane coordinate system. Let the rule of assignment be

$$(x,y) \longrightarrow (x^2,\ 2y).$$

Does every point in the plane have a unique assignment by this rule? If so (and it is so), this is a mapping. To determine whether or not this mapping is a transformation, we ask if every point serves as the image of exactly one point. Consider $P'(4,6)$. It serves as the image of $P_1(2,3)$ because $4 = 2^2$ and $6 = 2 \cdot 3$. And it also serves as image of $P_2(-2,3)$ because $4 = (-2)^2$ and $6 = 2 \cdot 3$. But $P_1 \neq P_2$. We conclude that this mapping is not a transformation because it is not one-to-one. For that matter, does

Q' (-9,3) serve as an image? If not (and it does not) then neither is this mapping <u>onto</u> the plane.

9.2 Exercises

1. Make a drawing like Figure 9.1 showing the images of points A, B and C in a plane rectangular coordinate system under the mapping whose rule is as follows: Given point O in the plane, then the image of O is O, and the image of any point $P \neq O$ in the plane is P', where O is between P and P' and $OP' = 2OP$. Determine whether or not this mapping is a transformation. If you think it is, describe how to locate a point whose image is known.
2. Repeat the instructions in Exercise 1 using all its data with the modification $OP' = \frac{1}{2}OP$.
3. Make a drawing showing points and their images, as is done in Figure 9.2 for the points A(-2, -1), B(0,4), C(3,2) and D(1,-3) if the rule of assignment is $(x,y) \longrightarrow (x + 3,\ y - 2)$. Is this procedure of assigning images to the points of a plane a mapping? a transformation?
4. Carry out the instructions of Exercise 3 and answer its questions for each of the following coordinate rules, making a different diagram for each.
   (a) $(x,y) \longrightarrow (x,-y)$
   (b) $(x,y) \longrightarrow (-x,y)$
   (c) $(x,y) \longrightarrow (-x,-y)$
   (d) $(x,y) \longrightarrow (x^2,y)$
   (e) $(x,y) \longrightarrow (-x^2,-y)$

(f) $(x,y) \longrightarrow (x^3, y)$

(g) $(x,y) \longrightarrow (x^3, y + 1)$

(h) $(x,y) \longrightarrow (x^2, y^2)$

(i) $(x,y) \longrightarrow (2x + 1, y - 3)$

(j) $(x,y) \longrightarrow (y, x)$

(k) $(x,y) \longrightarrow (x + y, x - y)$

(l) $(x,y) \longrightarrow (2x - y, x + 2y)$

5. Let O_1 and O_2 be two distinct points of a plane. Make a drawing that shows the images of two points A and B where the following assignment rule is used: For any point P find P_1 such that O_1 is the midpoint of $\overline{P_1P}$ and then find P' such that O_2 is the midpoint of $\overline{P_1P'}$. Take P' to be the image of P. Is this a mapping? a transformation?

6. Repeat Exercise 5 with the modification that O_1 is between P and P_1 with $P_1O_1 = 2 \cdot PO_1$, and O_2 is the midpoint of $\overline{P_1P'}$. Is this a mapping? a transformation? if you think it is a transformation, describe the rule of the inverse transformation.

7. Consider transformations f, g, and h. Suppose for point A in a plane

$$A \xrightarrow{f} A_1 \xrightarrow{g} A_2 \xrightarrow{h} A_3$$

Show that the image of A under gof, followed by h is the same as the image of A under f followed by hog. What does this show about composition of transformations?

8. Let A and B be distinct points, and let A be assigned to B, and B to A, while each other point of the plane is assigned to itself. Is this a transformation?

9.3 Reflections in a Line

For the greater part of this chapter we study *isometries*, which are transformations that preserve distance. We shall examine this property more carefully as we examine various kinds of isometries, the first of which is a *reflection in a line*. Let us agree that all transformations in the remainder of this chapter have a plane as domain and range.

We approach a reflection in a line with a paper folding exercise.

On paper π we make three ink dots, shown in Figure 9.3 as points A, B, C. If we fold the paper along line ℓ, a line that contains C, then the ink spot at A leaves its mark (image)

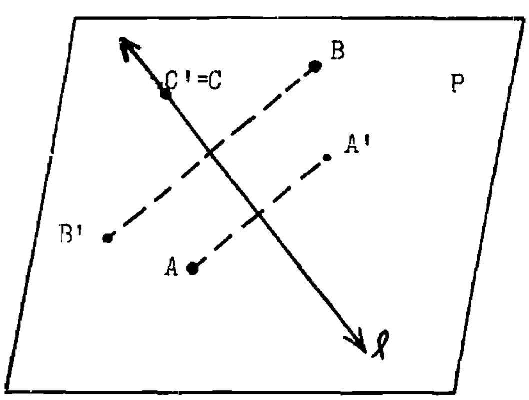


Figure 9.3

at A' and the spot at B leaves its image at B', while C is its own image. We call this correspondence a *reflection in a line*. We designate it R_ℓ, the subscript naming the *axis* ℓ of the reflection. We call A' the *reflection of A in* ℓ and B' the *reflection of B in* ℓ, while C, being its own reflection, is called a *fixed point*.

Let us describe some mathematical features of R_ℓ.

(1) For any point A not in ℓ, A is in one of the half planes determined by ℓ, A' is in the other.

(2) $\overline{AA'}$ intersects ℓ in a point, say M, and AM = MA'.

(3) $\ell \perp \overleftrightarrow{AA'}$.

We summarize these features in one statement:

<u>ℓ is the perpendicular bisector of $\overline{AA'}$</u>

It is easy to give a coordinate rule for a reflection in the x-axis of a rectangular coordinate system. In Figure 9.4 the x-axis is the perpendicular bisector of $\overline{AA'}$ and also of $\overline{BB'}$. Thus, under this reflection, $A \longrightarrow A'$ and $B \longrightarrow B'$. Since C is on the x-axis C = C'. If you study the coordinates of a point and its image you will see the rule

$$(x,y) \xrightarrow{R_x} (x, -y)$$

where R_x is read: the reflection in the x-axis.

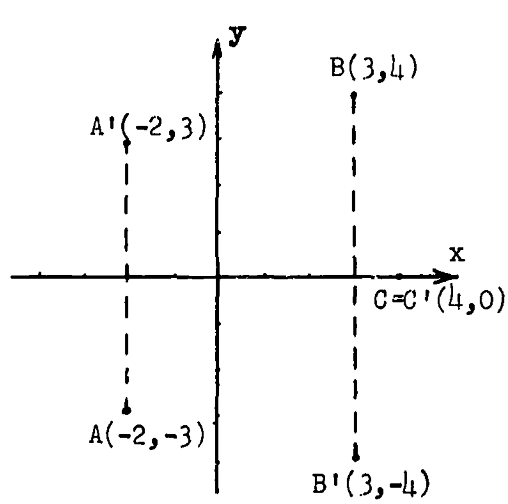


Figure 9.4

If you draw a diagram like the one in Figure 9.4, showing reflection in the y-axis, denoted R_y, you will see that its coordinate rule is: $(x,y) \xrightarrow{R_y} (-x,y)$.

Let us consider whether or not a line reflection is an isometry. It is, if it preserves distance. Suppose that for a given line ℓ, $A \xrightarrow{R_\ell} A'$, $B \xrightarrow{R_\ell} B'$. (See Figure 9.5)

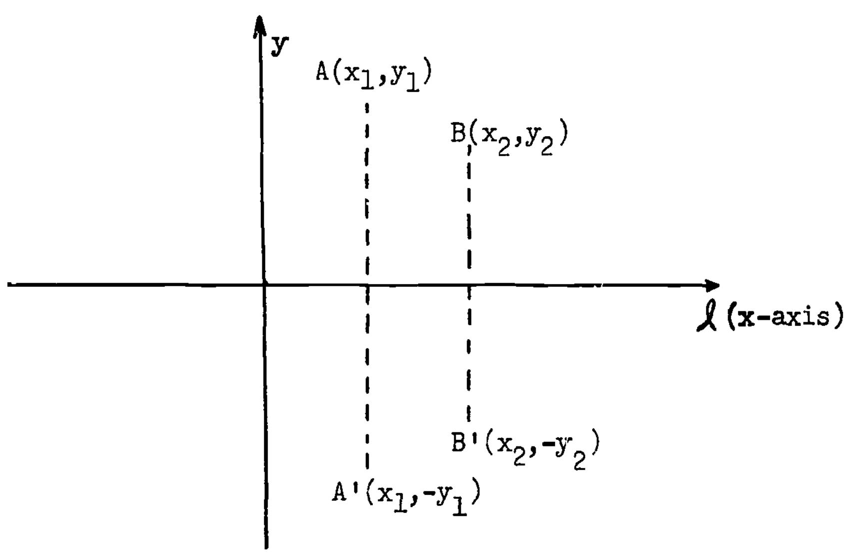


Figure 9.5

We take ℓ as the x-axis of a rectangular coordinate system. In this system, let A acquire coordinates (x_1, y_1) and $B(x_2, y_2)$. Then A' will have coordinates $(x_1, -y_1)$ and B' $(x_2, -y_2)$. Recalling the distance formula, we can write:

$$AB = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

$$\text{and } A'B' = \sqrt{(x_1 - x_2)^2 + ((-y_1) - (-y_2))^2}$$

Clearly, $AB = A'B'$ if $(y_1 - y_2)^2 = ((-y_1) - (-y_2))^2$.

We leave it to you to show that the last two terms are equal, and we conclude that <u>a line reflection is indeed an isometry</u>.

In the exercises that follow you will be asked to establish some of the properties of line reflections. Having done so you may use these properties in subsequent investigations; that is, they may be used as theorems.

9.4 Exercises

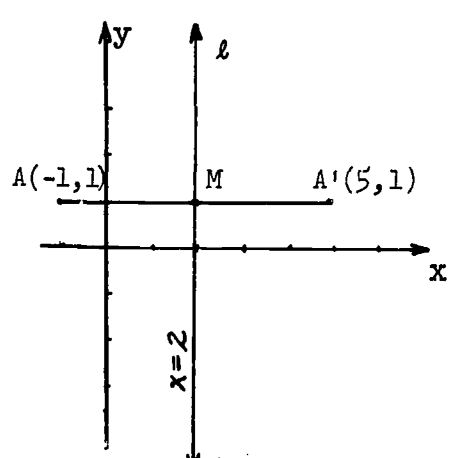
1. Find the coordinates of the image of each point listed below under the reflection in the x-axis.

   (a) $A(3,5)$ (b) $B(-3,5)$ (c) $C(-3,-5)$ (d) $D(\sqrt{2},-5)$

   (e) $E(3,0)$ (f) $F(0,-3)$ (g) $G(0,0)$ (h) $H(a,b)$.

2. Find the coordinates of the image of each point listed in Exercise 1 under the reflection in the y-axis.

3. You can see in the figure below that the reflection of $A(-1,1)$ in the line ℓ with equation $x = 2$, is $A'(5,1)$. To calculate the coordinates of A', knowing those of A, we use the fact that ℓ is the perpendicular bisector of $\overline{AA'}$.



This implies that M, the intersection of ℓ and $\overleftrightarrow{AA'}$, is the midpoint of $\overline{AA'}$, and that the y-coordinate of A' is the same as the y-coordinate of A. To calculate the x-coordinate of M we recall the midpoint formula $x_m = \frac{1}{2}(x_1 + x_2)$. In this problem $x_m = 2$ (because M is on ℓ) and $x_1 = -1$. Therefore $2 = \frac{1}{2}(-1 + x_2)$.

Solving for x_2, we get $x_2 = 5$, and A' has coordinates (5,1). Using this method of calculation, find the coordinates of the image of each point listed below under the line reflection in ℓ, the line with equation $x = 2$.

(a) A(3,-2) (b) B(-2,5) (c) C($\sqrt{2}$, 3) (d) D(a,b)

4. Adapt the method of calculating coordinates of images in Exercise 3 to find the coordinates of the images of A(3,-2) and B(-2,5) under a reflection in a line whose equation is:

   (a) $x = -3$ (b) $y = 1$ (c) $y = -2$

5. Verify that AB = A'B', where A' and B' are images of A and B respectively under R_y if A and B have coordinates:

   (a) (4,2) and (-1,5) (b) (0,5) and (4,-1) (c) (-2,0) and (0,-5)

6. Suppose distinct points A, B and C are collinear, and their images under R_ℓ, when ℓ is a given line, are respectively A', B' and C'. We can prove A', B' and C' are also collinear by using the fact that a line reflection is an isometry. Study the proof below and answer the questions. One of the points A, B or C is between the other two. Say it is B. Then AB + BC = AC. But AB = A'B', BC = B'C' and AC = A'C'. Why? Therefore A'B' + B'C' = A'C'. Why? But the last statement implies that B' is between A' and C', because otherwise

$A'B' + B'C' > A'C'$. Explain why this inequality is true if: (a) B' is not in $\overline{A'C'}$ or (b) B' is in $\overleftrightarrow{A'C'}$ but not in $\overline{A'C'}$. This proof also shows that a line reflection preserves the betweeness relation for points.

7. Show that the set of images of the points in a line, under a line reflection, forms a line. You may remember this basic property as follows: The reflection of a line in a line is a line.

8. Show that the reflection of a ray in line is a ray, and the reflection of a segment in a line is a segment.

9. (a) Show that the reflection of the sides of an angle, in a line, are the sides of an angle. While an isometry, by definition, preserves distance, the question whether or not it preserves angle measure is open. We shall assume that it does. We do not prove it because we do not have the mathematical machinery to do so. But it is quite plausible, so we are encouraged to assume this until we have the necessary machinery to prove it.

   <u>Isometries preserve angle measures.</u>

   (b) Of particular interest is the statement: Each side of an angle is the reflection of the other in the bisector of the angle, that is in the line containing the midray. This statement is equivalent to the statement: The x-axis bisects the angle determined by the two lines with equations $y = mx$ and $y = -mx$ for $m \neq 0$. Draw diagrams that illustrate the plausibility of this statement for $m = \frac{1}{2}$, $m = 3$.

10. Let ℓ_1 and ℓ be lines such that $\ell_1 \parallel \ell$. Show that $\ell_1 \parallel \ell_2$ if $\ell_1 \xrightarrow{R_\ell} \ell_2$. (You may use coordinates if you wish.)

11. Let ℓ_1 and ℓ be lines such that $\ell_1 \perp \ell$. Show that $\ell_1 = \ell_2$ if $\ell_1 \xrightarrow{R_\ell} \ell_2$.

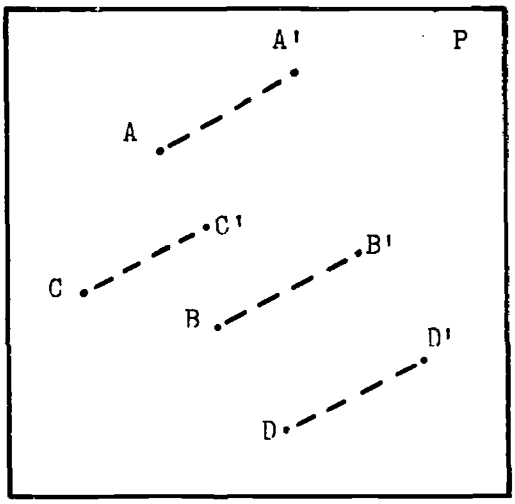
Note. Every point in ℓ is its own reflection in ℓ. Hence the axis is fixed under a line reflection. In this exercise we prove that a line perpendicular to the axis is also fixed, but not every point in the perpendicular is fixed. How many points are ? To distinguish between the two cases we say that the axis is fixed pointwise, while the perpendicular is fixed, but not pointwise.

12. Let ℓ be a line and P and point. If $P \xrightarrow{R_\ell} P'$ and $P' \xrightarrow{R_\ell} P_1$, show that $P = P_1$. A transformation that leaves every point fixed is called the identity transformation and is designated by the symbol i. Thus the composition of R_ℓ with itself is the identity transformation. A transformation f that has the property $f \circ f = i$, is called an involution if f is not itself the identity transformation. Verify that R_ℓ is an involution.

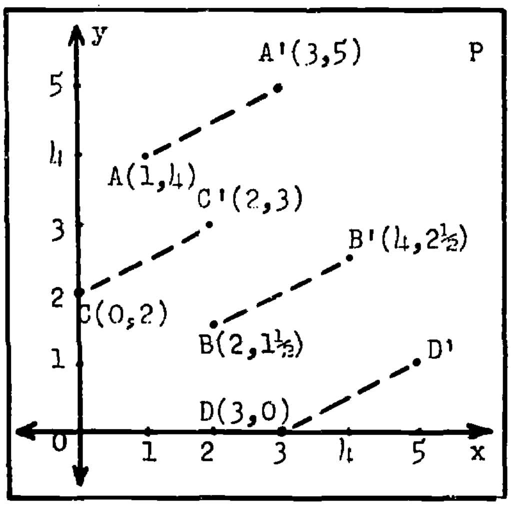
Before going on to the next section let us review some of the properties that were mentioned in the above exercises. A line reflection in ℓ preserves (a) distance (b) the betweeness relation for points (c) collinearity, (mapping lines onto lines, rays onto rays, segments onto segments, angles onto angles) (d) angle measures (e) ℓ, pointwise (the axis) (f) lines perpendicular to the axis, fixing them but not pointwise. Finally, a line reflection is an involution; that is $R_\ell \circ R_\ell = i$.

9.5 Translations

Let us mark some dots on a sheet of paper, say A, B, C, and D, as shown in Figure 9.6(a). Imagine a transparent paper placed over the sheet of paper and dots made on the transparent paper to locate the positions of A, B, C, D. Now imagine that the transparent paper is moved in a certain direction a certain distance. Then the dot over A is moved to a new position, called A' (see Figure 9.6(a)). Similarly the dot over B is moved to B', and this happens for all dots on the sheet of paper. If we think of the sheet of paper as a plane, we have described a method of assigning an image point to each point of the plane. Thus we have a mapping of the plane onto itself. Moreover every point of the plane serves as the image of exactly one point. This mapping therefore is a plane transformation. It is an example of a <u>translation</u>.



(a)



(b)

Figure 9.6

The rule of assignment of a translation is easily stated if we use coordinates. In some rectangular coordinate system let A

have coordinates (1,4) and let A' have coordinates (3,5). (See Figure 9.6(b).) B has coordinates $(2, 1\frac{1}{2})$ and B' has coordinates $(4, 2\frac{1}{2})$. If we add 2 to the x-coordinate of a point we get the x-coordinate of its image and if we add 1 to the y-coordinate of a point we get the y-coordinate of its image. Thus $(1,4) \longrightarrow (3,5)$ and $(0,2) \longrightarrow (2,3)$. What are the coordinates of D', the image of D(3,0)? The rule of this mapping is $(x,y) \longrightarrow (x + 2, y + 1)$. In general the coordinate rule of a translation is

$$(x,y) \longrightarrow (x + a, y + b)$$

where a and b are fixed numbers. For each different choice of a and b we have a different translation. Note when a = 0 and b = 0, the translation is the identity transformation.

It is startling that every translation can be expressed as the composition of two line reflections whose axes are parallel. We illustrate this fact for a particular translation.

Suppose in the particular translation, point A has image A' and that AA' = 8. We can take $\overleftrightarrow{AA'}$ (or any line parallel to $\overleftrightarrow{AA'}$) as the x-axis of a coordinate system, and let A have coordinates (-1,0) which implies that A' has coordinates (7,0). (See Figure 9.7.) Further, let the y-axis be ℓ, the axis of the first line reflection, and let the line m with equation x = 4 be the axis of the second line reflection.

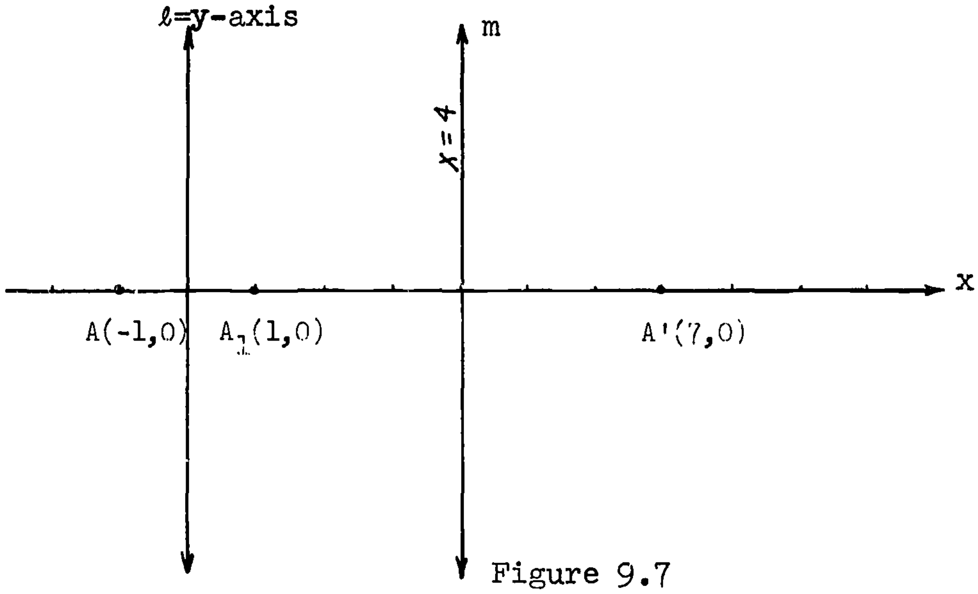


Figure 9.7

Let us follow the effects of R_ℓ on A, and of R_m on the image of A.

$$A(-1,0) \xrightarrow{R_\ell} A_1(1,0) \xrightarrow{R_m} A'(7,0)$$

We see that A has been assigned A' by $R_m \circ R_\ell$, just as by the original translation. Let us find the image of B(-3,1) under $R_m \circ R_\ell$.

$$B(-3,1) \xrightarrow{R_\ell} B_1(3,1) \xrightarrow{R_m} B'(5,1)$$

and again we see that $R_m \circ R_\ell$ assigns to B the same image as does the translation whose rule is

$$(x,y) \longrightarrow (x+8,y)$$

Note that the distance of the translation is twice the distance between the axes of reflection, and the translation is in the same direction as the direction from the first axis to the second. It is interesting that we might have chosen other pairs of axes that were 4 units apart. For instance, you might try taking the lines with equations x = 1 and x = 5 as the axes, in that order.

The reasoning used in this particular case can be used in any case, and we conclude with the following statement. <u>Every</u>

translation can be expressed as the composition of two line reflections whose axes are parallel. The distance of the translation is twice the distance between the axes, and the direction of the translation is the same as the direction from the first axis to the second. Any two axes may be chosen, if they are perpendicular to the direction of the translation, if the distance between them is half the distance of the translation, and if the direction from the first to the second is the same as the direction of the translation.

There is an important bonus in this statement: Every property of a line reflection that is retained by the composition of two line reflections is therefore a property of all translations.

Since each line reflection preserves distance, the composition of any number of line reflections also preserves distance, and hence is an isometry. It follows that a translation is an isometry. All the properties that follow logically from isometries therefore belong to translations. Among these properties are: translations preserve collinearity, the betweeness relation for points, and map lines onto lines, rays onto rays, segments onto segments, and angles onto angles of the same measure.

9.6 Exercises

1. The coordinate rule of a translation is $(x,y) \longrightarrow (x + 1, y + 1)$. Find the coordinates of the images of the following points under the translation:

   (a) $A(3,2)$ (b) $B(-3,2)$ (c) $C(-3,-2)$ (d) $D(8\frac{1}{2}, 3\frac{1}{3})$

(e) $E(\sqrt{2}, 1)$ (f) $F(-1, \sqrt{3})$ (g) $G(0,0)$ (h) $H(-1,1)$

2. Find the coordinate rule of a translation that maps:

   (a) $A(3,2)$ onto $A'(2,3)$ (b) $B(0,0)$ onto $B'(-3,5)$

   (c) $C(-3,5)$ onto $C'(0,0)$ (d) $D(2\frac{1}{3}, 3)$ onto $D'(-2\frac{1}{2}, 1\frac{1}{3})$

   (e) $E(a,b)$ onto $E'(0,0)$ (f) $F(a,b)$ onto $F'(2a,3b)$

3. Let A and B have coordinates $(0,2)$ and $(5,1)$ respectively in some rectangular coordinate system. Find the coordinates of A' and B', the images of A and B, under the translation whose rule is $(x,y) \longrightarrow (x - 1, y + 2)$.

   Justify each of the following statements:

   (a) $AB = A'B'$

   (b) $\overline{AB} \longrightarrow \overline{A'B'}$

   (c) $\overleftrightarrow{AB} \parallel \overleftrightarrow{A'B'}$ (You may use the slope formula $\frac{y_1 - y_2}{x_1 - x_2}$.)

   (d) $\overleftrightarrow{AA'} \parallel \overleftrightarrow{BB'}$

   (e) $AA' = BB'$

   (f) $ABB'A'$ is a parallelogram.

   (g) $\overline{AB'}$ and $\overline{A'B}$ bisect each other.

   This exercise suggests that in general, if A' and B' are the images of A and B under a translation, and A, B, A' are not collinear, then $ABB'A'$ is a parallelogram. Prove it.

4. Let translation T_1 have the rule $(x,y) \longrightarrow (x + 3, y - 2)$ and let translation T_2 have the rule $(x,y) \longrightarrow (x + 1, y + 3)$. Investigate the nature of $T_2 \circ T_1$, showing its effect on points $A(3,2)$, $B(-4,0)$, and $C(-2, -5)$.

5. Let translation T_1 have the rule $(x,y) \longrightarrow (x + 3, y - 2)$ and let translation T_2 have rule $(x,y) \longrightarrow (x - 3, y + 2)$. Let A have coordinates $(2,3)$. Find the coordinates of A' if

$A \xrightarrow{T_1} A'$. Show that $A' \xrightarrow{T_2} A$. Does this suggest that T_2 is the inverse of T_1? Verify that $T_2 \circ T_1 = i$ using points $B(-2,8)$ and $C(a,b)$. Also verify $T_1 \circ T_2 = i$.

6. In Exercise 5 we can represent T_1 as $T_{3,-2}$ and T_2 as $T_{-3,2}$. In general the translation with rule $(x,y) \longrightarrow (x + a,\ y + b)$ can be represented $T_{a,b}$. Using this notation represent:

   (a) The composition of $T_{a,b}$ and $T_{c,d}$.

   (b) The inverse of $T_{a,b}$.

   (c) The identity transformation.

7. Show that composition of translations is commutative.

8. Using the data in Figure 9.7, show that $R_m \circ R_\ell$ maps each of the following points onto the same point as does the translation that maps $A\ (-1,0)$ onto $A'\ (7,0)$.

   (a) $C(2,0)$ (b) $D(3,-4)$ (c) $F(10,-3)$.

9. Show that composition of line reflections having parallel axes is not commutative. If these line reflections are R_m and R_ℓ how is $R_m \circ R_\ell$ related to $R_\ell \circ R_m$?

10. Is the composition of line reflections associative? Is the composition of translations associative? (See Exercise 7 of Section 9.2.)

11. Is the composition of two translations a translation? If so how can you find the rule of the composition from the rules of the two translations?

12. Is the composition of two line reflections a line reflection? If so, how can you describe the rule of the composition?

13. Before answering this exercise, recall that a set of elements, together with a binary operation defined on that set, is a group if

(i) the set and binary operation constitute an operational system;

(ii) the operation on this set is associative,

(iii) there is an identity element in the set, and

(iv) each element in the set has an inverse, also in the set.

(a) Is the set of all translations in a plane, with composition of translations as the operation, a group?

(b) Is the set of all line reflections, with composition as the operation, a group?

14. $T_{o,o}$ is a translation. What are two line reflections from which it is composed?

9.7 Rotations and Half-Turns

We extend our study of isometries (so far consisting of reflections in a line and translations) to rotations. Let us select on a sheet of paper π a point O on π at which we attach a transparency (called π'). (See Figure 9.8).

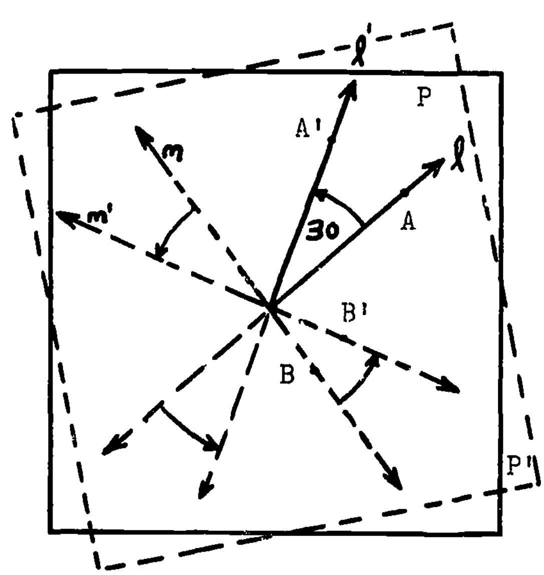


Figure 9.8

Select any two points (other than 0) in π, say A and B, and mark dots on the transparency over A and B. Now turn the transparency about 0 as pivot, in the counterclockwise direction, through an angle of 30°. The "A dot" on the transparency ends at a point A'. We assign A' to be the image of A under this rotation. The image of B is B'. If we think of the paper as a plane then we have a method by which each point of π is assigned exactly one point of π, and each point of π serves as the image of one point in π. Thus we have a transformation of the plane onto itself. The essential data that determine this transformation are point 0 and 30°, that is the center or pivot of the rotation, and the measure of the angle through which the rotation takes place, in the counterclockwise direction if the angle measure is positive. To indicate a clockwise rotation we use negative numbers as angle measures. The above transformation can be designated r(0,30). (Capital R denotes a line reflection, lower case r will denote a rotation.) What does r(0,-30) mean?

You should note the following about r(0,30) in Figure 9.8:

For any point A and its image A', 0A = 0A' and $m\angle AOA' = 30$. In particular, OB = OB' and $m\angle BOB' = 30$.

We saw that a translation is a composition of two reflections in parallel axes. It is surprising that every rotation is also the composition of two reflections in axes, but not in parallel axes.

We offer here a discussion, not a proof, of this fact. It is intended to illustrate, using Figure 9.9, that a rotation is the composition of two reflections whose axes intersect.

We examine a particular rotation, r(0,80). (See Figure 9.9.) Let r(0,80) map A onto A'. Then OA = OA' and $m\angle AOA' = 80$.

Let $\overrightarrow{OM}$ and $\overrightarrow{ON}$ be rays, with the angle of counterclockwise rotation from $\overrightarrow{OM}$ to $\overrightarrow{ON}$ measuring 40, that is $\frac{1}{2}$ of 80. Label $\overleftrightarrow{OM}$ = m, $\overleftrightarrow{ON}$ = n. We shall see that if $A \xrightarrow{R_m} A_1$, then $A_1 \xrightarrow{R_n} A'$, which shows that $A \xrightarrow{R_n \circ R_m} A'$, or that $R_n \circ R_m$ maps A onto the same point as does R(0,80).

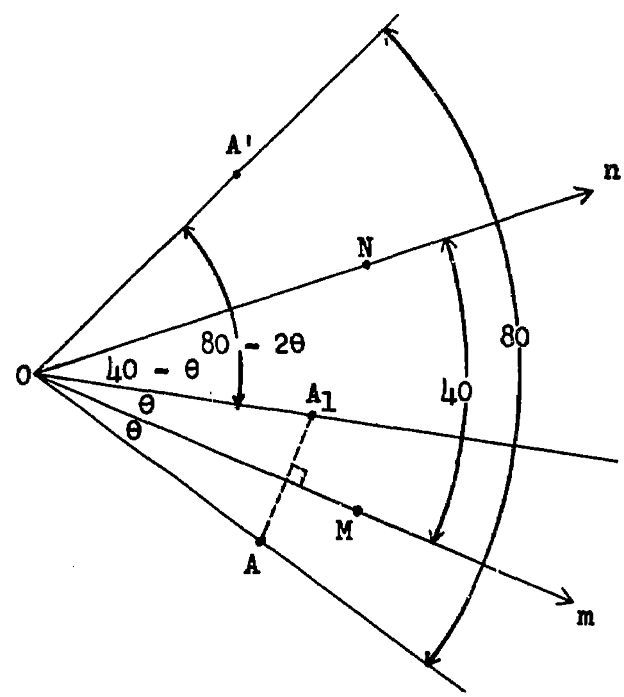


Figure 9.9

Since R_m maps A, O, M onto A_1, O, M respectively, R_m maps $\angle AOM$ onto $\angle A_1OM$; hence $m\angle AOM = m\angle A_1OM = \theta$, say, and $m\angle AOA_1 = 2\theta$. Since $m\angle AOA' = 80$, it follows that

$$(1) \quad m\angle A'OA_1 = 80 - 2\theta.$$

But $m\angle MON = 40$ and $m\angle MOA = \theta$ imply

$$(2) \quad m\angle A_1ON = 40 - \theta.$$

(1) and (2) yield $m\angle A'ON = 40 - \theta$. Thus $\overleftrightarrow{ON}$ is the bisector of

$\angle A_1OA'$, so that R_n maps $\overrightarrow{OA_1}$ onto $\overrightarrow{OA}$ (Exercise 9(b) of Section 9.4). Since $OA = OA_1 = OA'$, $A_1 \xrightarrow{R_n} A'$. This, and $A \xrightarrow{R_m} A_1$, yield $A \xrightarrow{R_n \circ R_m} A'$. A discussion like this can be given for all A and A', where $A \xrightarrow{R(0,80)} A'$. We conclude $r(0,80) = R_n \circ R_m$.

We conclude that every rotation, $r(0, 2\theta)$, can be regarded as the composition of two line reflections whose axes meet at 0 and determine an angle of measure θ, where θ is positive if the rotation is in the counterclockwise direction, and negative if in the clockwise direction. We will call 2θ (whether positive or negative) the <u>measure of the rotation</u>.

We deduce from this that a rotation is an isometry, and has all the properties possessed by the composition of two reflections in intersecting axes. Among them are the preservation of distances, lines, rays, segments, angles, and angle measures.

An important special case of a rotation is the composition of two reflections in perpendicular axes. In this case $\theta = 90$ and the rotation has measure 180. Then A, 0 and A' are collinear, and $OA = OA'$. This implies that 0 is the midpoint of $\overline{AA'}$. This transformation is called a <u>half-turn</u> with center 0, and is denoted H_0. For a diagram of a half-turn see Figure 9.1.

A coordinate rule for a half-turn is easily determined if the center of the half-turn is the origin of a rectangular coordinate system: $H_0 = R_y \circ R_x$ Since the rule for R_x is $(x,y) \longrightarrow (x,-y)$, and the rule for R_y is $(x,y) \longrightarrow (-x,y)$, the effect of the composition on (x,y) is:

$$(x,y) \xrightarrow{R_x} (x,-y) \xrightarrow{R_y} (-x,-y).$$

Therefore the rule of H_0 is $(x,y) \longrightarrow (-x,-y)$.

If we wish $A(a,b)$ to be the center of a half-turn that maps $P(x,y)$ onto $P'(x', y')$, then, because A is the midpoint of $\overline{PP'}$:

$$a = \tfrac{1}{2}(x + x') \text{ and } b = \tfrac{1}{2}(y + y')$$

this leads to

$$x' = 2a - x \text{ and } y' = 2b - y \quad \text{or}$$

$$(x,y) \longrightarrow (2a - x,\ 2b - y).$$

9.8 Exercises

1. Make a drawing which shows non-collinear points A, B, and C, and their images under a rotation with center O and measure 65, if O is in the interior of $\triangle ABC$. Use a protractor.
2. Suppose A, B, and C are three collinear points, and B is between A and C. Let $r(P,\theta)$ assign images A', B' and C' respectively. Are the images collinear? Is B' between A' and C'? Compare AB with $A'B'$, AC with $A'C'$ and BC with $B'C'$.
3. Let P be a given point. Is the composition $r(P,20) \circ r(P,30)$ a rotation? If so, what is its center and what is its measure?
4. Express as a single rotation:
   (a) $r(P,40) \circ r(P,20)$ (b) $r(Q,30) \circ r(Q,-20)$
   (c) $r(P,90) \circ r(P,80)$ (d) $r(Q,40) \circ r(Q,-40)$
5. Let $r(P,\theta)$ represent a rotation. Does it have an inverse? If so, how is its designated?
6. Is the set of rotations with center O and with the operation of composition a group? Justify your answer.

7. Let the origin O of a rectangular coordinate system be the center of a half-turn. Find the coordinates of the images of the following points:

   (a) $A(3,-2)$ (b) $B(-2,3)$ (c) $C(0,-2)$ (d) $D(\sqrt{2},-\sqrt{3})$

8. Find the coordinates of the images of the points in Exercise 7 under the half-turn where center is $(1,-2)$.

9. Repeat Exercise 8 for the center $(-1,3)$.

10. Prove that a half-turn is an involution.

11. Let P and Q be distinct points. Show that the composition $H_Q \circ H_P$ is a translation in the direction from P to Q a distance of 2PQ.

12. Using the data of Exercise 11 show that if $P \neq Q$, $H_Q \circ H_P \neq H_P \circ H_Q$. Further show that $H_Q \circ H_P$ and $H_P \circ H_Q$ are inverses of each other.

13. Are any lines fixed under a half-turn? If there are describe them. Are they fixed pointwise?

14. Given: H_O line ℓ not containing O, and $\ell \xrightarrow{H_O} \ell'$. Prove $\ell \parallel \ell'$. (You may wish to use coordinates in this proof. If you do, recall the formula for the slope of a line: $m = \frac{y_1 - y_2}{x_1 - x_2}$)

15. Let H_P be a half-turn, and A and B be two points not collinear with P. Let $A \xrightarrow{H_P} A'$, and $B \xrightarrow{H_P} B'$. Prove $ABA'B'$ is a parallelogram.

*16. We have given coordinate rules for line reflections in the x-axis and in the y-axis. There are also coordinate rules for reflections in other lines passing through the origin, which we give in this exercise. If ℓ has equation $y = \frac{b}{a+1}x$, where $a^2 + b^2 = 1$ and $a \neq -1$ (in a rectangular coordinate

system), the rule for R_ℓ is $(x,y) \longrightarrow (ax + by,\ bx - ay)$. Verify that this rule:

(a) maps $P(1,0)$ onto $Q(a,b)$.

(b) maps $Q(a,b)$ onto $P(1,0)$.

(c) maps any point of ℓ onto itself.

(d) has the property $R_\ell \circ R_\ell = i$.

If we assume that a transformation that is an involution and has a line fixed pointwise is a line reflection, then (c) and (d) verify that the mapping described above is a line reflection.

*17. Consider the mapping with the rectangular coordinate rule

$$(x,y) \longrightarrow (ax - by,\ bx + ay)$$

where $a^2 + b^2 = 1$. To help verify that this mapping is a a rotation with center O, the origin of the coordinate system, prove the following: If $(a,b) = (1,0)$ then the mapping is i. In all other cases:

(a) O is the only fixed point.

(b) If $A \longrightarrow A'$, then $OA = OA'$.

(c) For further verification show $(1,0) \longrightarrow (a,b)$, $(0,1) \longrightarrow (-b,a)$ and $(a,-b) \longrightarrow (1,0)$. You should draw a diagram to understand the significance of these results.

18. Let line ℓ have slope $m \neq 0$ relative to a rectangular coordinate system. Show that the slope of ℓ_1, the reflection of ℓ in the x-axis, is $-m$. Is the slope of ℓ_2, the reflection of ℓ in the y-axis, also m?

*19. Copy Figure 9.9. Select a point B in the interior of $\angle MON$. Use a protractor to find its image under the rotation $r(O,80)$.

Now find the image of B under the composition $R_n \circ R_m$. You should get the same image point by both methods. Repeat the above starting with point C, the image of A_1 under the half-turn H_O.

20. Show that $(x,y) \longrightarrow (y,x)$ is the coordinate rule for the reflection in the line with equation $y = x$. (Hint: See Exercise 16.)

9.9 Composing Isometries, Glide Reflections

We have seen that translations and rotations can be regarded as compositions of line reflections. It is natural to wonder what results if we compose a line reflection with a rotation or with a translation, or if we compose a rotation with a translation. (There are other possibilities.) No matter what isometries we compose we know that the composition will be an isometry. Why? Since translations and rotations (including half-turns) can be constructed out of reflections only, it would seem that the line reflection is the basic isometry and we might wonder if every isometry can be composed of line reflections only. The answer is yes. But this idea is worthy of careful attention and we pursue it in another section.

Meanwhile, we investigate the composition of a line reflection with a translation in a direction of the axis of the line reflection. In exercises you will be asked to investigate other compositions.

Let ℓ be a line. (See Figure 9.10).

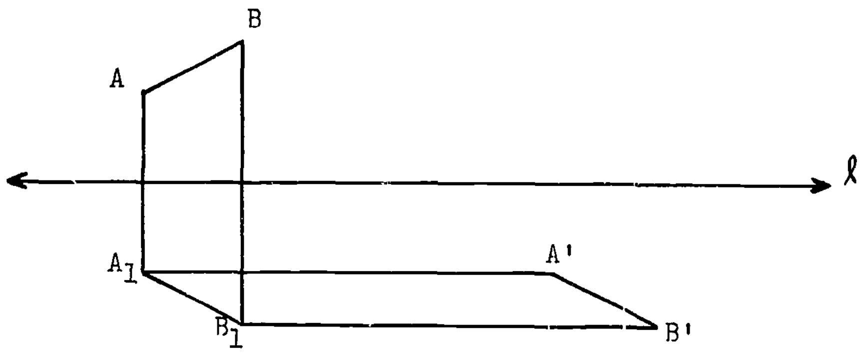


Figure 9.10

If R_ℓ and T represent the reflection and translation then:

$$A \xrightarrow{R_\ell} A_1 \xrightarrow{T} A', \text{ or } A \xrightarrow{T \circ R_\ell} A'$$

$$B \xrightarrow{R_\ell} B_1 \xrightarrow{T} B', \text{ or } B \xrightarrow{T \circ R_\ell} B'$$

Since R_ℓ and T are isometries, so is $T \circ R_\ell$. The composition is different in its apparent effect from any other isometry we have seen so far. It is called a glide reflection (in ℓ). ("Glide" is synonomous with "translation.") It has the properties of preserving distance, lines, betweeness for points, rays, segments, angles and angle measures. It is instructive to note, in Figure 9.10, that $A_1B_1B'A'$ is a parallelogram, or to say this another way the reflection of $\overline{AB}$ in ℓ is translated to $\overline{A'B'}$.

9.10 Exercises

1. We have defined a glide reflection as a line reflection followed by a translation. Show that the glide reflection may also be described as a translation followed by a line reflection, where the translation is in a direction of the axis of the reflection.

2. Show that a glide reflection may be regarded as the composition of three line reflections, where the first two have

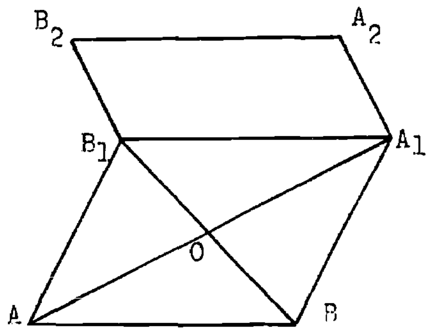
parallel axes, each perpendicular to the third axis. Will a glide reflection result if the last two have parallel axes, each perpendicular to the axis of the first reflection?

*3. Let ℓ be a line and P any point not in ℓ. Let F be a glide reflection in ℓ, and let $P \xrightarrow{F} P'$. Show that ℓ bisects $\overline{PP'}$.

4. Given a glide reflection. Describe how it may be regarded as the composition of two half-turns and a line reflection. (See Exercise 11 of Section 9.8 for a hint if you need one.)

5. Show that the composition of a glide reflection with itself is a translation. Describe the translation.

6. To show that a half-turn followed by a translation is a half-turn, you might start with two points A and B, and point O not in $\overleftrightarrow{AB}$. If $A \xrightarrow{H_O} A_1$ and $B \xrightarrow{H_O} B_1$, what kind of figure is ABA_1B_1? Let the translation that maps A_1 onto A_2 map B_1 onto B_2. In general what kind of figure is



$A_1A_2B_2B_1$? How are $\overline{AB}$ and $\overline{B_2A_2}$ related? What kind of figure is ABA_2B_2? What is the isometry that maps A onto A_2 and B onto B_2? What would your answers be if O is on $\overleftrightarrow{AB}$?

7. Investigate the composition of a translation followed by a half-turn.

8. Show that the composition of three line reflections in three parallel axes is a line reflection.

9. Show that the composition of three line reflections in lines a, b, and c, where $a \perp b$ and $b \perp c$, is a glide reflection.

*10. In this exercise we consider a coordinate rule for a glide reflection. It is the composition of the reflection in line ℓ with equation $y = \frac{b}{a+1}x$ where $a^2 + b^2 = 1$, and the translation that maps (0,0) onto $((a + 1)c, bc)$. Its rule is

$$(x,y) \longrightarrow (ax + by + (a + 1)c,\ bx - ay + bc)$$

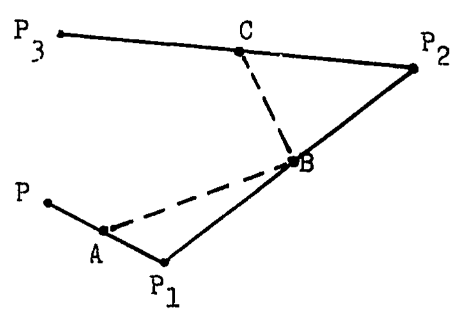
Verify the following:

(a) The direction of the translation is a direction of the axis of the reflection.

(b) The glide reflection assigns a point of ℓ to a point of ℓ. Explain why it should.

(c) The rule actually combines the rules of a reflection and a translation. (Hint: See Exercise 16 of Section 9.8.)

11. Let A, B, C be three noncollinear points as shown, and let P be any point in the plane. Consider $P \xrightarrow{H_A} P_1$, $P_1 \xrightarrow{H_B} P_2$, $P_2 \xrightarrow{H_C} P_3$. Show that there is a point D such that $P \xrightarrow{H_D} P_3$.



(Hint: Take D as the fourth vertex of parallelogram ABCD.) How does this show that $H_C o H_B o H_A = H_D$, that is, the composition of three half-turns is a half-turn?

12. Let F be the midpoint of $\overline{AB}$. Prove $H_A o H_F o H_B o H_F = i$, where i is the identity transformation.

9.11 The Three Line Reflection Theorem

In this section we will keep our promise to show that any isometry may be constructed as the composition of line reflections; and, in fact, no more than three. We do this in two stages. First we prove a preliminary theorem, which is called a lemma. In it, and in the theorem that follows, we make use of the following result, which was given in Course 1, Section 10.14, Exercises 4 and 5:

> A point is on the perpendicular bisector of a segment if and only if it is as far from one endpoint of the segment as it is from the other.

Lemma. Let A, B, C be three noncollinear points, and let F and G be isometries such that $(A, B, C) \xrightarrow{F} (A', B', C')$ and $(A, B, C) \xrightarrow{G} (A', B', C')$. Then F = G. (For isometries we need only prove: For any point X, if $X \xrightarrow{F} X'$, then $X \xrightarrow{G} X'$.)

The significance of the theorem is this: Given three noncollinear points A, B, C and their respective images A', B', C' under an isometry, then that isometry is the only one that effects the mapping of (A, B, C) onto (A', B', C'). Even though an isometry of the plane has infinitely many points in its domain, it is uniquely

determined by its effect on only three noncollinear points.

*$(A, B, C) \xrightarrow{F} (A', B', C')$ is an abbreviation of $A \xrightarrow{F} A'$, $B \xrightarrow{F} B'$, $C \xrightarrow{F} C'$.

Proof of Lemma.

We use the indirect method. Suppose $X \xrightarrow{G} X''$, where $X'' \neq X'$. Since F is an isometry, $AX = A'X'$, $BX = B'X'$, $CX = C'X'$. Since G is also an isometry, $AX = A'X''$, $BX = B'X''$, $CX = C'X''$. Thus $A'X' = A'X''$, $B'X' = B'X''$ and $C'X' = C'X''$. Put into words, this means that A' is as far from X' as from X'', also B' is as far from X' as from X'', and C' is as far from X' as from X''. We have assumed $X' \neq X''$. Thus $\overline{X'X''}$ is a segment and A', B', C' are on the perpendicular bisector of $\overline{X'X''}$. But this implies that A, B, C are collinear, contrary to the information given in the theorem. Therefore $X' = X''$. Since F and G make the same assignment to all points of the plane (X is <u>any</u> fourth point), $F = G$.

And now for the key theorem:

Theorem.

Every isometry may be constructed as the composition of at most three line reflections.

Proof.

We know from the preceding lemma that an isometry is uniquely

determined by its effect on three noncollinear points. Let the isometry f map A onto A', B onto B' and C onto C', where A, B, C are noncollinear points. Figure 9.11(a) shows points A, B, C and their images A', B', C'.

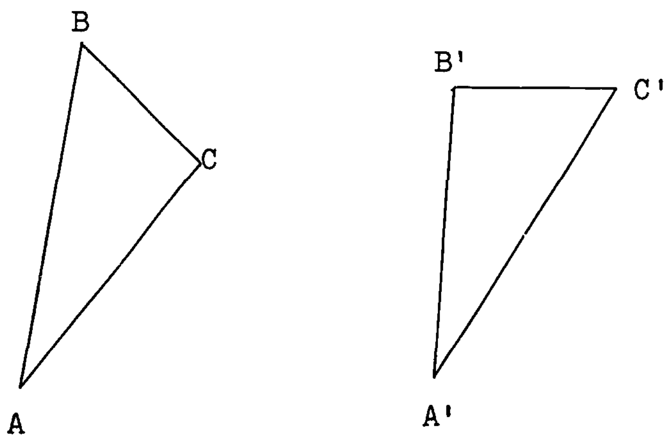


Figure 9.11(a)

$AB = A'B'$, $AC = A'C'$, $BC = B'C'$

We demonstrate in a sequence of figures the line reflections whose composition maps (A, B, C) onto (A', B', C'). First we consider the trivial case in which $A' = A$, $B' = B$, and $C' = C$. Here the identity transformation i maps (A, B, C) onto (A', B', C'). We may regard i as the composition of two reflections in the same line. This proves the theorem for this trivial case.

Now suppose that at least one of the points A, B, C is different from its image, say $A' \neq A$. The line reflection in ℓ, the perpendicular bisector of $\overline{AA'}$, maps A onto A'. This is shown in Figure 9.11(b), along with logical consequences stemming from [illegible] reflection. $(A, B, C) \xrightarrow{R_{\ell_1}} (A', B_1, C_1)$.

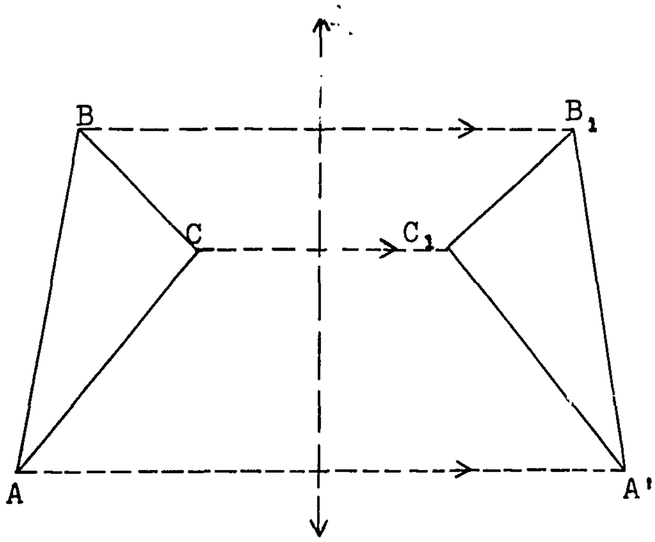


Figure 9.11(b)

$AB = A'B_1$ $AC = A'C_1$ $BC = B_1C_1$

If $B_1 = B'$ and $C_1 = C'$, only one line reflection is needed to prove the theorem. Suppose $B_1 \neq B'$. A second reflection in ℓ_2, the perpendicular bisector of B_1B' maps B_1 onto B' (See Figure 9.11(c)). Since $A'B_1 = A'B'$, A' is as far from B as from B', so that A' is on ℓ_2 and hence A' is mapped onto itself. So $(A', B_1, C_1) \xrightarrow{R_\ell} (A', B', C_2)$. Again note the logical consequences of this reflection.

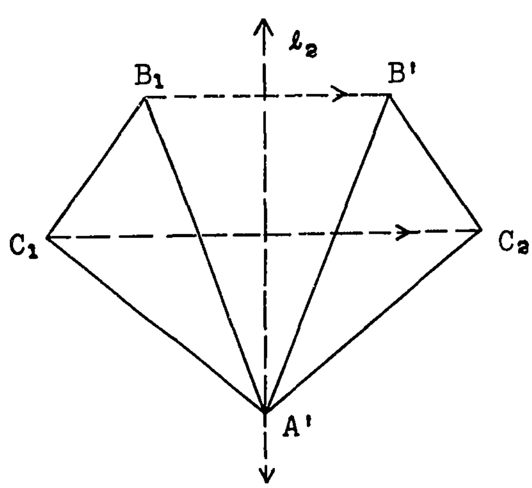


Figure 9.11(c)

If $C_2 = C'$, then two line reflections effect the desired isometry. If $C_2 \neq C'$, then the line reflection in $\ell_3 = \overleftrightarrow{A'B'}$ effects the isometry. (See Figure 9.11(d)), for $A'C_2 = A'C'$ and $B'C_2 = B'C'$; that is, each of A' and B' is equidistant from C_2 and C'. Hence ℓ_3 is the perpendicular bisector of $\overline{C_2C'}$ and $(A', B', C_2) \xrightarrow{R_{\ell_3}} (A', B', C')$.

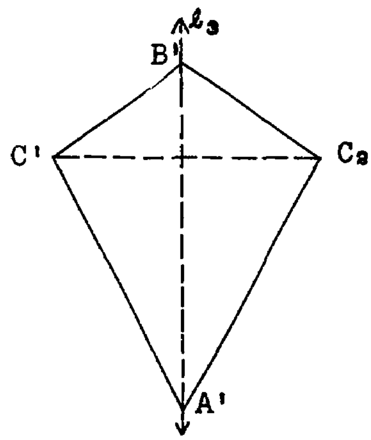


Figure 9.11(d)

You may find it instructive to see the three line reflections in one diagram. They are shown together in Figure 9.11(e).

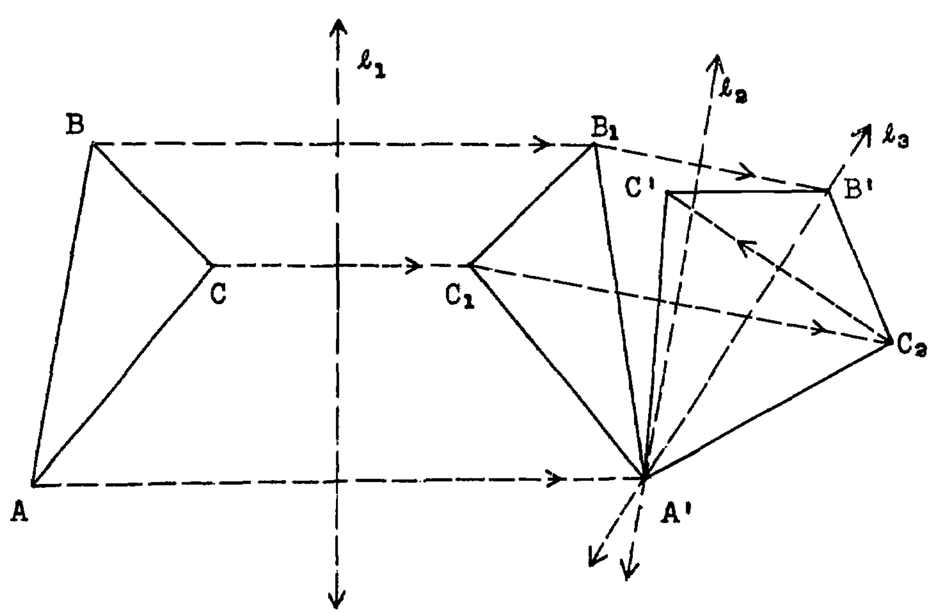


Figure 9.11(e)

9.12 Exercise

Our artist took great pains to get Figure 9.11 to convey clearly the sequence of line reflections that map (A, B, C) onto (A', B', C'). We issue this challenge to you. Using a cardboard triangle trace it in two different positions and then show clearly the sequence of line reflections that map one of these triangles onto the other. As you do this follow the proof given above.

9.13 Directed Isometries

In Figure 9.12 you see a pennant, a triangle, and a ray, each reflected in ℓ. Each has been flipped over.

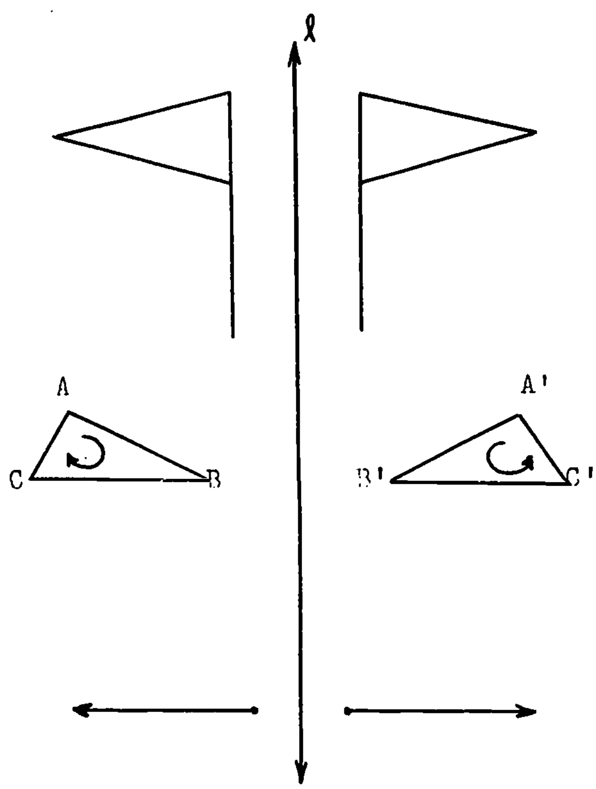


Figure 9.12

How can we describe this mathematically? We can see that the rays are directed in opposite directions. So are the pennants. To describe the "flipped over" position of the triangle we look at ΔABC reading the vertices from A to B, then to C and note that the direction of reading the vertices in that order resembles the motion of the hands of a clock. But if we read the image of ΔABC, from A' to B' to C' we note the direction to be opposite that of the hands of a clock. We see that the isometry has reversed the direction, or sense, of the three vertices. If a line reflection changes the sense of ΔABC, it will change the sense of any three noncollinear points. Because a line reflection reverses the sense of three noncollinear points we call it an opposite mapping.

Definition 1. If a mapping preserves collinearity and non-collinearity it is called opposite if it reverses the sense of three non-collinear points; otherwise it is called a direct mapping.

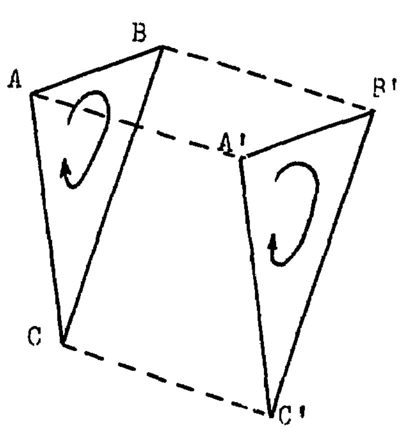


Figure 9.13

It is clear from Figure 9.13 that a translation is a direct isometry. Figure 9.14 shows that a half-turn is also a direct isometry.

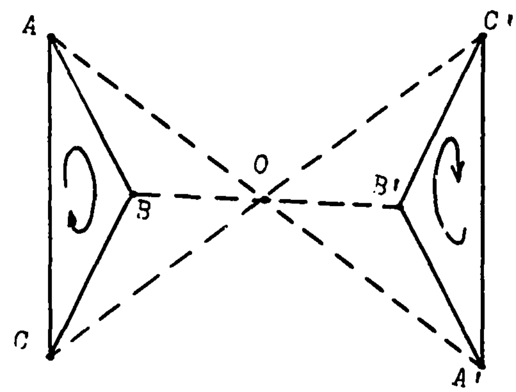


Figure 9.14

These observations however should not surprise us, for a translation is the composition of two line reflections, each of which is an opposite isometry. The second reflection reverses the reversal, and so restores the sense of the original triangle. Hence a translation <u>must</u> be a direct isometry.

Since a half-turn or for that matter, any rotation, is also the composition of two line reflections, then half-turns and rotations in general are direct isometries.

9.14 <u>Exercises</u>

1. Prove that the composition of three line reflections is an opposite isometry.
2. Investigate the nature of the composition of an even number of line reflections and of the composition of an odd number of line reflections. For verification see Figure 9.11(e).
3. What kind of isometry is a glide-reflection?
4. What kind of isometry is the composition of any number of:
   (a) translations? (b) rotations? (c) half-turns?
5. (a) Show that the composition of two line reflections and a half-turn, in any order, is a direct isometry.

(b) Should we use an even number or an odd number of line reflections in composition with a half-turn to produce an opposite isometry?

6. Prove that the composition of half-turns cannot be a line reflection.

7. Prove that the composition of rotations is never a line reflection.

8. What should you compose with a glide-reflection to obtain a direct isometry? An opposite isometry?

9. Is the identity isometry direct or opposite? Show that an isometry and its inverse isometry are either both direct or both opposite. (An isometry f is the inverse of isometry g if f∘g is the identity isometry.)

10. Prove: A direct isometry with two fixed points is the identity isometry.

9.15 Groups of Isometries

You know that a group is an operational system (S,∘) with the three properties of associativity, existence of an identity element in S, and the existence in S of an inverse for each element in S. The set of all isometries in a plane with composition as the operation is, as we shall see, a group.

Let us first see whether this system is an operational system. To do this we must convince ourselves that the composition of any two isometries is an isometry. Remembering that all iso-[illegible]ies are one-to-one mappings of the plane onto itself, it is

clear that the composition of two isometries is also a one-to-one mapping of the plane onto itself. Since distance is preserved under under an isometry, it continues to be preserved under the composition of two isometries. Thus composition of isometries preserves distance and hence is an isometry.

Now let us check whether or not the operational system has the group properties. First the associative property. Since composition of transformations in general is associative, so is composition of isometries.

The isometry that leaves all points of the plane fixed is the identity isometry, designated i. It is clear that, if f is an isometry, $i \circ f$ or $f \circ i$ has the same effect as f alone. So $i \circ f = f \circ i = f$. Thus the identity requirement is satisfied.

Let f be any isometry. Since f is a one-to-one mapping of the plane onto itself, it follows that the inverse mapping f^{-1} exists, and is a one-to-one mapping of the plane onto itself. Now let A, B be distinct points, $A \xrightarrow{f^{-1}} C$ and $B \xrightarrow{f^{-1}} D$. Then $C \xrightarrow{f} A$, $D \xrightarrow{f} B$. Since f is an isometry, CD = AB. Thus f^{-1} has the property that if $A \xrightarrow{f^{-1}} C$ and $B \xrightarrow{f^{-1}} D$, then AB = CD. Thus f^{-1} is an isometry, and clearly $f \circ f^{-1} = f^{-1} \circ f = i$. Thus the set of all isometries is a group.

As an example of a subgroup (see Chapter 2 Section 2.1) of the group of isometries we offer T, the set of translations. Let $T_{a,b}$ represent the translation with rule $(x,y) \longrightarrow (x+a, y+b)$. Then for all $T_{a,b}$, $T_{c,d}$

(1) $T_{c,d} \circ T_{a,b} = T_{a+c,\ b+d}$ which is in T, proving $(T,\circ)$ is an operational system.

(2) $T_{o,o} \circ T_{a,b} = T_{a,b} \circ T_{o,o} = T_{a,b}$. Hence $T_{o,o} = i$.

(3) $T_{-a,-b} \circ T_{a,b} = T_{a,b} \circ T_{-a,-b} = T_{o,o} = i$.

Thus T is a subgroup of the group of isometries.

9.16 Exercises

1. Consider the set of translations in a plane. Let $T_{(A,B)}$ represent the translation that maps A onto B.

   Interpret: (i) $T_{(A,A)}$ (ii) $T_{(B,A)}$ (iii) $T_{(B,C)} \circ T_{(A,B)}$.

2. Prove the set of direct isometries is a group under composition.

3. Does the set of opposite isometries form a group?

4. Is the set of half-turns a subgroup of the group of isometries? (Hint: Consider the composition of two half-turns.)

5. Is the set consisting of half-turns and translations a subgroup of the group of isometries?

6. Does the set of rotations with the same center form a subgroup? In your investigations designate a rotation with center P and measure a by r(P,a). Interpret r(P,a) to be in the counter-clockwise direction when $a > 0$, in the clockwise direction when $a < 0$.

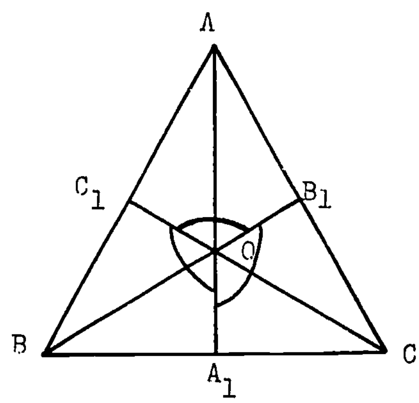
7. Prove: If f and g are isometries, then

   $$(f \circ g)^{-1} = g^{-1} \circ f^{-1}.$$

8. Generalize the theorem in Exercise 7 for n isometries.

9. The groups we have considered so far are infinite. In this exercise we consider a finite group of isometries. Let $\triangle ABC$

be equilateral. Observe that each median* of an equilateral triangle lies in the perpendicular bisector of a side and in the bisector of an angle. Let the medians of $\triangle ABC$ be $\overline{AA_1}$,



$\overline{BB_1}$, and $\overline{CC_1}$. The medians meet at a point; call it O. $m\angle AOB = m\angle BOC = m\angle COA = 120$, and $OA = OB = OC$. Consider the rotation $r(O,120)$, which we abbreviate r_1. $A \xrightarrow{r_1} B$, $B \xrightarrow{r_1} C$, $C \xrightarrow{r_1} A$. Therefore $\overline{AB} \xrightarrow{r_1} \overline{BC}$, $\overline{BC} \xrightarrow{r_1} \overline{CA}$, and $\overline{CA} \xrightarrow{r_1} \overline{AB}$. In short, $\triangle ABC \xrightarrow{r_1} \triangle BCA$. Therefore the image of $\triangle ABC$ under r_1 is $\triangle ABC$ itself. We describe this by saying that r_1 leaves the triangle invariant.

(a) Let $r_2 = r(O, -120)$. Show that r_2 leaves $\triangle ABC$ invariant.

(b) Let R_1 be the reflection in $\overleftrightarrow{AA_1}$. Show that R_1 leaves $\triangle ABC$ invariant.

(c) Let R be the reflection in $\overleftrightarrow{BB_1}$. Show that R_2 leaves $\triangle ABC$ invariant.

(d) Let R_3 be the reflection in $\overleftrightarrow{CC_1}$. Show that R_3 leaves $\triangle ABC$ invariant.

* A median of a triangle is a line segment whose endpoints are a vertex of the triangle and the midpoint of the side of the triangle opposite that vertex.

(e) Make a table showing all possible compositions of two isometries in the set $S = \{i, r_1, r_2, R_1, R_2, R_3\}$.

(f) Show that (S, o) is a group.

(g) Find all the subgroups of (S, o)

9.17 Isometry, Congruence and Symmetry

A congruence is a relation between figures. In everyday life we describe two such related figures as being exact copies of each other, or as having the same shape and size. It is quite difficult to give a precise mathematical definition of "same shape and size," as you might convince yourself if you were to try. However, the notion of an isometry is helpful.

Definition 2. Two figures are called congruent if there is an isometry that maps one of these figures onto the other. If there is an isometry that maps figure F onto figure F' we designate the congruence

$$F \cong F',$$

and read it: F is congruent to F'.

In Figure 9.15 we illustrate such a congruence for two triangles. Note first that the isometry is a translation.

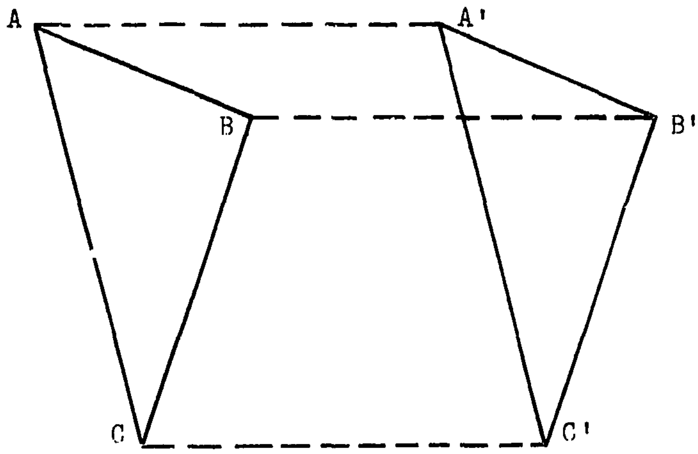


Figure 9.15

Under this isometry $A \longrightarrow A'$, $B \longrightarrow B'$, and $C \longrightarrow C'$. Since isometries preserve segments:

$$\overline{AB} \longrightarrow \overline{A'B'}, \qquad \overline{BC} \longrightarrow \overline{B'C'}, \text{ and } \overline{CA} \longrightarrow \overline{C'A'}$$

In short, $\triangle ABC \longrightarrow \triangle A'B'C'$. Therefore $\triangle ABC \cong \triangle A'B'C'$.

Each part (side or angle) of $\triangle ABC$ is mapped onto a part of $\triangle A'B'C'$. These parts are called <u>corresponding</u> parts. In particular, each side of $\triangle ABC$ corresponds to a side of $\triangle A'B'C'$. Since isometries preserve both distance and angle measures, we conclude that corresponding sides and corresponding angles of these congruent triangles, or congruent figures in general, have the same measure.

It is customary in designating a congruence to indicate the correspondence of parts by the order of the vertices. Thus $\triangle ABC \cong \triangle DEF$ indicates $A \longrightarrow D$, $B \longrightarrow E$ and $C \longrightarrow F$. These correspondences determine all other correspondences that involve vertices, such as $\overline{AB} \longrightarrow \overline{DE}$.

An important instance of a congruence occurs in a parallelogram. We shall use it often.

If ABCD is a parallelogram then $\triangle ABD \cong \triangle CDB$. To prove this statement we have to find an isometry under which $\triangle ABD \longrightarrow \triangle CDB$. (See Figure 9.16.)

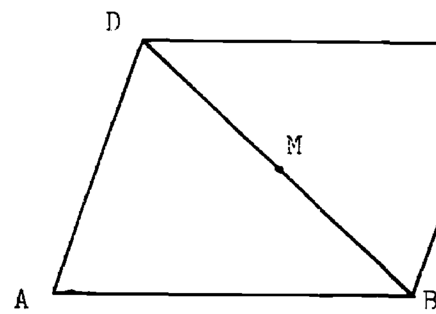


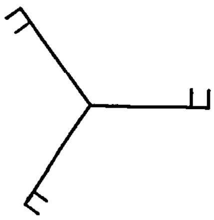
Figure 9.16

The isometry is H_M, the half-turn about M, the midpoint of $\overline{BD}$; $B \xrightarrow{H_M} D$, $D \xrightarrow{H_M} B$, and, since the diagonals of a parallelogram bisect each other, $A \xrightarrow{H_M} C$. Thus $\triangle ABD \xrightarrow{H_M} \triangle CDB$, and finally, $\triangle ABD \cong \triangle CDB$.

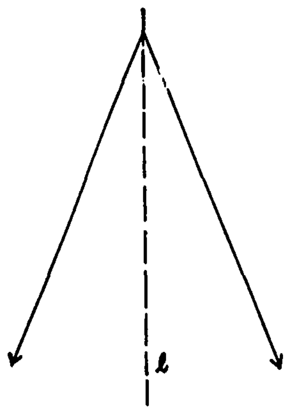
Symmetry is a property of a figure.

<u>Definition</u> 3. A figure is <u>symmetric</u> if there is an isometry, other than the identity, that transforms the figure onto itself.

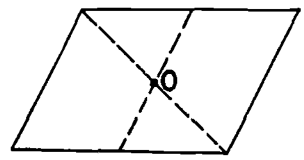
If the isometry is a line reflection we say the symmetry is a line symmetry, or symmetry in a line. See Figure 9.17 for symmetries related to different isometries.



(a)
rotation symmetry
(through 120^{o})



(b)
line symmetry
in ℓ



(c)
point symmetry
in O

Some figures, like the rectangle, have a variety of symmetries because there are more isometries than one that map the the figure onto itself.

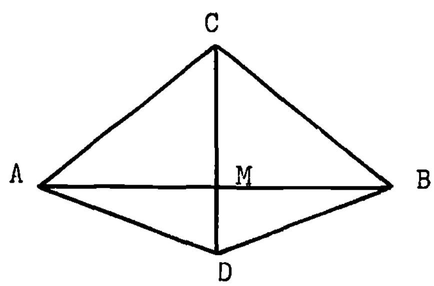
9.18 Exercises

1. Draw a diagram of $\overline{AC}$ and $\overline{DE}$, bisecting each other at B. Prove $\Delta ABD \cong \Delta CBE$.

2. Let $\overleftrightarrow{CD}$ be the perpendicular bisector of $\overline{AB}$. Prove:

   (a) $\Delta ACD \cong \Delta BCD$

   (b) $\angle CAD$ and $\angle CBD$ have the same measure.



(c) $AC = BC$.

(d) If $\overline{AB} \cap \overleftrightarrow{CD} = M$, prove $\triangle ACM \cong \triangle BCM$ and $m\angle CAB = m\angle CBA$.

(e) Show that ADBC has line symmetry.

3. Would your proof in Exercise 2 be different if D were between C and M in the diagram?

4. Suppose two circles have radii of the same length. Do you think they are congruent? Support your answer with a drawing.

5. (a) Let ABCD be a rectangle. Prove $\triangle ABD \cong \triangle BAC$. (Assume that the perpendicular bisector of $\overline{AB}$ is also the perpendicular bisector of $\overline{DC}$.)

   (b) List all the symmetries in ABCD.

6. Let $\triangle ABC$ be equilateral, and let its medians intersect at O. Assuming that $m\angle AOB = m\angle BOC = m\angle COA = 120$, prove $\triangle AOB \cong \triangle BOC$, and $\triangle AOB \cong \triangle COA$.

7. Prove:

   (a) Any figure is congruent to itself.

   (b) If $F \cong F'$, then $F' \cong F$.

   (c) If $F_1 \cong F_2$ and $F_2 \cong F_3$, then $F_1 \cong F_3$.

   Is the congruence relation an equivalence relation?

8. In how many ways can one consider an equilateral triangle to be symmetric? An isosceles triangle? A circle? (See Definition 3 and Figure 9.17.)

9.19 Other Transformations: Dilations and Similarities

We would not like to leave the impression that all transformations are isometries. Indeed there are many more that are not.

In this section we examine just one set of transformations that are not isometries, although it bears many resemblances to the set of isometries.

Let point O be given, and also a fixed non-zero number, say 2, for purposes of illustration. Then to point A (see Figure 9.18) we assign A', where A' is in $\overrightarrow{OA}$ and OA' is 2OA. B' is assigned to B if B' is in $\overrightarrow{OB}$ and OB' = 2OB, and so on.

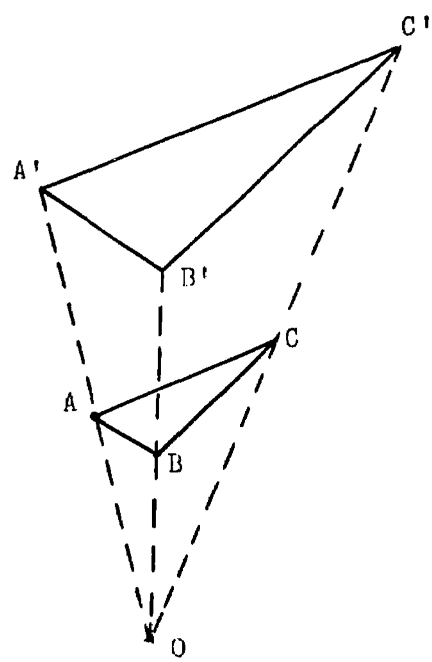


Figure 9.18

If the given number is negative we take the image of A in the ray opposite $\overrightarrow{OA}$. Clearly every point in the plane is assigned a point of the plane, and every point serves as the image of one point. Thus this method of assignment is a transformation; but as you can readily see by comparing AC with A'C', it is not an isometry. This transformation is called a *dilation* with *center O* and *scale factor 2*.

We can form a composition of a dilation with any isometry, and the result is called a *similarity*. Note that if the scale

factor of the dilation component is 1 or -1, a similarity is an isometry. When it is -1, then the dilation is a half turn. (See Figure 9.19.)

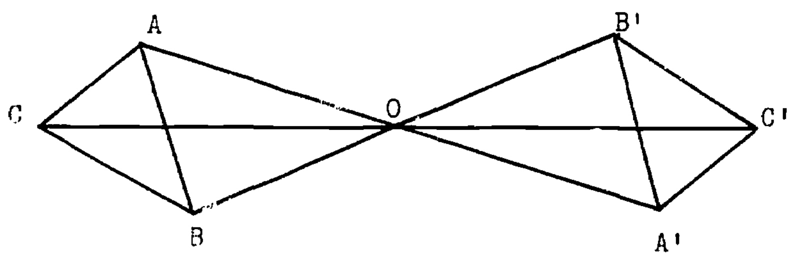


Figure 9.19

9.20 Exercises

1. Make a drawing of two triangles related by a dilation with scale factor

   (a) 3 (b) $\frac{1}{2}$ (c) -1 (d) -2

2. Show that a dilation with scale factor -1 is a half-turn.

3. Using a drawing show that if a dilation with scale factor 2 maps points A and B onto A' and B', then A'B' = 2AB. Generalize this result.

4. Show that a dilation maps three collinear points onto three collinear points. (Hint: Use the generalization in Exercise 3. If necessary review the proof of the analogous statement in Exercise 6, Section 9.4.)

5. Show that the set of dilations with center O is a group under the operation of composition.

6. Draw a picture of a dilation with scale factor $\frac{1}{2}$, followed by a translation, as it affects three noncollinear points.

7. Draw a picture of a dilation with scale factor $1\frac{1}{2}$, followed by a reflection in a line, as it affects three noncollinear points.
8. Suggest a definition for direct and opposite similarities.
9. You may have observed that a similarity preserves the ratio of distances. Does it seem to preserve the measure of angles? Base your answer on the drawings you have made in these exercises.

9.21 <u>Summary</u>

1. A plane transformation is a mapping of the plane onto itself that is one-to-one. A transformation is an isometry if it preserves distance.
2. Among isometries are compositions of line reflections, translations, glide-reflections, and rotations. The half-turn is a special case of a rotation.
3. The line reflection is the basic isometry in the sense that any isometry is the composition of line reflections of which no more than three are needed. A rotation is a composition of two reflections with intersecting axes. In the case of a half-turn the axes are perpendicular. A translation is the composition of two line reflections in parallel axes. A glide reflection is composition of a line reflection and a translation with a direction of the axis of the reflection.
4. All isometries preserve collinearity of points, the betweeness relation among points, rays, segments, angles, and angle measure, in addition to distance.

5. Under the identity transformation all points are fixed. Under a line reflection the axis is fixed pointwise, and lines perpendicular to the axis are fixed, but not pointwise. Under a rotation the center is fixed; under a translation no points are fixed; under a glide reflection the axis is fixed, but not pointwise. Under a half-turn all lines through the center are fixed, but not pointwise.
6. Half-turns, rotations and translations are direct isometries, that is, they preserve the sense of three non-collinear points. All other isometries we discussed are opposite.
7. The set of isometries, with the operation of composition, is a group. The set of translations and half-turns is a subgroup. The set of translations is a subgroup.
8. Two figures are congruent if an isometry maps one onto the other. A figure is symmetric if an isometry maps the figure onto itself.
9. Not all plane transformations are isometries. For instance, the composition of a dilation with an isometry, called a similarity, is not in general an isometry. Similarities preserve collinearity, betweeness, ratio of distances and angle measure.

9.22 Review Exercises

1. Given three noncollinear points A, B, C. For each part below make a different diagram showing how to find:
   (a) the reflection of A in $\overleftrightarrow{BC}$.
   (b) the image of A under the translation that maps B onto C.
   (c) the image of $\overline{AB}$ under the half-turn in C.
   (d) the image of A under the glide reflection whose axis is $\overleftrightarrow{BC}$, and that maps B onto C.

2. (a) Given parallelogram ABCD. Show that the parallelogram is preserved under the half-turn whose center is the midpoint of $\overline{AC}$.
   (b) Is the parallelogram preserved under the translation that maps A onto B?
   (c) Is the parallelogram preserved under the line reflection whose axis contains the midpoints of $\overline{AB}$ and $\overline{CD}$?

3. Given ABCD is a square, and let $\overline{AC}$ intersect $\overline{BD}$ in E. Show that the square is invariant under:
   (a) a rotation with center E and measure 90 (counterclockwise), (r_1).
   (b) a half-turn with center E(r_2).
   (c) a rotation with center E and measure -90 (clockwise), (r_3).
   (d) a reflection in the axis through the midpoints of $\overline{AB}$ and $\overline{CD}$ **(R_1)**.
   (e) a reflection in $\overleftrightarrow{AC}$ (R_3).
   (f) Name three more isometries that preserve the square.

(Did you remember that the identity transformation is an isometry?)

(g) Let the symbols in parentheses name the respective isometries. They are r_1, r_2, r_3 for rotations; R_1, R_2 for the reflections in axes passing through midpoints of sides; R_3, R_4 for reflections in diagonals. Show that $(\{i, r_1, r_2, r_3, R_1, R_2, R_3, R_4\}, \circ)$ is a group, where $\circ$ denotes composition of transformations, by displaying the table showing all possible compositions.

(h) Find a subgroup containing 4 elements.

(i) Find five subgroups each containing 2 elements.

4. Given ABCD is a parallelogram. Show $H_P \circ H_A = H_C \circ H_D$.

5. Given point P is on line a. Show $H_P \circ R_a = R_a \circ H_P$.

6. Prove that the composition of four line reflections cannot be a glide reflection.

7. Given line ℓ. Under which isometries is the image of ℓ parallel to ℓ? Do not consider such cases as line reflections where ℓ is parallel to the axis, nor half-turns whose centers are on ℓ.

8. Describe isometries under which each of the following conditions is satisfied:

(a) All points are fixed.

(b) All points in one line are fixed.

(c) There are no fixed points.

(d) There is exactly one fixed point.

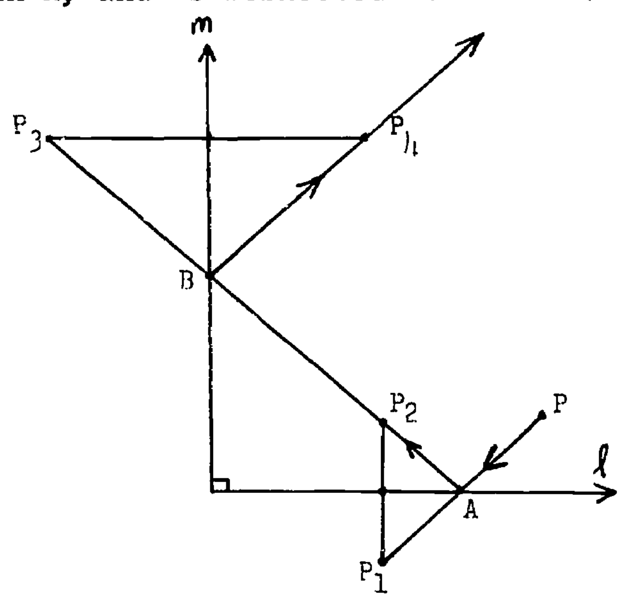
9. (a) Which isometries are direct? Which are opposite?

(b) An isometry is an involution. Is it necessarily an opposite isometry? May it be?

10. Let line b be the perpendicular bisector of $\overline{PQ}$.

Show $H_p \circ R_b = R_b \circ H_Q$.

*11. Let ℓ and m represent mirrors at right angles. A ray of light issues from point P in the direction indicated in the figure, strikes ℓ in A, and is reflected as follows: Let $P \xrightarrow{R_\ell \circ H_A} P_2$.



Then the beam of light follows the path $\overrightarrow{AP_2}$ until it strikes m, say at B. Let $P_2 \xrightarrow{R_m \circ H_B} P_4$. Then the beam of light follows the path $\overrightarrow{BP_4}$. Show that $\overleftrightarrow{AP} \parallel \overleftrightarrow{BP_4}$.

12. What is meant by saying that two figures are congruent? Give an example, in a drawing, of:

(a) two triangles that are congruent under a line reflection.

(b) two parallelograms that are congruent under a half-turn.

13. What is meant by saying that a figure has symmetry?

(a) Give an example of a figure that has line symmetry.

(b) Give an example of a figure that has both line and point symmetry.

(c) Give an example of a figure that has line, point, and rotation symmetry.

CHAPTER 10

LENGTH, AREA, VOLUME

10.1 Introduction

Kepler (1571-1630) said: "To measure is to know," and scientists and technicians have worked successfully by this dictum. As a result of their experiences in measuring various things, a number of questions have arisen for which they look to mathematicians for answers. Some of these questions are:

(1) What is the mathematical meaning of a measurement?

(2) What kind of numbers are needed for measuring?

(3) How are the operations of addition and multiplication related to measurement?

We look for answers to these questions in this chapter as we measure segments, regions, and solids.

10.2 Measures on Sets

The first mathematical process we learn to use is that of counting. Counting is not only the first process we learn, but it is one we continue using as we progress in mathematics, developing new uses and new techniques at various stages of our development. In this section we use counting as a tool for assigning measures to finite sets and unions, intersections, and cartesian products of such sets. Later in the chapter we [illegible]ly these measures of sets to the measurement of line segments,

planar regions, and solids.

Two sets, X and Y, are said to be equipotent (denoted $X \approx Y$) if there exists a one-to-one mapping of X onto Y. If X and Y are finite sets then $X \approx Y$ if and only if X and Y have the same number of elements.

We define the <u>counting measure of a finite set</u>, X, to be the number of elements in X. We denote this by n(X). If $X = \{1, 2, 3\}$ then $n(X) = n(\{1, 2, 3\}) = 3$; $n(\emptyset) = 0$. If $A = \{1, 2, 3\}$ and $B = \{2, 3, 5\}$, then $n(A) = 3$ and $n(B) = 3$ and $A \approx B$. In general,

$$n(X) = n(Y) \text{ if and only if } X \approx Y.$$

With sets A and B as given above, consider $A \cup B$ and $A \cap B$. $A \cup B = \{1, 2, 3, 5\}$. $n(A \cup B) = 4$. $A \cap B = \{2, 3\}$ and $n(A \cap B) = 2$. Thus $n(A) + n(B) = n(A \cup B) + n(A \cap B)$, or $n(A \cup B) = n(A) + n(B) - n(A \cap B)$.

If the set $C = \{4, 5, 6, 7\}$, then $n(C) = 4$. What is the counting measure of $A \cup C$? $A \cup C = \{1, 2, 3, 4, 5, 6, 7\}$, so $n(A \cup C) = 7$ and $n(A \cup C) = n(A) + n(C)$. Does this result differ from the one we obtained using sets A and B? No, because $A \cap C = \emptyset$ and $n(\emptyset) = 0$. This illustrates:

If X and Y are <u>disjoint</u> sets, then $n(X \cup Y) = n(X) + n(Y)$. This property is called <u>additivity</u> of measures, and is used only with reference to disjoint sets.

If $P \subset Q$ it follows that $n(P) \leq n(Q)$. (When is $n(P) = n(Q)$?)

Now let us consider the set $A \times B = \{(1, 2), (1, 3), (1, 5), (2,2), (2, 3), (2, 5), (3, 2), (3, 3), (3, 5)\}$. This set has 9 elements, each of which is an ordered pair of elements, the

first from A, the second from B. $n(A) = 3$, $n(B) = 3$, $n(A \times B) = 3 \cdot 3 = 9$. In general, if X and Y are finite sets, then

$$n(X \times Y) = n(X) \cdot n(Y).$$

Note that in the case of cross products, we are not concerned with disjointness of the original sets. We shall extend these counting techniques in Chapter 11 on Combinatorics.

In any section dealing with sets, we select one set, which we call the universal set (for that section), and restrict the discussion to subsets of the universal set. In this section we might have selected $U = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$ for our universal set.

10.3 Exercises

Let $X = \{\text{positive integers less than } 50\}$ be our universal set with $A, B, C, D \subset X$ where:

$A = \{x \in X: X \text{ is a multiple of } 3\}$

$B = \{x \in X: X \text{ is a multiple of } 5\}$

$C = \{x \in X: X \text{ is a multiple of } 6\}$

$D = \{x \in X: X \text{ is a multiple of } 11\}$

Find the number of elements in:

1. A, B, C, D.
2. $A \cap B$, $A \cap C$, $A \cap D$, $B \cap C$, $B \cap D$, $C \cap D$.
3. $A \cup B$, $A \cup C$, $A \cup D$, $B \cup C$, $B \cup D$, $C \cup D$.
4. $A \times B$, $A \times C$, $A \times D$, $B \times C$, $B \times D$, $C \times D$.
5. $A \times B \times D$.
6. $A \cap B \cap D$, $A \cup B \cup D$.

10.4 Lengths of Line Segments

You know how to measure a line segment with a ruler. Nevertheless, let us review the procedure with the hope of finding clues that will suggest answers to the questions of Section 10.1 as they apply to segments.

Suppose we are to measure the segment in Figure 10.1 with a ruler marked only in inches. Our first assumption is that each segment on the ruler between consecutive marked points is congruent to a unit segment whose measure, we say, is 1, and therefore each segment has a length of 1 inch.

This observation, though obvious, embodies two principles of measure, important enough to formulate:

The unit principle: To measure a segment we must start with a unit segment. (To measure an angle we must start with a unit angle. To measure anything we must start with a unit of that thing.) The measure of that unit is 1.

The congruence principle: Segments congruent to the unit segment have measure 1. In general, congruent segments have the same measure. So do congruent rectangles, congruent triangles, congruent cubes; and so any congruent figures have the same measure.

We have used the term "measure" in a precise way. It means a number. However the number is associated with a unit. When we say that the length of $\overline{AB}$ is 3 inches, the "3" is the measure, and the "inch" is the unit. When the unit is clearly understood we say: "The measure of $\overline{AB}$ is 3", and write $m(\overline{AB}) = 3$, or $AB = 3$.

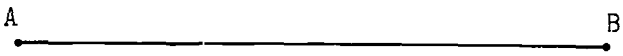


Figure 10.1

Returning to the measurement of segment $\overline{AB}$ in figure 10.1, we place the zero point of the ruler on A, and see if B is at one of the points marked on the ruler. If B is at 3, we say that the measure or length of $\overline{AB}$ is 3; or AB = 3.

In passing we note that if AB = 3, then $\overline{AB}$ is made up of three unit segments, each having measure 1. To get 3, we add 1 + 1 + 1. This illustrates the next measure principle.

<u>The additive principle</u>: The measure of a segment is the sum of the measures of the parts into which it is subdivided. For example, if E is between C and D, as in Figure 10.2 then CE + ED = CD.

Figure 10.2

Going back to our measurement of $\overline{AB}$, suppose B does not fall on one of the inch marks of our ruler. Does this mean we have failed in our measurement? Not entirely; for suppose B falls somewhere between 3 and 4, as it actually does in Figure 10.3. Then we know that the measure of $\overline{AB}$, is somewhere between 3 and 4. We can use 3 as a first approximation to $m(\overline{AB})$, and $3 < m(\overline{AB})$. We recognize that 4 is an upper bound for $m(\overline{AB})$.

Suppose the 3" mark on our ruler falls on a point D of $\overline{AB}$. Then D is between A and B, with $m(\overline{AD}) = 3$, and AD + DB = AB.

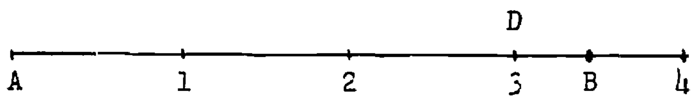


Figure 10.3

To get a better approximation to $m(\overline{AB})$, we examine the segment $\overline{DB}$ which we have not yet measured. A part of Figure 10.3 containing $\overline{DB}$ is magnified and shown in Figure 10.4, in which a ruler marked in tenths of an inch is used.

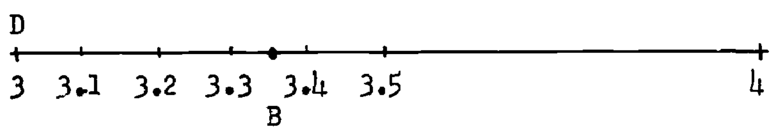


Figure 10.4

If B is at one of the points marked on the ruler, say at 3.3, we take $m(\overline{AB}) = 3.3$ If B falls between two of the marks say between 3.3 and 3.4, as in Figure 10.4, we get a second approximation, 3.3, for $m(\overline{AB})$, and $3.3 < m(\overline{AB})$. We next use a ruler marked in hundredths of an inch, etc. This procedure is continued, and it may happen that ultimately B coincides with a point of division of a ruler. We then get a measure for $\overline{AB}$ such as 3.37 or 3.372. In this case $m(\overline{AB})$ is a rational number. It may however happen that B never coincides with a point of division of a ruler, and we get a sequence of approximations to $m(\overline{AB})$: 3, 3.3, 3.37, 3.372,..., each less than $m(\overline{AB})$. Since $m(\overline{AB})$ is bounded above (by 4), the completeness property of the real number system assures us that this sequence has a least upper bound, which is a real but not necessarily rational number. The way the sequence was constructed suggests that this least upper bound be called the measure (or length) of the line segment. This leads to:

<u>The real number principle</u>: The measure of a segment is a positive real number.

Our examination of the procedure for measuring a line segment has led us to a number of measurement principles. We stated them as applying to segments. It is remarkable that the same principles operate when we measure other geometric figures. You are asked to be alert in noting that this is so when we examine areas and volumes. To review these principles we list their names: the unit principle, the congruence principle, the additive principle, the real number principle.

10.5 Areas of Rectangular Regions

The next object for measure study is the union of a rectangle with its interior, which we call a <u>rectangular region</u>. Such a region is <u>not</u> a set of collinear points, and so we cannot use a segment as the unit of measurement. It <u>is</u> a set of coplanar points. A unit principle for measuring regions requires a unit that is a region. A convenient unit region is a square region; that is, the union of a square with its interior. To be definite we can use a square region each of whose sides is 1 inch long. We call this unit a <u>square inch</u>. It is also possible to use other units such as a square foot or a square meter.

If we put six square inches together as shown in Figure 10.5, we form rectangular region ABCD.

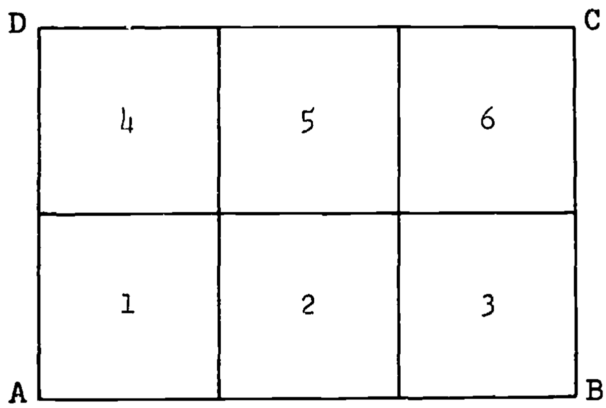


Figure 10.5

By a congruence principle for regions corresponding to the one used for line segments, the measure of each square inch is 1. By the corresponding additive principle for r-gions, the measure of the rectangular region ABCD is 6. It is convenient to have a brief symbol for the measure of the region bounded by rectangle ABCD. We use K_{ABCD}. The combination of the measure with the unit region, like 6 square inches, is called the area of the region. The word "area" is synonymous with "measure of region".

As a second example of the way in which the additive principle works for regions, look at Figure 10.6. Let D be a point in the interior of $\triangle ABC$. The union of $\triangle ABC$ with its interior is called a triangular region ABC. We have not yet shown how to assign measures to triangular regions.

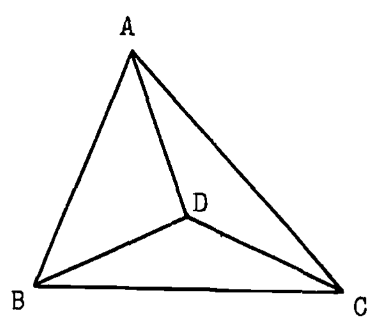


Figure 10.6

Nevertheless, we can be sure that by the additive principle for regions, in this case triangular regions, the measure of triangular region ABC is equal to the sum of the measures of triangular regions ABD, BDC, and CDA.

To develop the technique for finding the measure of a rectangular region, called the <u>area</u> of the region, let us use PQRS (see Figure 10.7 (a)) as our unit square.

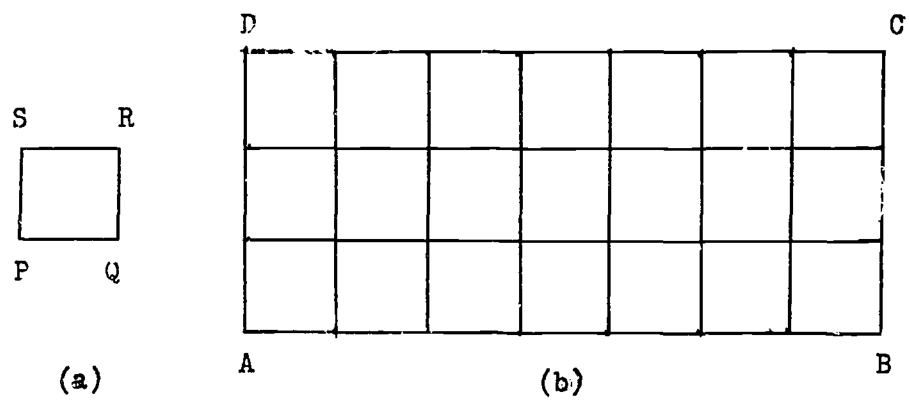


Figure 10.7

Since we have a unit square, it seems reasonable for each of its sides to have unit measure of length, and we assume this to be so.

<u>Case</u> 1: The sides of ABCD have whole number measures. In this case we subdivide ABCD, as in Figure 10.7 (b), into squares each of which is congruent to PQRS. If $m(\overline{AB}) = p$ and $m(\overline{AD}) = n$, then we count $p \cdot n$ of these squares and we see that the measure of ABCD, denoted K_{ABCD}, is $p \cdot n = m(\overline{AB})m(\overline{AD})$.

<u>Case</u> 2: $m(\overline{AB})$ and $m(\overline{AD})$ are rational numbers, say $\frac{p}{q}$ and $\frac{s}{q}$. Here p, s, and q are positive integers.

In this case we introduce a square, whose side measures $\frac{1}{q}$, as our new unit square. As in Case 1 there are $p \cdot s$ such unit squares. Since our original unit square can be subdivided into q^2 of these new unit squares we have by additivity the area of each of the new unit squares to be $\frac{1}{q^2}$. (Confirm this.) Therefore (again by additivity) the area measure of ABCD is $p \cdot s \cdot (\frac{1}{q^2}) = \frac{p}{q} \cdot \frac{s}{q}$ $= m(\overline{AB})m(\overline{AD})$.

Case 3: $m(\overline{AB})$ and $m(\overline{AD})$ are real but at least one of them not rational. In this case we define the area measure of ABCD to be $m(\overline{AB}) \cdot m(\overline{AD})$. That such a definition is reasonable is illustrated in Figure 10.8, namely for any rational approximations $m(\overline{AB'})$ and $m(\overline{AD'})$ with $m(\overline{AB'}) \leq m(\overline{AB})$ and $m(\overline{AD'}) \leq m(\overline{AD})$ the area of ABCD is larger than the area of AB'C'D' and as B' gets closer to B and D' closer to D the area of AB'C'D' which is equal to $m(\overline{AB'}) \cdot m(\overline{AD'})$ gets closer to the area of ABCD.

Since in all three cases the area is given by $m(\overline{AB}) \cdot m(\overline{AD})$ we can state, if we write $m(\overline{AB}) = \ell$, $m(\overline{AD}) = w$ that:

$$K = K_{ABCD} = \ell w$$

ℓ is called the length of ABCD, w the width.

ℓ and w are called the dimensions of the rectangular region.

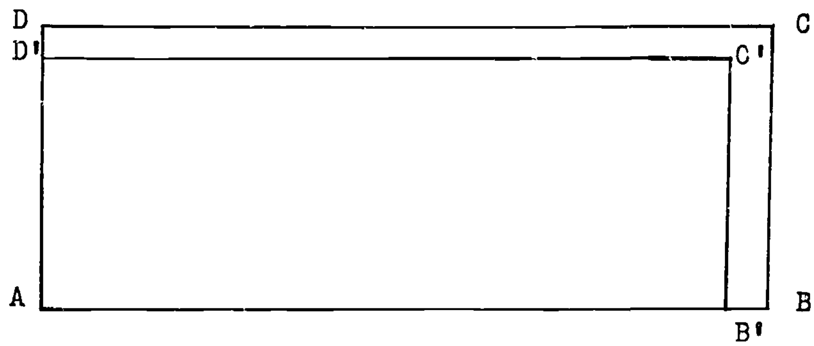


Figure 10.8

You should observe that l and w in this formula are measures that refer to the same segment (or linear) unit. If both are in inches, then the area measure is in square inches. However, if the length of one side of a rectangle is given as 4 inches, and another as 2 feet, we convert one unit into the other, say 2 feet into 24 inches, and then proceed by the formula $K = lw$.

Note that measures of two sides of a rectangle can also be used to find the perimeter of a rectangle, that is, the sum of the measures of all sides, the sum being given in the common unit. Thus the perimeter of a rectangle, whose dimensions are 7 inches by 8 inches, is $7 + 8 + 7 + 8$ or 30 inches. Here too, we cannot find a perimeter unless all measures of sides are given in the same unit.

In this discussion we have an answer to question (3) of Section 10.1: How are the operations of addition and multiplication of numbers related to measurement?

In the exercises that follow, and thereafter, we will

e K_{ABC} to denote the measure of the region bounded by $\triangle$ ABC;

and if P, Q, R and S are the successive vertices of a quadrilateral, K_{PQRS} will denote the measure of the region bounded by quadrilateral PQRS.

10.6 Exercises

1. Let h be the inch-measure of one side of a rectangle and b the inch-measure of an adjacent side. For each pair of values listed below find the area of the related rectangular region:

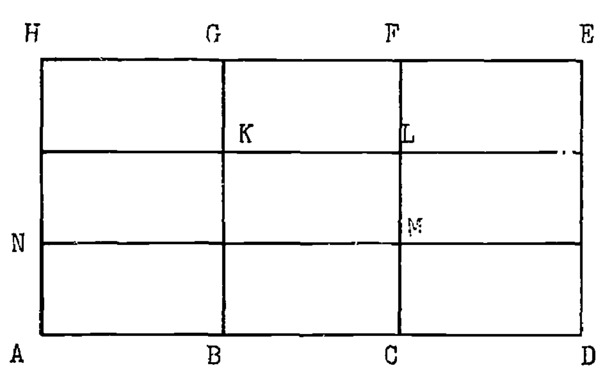
   (a) $h = 3$, $b = 7$ (b) $h = 7$, $b = 3$ (c) $h = 3\frac{1}{2}$, $b = 6$

   (d) $h = 4.1$, $b = 3.2$ (e) $h = \sqrt{2}$, $b = \sqrt{3}$ (f) $h = 2\sqrt{5}$, $b = 3\sqrt{5}$

   (g) $h = 6$, $b = 1\frac{1}{3}$ (h) $h = 1\frac{1}{2}$, $b = 2\frac{1}{5}$ (i) $h = 2-\sqrt{2}$, $B = 2+\sqrt{2}$

2. Prove that the area of a square region, each of whose sides has length s is given by the formula $K_{square} = s^2$.

3. The rectangular region ADEH is subdivided into 9 congruent rectangular regions. Let $m(\overline{AD}) = 15$, and $m(\overline{AH}) = 6$. Find the area of each of the following rectangular regions:



   (a) ADEH (b) ACMN (c) ABGH (d) BCLK

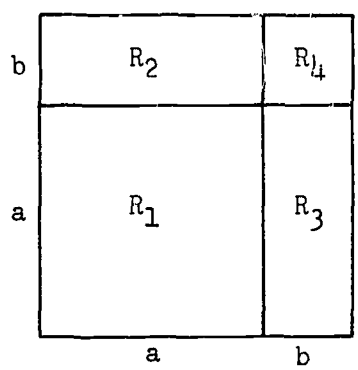
   (e) The rectangular region with diagonal $\overline{AL}$.

4. Using the figure in Exercise 3, tell why each of the following is true:

(a) $K_{ABGH} = K_{BCFG}$ (b) $K_{BCLK} + K_{KLFG} = K_{BCFG}$

(c) $K_{BCLK} = \frac{2}{9}K_{ADEH}$ (d) $K_{ADEH} - K_{CDEF} = K_{ACFH}$

5. In the figure, the regions R_1 and R_4 are bounded by squares, and R_2 and R_3 are bounded by rectangles. Let a be the measure of each side of R_1, and b the measure of each side of R_4.



In terms of a and b, express:

(a) the area of R_1 (b) the area of R_2

(c) the area of R_3 (d) the area of R_4

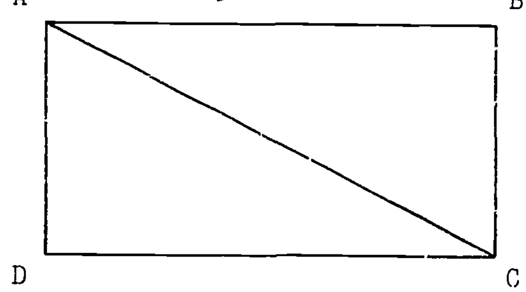
(e) The entire figure which is subdivided into R_1, R_2, R_3 and R_4, is itself a square region. Show that the area of the entire square region is equal to the sum of the areas of the subdivisions.

6. Let d represent the inch-measure of a diagonal of a rectangle and s the inch-measure of one side of that rectangle. Find the area of the corresponding rectangular region for the values of d and s given below.

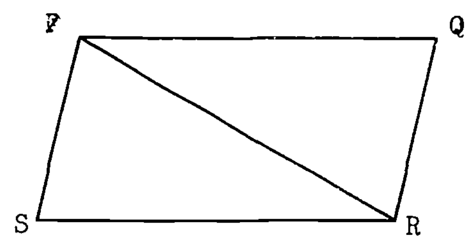
(a) $d = 5$, $s = 4$ (b) $d = \sqrt{41}$, $s = 5$ (c) $d = 13$, $s = 12$

(d) $d = 10$, $s = 6$ (e) $d = \sqrt{11}$, $s = \sqrt{5}$ (f) $d = 25$, $s = 15$

7. ABCD is a rectangle. Prove: $K_{ACD} = K_{CAB}$, and $K_{ACD} = \frac{1}{2}K_{ABCD}$. (Hint: Use a halfturn about midpoint of $\overline{AC}$ to prove triangle ACD congruent to triangle CAB.)



8. PQRS is a parallelogram. Prove: $K_{PRS} = K_{RPQ}$, and $K_{PQRS} = 2K_{PQR}$.



9. In the figure of Exercise 8, let $\overline{SQ}$ intersect $\overline{PR}$ in T. Compare K_{PTQ} with K_{RTS}.

10. Discuss the validity of the statement: If two rectangular regions have the same area they are congruent.

11. Show: (a) The area of a square foot is 144 square inches.
    (b) The area of a square yard is 9 square feet.
    (c) The area of a square centimeter is 100 square millimeters.

12. Compare the areas of two rectangular regions if each of the dimensions of one is twice a dimension of the other.

13. Let ℓ_1 and w_1 be the dimensions of a rectangular region, and ℓ_2 and w_2 the dimensions of a second rectangular region. Find the ratio of the area of the first region to that of the second region if:

(a) $\ell_2 = \ell_1$ and $w_2 = 2w_1$ (b) $\ell_2 = 2\ell_1$ and $w_2 = 3w_1$

(c) $\ell_2 = \frac{1}{2}\ell_1$ and $w_2 = 2w_1$ (d) $\ell_2 = \frac{1}{3}\ell_1$ and $w_2 = 4w_1$

14. Find the area of a square region whose diagonal is:

    (a) 12 inches long (b) $8\sqrt{2}$ feet long

    (c) $6\sqrt{2}$ miles long

15. Prove the area of a square region is $\frac{d^2}{2}$ if d represents the measure of a diagonal.

10.7 Volumes of Rectangular Solids

You know the shape of a box. It has 6 faces or regions each of which is rectangular. The union of such a figure with its interior is called a rectangular solid. It is a set of points, not all in one plane. Some of its subsets are in a plane and these have areas. But the entire set can be measured. Such a measurement is called a volume.

The unit principle demands that the unit of volume be a solid. We take a cubic solid for convenience. In Figure 10.9 you see a cubic solid. Each of its six faces is a square region. The solid itself consists of all points in the faces and all points

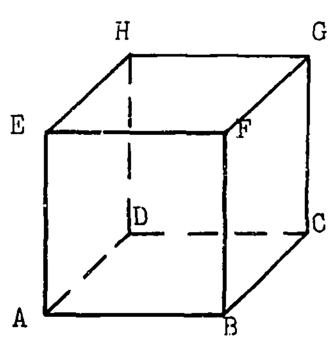


Figure 10.9

in the interior. If each of its edges measures 1 inch, it is called a cubic inch; and the measure of its volume is 1.

No doubt you have observed that the units we have dealt with are an inch (a segment), a square inch (a region) and now a cubic inch (a solid).

In Figure 10.10 we put together 6 cubic inches.

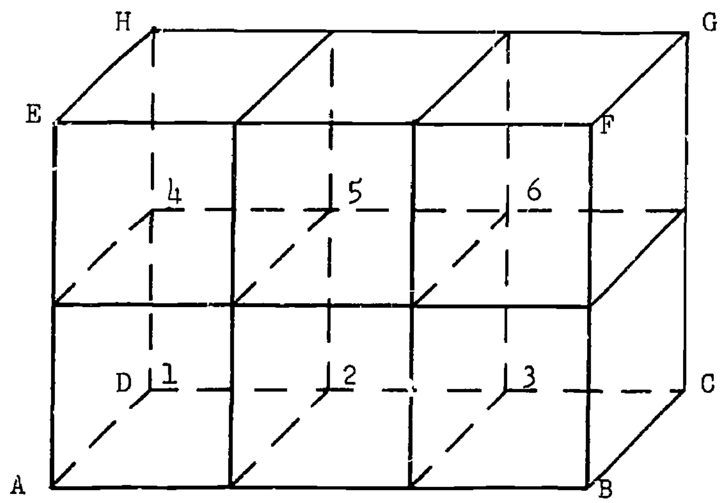


Figure 10.10

The result is a rectangular solid. The six cubic inches are congruent to each other, so the measure of each is 1. The rectangular solid is subdivided into the six cubes. By the additive principle its volume is 6 cubic inches.

In this rectangular solid the region ABCD can be taken as the base of the rectangular solid. The length and width of this region is also the length and width of the rectangular solid. For the base ABCD, the height of the solid is BF. Note, in this case, if ℓ is the measure of $\overline{AB}$, w the measure of $\overline{BC}$, and h the measure of $\overline{BF}$, then the measure of the solid is $\ell \cdot w \cdot h$, that is $3 \cdot 1 \cdot 2 = 6$.

Can we use this method for finding the volume of any rectangular solid? The answer is yes. In a method analogous

to that used in Section 10.5 for finding the area of a rectangular region, we work with the measures of three edges of a rectangular solid that meet in a point (vertex), say $\overline{AB}$, $\overline{BC}$, and $\overline{BF}$ in Figure 10.11, and arrive at an entirely analogous result;

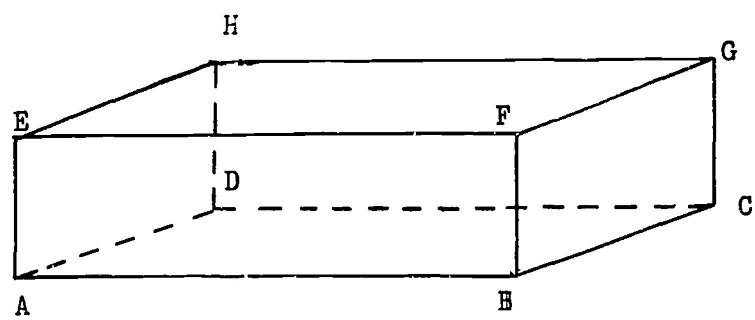


Figure 10.11

namely, if $m(\overline{AB}) = \ell$, $m(\overline{BC}) = w$, $m(\overline{BF}) = h$, and if V denotes the volume of the rectangular solid, then:

$$V = m(\overline{AB}) \cdot m(\overline{BC}) \cdot m(\overline{BF}) = \ell wh$$

The measures of three edges of a rectangular solid that meet in a vertex are called the <u>dimensions</u> of the solid.

10.8 <u>Exercises</u>

1. Find the volume of a rectangle solid if its dimensions, in feet, are given below:

   (a) $\ell = 3$, $w = 4$, $h = 2$ (b) $\ell = 2\frac{1}{2}$, $w = 4$, $h = 5$

   (c) $\ell = \sqrt{2}$, $w = \sqrt{3}$, $h = 2$ (d) $\ell = 3.1$, $w = 2.3$, $h = 4$

2. The dimensions of a box are 2', $1\frac{1}{2}$', 1'. The dimensions of a second box are $1\frac{3}{4}$', $1\frac{3}{4}$', 1'. Which has the greater volume? How much larger?

3. I want to have a measure that tells me how much water a tank can hold. Is the measure length, area, or volume?

4. Assume the question asked in Exercise 3 if I want to know how much:

(a) fencing I need to enclose a yard.

(b) seed I need to plant a lawn.

(c) air there is in a room.

5. Prove: The volume of a cube each of whose edges is e units long, is e^3 cubic units.

6. (a) Prove there are 1728 cubic inches in one cubic foot.

   (b) How many cubic feet are there in one cubic yard?

   (c) How many cubic decimeters are there in one cubic meter?

   (d) Show that the number of cubic yards in a cubic mile is about $5 \cdot 10^9$.

7. The coordinates in a rectangular space coordinate system, listed below in each part are those of vertices of a rectangular solid. Find the volume of each.

   (a) (0,0,0), (3,0,0), (3,2,0) , (0,2,0), (0,0,5), (3,0,5), (3,2,5), (0,2,5)

   (b) (-2,1,-3), (2,1,-3), (2,4,-3), (-2,4,-3), (-2,1,2), (2,1,2), (2,4,2), (-2,4,2)

10.9 Areas of Triangular Regions

Using the formula for finding the area of a rectangular region, we can derive formulas for areas of triangular regions.

Formula 1: If a and b are the measures of the legs of a right triangle, and K the measure of the related triangular region, then $K = \frac{1}{2}ab$.

Derivation: Let the triangle have vertices A,B,C, with the right angle at C (See Figure 10.12). Let AC = b, CB = a.

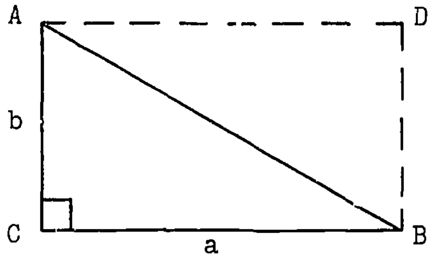


Figure 10.12

If D is the image of C under a half turn in the midpoint M of $\overline{AB}$, we can show (using the sum of the degree measures of the angles of a triangle is 180), that ACBD is a rectangle and that $\triangle ACB \cong \triangle BDA$. This implies:

$K_{ACB} = K_{BDA}$ (by the congruence principle)

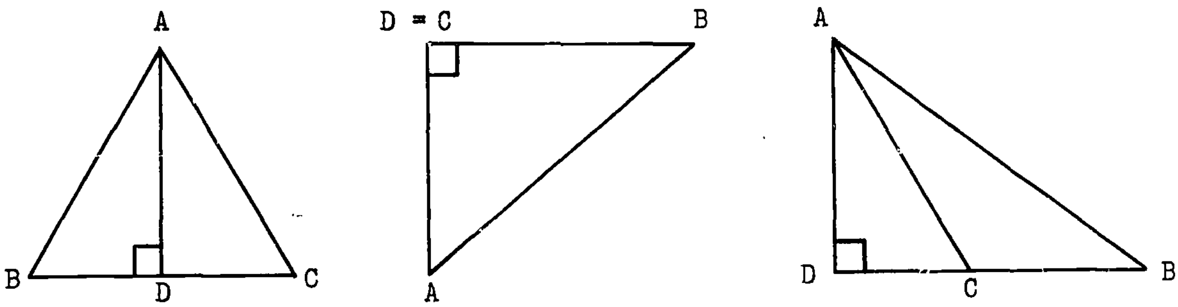
But $K_{ACB} + K_{BDA} = K_{ACBD}$ (by the additive principle)

$K_{ABCD} = ab$ (since ACBD is a rectangle)

Therefore: $K_{ACB} = \frac{1}{2}ab$

To be able to state Formula 2 we define altitude of a triangle.

In each triangle of Figure 10.13, the perpendicular from A to BC meets BC in D. Note in Figure 10.13 (a) that D is in the interior of $\overline{BC}$, in Figure 10.13 (b) D = C, and in Figure 10.13 (c) D is not in $\overline{BC}$. For each figure $\overline{AD}$ is the altitude of $\triangle ABC$ from A to BC. The word altitude is also used to mean the measure of $\overline{AD}$.



(a) (b) (c)

Figure 10.13

Definition 1. If A is a point and ℓ a line, the point in which the perpendicular from A to ℓ intersects ℓ is called the foot of the perpendicular from A to ℓ.

In each of the triangles in Figure 10.13, D is the foot of the perpendicular from A to BC.

Definition 2. The segment that joins a vertex of a triangle to the foot of the perpendicular from that vertex to the line containing the opposite side, is called the altitude of the triangle from that vertex.

How many altitudes does a triangle have?

Formula 2: If a is the measure of an altitude of a triangle from one vertex, b the measure of the side opposite that vertex, and K the measure of the triangular region, then

$$K = \frac{1}{2}ab.$$

Derivation: We have to consider three cases: an acute triangle (each of its angles has measure less than 90), a right triangle (one of its angles has measure 90), and an obtuse triangle (one of its angles has measure greater than 90). In all three cases we consider $\triangle ABC$ with altitude $\overline{AD}$ from A.

Case 1. $\triangle ABC$ is an acute triangle. In this case we assume, from Figure 10.13 (a), that D is between B and C. We use the preceding theorem to calculate the areas K_{ABD} and K_{ACD}. For each AD = a; and by the additive

principle of areas, their sum is K_{ABC}. Thus

$$K_{ABD} = \frac{1}{2}a \cdot BD, \; K_{ACD} = \frac{1}{2}a \cdot DC.$$

Therefore

$$K_{ABC} = \frac{1}{2}a \cdot BD + \frac{1}{2}a \cdot DC$$
$$= \frac{1}{2}a(BD + DC).$$

Since D is between B and C, BD + DC = BC or b, and

$$K_{ABC} = \frac{1}{2}ab$$

<u>Case</u> 2. $\triangle ABC$ is a right triangle. D must coincide with C since $\angle C$ is a right angle. (See Figure 10.13 (b).) This shows Formula 1 is a special case of Formula 2.

<u>Case</u> 3. $\triangle ABC$ is an obtuse triangle with obtuse angle at C. In this case we assume, from Figure 10.13 (c) that C is between D and B, so that DC + CB = DB.

$$K_{ADB} = \frac{1}{2}a \cdot DB, \; K_{ADC} = \frac{1}{2}a \cdot DC,$$

and

$$K_{ADB} = K_{ADC} + K_{ACB}$$

or

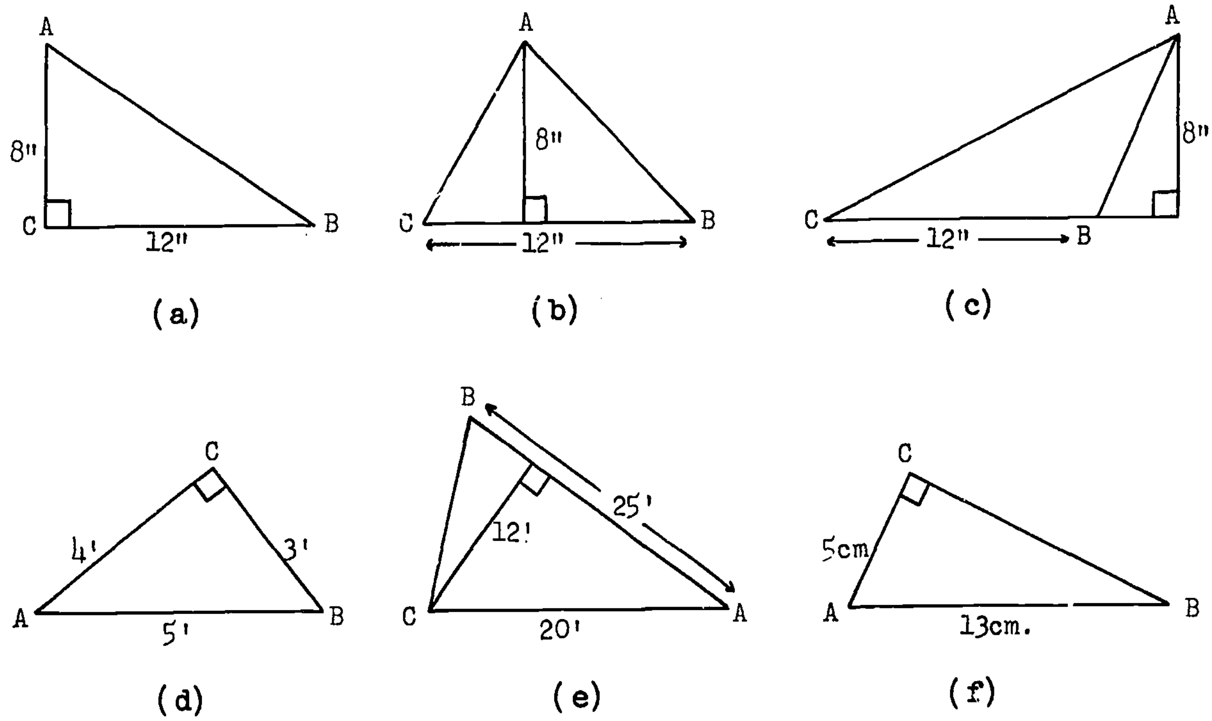
$$K_{ACB} = K_{ADB} - K_{ADC}$$
$$= \frac{1}{2}a \cdot DB - \frac{1}{2}a \cdot DC$$
$$= \frac{1}{2}a(DB - DC).$$

Since (in Figure 10.13 (c)) C is between D and B, DB - DC = CB or b. Therefore,

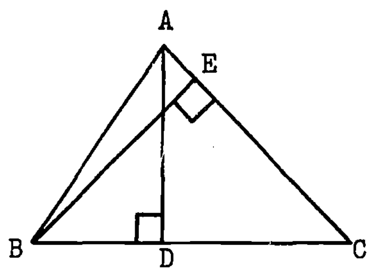
$$K_{ACB} = \frac{1}{2}ab.$$

10.10 Exercises

1. For each of the following figures use the indicated information to find K_{ABC}.

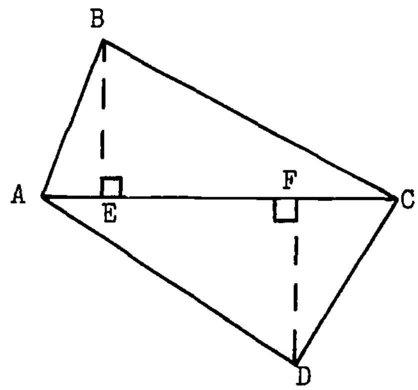


2. In finding K_{ABC} one can use $\overline{BC}$ as base or $\overline{AC}$ as base. Draw a large triangle with two altitudes $\overline{AD}$ and $\overline{BE}$, and measure $\overline{BC}$, $\overline{AC}$, $\overline{AD}$, and $\overline{BE}$ (preferably in millimeters or tenths of an inch).



Using these measures, calculate K_{ABC} in two ways and compare the results.

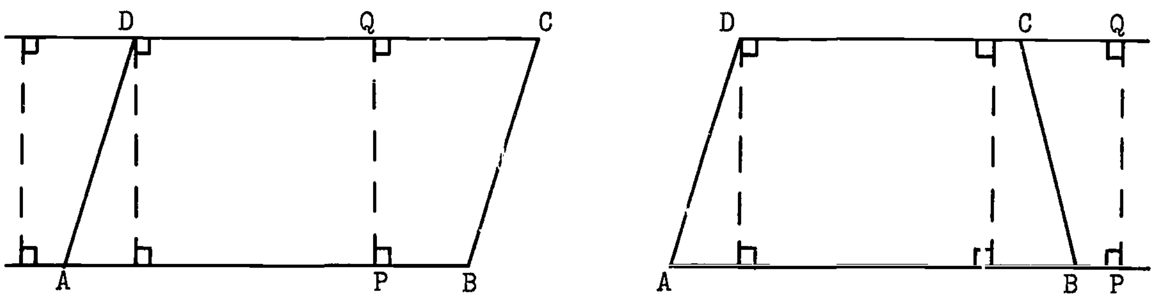
3. Find the area of a right triangular region if its hypotenuse is 26 inches long and one of its legs is 10 inches long.
4. The length of the hypotenuse of a right triangle is h and the length of one of its legs is ℓ. Prove the area of the triangular region is $\frac{1}{2}\ell\sqrt{h^2 - \ell^2}$.
5. The legs of a right triangle are 6 and 8 inches long. How long is the altitude to the hypotenuse?
6. A field has the shape of ABCD, as shown. If $\overline{AC}$ is 40 yards long, $\overline{BE}$ an altitude of $\triangle ABC$, is 20 yards long and $\overline{DF}$, an altitude of $\triangle ABC$, is 30 yards long, what is K_{ABCD}?



7. The measure of each leg of an isosceles triangle is 10 and that of the base is 8. Find the area of the triangular region.
8. Each side of an equilateral triangle measures 12. Show
   (a) the length of any of its altitudes is $6\sqrt{3}$, and
   (b) the area of the triangular region is $36\sqrt{3}$.
9. The measure of each side of an equilateral triangle is s. Prove the area of the triangular region is $(\frac{s}{2})^2\sqrt{3}$ or $\frac{s^2}{4}\sqrt{3}$.

10.11 <u>Areas of Parallelogram and Trapezoidal Regions</u>

Now that we have a formula for the area of a triangular region, we can use it to find the area of any region that can be subdivided into triangular regions. Simple examples of such regions are those bounded by parallelograms and trapezoids. A trapezoid is a quadrilateral ABCD for which one of AB $\parallel$ DC, AD $\parallel$ BC holds. If both hold, the quadrilateral is a parallelogram. In trapezoid (or parallelogram) ABCD let $\overleftrightarrow{AB} \parallel \overleftrightarrow{DC}$, $P \in \overleftrightarrow{AB}$, $Q \in \overleftrightarrow{DC}$, $\overline{PQ} \perp \overleftrightarrow{AB}$, and $\overline{PQ} \perp \overleftrightarrow{DC}$; then $\overline{PQ}$ is called an <u>altitude</u> of the trapezoid (or parallelogram). In Figure 10.14 (a) we show a parallelogram with three altitudes, and in Figure 10.14 (b) we show a trapezoid, also with three altitudes. The altitudes are all shown as dotted lines.



(a) (b)

Figure 10.14

Observe that the altitudes of a parallelogram or a trapezoid are congruent, since any two altitudes are opposite sides of a rectangle.

<u>Definition</u> 3. The sides of a parallelogram or a trapezoid that contain the endpoints of an altitude are called <u>bases</u> for that altitude. (In the case of a parallelogram, there are two pairs of bases,

and the bases in each pair are congruent.)

Formula 1: If the measure of a base of a parallelogram is b, the measure of an altitude to that base is a, and the area of the parallelogram region is K, then:

$$K = ab$$

Derivation: Let ABCD be the parallelogram, with $\overline{AB}$ taken as base, and $\overline{DE}$ as altitude to AB. (We assume from Figure 10.15 that $\overline{DB}$ divides ABCD into two triangles to which we may apply the additive principle for areas.) Then $DE = a$ and $AB = b$. Since $\triangle ABC \cong \triangle CDB$, by the congruence principle for areas,

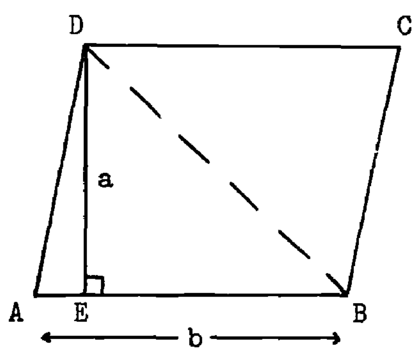


Figure 10.15

$K_{ABD} = K_{CDB}$. Also by the additive principle for areas, $K_{ABCD} = K_{ABD} + K_{CDB}$. However, $K_{ABD} = \frac{1}{2}ab = K_{CDB}$. Therefore $K_{ABCD} = \frac{1}{2}ab + \frac{1}{2}ab = ab$.

Formula 2: If the measures of the bases of a trapezoid are b and c, the measure of its altitude is a, and the measure of its region is K, then

$$K = \frac{1}{2}a(b + c).$$

<u>Derivation</u>: Let the trapezoid be ABCD, with bases $\overline{AB}$ and $\overline{CD}$. (We make the same assumption here from Figure 10.16, that we made in the proof of Formula 1.) Then AB = b, CD = c, and each altitude drawn, $\overline{DE}$ and $\overline{BF}$ (see Figure 10.16), has measure a.

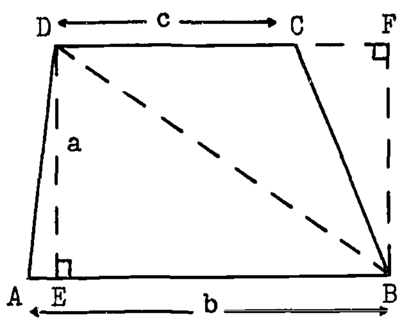


Figure 10.16

$$K_{ABCD} = K_{ABD} + K_{BDC}$$

$$K_{ABD} = \frac{1}{2}ab, \quad K_{BDC} = \frac{1}{2}ac$$

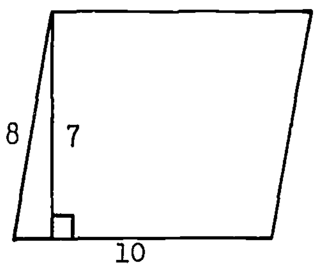
Therefore:

$$K_{ABCD} = \frac{1}{2}ab + \frac{1}{2}ac$$

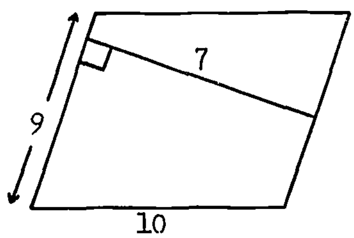
$$= \frac{1}{2}a(b + c)$$

10.12 <u>Exercises</u>

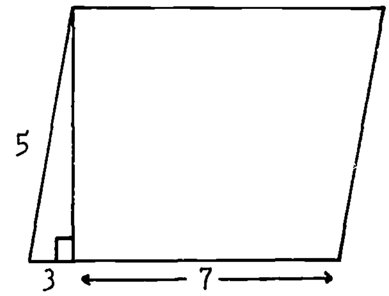
1. For each of the parallelograms, assume that all indicated segment measures have the same unit, and find the area of the region bounded by each.



(a)

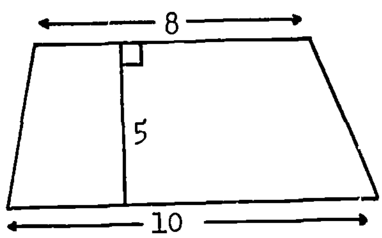


(b)

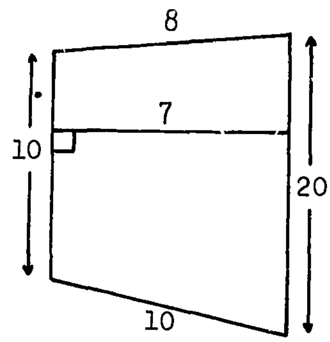


(c)

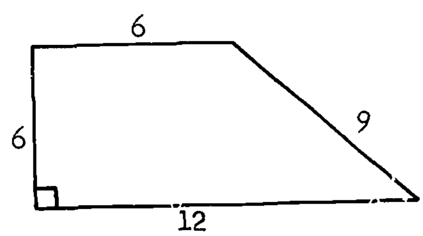
2. For each trapezoid below, assume all indicated segment measures have the same unit and find the area of the trapezoidal region.



(a)



(c)



(b)

3. In a certain plane rectangular coordinate system, the vertices of a quadrilateral have coordinates as listed below. Find the area of the region bounded by each such quadrilateral.

(a) (0,0), (8,0), (7,5), (2,5)

(b) (0,0), (8,0), (10,6), (2,6)

(c) (-3,0), (0,-3), (3,0), (0,3)

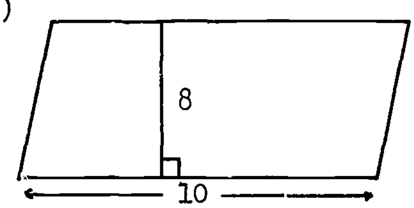
(d) (-3,-2), (5,-2), (6,3), (0,3)

(e) (0,0), (5,0), (5,8), (0,4)

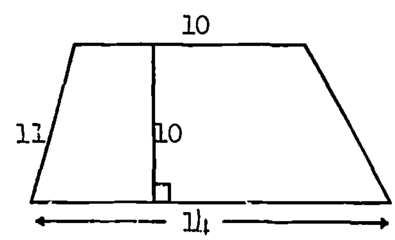
(f) (-3,-1), (5,-3), (5,4), (-3,3)

4. For each figure below a piece of metal is to be made having inch measures as indicated. If this metal costs 15 cents per square inch, what is the cost of each?

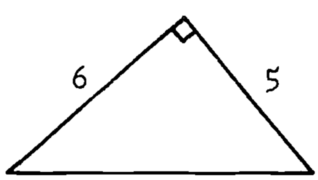
(a)



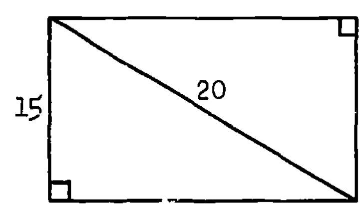
(b)



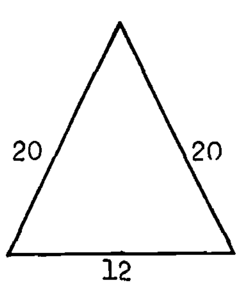
(c)



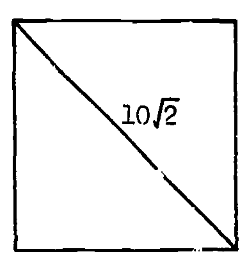
(d)



(e)



(f)



a square

5. Find the ratio of the areas of two regions, each bounded by a square, if the length of the side of the first square is n times the length of a side of the second, and n is equal to:

(a) 2 (b) 3 (c) $\frac{2}{3}$ (d) k

6. Find the ratio of the areas of two regions, each bounded by an equilateral triangle, if the length of a side of the first is n times the length of a side of the second, and n is equal to:

   (a) 2 (b) 3 (c) $\frac{2}{5}$ (d) k

10.13 Areas of other Regions

Having seen how to measure some simple regions, we get some notions about area which apply to more general figures.

First of all we see that area is a function that maps regions into the set of positive real numbers. It is determined as soon as a unit region is selected. The domain has been vaguely defined, but it certainly includes rectangular regions and such other regions that can be subdivided into triangular regions. Moreover, if figure A is congruent to figure B, and they have areas, then they have the same area. Also, if a figure can be subdivided into a finite number of regions, each having an area, then the area of that figure is the sum of the areas of the subdivisions.

Now we go on to see how these notions can be used to determine the area of a plane figure that cannot be subdivided into triangular regions, say a map of Africa. Our first step is to choose a unit region, preferably a square region. We have chosen a square each of whose sides is $\frac{1}{2}$ inch long, and shall call it a square unit. Our next step is to overlay the map of Africa with a grid whose squares are the square units we have chosen. See Figure 10.17. Observe that every square unit region [illegible]n be classified as belonging to one of three sets: First,

there are square units all of whose points are points of the map. Let us call these inner square units. Second, there are square

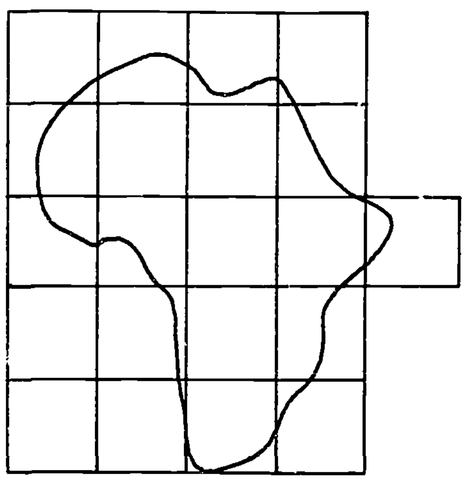


Figure 10.17

units that have some points in common with the map, but not all. Let us call these border square units. Finally, there are square units that have no points in common with the map. We disregard these.

The unit squares are congruent to each other. Hence the area of each is 1. Now consider the union of the inner square units. It consists of 4 such regions and hence its area is 4. We call this an approximation of the area of the map. Now consider the union of all inner and border square units. There are 19 of these. The area of this union is therefore 19. This is an upper bound of the area of the map. In finding the approximation and the upper bound we used the additive property of areas. Let M be the area of the map. We can then say

$$4 \leq M \leq 19.$$

To get a closer approximation than 4 we use a smaller square region. We can bisect the sides of the square units by lines midway between the parallel lines in the grid. We get a new grid where square regions have areas of $\frac{1}{4}$ square unit. See Figure 10.18.

A count of the inner squares shows them to number 23. Hence, the approximation by the grid is $23\ \frac{1}{4}$ square units, or $5\frac{3}{4}$ square units.

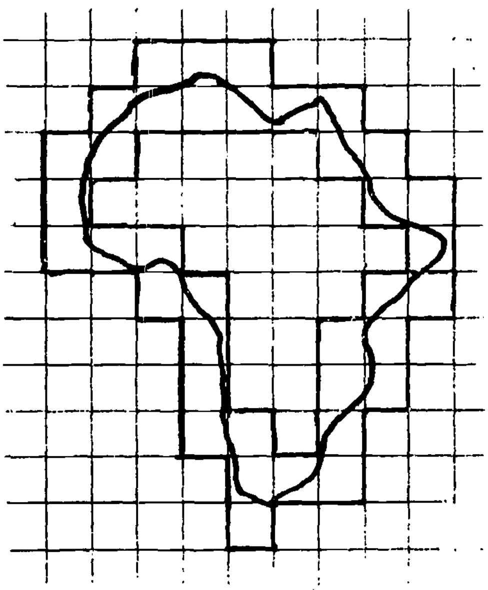


Figure 10.18

We continue with a grid of still smaller square regions, each of which has an area of $\frac{1}{16}$ square unit. See Figure 10.19.

By actual count we find the approximation to be $121\ \frac{1}{16}$ square units or $7\frac{9}{16}$ square units.

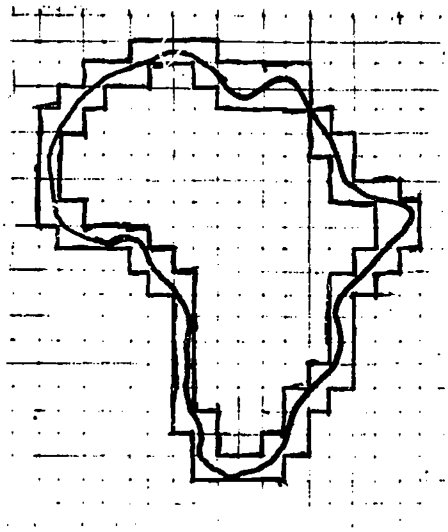


Figure 10.19

If we continue improving our approximations by using a finer grid, we get a sequence of rational numbers which has an upper bound. Therefore this sequence has a least upper bound, and this least upper bound we define to be M, the area of the map of Africa.

10.14 Circumference of a Circle and π

You may recall that a circle C is a set of points in a plane (see Figure 10.20) such that the distance OP, from any point P of the set to a fixed point O, is the same for all choices of P. The point O, though not a point of C, is called the center of C. $\overline{OP}$ is called a radius of C, and all radii (plural of radius) of a circle have the same length, and hence are congruent.

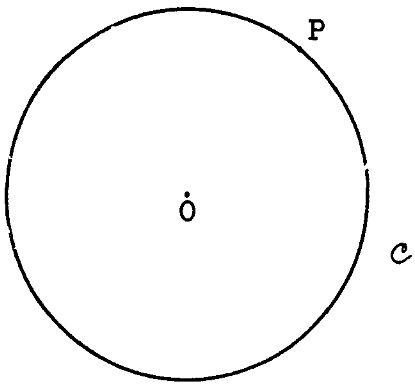


Figure 10.20

The term "radius" is also used to denote the length of segment $\overline{OP}$. The context will make clear whether "radius" means a segment or a measure. It is clear that if we place a string along a circle, and cut the string when it makes one turn of the circle, we can find the length of the section of string that was cut off. It would certainly seem reasonable to take this length as the length of the circle. The length of a circle is called its <u>circumference</u>.

In this section we will outline a method for finding the circumference of a circle. Since a circle is not a set of collinear points, nor is it a union of segments, we cannot measure a circle as we measure segments. But, as you might anticipate, we can use the method of approximations, which have an upper bound.

In Figure 10.21 we have drawn square ABJD so that its vertices are points of circle C with center O, and radius r and a second square, EFGH whose sides touch the circle at A, B, J and D. Let C present the circumference of C. The perimeter of ABJD is

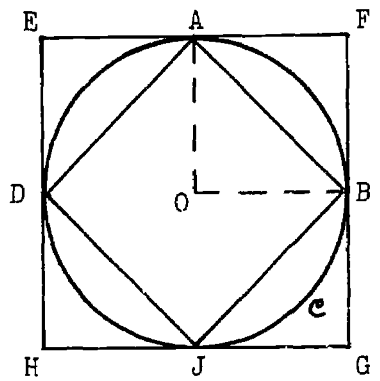


Figure 10.21

an approximation of C and the perimeter of EFGH is taken as upper bound of C. We can see that ΔAOB is a right triangle. By the Pythagorean property $(AB)^2 = r^2 + r^2$, or $AB = r\sqrt{2}$. We can also see that EF = 2r. So the perimeter of EFGH is 8r. Thus

$$4r\sqrt{2} < C < 8r \quad \text{or} \quad 5.656r < C < 8r$$

To improve our approximation of C, we take additional points on the circle (see Figure 10.22) and form a new polygon by drawing line segments joining consecutive points.

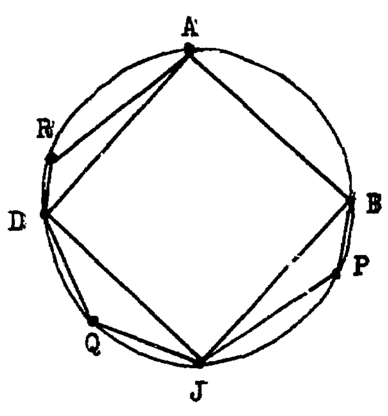


Figure 10.22

P, Q, and R are such points in Figure 10.22. Since P, B, and J are the vertices of a triangle we know that $BP + PJ > BJ$. Thus each time we choose another point on the circle and use it to form a new polygon, the new polygon has a greater perimeter than did the old one. This process of choosing additional points is restricted in only one way. The sides of the polygon must all become arbitrarily small as the process goes on. This process gives us an increasing sequence of numbers, bounded above by $8r$, the perimeter of the circumscribed square. Therefore, there is a least upper bound for the sequence and this least upper bound we define to be the *circumference of the circle*.

No matter how the points are chosen, except for the restriction above, it turns out that the least upper bound of the resulting sequence of numbers will be the same, a constant multiple of the radius of the circle. It is convenient to represent half of this number by the symbol π (Greek letter pi) and we conclude

$$C = 2\pi r.$$

We saw above:

$$4r\sqrt{2} < C = 2\pi r < 8r$$

Dividing the inequalities by $2r$ yields:

$$2\sqrt{2} < \pi < 4$$

Since $1.4 < 2$ (why?), it follows that:

$$2.8 < \pi < 4$$

It is known that π is not a rational number. A better approximation to π than the above is $3\frac{1}{7}$ or 3.14. A still better approximation is 3.1416. The formula relating π to C and r is $C = 2\pi r$.

To illustrate: The circumference of a circle with radius 5 is $C = 2\pi \cdot 5 = 10\pi$. This is the exact circumference. An approximation to C is (10)(3.14) or 31.4.

10.15 Exercises

1. For each radius below find, in terms of π, the circumference of the circle of which it is the radius.

   (a) 10 (b) 8 (c) 1 (d) $\frac{1}{2}$ (e) $\sqrt{3}$

2. Using 3.14 as an approximation for π, find to the nearest unit, the circumference of a circle whose radius is:

   (a) 6 inches (b) 12 yards (c) 100 miles (d) $\sqrt{3}$ cm.

3. Using $3\frac{1}{7}$ as an approximation for π find the circumference of a circle whose radius is:

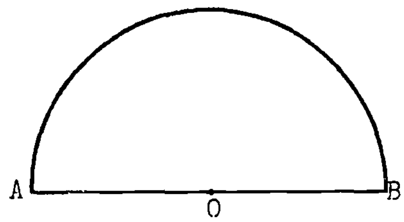
   (a) 7 inches (b) $3\frac{1}{2}$ feet (c) 1400 miles (d) 28 miles

4. You probably know that the segment that joins two points of a circle, and contains the center of the circle, is called a diameter of the circle. The term diameter is used also to denote the length of this segment. If d is the diameter and r is the radius of a circle, what is the ratio of d : r?

5. Find the circumference of a circle whose diameter is:

   (a) 12 inches (b) 50 feet (c) $\frac{1}{2}$ yard (d) .1 foot

6. Only that part of a circle is shown below that lies on one side of $\overleftrightarrow{AB}$, a line that contains diameter $\overline{AB}$. That part, including the end points A and B, is called a semicircle, and it too has a length.

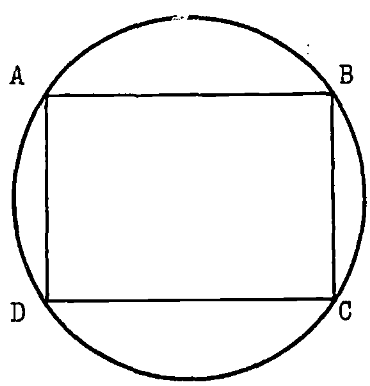


Find the length of a semicircle on a diameter whose length is:

(a) 20 inches (b) 50 feet (c) d

7. A square is drawn inside a circle with its vertices on the circle. If the radius of the circle is 10 inches, by how many inches, to the nearest inch, does the circumference exceed the perimeter of the square? (Assume a diagonal of the square is a diameter of the circle.)

8. ABCD is a rectangle with its vertices on a circle. Assume that a diagonal of this rectangle is also a diameter of the circle. Find the circumference of the circle if AD = 3 and DC = 5.



9. Find the radius of a circle if its circumference is:

(a) 24π (b) 33π (c) 24 (d) $2\pi k$

10. Find the radius of a circle if a semicircle of the circle has length:

(a) 18π (b) 4π (c) 18 (d) $2\pi k$

10.16 Areas of Circular Regions

The union of a circle with its interior is a circular region. A model of a circle is a ring; a model of a circular region is a coin or a phonograph record.

Once again we outline the method of approximations and upper bounds, this time to derive a formula for the area of a circular region.

For our first approximation we use the two square regions in Figure 10.21. Since AB = $r\sqrt{2}$, $K_{ABCD} = (r\sqrt{2})^2$ or $2r^2$. Since EF = 2r, $K_{EFGH} = (2r)^2$, or $4r^2$. Hence if we let $K_{\mathcal{C}}$ denote the area of the circular region:

$$2r^2 < K_{\mathcal{C}} < 4r^2$$

Here $2r^2$ is a first approximation to $K_{\mathcal{C}}$ and $4r^2$ an upper bound for $K_{\mathcal{C}}$.

We improve our approximation in the same way as before, except that we use regular polygons. Since the computations are difficult and not important at this time, we state the results.

If a regular octagon is inscribed in the circle, it encloses a region with area about $2.828r^2$, which is our second approximation to $K_{\mathcal{C}}$. If a regular polygon with 180 sides is used we get the approximation $3.141r^2$ for the area enclosed by the circle.

It seems that $K_{\mathcal{C}}$ is related to π. Indeed it has been proved that

$$K_{\mathcal{C}} = \pi r^2$$

To illustrate the use of this formula, we find the area of a circular region with radius 8. It is $\pi \cdot 8^2$ or 64π. To the nearest .1 this is 201.1 (using 3.1416 as an approximation for π.)

10.17 Exercises

1. For each radius listed below find, in terms of π, the area of a circular region of which it is the radius.

   (a) 10 (b) 8 (c) 1 (d) $\frac{1}{2}$ (e) $\sqrt{3}$

2. Using 3.14 as an approximation for π, find the area of a circular region whose radius is:

   (a) 6 inches (b) 8 yards (c) 10 miles (d) $\sqrt{5}$ cm.

3. Using $3\frac{1}{7}$ as an approximation of π find the area of the circular region whose radius is:

   (a) 7 inches (b) $3\frac{1}{2}$ feet (c) 140 yards (d) $\sqrt{7}$ miles

4. Find the area of a circular region whose diameter is:

   (a) 2 (b) 1 (c) 5 (d) $\sqrt{12}$

5. Find the radius of a circular region whose area is:

   (a) 25π (b) 64π (c) 49 (d) 20

6. Find the circumference of a circle whose region has area:

   (a) 25π (b) 4π (c) π (d) $\frac{1}{4}\pi$

7. Find the area of a circular region if the circumference of its circle is:

   (a) 16π (b) 26π (c) 8π (d) 8

8. The region bounded by a diameter and a semicircle on one side of the diameter is a semicircular region. Find the area of a semicircular region to the nearest unit if its diameter is:

   (a) 10 (b) 8 (c) 100 (d) 6

9. The ratio of the radii of two circles is 2 : 1.

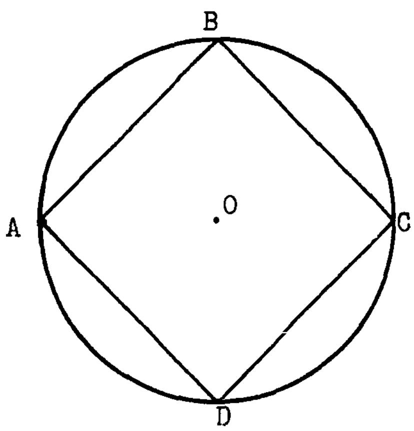
   (a) Find the ratio of their circumferences.

   (b) Find the ratio of the areas of their regions.

10. Find the two ratios called for in Exercise 9 if the ratio of the radii is:

    (a) 3 : 1 (b) 3 : 2 (c) 4 : 3 (d) 5 : 1

11. Points A, B, C, D are on a circle with center O and radius 10; and ABCD is a square.



(a) Find K_{ABCD}.

(b) Find the area of the circular region.

(c) Find the area of the region bounded by $\overline{AD}$ and that part of the circle that is on the opposite side of $\overleftrightarrow{AD}$ from B.

(d) Find the area of the region bounded by $\overline{AD}$ and that part of the circle that is on the same side of $\overleftrightarrow{AD}$ as B.

10.18 Summary

1. We have discussed four basic measurement principles.

   (a) The unit principle: The unit of measurement should be of the same kind as the object being measured. The measure of the unit is 1.

   (b) The congruence principle: If two figures are congruent and have measures, they have the same measure.

(c) The additive principle: If a figure is subdivided and the subdivisions have measures, then the measure of the figure is the sum of the measures of the subdivisions.

(d) The real number principle: The measure of a figure is a non-negative real number.

2. The basic method for measurement is the approximation and least upper bound method. If the approximations have a least upper bound, then that least upper bound is the measure of the figure.

3. Formulas for regions.

   Rectangular: $K = \ell w$ (ℓ and w are the dimensions of the rectangular region.)

   Triangular: $K = \frac{1}{2}ab$ (a is the measure of the altitude to the base with measure b.)

   Parallelogram: $K = ab$ (a is the measure of the altitude to the base with measure b.)

   Trapezoidal: $K - \frac{1}{2}a(b + c)$ (a is the measure of an altitude; b and c the measures of the bases.)

   Circular: $K = \pi r^2$

4. Formula for circumference of a circle:

$$C = 2\pi r$$

5. Formula for volume of a rectangular solid:

$V = \ell wh$ (ℓ, w, h are the dimensions of the rectangular solid.)

10.19 Review Exercises

1. State the additive principle as it applies to:

   (a) segments (b) regions (c) volumes

2. Find the area of the region bounded by an isosceles right triangle whose hypotenuse is 8 inches long.

3. Find the circumference of a circle, and the area of the circular region if the radius of the circle is:

   (a) 8 (b) 5 (c) 12

4. Find the circumference of a circle if the area of its region is 100π.

5. Find the perimeter of a square, and the area of its region, if the measure of one of its diagonals is 12.

6. Show that a median of a triangle subdivides it into two triangles whose regions have the same area.

7. In Δ ABC, median $\overline{AD}$ meets median $\overline{BE}$ in G. Find $K_{ADB} : K_{ADC}$.

8. The vertices of quadrilaterals have coordinates in a plane rectangular coordinate system as listed below, in the order given. Find the area of the region bounded by the quadrilateral.

   (a) (0,0), (3,0), (3,7), (0,2)

   (b) (-2,-2), (4,-2), (5,3), (-1,3)

   (c) (-2,0), (1,-3), (3,0), (1,3)

   (d) (-3,0), (-1,-4), (4,0), (3,5)

9. Find the volume of a rectangular solid whose dimensions are 3", 6", 2".

10. Find the volume of a rectangular solid whose vertices have the coordinates listed below, in some rectangular space coordinate system.

(3,-1,-2), (3,4,-2), (0,4,-2), (0,-1,-2)

(3,-1,4), (3,4,4), (0,4,4), (0,-1,4)

Chapter 11

Combinatorics

11.1 Introduction

The study of combinatorics had its origin in problems involving counting. Problems such as finding the number of one-to-one mappings of a set onto itself, and finding, for a given set, the number of subsets that have some specified number of members are examples.

The above mentioned types of problems come from a class of mathematical ideas known generally as combinatorial counting. Although combinatorics today encompasses a much wider range of ideas and overlaps such studies as group theory, graph theory and topology as well as others, we will restrict our interest in this chapter to combinatorial counting. Sometimes combinatorial counting is referred to as sophisticated counting. This means that instead of counting each member of a set individually, when counting its members, it is sometimes possible to find this number more efficiently.

11.2 Counting Principle and Permutations

Example 1. Suppose that A, B and C are three cities, and you wish to travel from City A to City C by passing through City B. There are exactly three roads from A to B -- the red road, the blue road, and the yellow road. There are exactly two roads from City B to City C --

the green road and the orange road. How many ways are there to make the trip from A to C? (See Figure 11.1.)

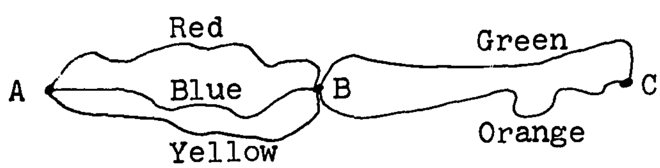


Figure 11.1

One way is to take the red road from A to B, and then the green road from B to C; we shall call this route the red-green route. All the possible routes are shown in Table 11.1.

Roads from A to B	Roads from B to C	Routes from A to C
red	green	red-green
blue	orange	red-orange
yellow		blue-green
		blue-orange
		yellow-orange
		yellow-green

The total number of routes is 6. Notice that $6 = 3 \cdot 2$, where 3 is the number of ways to make the first part of the trip, and 2 is the number of ways to make the second part of the trip.

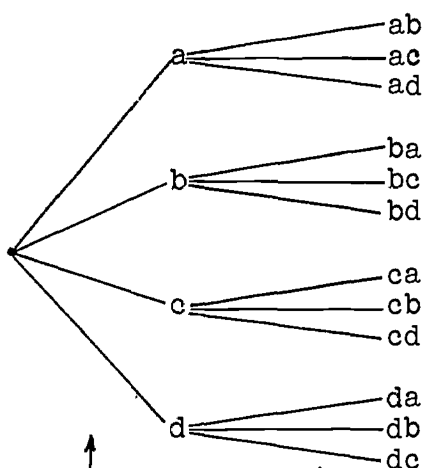
Example 2. Let S be the set $\{a,b,c,d\}$ consisting of four different letters of the alphabet. How many two-letter "words" can you make using the letters in this set? Before answering the question, we must agree to certain rules. One rule is that the "word" does not necessarily have any meaning; another rule is that a letter may not be used more than once in the same "word." Thus, while we accept "bd" as a "word" we do not accept "dd."

All possible "words" follow:
ab, ac, ad, ba, bc, bd, ca, cb, cd, da, db, dc. There is a total of 12 words, As in Example 1, there are two choices to be made in forming a word. First, choose the first letter of the word. There are 4 choices, since you may use any one of the four letters in the set. Next, choose the second letter of the word. How many choices are there in this case? Not 4, since the second letter cannot be the same as the first. Therefore, there are just 3 choices for the second letter, once the first letter has been selected. Do you see from Table 11.2 that we have the same sort of situation as we had in the earlier example about the roads?

Number of Choices for First Letters	Number of Choices for Second Letter	Total Number of Words
4	3	$12 = 4 \cdot 3$

Table 11.2

Specifically, in this case we have 12 = 4 · 3 words. The "tree" diagram, Figure 11.2 is another way to make this clear.



The total number of words is 12.

There are 4 ways to make the first choice.

For each of these, there are 3 ways to make the second choice.

Figure 11.2

The two examples just discussed illustrate a principle called the <u>counting principle</u>. It may be stated as follows:

<u>CP</u> If an activity can be accomplished in <u>r</u> ways, and after it is accomplished, a second activity can be accomplished, in <u>s</u> ways, then the two activities can be accomplished, one after the other, in r·s ways.

Example 3. Suppose in Example 2 we lift the restriction that no letter can be selected twice. If we do so we would have four ways to select the first letter, and then the second letter

could be selected in four different ways. Therefore we would have $4 \cdot 4 = 16$ distinct possible words. This result is illustrated in the tree diagram of Figure 11.3 and suggests we state a more general counting principle.

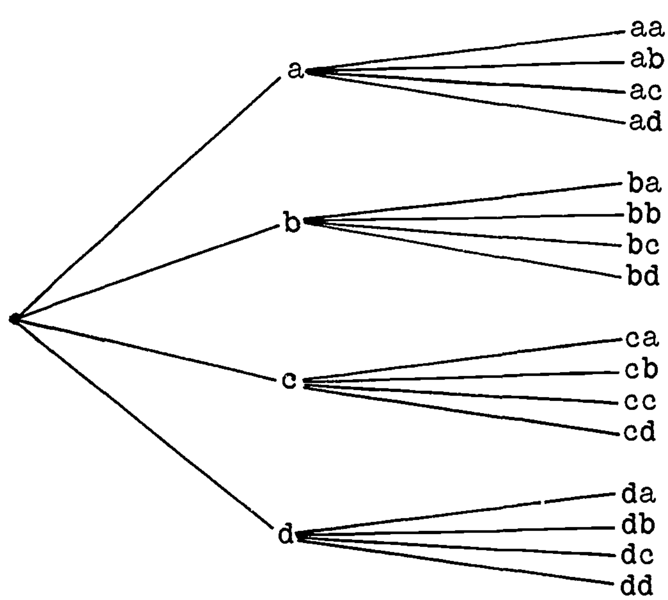


Figure 11.3

CP' Let A_1 and A_2 be sets with r_1 and r_2 elements respectively, where $r_1, r_2 \in Z^+$. Then $A_1 \times A_2 = \{(a_1, a_2): a_1 \in A_1 \text{ and } a_2 \in A_2\}$ contains $r_1 \cdot r_2$ elements.

Example 4. Given the set of letters $\{a,e,i,o,u\}$, how many two letter "words" can be formed, using the same rules as in (a) Example 2? The first letter may be chosen in 5 ways ($r_1 = 5$). The second letter may then be chosen in 4 ways ($r_2 = 4$). The total number of words is $5 \cdot 4 = r_1 \cdot r_2 = 20$ (b) Example 3? Here $r_1 = r_2 = 5$ and thus the total is $25 = 5 \cdot 5$.

One might well wonder if the counting principle CP and its generalization CP' can be extended to more than two sets A_1 and A_2. For instance suppose, in Example 4 (a) we wanted to form 3 letter words. Is the number of such words $5 \cdot 4 \cdot 3 = r_1 \cdot r_2 \cdot r_3 = 60$? Would the number in Example 4(b) be $5 \cdot 5 \cdot 5 = r_1 \cdot r_2 \cdot r_3 = 125$? The answer is yes, to both questions. Perhaps you might confirm this with a tree diagram. Suppose in Example 4 (b) we ask how many words 15 letters long can you form? Is the answer $5 \cdot 5 \cdot \ldots \cdot 5 = 5^{15} = r_1 \cdot r_2 \cdot \ldots \cdot r_{15}$. The answer again is yes. You could of course prove it by drawing a tree diagram and counting the 30,517,578,125 possible words. However to prevent you from tiring we state as Theorem 1 our general counting principle for a finite number of non-empty sets, each with a finite number of elements. The proof would require the principle of mathematical induction, which is not yet available to us. To facilitate the writing of the theorem and subsequent statements we adopt the following notation. If a set S contains s elements we will write $n(S) = s$.

<u>Theorem 1</u>. <u>CP</u> Let A_1, ---, A_k be non-empty sets and let $n(A_i) = r_i$ for $i = 1, 2, ---, k$. where each $r_i \in Z^+$. Let $A_1 \times A_2 \times --- A_k = \{(a_1, a_2, \ldots, a_k): a_i \in A_i \; i = 1, 2, \ldots, k\}$ Then $n(A_1 \times A_2 \ldots \times A_k) = r_1 \cdot r_2 \cdot --- \cdot r_k$.

Example 5. A direct mail firm plans to send out one million letters, each containing the same four pieces of literature, one piece each from the four companies this firm represents. Company A_1 has made available six different pieces of

literature, Company A_2 three pieces, Company A_3 two pieces and Company A_4 eight pieces. How many different mailings are possible? We have $n(A_1) = r_1 = 6$, $n(A_2) = r_2 = 3$, $n(A_3) = r_3 = 2$, $n(A_4) = r_4 = 8$. Therefore $n(A_1 \times A_2 \times A_3 \times A_4) = r_1 \cdot r_2 \cdot r_3 \cdot r_4 = 6 \cdot 3 \cdot 2 \cdot 8 = 288$.

Example 6. In a certain school, the student council decides to give each student an ID number consisting of a letter of the alphabet followed by two digits. Will there be enough ID numbers so that each student in the school may have one?

Let $A_1 = \{\text{all letters in the alphabet}\}$, $A_2 = A_3 = \{\text{all digits}\}$. Therefore $n(A_1) = 26$, $n(A_2) = n(A_3) = 10$. Therefore the number of ID numbers is $n(A_1 \times A_2 \times A_3) = 26 \cdot 10 \cdot 10 = 2600$.

So, unless the school has more than 2600 students, there will be enough ID numbers to go around.

In Chapter 2, Section 2.3 we defined a permutation of a set S as a one-to-one mapping of the set onto itself; and saw that if the set contains n elements, then there are $n! = n(n-1) \cdot --- 2 \cdot 1$ such permutations. In Example 7, we shall see that the counting principle may be used to get the same result.

Example 7. How many permutations are there of the set

$$S = \{a,b,c\}?$$

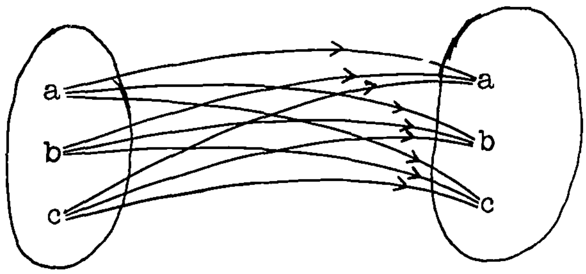


Figure 11.4

As illustrated in Figure 11.4 we may choose any one of the 3 arrows starting at *a*; that is, there are 3 choices. Next, we move to *b*. We do not have 3 choices, since we cannot choose the arrow that goes to the same image we chose before, if we want a one-to-one mapping. So, the number of choices here is 2. Next, we move to *c*. Two of the images have now been used. So here we have only 1 choice.

To summarize: At *a* we have 3 choices; at *b* we have 2 choices; at *c* we have 1 choice. The total number of one-to-one mappings is $3 \cdot 2 \cdot 1 = 6 = 3!$ In the language of our theorem, $n(A_1) = 3$, $n(A_2) = 2$, $n(A_3) = 1$, and therefore $n(A_1 \times A_2 \times A_3) = 3 \cdot 2 \cdot 1 = 3!$.

Example 8. Given the sets in Figure 11.5 how many ways are there to make a one-to-one mapping from set a to set B?

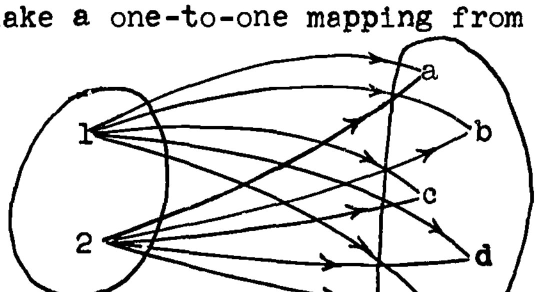


Figure 11.5

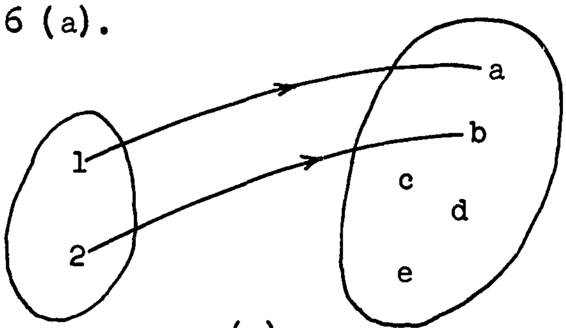
We may choose any one of the 5 arrows starting at 1; there are 5 choices. Then we may choose any one of 4 arrows, starting at 2; we cannot choose the arrow which goes to the same image as our first arrow. Therefore, the total number of one-to-one mappings from A to B is $5 \cdot 4 = 20$.

We often use the word permutation also to describe a situation such as that in example 8. Specifically, we would say that the number of permutations of 5 elements taken 2 at a time is 20. In Example 8, the 5 elements are a, b, c, d, and e. And the 20 permutations of these elements taken 2 at a time are listed in Table 11.3.

ab	ac	ad	ae
ba	bc	bd	be
ca	cb	cd	ce
da	db	dc	de
ea	eb	ec	ed

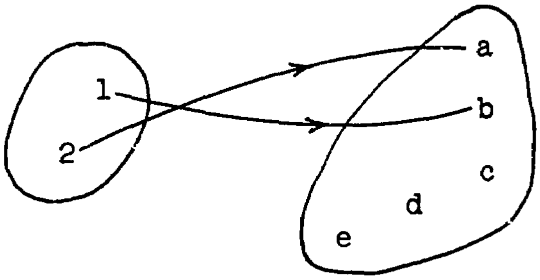
Table 11.3

Each of these, of course, corresponds to one of the 20 mappings mentioned in Example 8. For instance, "ab" refers to the mapping in Figure 11.6 (a).



(a)

Figure 11.6



(b)

Figure 11.6

On the other hand "ba" refers to the mapping in Figure 11.6 (b). Thus, "ab" and "ba" are different permutations (i.e., they are different mappings).

Example 9. What is the number of 4-letter words that can be formed from the set {a,b,c,d,e,f,g}? The number is $7 \cdot 6 \cdot 5 \cdot 4$. (Express in the language of theorem 1.) This is the number of permutations of 7 elements taken 4 at a time.

Example 10. What is the number of permuations of 10 elements taken 3 at a time?

$$10 \cdot 9 \cdot 8 = 720$$

This is the number of one-to-one mappings from a set containing 3 elements to a set containing 10 elements.

Example 11. What is the number of permutations of 5 elements taken 5 at a time?

This is the number of one-to-one mappings from set A to set B, where both A and B have 5 elements. (See Figure 11.7)

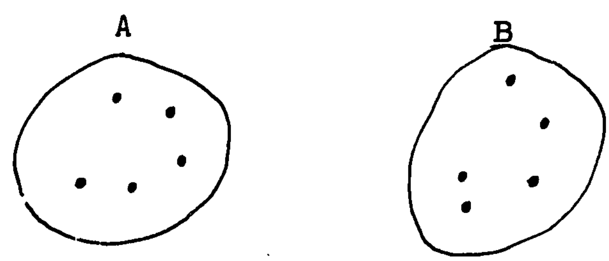


Figure 11.7

But the number of such mappings is the same as the number of mappings of A onto itself. Therefore, the answer is 5! or 120.

Example 12. Suppose you had five colored flags, one in each of the following colors: red, white, blue, green, yellow. If you agree that a given signal is to be represented by a particular arrangement of three colored flags, how many different signals could you devise using the five flags? For example, the arrangement

RED YELLOW BLUE

might mean "Help." This problem really asks for the number of one-to-one mappings from a set containing 3 elements to a set containing 5 elements. This number is:

$$5 \cdot 4 \cdot 3 = 60$$

In Examples 8 to 12 we have been considering the number of one-to-one mappings from a set A, with $\underline{r}$ members, to a set B, with $\underline{n}$ members, where $\underline{r} < \underline{n}$. Another way to describe the number of one-to-one mappings from a set with $\underline{r}$ members to a set with

n members ($r \leq n$) is the <u>number of permutations of n elements taken r at a time</u>.

We found that there were n ways of finding an image in B for which every member of A is selected first, (n - 1) ways to find an image for the second selection from A, and so on until each of the r members in A was selected and assigned an image. We then used the counting principle to compute the desired number of permutations by finding the product of r numbers starting with n;

The first factor was n.

The second factor was 1 less than n (or $n - 1$).

The 3rd factor was 2 less than n (or $n - 2$).

and so on until

the rth factor was (r-1) less than n (or $n - (r - 1)$.

In brief $n(A_k) = n - (k - 1)$, where $k = 1, \ldots, r$.

Since $n - (r - 1) = (n - r + 1)$ the product number

$n(A_1 x \cdots x A_r) = n(n - 1)(n - 2) \cdots (n - r + 1)$.

The symbol $(n)_r$ is used to represent the number of permutations of n elements taken r at a time. We write:

$$(n)_r = n(n - 1) \cdots (n - r + 1)$$

Example 13. (a) $(8)_5 = 8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 = 6720$

(b) $(4)_4 = 4 \cdot 3 \cdot 2 \cdot 1 = 4! = 24$

The exercises in Section 11.3 will contain specific examples of permutations of n elements taken r at a time. An alternative form of the general formula for $(n)_r$ will be developed in Exercise 17 of Section 11.3.

11.3 Exercises

1. Given the set of letters $\{r,s,t,u,v,w,x\}$, how many "words" can be formed having:

   (a) one letter
   (b) two letters
   (c) three letters
   (d) four letters
   (e) five letters
   (f) six letters
   (g) seven letters

2. If set B contains seven elements, how many one-to-one mappings are there from set A to set B if set A contains:

   (a) one element
   (b) two elements
   (c) three elements
   (d) four elements
   (e) five elements
   (f) six elements
   (g) seven elements

3. Use the results of Exercises 1 or 2 to answer the following:

   (a) What is $(7)_1$? (b) What is $(7)_2$? (c) What is $(7)_3$?
   (d) What is $(7)_4$? (e) What is $(7)_5$? (f) What is $(7)_6$?
   (g) What is $(7)_7$?

4. How many permutations are there of the set $\{a,b,c,d,e,f,g,h\}$? taken 5 at a time?

5. Suppose you have 5 books to put on a shelf. In how many orders can the 5 books be arranged?

6. In Exercise 5, suppose there is room for only 3 of the books on the shelf, but you may use any 3. How many arrangements are possible? That is, what is the number of permutations of 5 elements taken 3 at a time?

7. In a certain state, the license tags consist of two letters of the alphabet followed by three digits.

(a) How many different license "numbers" are possible?

(b) How many are possible if the letters O and ℓ are not used?

8. A telephone number consists of 10 digits.

(a) How many numbers are possible if there are no restrictions?

(b) How many are possible if the digit "0" cannot be used as the first digit?

(c) How many are possible if the digit "0" cannot be used as the first digit and also cannot be used as the fourth digit?

9. If a baseball team has 10 pitchers and 4 catchers, how many batteries (pitcher-catcher pairs) are possible?

10. If a girl has 5 blouses and 4 skirts, how many blouse-skirt combinations can she get?

11. If you toss one die for a first number, then toss a second die for a second number, how many results (ordered number pairs) are possible?

12. Find:

(a) $(5)_4$ (b) $(8)_3$ (c) $(8)_5$ (d) $(20)_2$ (3) $(9)_5$

13. (a) What is $(8)_3$ (b) What is $8!$?

(c) What is $(8-3)!$? (d) What is $\dfrac{8!}{(8-3)!}$?

14. What is: (a) $(6)_4$? (b) $6!$? (c) $(6-4)!$?

(d) $\dfrac{6!}{(6-4)!}$?

15. What is: (a) $(10)_3$ (b) $10\,!$ (c) $(10 - 3)\,!$

(d) $\dfrac{10!}{(10-3)!}$?

16. Let n and r be positive integers and $r \leq n$. Give an argument to justify:

$$n\,! = n(n-1)\,(n-2)\,\ldots\,(n-r+1)\quad[(n-r)]$$

17. Using the result in Exercise 16, give an argument to justify:

$$(n)_r = \frac{n\,!}{(n-r)\,!}$$

18. Use the formula in Exercise 17 to find:

(a) $(11)_3$ (b) $(7)_5$ (c) $(15)_3$ (d) $(100)_2$

19. Make up permutation problems for each of the following answers:

(a) $\dfrac{8\,!}{(8-2)\,!}$ (b) $\dfrac{9\,!}{(9-3)\,!}$ (c) $\dfrac{9\,!}{5\,!}$

(d) $\dfrac{15\,!}{13\,!}$ (e) $\dfrac{7\,!}{6\,!}$

20. Use the formula in Exercise 17 to find the number of permutations of 5 elements taken 5 at a time. Do you see that the denominator is $0\,!$? $0\,!$ has no meaning. We define $0\,! = 1$

so that the formula in Exercise 17 holds for all whole numbers n,r with $r \leq n$ without exception.

21. Find a standard name for each of the following:

(a) $\dfrac{8\,!}{(8 - 8)\,!}$ (c) $3! + 2! + 1! + 0!$

(b) $\dfrac{12\,!}{(12 - 12)\,!}$ (d) $\displaystyle\sum_{i=1}^{4} i!$

11.4 Number of Subsets of a Given Size

Before considering the number of subsets of a set S that are of a given size (here $n(s) \in Z_0^+$) we will first consider another set, the set whose elements are all of the subsets, of S. This is called the power set of S and contains the empty set, S itself, all of the one-membered subsets, two-membered subsets and so on to include every subset of S.

Definition. The power set of a set S, denoted P(S), is the set whose elements are the subsets of S. (Thus $A \in P(S)$ if and only if $A \subset S$.)

We summarize in Table 11.4. Copy and complete this table.

S	n(S)	$\mathcal{P}(S)$	$n(\mathcal{P}(S))$
$\{\} = \emptyset$	0	$\{\emptyset\}$	1
$\{a\}$	1	$\{\emptyset, \{a\}\}$	2
$\{a,b\}$	2	$\{\emptyset, \{a\}, \{b\}, \{a,b\}\}$	4
$\{a,b,c\}$	3		
$\{a,b,c,d\}$	4		

Table 11.4

Once again the counting principle is useful in the general case of finding the number of subsets of a set S with $n(S) = n$

$\in Z^+$. Selecting any subset of S can be thought of as a set of n tasks. Tasks consist of a decision for each member of S; either you select the first member or reject it, and likewise for the second, third and so on. In other words, there are two possibilities for each member of S. Then, since S has n members, the counting principle tells us that the product of n factors, each equal to 2, is the number of ways of performing these tasks one after the other. Each subset of S is the result of exactly one performance of the tasks, and each performance of the tasks results in exactly one subset of S. Accordingly the number of subsets of a set S with n elements is:

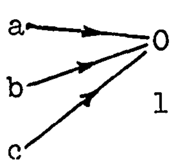
$$\underbrace{2 \cdot 2 \cdots\cdot 2}_{\text{n factors}} = 2^n$$

Is this the conclusion you drew when you completed Table 11.4? In the language of Theorem 1, for each $i \in S, i = 1,\cdots, n$ let A_i = {select, reject}. Therefore $r_1 = r_2 = \cdots = r_n = 2 = n(A_i)$. Thus $n(A_1 \times A_2 \times \cdots \times A_n) = r_1 \cdot r_2 \cdots \cdot r_n = 2^n$. If we replace the word select by the digit 1 and the word reject by the digit 0 then $A_i = \{1,0\}$ and we can reason as follows:

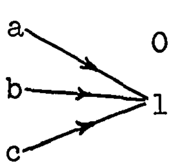
The number of elements in the power set of S is equal to the number of mappings with domain S and codomain {0,1}. The elements in S that map onto 1 are selected and those that map onto 1 are selected and those that map onto 0 are rejected for the subset selected by that particular mapping. Here we do not require that the mappings be one-to-one, nor do we require that they be onto. For example, each member of S may be mapped onto 1 and the set S itself would be the selected subset. Likewise

each member of S may be mapped onto 0 and then the empty set would be selected.

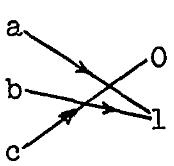
Example 1. Figure 11.8 exhibits some mappings from $\{a,b,c\}$ to $\{0,1\}$ and the sets they generate.



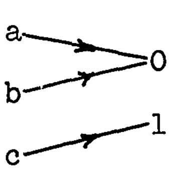
generates $\{\ \} = \emptyset$



generates $\{a,b,c\}$



generates $\{a,b\}$



generates $\{c\}$

Figure 11.8

Complete the rest of the mapping diagrams from $\{a,b,c\}$ to $\{0,1\}$ as an exercise.

We will now turn our attention to the number of subsets of S that have some given number of elements; for example the number of subsets of $\{a,b,c\}$ that have exactly two elements. From your mapping diagrams you can see that this number is 3. In general we will be concerned with the number of $\underline{r}$-member subsets of a set S with $\underline{n}$ members.

Example 2. Suppose that {a,b,c,d,e} is a set of club members. How many committees can be formed which have exactly two members? The committees are listed below:

{a,b}, {a,c}, {a,d}, {a,e}, {b,c}, {b,d},
{b,e}, {c,d}, {c,e}, {d,e}

The number in this case is 10. This question is the same as asking how many subsets of 2 elements can be formed from a set of 5 elements.

In general questions such as this may be phrased as follows: Given a set containing n elements, how many of its subsets contain exactly r elements?

In order to answer the general question, let's look again at the original question, a question whose answer we already know. Given the set {a,b,c,d,e}, how many different subsets of 2 elements can be formed? We introduce the symbol

$$\binom{5}{2}$$

to represent this number. That is, $\binom{5}{2}$ is the number of subsets of 2 elements that can be formed from a set of 5 elements.

Figure 11. 9 shows a one-to-one onto mapping from the set {1,2} to the subset {a,b}. The set {1,2} is used since

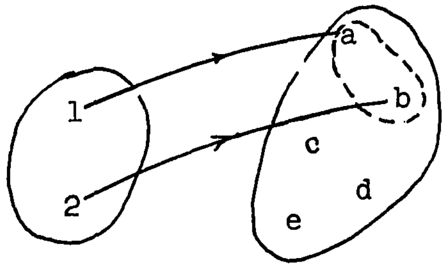


Figure 11.9

We want a subset having two elements. However, the diagram shows only one such mapping. How many one-to-one onto mappings are there from $\{1,2\}$ to the subset $\{a,b\}$? Since $\{a,b\}$ has the same number of elements as $\{1,2\}$, this is the same as the number of permutations of a set of 2 elements -- that is $2\,!$. So there are 2 different one-to-one onto mappings from $\{1,2\}$ to $\{a,b\}$. (Be sure that you can draw a diagram for each.)

Also there are $2\,!$ different one-to-one onto mappings from $\{1,2\}$ to the subset $\{a,c\}$, to the subset $\{a,d\}$, etc. In fact, there are $2\,!$ different one-to-one onto mappings from $\{1,2\}$ to <u>every</u> subset of S containing two elements. Now how many such subsets are there? We have agreed to let $\binom{5}{2}$ represent this number. Thus if we form the product

$$2!\binom{5}{2}$$

we should get the total number of ways to form a one-to-one mapping from $\{1,2\}$ to the set S. However, from CP we know this number is:

$$(5)_2$$

Therefore we have:

$$2\,!\binom{5}{2} = (5)_2$$

Then dividing by $2\,!$ we get :

$$\binom{5}{2} = \frac{(5)_2}{2\,!} = \frac{5 \cdot 4}{2 \cdot 1} = 10$$

Of course this agrees with our earlier observation that there are 10 possible subsets, each with 2 persons that can be formed from a club of 5 persons.

Example 3. Consider the problem of finding how many subsets of 3 elements can be formed from a set of 7 elements. Again, let $\binom{7}{3}$ represent this number. To find the standard name for $\binom{7}{3}$ we begin by examining the mapping of Figure 11.10.

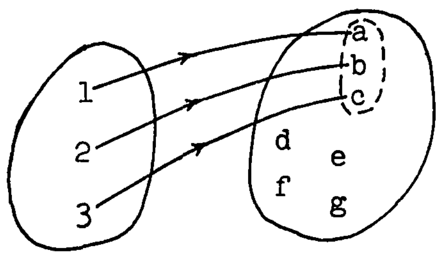


Figure 11.10

The diagram shows a one-to-one onto mapping from {1,2,3} to the subset {a,b,c}. The diagram shows only one such mapping, but there are 3! of them. (Why?) Furthermore, there are 3! different one-to-one onto mappings from {1,2,3} to every one of the $\binom{7}{3}$ subsets having 3 elements. Therefore,

$$3!\binom{7}{3} = (7)_3$$

where $(7)_3$ is obtain from the counting principle. Dividing by 3! gives,

$$\binom{7}{3} = \frac{(7)_3}{3!} = 35$$

Therefore, a set of 7 elements has 35 different 3-element subsets.

The two preceding examples suggest a perfectly general argument for finding the number of subsets having $\underline{r}$ elements that can be formed from a set having $\underline{n}$ elements, where

$r \leqq n$. Using $\binom{n}{r}$ to represent this number, we have,

<u>Theorem</u> 2.

$$r! \binom{n}{r} = (n)_r$$

Proof: Exercise

From Theorem 2, dividing by $r!$ we obtain

$$\binom{n}{r} = \frac{(n)_r}{r!} = \frac{\frac{n!}{(n-r)!}}{r!} = \frac{n!}{r!\,(n-r)!}$$

Example 4. In a club with 12 members, how many 5 member subsets are there::

$$\binom{12}{5} = \frac{(12)_5}{5!} = \frac{\frac{12!}{(12-5)!}}{5!}$$

$$= \frac{12 \cdot 11 \cdot 10 \cdot 9 \cdot 8\ (7!)}{7!\ 5!}$$

$$= \frac{12 \cdot 11 \cdot 10 \cdot 9 \cdot 8}{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}$$

$$= 792$$

Notice that in Example 4 each time you selected a subset of 5 elements from the set of 12 elements, there were 7 elements remaining that were not selected. In general, whenever you select a subset of <u>r</u> elements from a set of <u>n</u> elements there are $n - r$ elements remaining that are not selected. This means that there are just as many subsets with $n - r$ elements as there are subsets with r elements. This is expressed mathematically:

Theorem 3.

$$\binom{n}{r} = \binom{n}{n - r}$$

Proof: Exercise

Example 5. (a) Compute $\binom{7}{5}$ and $\binom{7}{2}$

(b) Did you get the same number for each of the computations in part (a)?

(c) If the answer to (b) is yes explain why. If not, do your computations again.

(d) Which of the two computations in (a) was easier? Why?

11.5 Exercises

1. In a voting body of 7 members, how many 3-man subsets are there?
2. In a voting body of 12 persons, how many 5-man subsets are there?
3. If set S has 6 elements, how many elements are in $\mathcal{P}$ (S)? How many of these subsets have exactly 3 elements?
4. Find a standard name for each of the following:

   (a) $\binom{7}{3}$ (b) $\binom{12}{5}$ (c) $\binom{6}{3}$
5. There are 8 books lying on the table, and you are to choose 3 of them.

   (a) How many ways are there to choose 3 books from 8?

(b) How many ways are there to choose the 3 books and arrange them on a shelf?

6. (a) Verify thefollowing formula for special cases of n and m (e.g.n = 5 and m = 3):

$$\binom{n}{m-1} + \binom{n}{m} = \binom{n+1}{m}$$

*(b) Now show by using the formula,

$$\binom{n}{r} = \frac{(n)_r}{r\,!}$$

that formula in 6(a) is true when $m \leq n$.

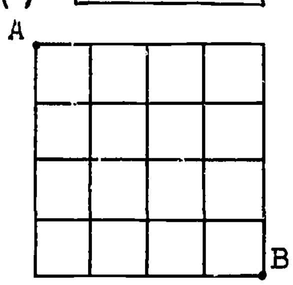
7. Use the fact that the formula in exercise 6 is true for all natural number replacements for m and n, $m \leq n$, to complete the following:

$$\binom{x-1}{y} + \binom{x-1}{y+1} = \boxed{}$$

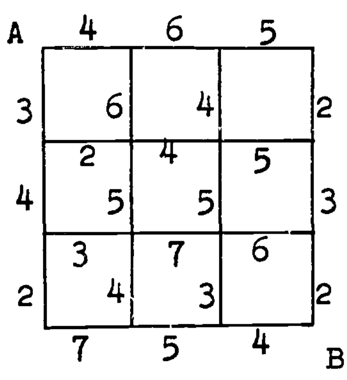
what relation must hold between x and y?

8. If n is a non-negative integer, then $\binom{n}{0} = \boxed{}$

9. If you can move only along the drawn segments down and to the right, how many paths are there from A to B? (Do this by figuring the number of paths to each point.)



10. If the numerals recorded at right indicate the length of the segments, find the shortest distance from A to B. (Travel rules are those of Exercise 9.)



11. If $\underline{n}$ is a positive integer, then $\binom{n}{1} = \boxed{}$.

12. For $\underline{n} = 4$, expand

$$\sum_{k=0}^{n} \binom{n}{k}$$ (Hint: The first two terms of the summation are $\binom{4}{0}$ and $\binom{4}{1}$.)

into a sum where each term makes use of the formula for $\binom{n}{r}$; then evaluate the sum and express the result in standard form.

*13. Prove: $$\sum_{k=0}^{n} \binom{n}{k} = 2^n$$ (Hint: For a set with n elements count the number of subsets in two different ways.)

for any positive integral $\underline{n}$.

14. If $\underline{n}$ is a non-negative integer, then $\binom{n}{n} = \boxed{}$.

15. What meaning can we give to $\binom{3}{5}$? From a set of 3 elements, how many 5-sets can be formed? Obviously there are none. Therefore, we shall define $\binom{3}{5} = 0$. What standard name would you suggest for each of the following?

(a) $\binom{2}{8}$ (b) $\binom{7}{8}$ (c) $\binom{3}{9}$ (d) $\binom{0}{4}$ (e) $\binom{1}{3}$

*16. In a deck of 52 playing cards, how many 13-card hands are possible?

17. Draw diagrams for each of the possible mappings from a set of 3 elements to a set of 2 elements. Don't restrict the mappings to one-to-one or onto.

18. Use the counting principle to suggest a way of expressing the number of mappings in Exercise 17 in exponential form.

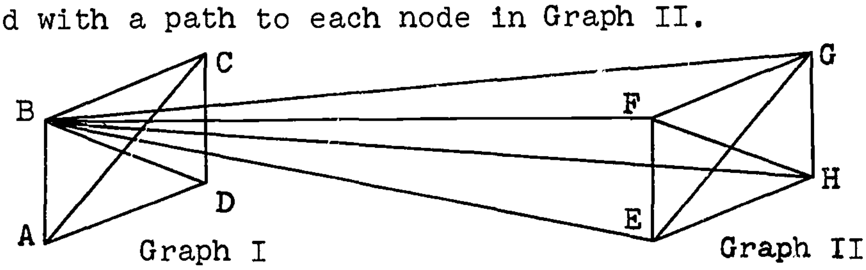
19. Use the counting principle to construct an argument that justifies the following:

The number of mappings from a set of $\underline{b}$ elements to a set of $\underline{a}$ elements is a^b.

20. In the diagram below there are two graphs each consisting of four nodes (points) and paths connecting the nodes by pairs:

(a) Explain why each graph has $\binom{4}{2}$ paths, and the total number of paths for the two graphs is $2\binom{4}{2}$.

In the next diagram node B is connected with each node in Graph II to illustrate how each node of Graph I may be connected with a path to each node in Graph II.



(b) Use the counting principle to explain why there are 16 or 4^2 paths required to connect each node of Graph I with each node of Graph II (that is to complete it).

(c) Assuming that the above graph is completed, explain why the number of paths is $\binom{8}{2}$ or $\binom{2 \cdot 4}{2}$.

(d) Use an argument concerning the above graphs to justify the statement: $2\binom{4}{2} + 4^2 = \binom{2\cdot 4}{2}$. Use computation to justify the statement.

*21. Use the graphs and explanations in Exercise 20 for this exercise.

(a) Suppose that you repeated the procedures in Exercise 20 using 5 nodes in each graph. Write the statement in Exercise 20(d) for the case of 5 nodes.

(b) Revise the statement in Exercise 20(d) for $\underline{n}$ nodes.

(c) Revise the statement in Exercise 20(d) for the case where Graph I has 6 nodes and Graph II has 4 nodes.

(d) Repeat part (c) where Graph I has $\underline{n}$ nodes and Graph II has m nodes.

*22. Show that the following statements (a) and (b) are equivalent:

(a) $2\binom{n}{2} + n^2 = \binom{2n}{2}$

(b) $n(n - 1) + n^2 = n(2n - 1)$

23. Use what you have learned in this chapter on combinatorics in addition to what you learned in the chapter on affine geometry to justify the following:

(a) If each line in the affine plane π contains k points, then π contains k^2 points.

(b) If the affine plane π contains k^2 points, then it contains $k\cdot(k + 1)$ lines.

*24. Prove Theorem 2.

25. Prove Theorem 3. (Hint: Use the formula developed from Theorem 2.)

11.6 The Binomial Theorem

Example 1. Suppose that you were given the problem of expanding the following power of a binomial:

$(a+b)^5 = (a+b)(a+b)(a+b)(a+b)(a+b)$.

After some labor you would find that the expansion of the above expression is:

$a^5 + 5a^4b + 10a^3b^2 + 10a^2b^3 + 5ab^4 + b^5$.

The symmetry of the coefficients in the above terms (1,5,10,10,5,1), and the decreasing powers of a (5,4,3,2,1,0) with the corresponding increasing powers of b (0,1,2,3,4,5) leads us to suspect that there might be a more efficient way to get the result without resorting to brute force multiplication of binomials. Note also that the sum of the exponents of a and b in each term is 5.

In this section, we are going to develop a theorem, known as the Binomial Theorem, which will be useful in expanding powers of binomials. It also has other applications in mathematics, for example, to probability theory. The development of the Binomial Theorem will make use of many ideas which you have learned such as the power set of a given set, the number of r-member subsets of a set with n elements, and the use of the symbol Σ to indicate summation.

Example 2. To illustrate the general theorem we expand $(a + b)^3$ by using the distributive property:

$$
\begin{aligned}
&(1)\quad (a+b)(a+b)(a+b) &&= a(a+b)(a+b) + b(a+b)(a+b)\\
&(2) &&= a[a(a+b) + b(a+b)] + b[a(a+b)\\
& && \quad + b(a+b)]\\
&(3) &&= a(aa + ab + ba + bb)\\
& && \quad + b(aa + ab + ba + bb)\\
&(4) &&= aaa + aab + aba + abb\\
& && \quad + baa + bab + bba + bbb\\
&(5) &&= a^3 + a^2b + a^2b + ab^2 + a^2b\\
& && \quad + ab^2 + ab^2 + b^3.\\
&(6) &&= a^3 + 3a^2b + 3ab^2 + b^3\\
&(7) &&= \binom{3}{0}a^3 + \binom{3}{1}a^2b + \binom{3}{2}ab^2\\
& && \quad + \begin{matrix}3\\3\end{matrix}\, b^3\\
&(8) &&= \sum_{r=0}^{3}\binom{3}{r} a^{3-r}b^r
\end{aligned}
$$

We can get the same result using the following combinational argument. We could get the terms in (4) directly from the left side of (1) by selecting just one of a or b from each of the the binomial factors and recording them in the order of the factors from which they were chosen. The mapping diagrams in Figure 11.11 show all the ways that this selection can be made, where 1,2 and 3 stand for the 1st, 2nd and 3rd factors respectively and the mapping is from {1,2,3} to {a,b}.

Note that the total number of mappings is $2^3 = 8$.

(CP)

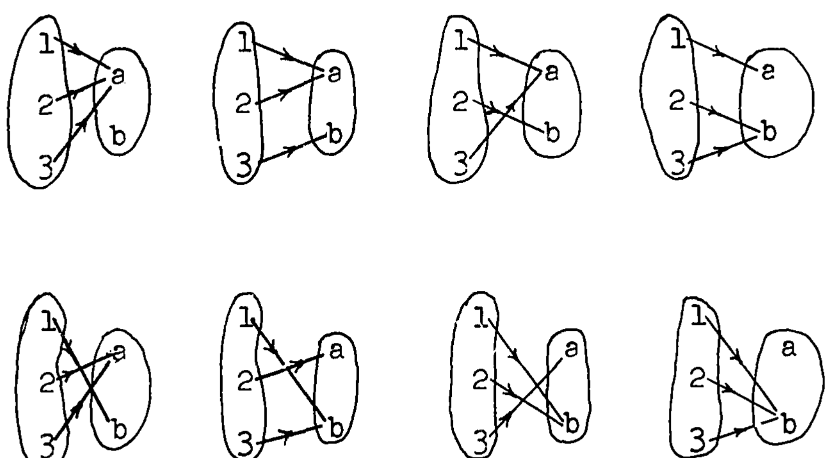


Figure 11.11

The number of times that b is selected as an image in a mapping determines the number of times that a is selected. If b is chosen r times, then a is chosen (3-r) times. Check this in the diagrams. Each mapping then is determined by the assignments of b.

The number of mappings in which

b is the image of 0 elements is 1.	$\binom{3}{0} = 1.$
b is the image of 1 element is 3.	$\binom{3}{1} = 3.$
b is the image of 2 elements is 3.	$\binom{3}{2} = 3.$
b is the image of 3 elements is 1.	$\binom{3}{3} = 1.$
	Total $8 = 2^3$

If b is the image of zero elements then a is the image of three elements, and thus the term which has coefficient $\binom{3}{0}$ is a^3.

If b is the image of one element then a is the image of two elements, and thus the term with coefficient

$\binom{3}{1}$ is a^2b.

If b is the image of two elements then we deduce as above that the term with coefficient $\binom{3}{2}$ is ab^2.

Similarly if b is the image of three elements then the term with coefficient $\binom{3}{3}$ is b^3. Multiplying each term by its coefficient and adding again yields

$$\sum_{r=0}^{3} \binom{3}{r} a^{3-r} b^r = (a+b)^3$$

You should recognize the above as a special case of ideas presented in this chapter:

(a) The number of subsets of a set with n elements is 2^n.

(b) The number of r-member subsets of a set with n elements is $\binom{n}{r}$. The binomial theorem can now be expressed.

Theorem 4. For any pair of real numbers, a and b, and any whole number n:

$$(a+b)^n = \sum_{r=0}^{n} \binom{n}{r} a^{n-r} b^r =$$

$$\binom{n}{0} a^n + \binom{n}{1} a^{n-1} b + \dots + \binom{n}{n} b^n$$

Example 3. Expand $(a+b)^5$.

$$(a+b)^n = \binom{5}{0} a^5 + \binom{5}{1} a^4 b + \binom{5}{2} a^3 b^2 + \binom{5}{3} a^2 b^3 +$$

$$\binom{5}{4} ab^4 + \binom{5}{5} b^5$$

$$= a^5 + 5a^4b + 10a^3b^2 + 10a^2b^3 + 5ab^4 + b^5$$

Example 4. Expand $(p+q)^1$

$$(p+q)^1 = \sum_{r=0}^{1} \binom{1}{r} p^{1-r} q^r = \binom{1}{0} p^1 + \binom{1}{1} q^1 = p+q$$

Example 5. Expand $(1+k)^3$.

$$(1+k)^3 = \binom{3}{0} 1^3 + \binom{3}{1} 1^2 k + \binom{3}{2} 1k^2 + \binom{3}{3} k^3$$

$$= 1 + 3k + 3k^2 + k^3$$

Example 6. Expand 1.03^4.

$$(1+.03)^4 = \binom{4}{0} 1^4 + \binom{4}{1} 1^3 (.03) + \binom{4}{2} 1^2$$

$$(.03)^2 + \binom{4}{3} 1(.03)^3 + \binom{4}{4}(.03)^4 .$$

$$= 1 + .12 + .0054 + .000108 + .00000081$$

$$= 1.12550881$$

Example 7. Expand $(a - b)^5$.

$(a - b)^5 = (a + (-b))^5$. Then apply example 3.

11.7 Exercises

1. Show that $(a+b)^2 = a^2 + 2ab + b^2$ is correct when $a = 3$ and $b = 2$.

2. Show that $(x+y)^3 = x^3 + 3x^2y + 3xy^2 + y^3$ is correct when $x = 1$ and $y = 2$.

3. Expand the following:

   (a) $(a+b)^4$ (b) $(x+y)^6$

   (c) $(c+d)^7$ (d) $(a+b)^{10}$

4. $(a-b)^2 = (a + (-b))^2 = a^2 + 2a(-b) + (-b)^2 = a^2 - 2ab + b^2$

   Using a similar approach, expand the following:

   (a) $(a-b)^3$ (b) $(x-y)^4$

   (c) $(a-b)^5$ (d) $(x-y)^6$

5. The coefficients in the expansion of $(a+b)^n$ are as follows:

   1 11 55 165 330 462 462 330 165 55 11 1

   What is n?
6. Expand $(x+1)^3$.
7. Expand $(x-1)^3$.
8. Expand $(x+2)^4$.
9. Expand $(x-2)^4$.
10. Expand $(1+1)^n$ to show that it equals $\sum_{k=0}^{n} \binom{n}{k} = 2^n$.
11. Use the binomial expansion to find $(1.01)^5$; also $(.99)^5$.

*12. Show that $(1+x)^n \geq 1 + nx$, for $x > 0$ and $n \in Z^+$.

*13. Use the combinational argument to prove $(a+b)^6 = \sum_{r=0}^{6} \binom{6}{r} a^{6-r} b^r$.

11.8 Summary

1. The counting principle was illustrated for two and three finite sets and stated as a theorem for any finite number of sets.
2. If a set A contains a elements and set B contains b elements ($a \leq b$), the number of different one-to-one mappings from A to B is called the number of permutations of b elements taken a at a time (a and b are whole numbers)

   If $a = b$, then the number of permutations is $b!$.

   If $a < b$, then the number of permutations is $b(b-1)(b-2)\ldots(b-a+1)$.

3. 0 ! is defined to be 1.
4. $\binom{n}{r}$ represents the number of subsets with r elements which can be formed from a set of n elements, where n and r are whole numbers.

   If $n < r$, then $\binom{n}{r} = 0$.

   If $n = r$, then $\binom{n}{r} = 1$.

   If $r = 0$, then for any n, $\binom{n}{r} = 1$. In general $\binom{n}{r} = \frac{(n)_r}{r!}$, for $n \geq r$.

5. (The Binomial Theorem). If a and b are real numbers and n is a whole number then

$$(a+b)^n = \binom{n}{0} a^n + \binom{n}{1} a^{n-1}b + \dots + \binom{n}{n} b^n$$

$$= \sum_{r=0}^{n} \binom{n}{r} a^{n-r}b^r.$$

11.9 Review Exercises

1. How many six-letter "words" can be formed from the set $\{t,h,e,o,r,y\}$ if
   (a) letters may not be repeated?
   (b) letters may be repeated?
2. A man conducts a probability experiment in which he does the following three things: he draws a marble from a bag of five differently colored marbles and records its color; then he tosses a die, recording the number the die shows; then he tosses a coin, recording the result "head" or "tail." How many possible outcomes are there in this experiment?

*3. In Exercise 2, what is the probability he will get an even number and a head?

4. If the call letters of a radio station must begin with "W" and contain three other letters (repetitions allowed) how many such arrangements of letters are there?

5. What is the answer to Exercise 4 if the call letters may begin with either "W" or "K"?

6. A person wishes to select 2 books from a set of 6 books. How many possible selections are there?

7. There are 5 points in a plane, no three of them in a straight line. How many lines can be drawn, with each line passing through exactly 2 of the points?

8. How many ways are there to arrange 3 books on a shelf if you have 8 books to choose from?

9. How many possible committees of 3 are there in a class of 8 persons?

10. Draw a "tree" diagram showing all the 2-letter words (no repetition) which can be formed from the set {a,e,i,o,u}. (See Section 11.2.)

11. If, from a set of 7 mathematics books and 5 history books, you must choose 1 mathematics book and 1 history book, in how many ways can you make your choice?

12. How many fractions can be formed having a numerator greater than 0 and less than 10, and a denominator greater than 0 and less than 15?

13. How many 3-digit numbers are there? (There are 10 digits to choose from, but the first digit cannot be 0.)

14. Referring to Exercise 13:

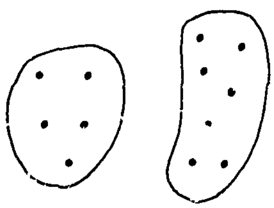
(a) How many 3-digit numbers have no two digits alike?

(b) How many 3-digit numbers have 3 digits alike?

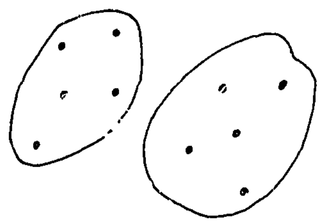
*(c) How many 3-digit numbers have exactly 2 digits alike?

15. For each of the following, tell how many one-to-one mappings are possible from set A to set B.

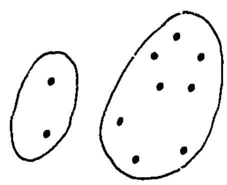
(a)



(b)



(c)



16. What is the number of permutations of 8 elements taken 2 at a time?

17. What is the number of permutations of 10 elements taken 6 at a time?

18. A set S has 10 elements.

(a) How many of its subsets have exactly 3 elements?

(b) How many of its subsets have exactly 7 elements?

(c) How many of its subsets have exactly 10 elements?

(d) How many of its subsets have exactly 1 element?

(e) How many of its subsets have exactly 0 elements?

19. Find a standard name for each of the following:

(a) $\binom{9}{2}$ (b) $\binom{11}{8}$ (c) $\binom{7}{6}$ (d) $\binom{6}{7}$ (e) $\binom{16}{0}$

20. A student is instructed to answer any 8 of 10 questions on a test. How many different ways are there for him to choose the questions he answers?

21. A basketball squad consists of four centers, five forwards, and six guards. How many different teams may the coach form if players can be used only at their one position? (A basketball team consists of 1 center, 2 forwards and two guards.)

22. A sample of five light bulbs is to be taken from a set of 100 bulbs. How many different samples may be formed?

23. Complete the following: $\binom{8}{6} + \binom{8}{7} = \boxed{}$

24. Expand $(a+b)^4$.

25. Expand $(a-b)^4$.

26. Write the first 6 terms in the expansion of $(a+b)^n$, where $\underline{n}$ is a positive integer greater than 6.

*27. Expand $(2u + v)^6$.

APPENDIX A

MASS POINTS

You have studied geometry from a number of viewpoints during Course I and Course II. First there were mappings on a line, then lattice points. This geometric study was enlarged by considering segments, angles, and isometries of the plane and transformation geometry of translations, reflections, rotations and dilations. Next you had an introduction to axiomatic affine geometry, followed by a more formal study of transformations using coordinate geometry. This appendix gives yet another kind of geometry, combining numbers and ratios with points.

A.1 Mass Points

What is a mass point? We get our initial ideas of such an object by looking at physical examples in the world around us. For example, a girl poised at the end of a see-saw, the earth at a particular point in its orbit, a carbon atom at a particular position inside a complicated molecule.

To establish something of the essential nature of each of these interpretations, we note that in each case a number and a position can be associated. For the girl it could be her weight and her position on the see-saw. For the earth it could be its mass and its position in orbit. For the carbon atom it could be a number, perhaps its electrical charge, and its location. of these cases has the property that a *number* and a *point*

are associated. This is what we mean by a mass point.

Definition 1. A mass point is an ordered pair consisting of a positive number and a point.

Can you find additional illustrations of mass points?

As you see, different physical interpretations have some properties in common and some that differ. Faced with such a situation a mathematician lists what he thinks are the basic properties common to all and proceeds to make deductions from this list. The basic property statements are called axioms. Those that are deduced are called theorems.

A.2 Notations and Procedures

First, it is convenient to have a concise way of referring to a mass point. The mass point with number 4 at point A could be written (4,A), since it is an ordered pair. But we find it convenient to designate it "4A," keeping in mind that this does not mean 4 times A, but represents the ordered pair (4,A). In general the mass point with number a at point P will be designated "aP." If in the course of deductions we conclude that aP = bQ, this will mean two things: a and b name the same number, and P and Q name the same point; that is, a = b and P = Q. If A and B name different points then 3A = 3B must necessarily be false; and also 4A = 2A must be false since $4 \neq 2$. We sometimes refer to the number of a mass point as its weight, from the idea of a girl at one end of a see-saw.

We now have a set of objects, mass points, very much as we had elements in clock arithmetic and in various operational sys-

tems. The question arises whether we can define some operations on these elements. What could we mean by "adding" two mass points?

We get an idea by examining the see-saw illustration. Suppose in Figure A.1 two weights are placed in the positions shown.

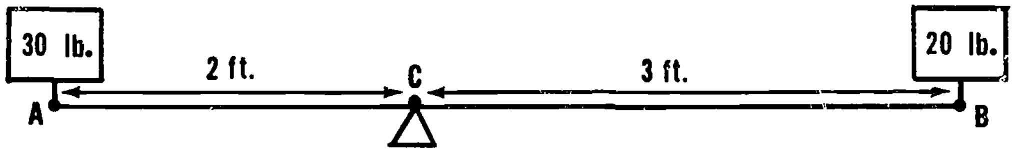


Figure A.1

It is an experimentally verifiable fact that they will balance at the point shown in the diagram; that is, if the weight of one object in pounds multiplied by its distance to the balancing point in feet is equal to the product of the weight of the other object in pounds and its distance to the balancing point in feet. In our example the first product is 30 x 2, the second product is 20 x 3 and they are equal.

This suggests what we mean by <u>adding</u> two mass points. (Don't confuse this with adding two numbers!) We will illustrate addition for two mass points as follows:

Suppose 3A and 2B are two mass points, points A and B, as in Figure A.2.

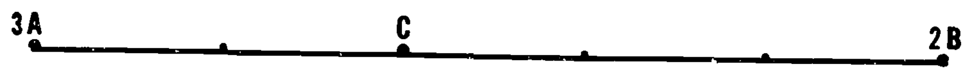


Figure A.2

To add them and to represent $3A + 2B$ as a single mass point we will do two things:

(1) Add the weights 3 and 2; $3 + 2$ or 5 will be the <u>weight</u> of $3A + 2B$.

(2) Find point C in $\overline{AB}$ such that $AC:CB = 2:3$. (Note the reversal of 3 and 2 in the ratio 2:3.) If on measuring $\overline{AB}$ we find its measure in inches to be 5, then $AC = \frac{2}{5} \cdot 5 = 2$ and $CB = \frac{3}{5} \cdot 5 = 3$. C is therefore two inches from A and 3 inches from B. C is the <u>point</u> in $3A + 2B$. Thus $3A + 2B$ has weight 5 and is at C, or $3A + 2B = 5C$. The sum is represented in Figure A.3 as follows:

Figure A.3

(The equally spaced marks should help you to see that $AC = 2$ and $CB = 3$.)

We call C the <u>center</u> <u>of</u> <u>mass</u> of the mass points at A and B. (In physics such a "balancing" point is also called the center of mass of the masses at A and B.)

Consider a second illustration (Figure A.4).

Figure A.4

Suppose the measure of $\overline{QP}$ in yards is 4. As in the first illustration we find the weight of $4Q + 3P$ to be 7. If R is the center of mass then $QR\colon RP = 3\colon 4$; that is $QR = \frac{3}{7}\cdot 4$ or $\frac{12}{7}$ and $RP = \frac{4}{7}\cdot 4$ or $\frac{16}{7}$. Thus $QR = 1\frac{5}{7}$ and we can approximate the location of R with a yardstick.

Returning to our first example of the see-saw, we would find the <u>sum</u> of the two mass points to be $30A + 20B$, for which $AB = 5$. The point C, the center of mass will be $\frac{20}{50}\cdot 5$ feet from A toward B. This is the point at which the see-saw will balance

for weights of 30 and 20 pounds respectively, at A and B. The sum is 50C.

In this way we are led to define the sum of two mass points.

<u>Definition</u> 2. If A and B are two points and a, b positive numbers, then by aA + bB (the sum of mass points aA and bB), we mean the mass point cC such that a + b = c, and C is the point in $\overline{AB}$ such that AC: CB = b: a. We then write aA + bB = cC. Furthermore, aA + bA = (a + b)A.

The definition aA + bA = (a + b)A (for example 4A + 3A = 7A), turns out to be the most useful way of having aA + bB defined for <u>all</u> possible mass points aA, bB.

We emphasize that C is in $\overline{AB}$. Furthermore, we might guess that each interior point of $\overline{AB}$ (that is, a point of $\overline{AB}$ distinct from A and B) can be determined by a correct choice of a and b. Thus, whenever we add two mass points, the center of mass of the sum will be found in the segment determined by the mass point addends.

We have defined the addition of two mass points. The question arises whether we can add three or more mass points. We will explore the addition of three or more mass points on the same line in the exercises below. In Section A.4 we will investigate the addition of non-collinear mass points in the plane, and in Section A.12 the addition of non-coplanar mass points in space.

A.3 <u>Exercises</u>

1. In each part below you are given the length in inches of

a segment for which you are to draw a diagram. On this diagram represent the sum of the two mass points at a single point.

(a) AB = 6, 5A + 1B

(b) AB = 6, 1A + 5B

(c) CD = 3, 2C + 1D

(d) CD = 3, 1C + 2D

(e) EF = 5, 1E + 1F

(f) GH = 3, 2G + 4H

(g) GH = 3, 3G + 2H

(h) KL = 5, 2K + 4L

(i) KL = 5, 1K + 2L

(j) KL = 5, $1\frac{1}{2}$K + 1L

(k) AB = 7, 3A + 4B

(l) CD = 10, 2C + 3D

(m) EF = 15, 5E + 2F

(n) GH = 7, 2G + 4H

(o) KL = 6, 5K + 4L

2. (a) You are given mass points 3A and 4B, where A and B are distinct points. Is their center of mass nearer to A or to B? Try to answer without calculating the position c the center of mass.

(b) Answer the same question for mass points 8A and 5B.

(c) Is the center of mass nearer the point with the greater or lesser weight?

3. For each of the following compute AG:GB, if A $\neq$ B.

(a) $3A + 2B = 5G$

(b) $1A + 6B = 7G$

(c) $2A + 1B = 3G$

(d) $5A + 5B = 10G$

4. In this exercise you are given one of two mass points and the sum. You are to find the other mass point. To illustrate, suppose xX is the missing mass point and $3A + xX = 5B$. Thus $3 + x = 5$, from which we deduce $x = 2$. (The weights of $3A$ and xX are 3 and 2.) B is the point in $\overline{AX}$ such that $AB : BX = 2 : 3$ and X is in AB with B in between A and X, and with $BX = \frac{3}{2}AB$, as shown below.

Solve for x and locate X from each of the following equations.

(a) $3A + xX = 4B$

(b) $4A + xX = 6B$

(c) $xX + 4A = 6B$

(d) $1A + xX = 3B$

(e) $2A + xX = 3B$

(f) $xX + 9A = 12B$

5. Suppose $12A + bB = cC$. What must be true about b and c in the following cases?

(a) C is the midpoint of $\overline{AB}$

(b) C is the trisection point of $\overline{AB}$ nearer A.

(c) C is the trisection point of $\overline{AB}$ nearer B.

(d) C is the point of division of $\overline{AB}$ such that $AC : CB = 3 : 4$.

6. Let weight 3 be assigned to A in $\overline{AB}$.

   (a) If C is the midpoint of AB, what weights should one assign to A and B so that C is then the center of mass?

   (b) If C is the trisection point of $\overline{AB}$ nearer B, what weight should one assign to A and B so that C is the center of mass?

7. Given segment $\overline{AB}$ and point C in it, so that C is the center of mass.

   (a) If AC: CB = 2: 3 and C has weight 5, what weights should be assigned to A and B?

   (b) If AC: CB = 2: 3 and C has weight 7, what weights should be assigned to A and B?

   (c) If AC: CB = 3: 4 and C has weight 10, what weights should be assigned to A and B?

   ★(d) If AC: CB = x: y and C has weight 5, what weights should be assigned to A and B?

   ★(e) If AC: CB = x: y and C has weight z, what weights should be assigned to A and B?

8. Draw a segment $\overline{AB}$ 3 inches long and take C in $\overline{AB}$ such that $\overline{AC}$ is $\frac{1}{2}$ inch long.

   (a) Represent 1A + 2B at one point. Name it D.

   (b) Represent 3D + 3C at one point. Name it E.

   (c) Represent 2B + 3C at one point. Name it F.

   (d) Represent 1A + 5F at one point. Name it G.

   (e) Do E and G name the same point?

   (f) If so, how does this exercise show

$$(1A + 2B) + 3C = 1A + (2B + 3C)$$

*9. Given mass points 2A, 3B, 5C on a line, with B between A and C.

(a) Represent $2A + 3B$ at one point. Call it D.

(b) Represent $5D + 5C$ at one point. Call it E.

(c) Represent $3B + 5C$ at one point. Call it F.

(d) Represent $2A + 8F$ at one point. Call it G.

(e) Prove that E and G name the same point.

(f) What does the proof in (e) show?

A.4 Axioms for Mass Points

We now investigate the system $(M,+)$, where M is the set of mass points and + denotes mass point addition, to see if it is an operational system. The basic requirement is that the sum of two mass points be a unique mass point. Otherwise such a sum as $5A + 6B$ may be assigned more than one mass point, and any computation with mass points would become impossible. We know that $5A + 6B$ must have the unique weight $5 + 6$ or 11. But is there exactly one location for the center of mass? From our definition $5A + 6B = 11C$, where C is a point in $\overline{AB}$ such that $AC:CB = 6:5$. Stated another way, C is on $\overline{AB}$, $\frac{6}{11}$ of the way from A to B. Since it seems clear that there is one and only one point of $\overline{AB}$ which is $\frac{6}{11}$ of the way from A to B, we are led to conclude that there is exactly one location for the center of mass. Note in the special case $aA + bB = (a + b)A$, the point A is assigned as the center of mass for mass points aA and bA. We thus take as our first axiom

Axiom 1. For any mass points aA and bB there is exactly one mass point cC such that aA + bB = cC.

In effect we are saying that the set M of mass points, with the operation of addition of mass points defined above, is an operational system (M,+).

Our construction of aA + bB leads us to accept that aA + bB = bB + aA. We will state this property as an assumption and call it the Commutativity Axiom.

Axiom 2. For any mass points aA and bB, aA + bB = bB + aA.

With these two assumptions (M,+) is beginning to look like some other operational systems we know. Another characteristic of these other systems was the associative property. Before we can raise that question here we need to explore what we mean by the addition of three mass points.

In Section A.3 Exercises 8 and 9 we examined the problem of adding aA, bB, and cC when A, B, and C are collinear. We even found it plausible that

$$(aA + bB) + cC = aA + (bB + cC).$$

We now examine the problem of adding aA, bB, and cC when A, B, and C are now collinear.

Suppose we have three mass points aA, bB, and cC, with A, B, C not collinear, as in the diagram below.

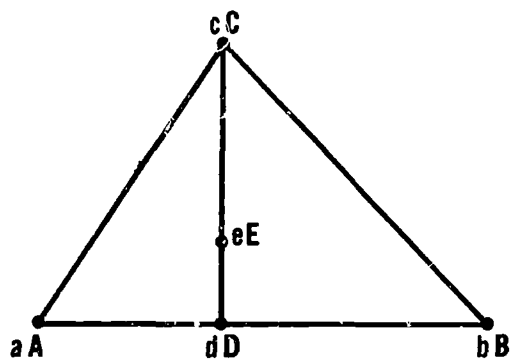


Figure A.5

We can find the sum aA + bB, and we know that we would get a mass point, say dD, with D in $\overline{AB}$. Now we can obtain dD + cC, and this would give us a mass point eE, with E in $\overline{CD}$. So the sum of the three mass points, constructed as (aA + bB) + cC would give us a mass point eE, with E in the interior of the triangle ABC. What would happen if we considered the sum aA + (bB + cC)? It is reasonable that we would again get a mass point, say fF, with F an interior point in triangle ABC. But would it be the same mass point as eE?

We shall perform an experiment. We want to see for instance whether (3A + 2B) + 1C = 3A + (2B + 1C), where A, B, C are points, not necessarily collinear, as shown in Figure A.6.

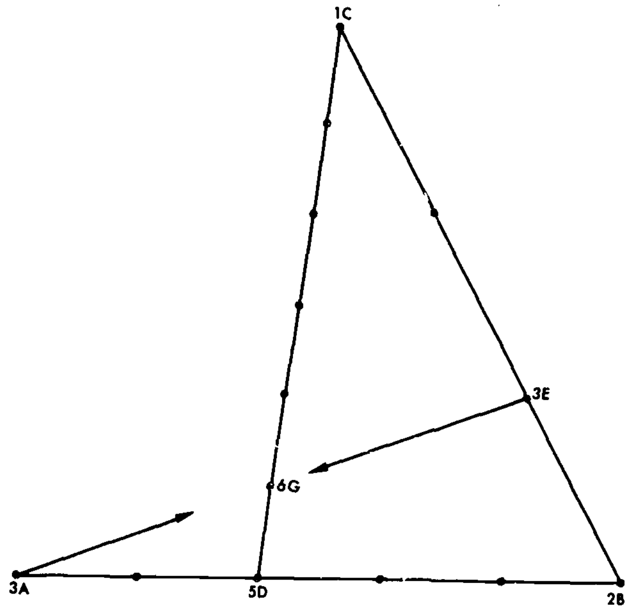


Figure A.6

To facilitate this experiment we have subdivided $\overline{AB}$ into 5 segments of the same length and $\overline{BC}$ into 3 segments of the same length.

First we find 3A + 2B to be 5D, where D is in $\overline{AB}$, with AD: DB = 2: 3 as shown in Figure A.6. Then subdividing $\overline{DC}$ into 6 segments of the same length we see (again in Figure A.6) that 5D + 1C = 6G, where G is on $\overline{DC}$, with DG: GE = 1: 5.

On the other hand we first find 2B + 1C, and find it to be 3E where E is on $\overline{BC}$, with BE: EC = 1: 2 (see diagram). We have only to test whether 3A + 3E = 6G. To convince ourselves that this is true, or false, we place our ruler on $\overline{AE}$ and see whether G is in $\overline{AE}$ with AG: GE = 3: 3. A test shows it to be true. Try it. Note that this experiment gave us

$$(3A + 2B) + 1C = 5D + 1C = 6G,$$

and

$$3A + (2B + 1C) = 3A + 3E = 6G.$$

This experiment and our experience with collinear points lead us to state the Associativity Axiom:

<u>Axiom</u> 3. For all mass points aA, bB, and cC

$$(aA + bB) + cC = aA + (bB + cC).$$

The axiom means that aA + bB + cC represents the same mass point no matter how we associate the individual mass points. This mass point has weight a + b + c and its point is called the center of mass of the three mass points at A, B and C.

We have not proved the associativity axiom. We have not deduced it. The purpose of the experiment was not to prove the axiom. It was to make it easier to accept it as an axiom. (Mathematicians may even accept as axioms statements which cannot be experimentally tested as being either true or false.)

In adding mass points we are also adding positive numbers. It should be understood that we are allowing ourselves to use those properties of (Q,+) which we need.

A.5 <u>Exercises</u>

1. Make an exact copy of the three mass points 3A, 2B and 1C used in the experiment above. (See Figure A.6.) Show by experiment that 3A + 2B + 1C can also be found by any of the following procedures:

   (a) Find 2B + 1C first; then (2B + 1C) + 3A.

   (b) Find 3A + 1C first; then (3A + 1C) + 2B.

2. Justify each of the following statements by citing the appropriate axiom or axioms:

   (a) $(2B + 1C) + 3A = (1C + 2B) + 3A$

   (b) $(2B + 1C) + 3A = 1C + (2B + 3A)$

   (c) $2B + 3A + 1C = 3A + 2B + 1C$

3. Represent $aA + bB + cC$ in 6 different ways.

4. Make a diagram which shows $2A + 1B + 2C$ at a single point. Take A, B, C as any three noncollinear points.

A.6 A Theorem

As you recall, we called a statement that is deduced from other statements a theorem. Our first theorem for mass points, is about any triangle and it may come to you as a surprise. Suppose the triangle is ABC. Let D be the midpoint of $\overline{AB}$, E the midpoint of $\overline{BC}$ and F the midpoint of $\overline{CA}$. Make such a diagram and draw $\overline{CD}$, $\overline{BF}$ and $\overline{AE}$. Do they meet in one point? We shall prove that they do; that is, we shall deduce this from our axioms. To make it easier to talk about the segments $\overline{CD}$, $\overline{BF}$, and $\overline{AE}$, we shall call them medians.

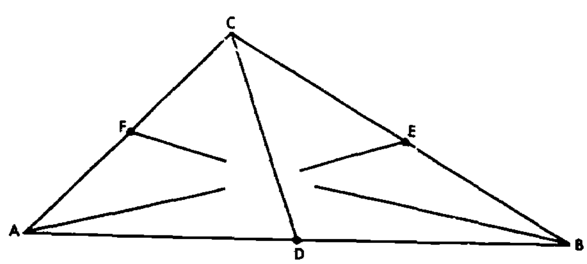


Figure A.7

<u>Definition</u> 3. A segment is a median of a triangle if it connects a vertex to the midpoint of the side opposite the vertex.

<u>Theorem</u> 1. The three medians of a triangle meet in one point.

To prove this theorem let us start by assigning weights to vertices, thus converting them to mass points. Let us assign 1 to A, 1 to B and also 1 to C. (You will see why we choose 1 as the weight of each point as the proof develops.) We remind you that D is the midpoint of $\overline{AB}$, E is the midpoint of $\overline{BC}$, and F is the midpoint of $\overline{CA}$. (See Figure A.7.)

By the Associativity axiom $(1A + 1B) + 1C = 1A + (1B + 1C)$. Let us first calculate $(1A + 1B) + 1C$. The mass of $1A + 1B$ is clearly 2, and the point in $1A + 1B$ is the point in $\overline{AB}$, $\frac{1}{2}$ the way from A to B. This is, of course, the midpoint D of $\overline{AB}$. Thus $1A + 1B = 2D$, so that $(1A + 1B) + 1C = 2D + 1C$. By the same reasoning, the mass of $2D + 1C$ is 3, and the point of $2D + 1C$ is a point G in $\overline{DC}$, with $DG{:}GC = 1{:}2$. Thus $2D + 1C = 3G$. In summary

$$(1A + 1B) + 1C = 2D + 1C = 3G,$$

and G divides $\overline{DC}$ in the ratio 1:2 from D to C.

Now we calculate $1A + (1B + 1C)$. First, the point of $1B + 1C$ is the midpoint E of $\overline{BC}$, so that $1B + 1C = 2E$. Then, $1A + (1B + 1C) = 1A + 2E$. The point of $1A + 2E$ is the point H in $\overline{AE}$ such that $AH{:}HE = 2{:}1$. Thus

$$1A + (1B + 1C) = 1A + 2E = 3H,$$

and H divides $\overline{AE}$ in the ratio 2:1 from A to E.

But by the Associativity Axiom,

$$(1A + 1B) + 1C = 1A + (1B + 1C), \text{ or } 3G = 3H,$$

and we conclude $G = H$. Thus $AG:GE = 2:1$. (Why?)

Now we calculate $(1A + 1C) + 1B$. As above $1A + 1C = 2F$, where F is the midpoint of $\overline{AC}$.

$$(1A + 1C) + 1B = 2F + 1B = 3K$$

where K is the point in $\overline{BF}$ that divides $\overline{BF}$ in the ratio 2:1 from B to F.

On the other hand

$$\begin{aligned} (1A + 1C) + 1B &= 1A + (1C + 1B) && \text{by Axiom 3} \\ &= 1A + (1B + 1C) && \text{by Axiom 2} \\ &= (1A + 1B) + 1C && \text{by Axiom 3} \end{aligned}$$

But, by above, $(1A + 1B) + 1C = 3G$. Thus $3G = 3K$, and we conclude $G = K$, so that G is also in $\overline{BF}$, and $BG:GF = 2:1$.

We have not only proved that the three medians meet in a point (the point G), but that this point divides each median in the ratio 2:1 from vertex to midpoint of opposite side.

We can also use the axioms to solve problems. This means we will discover other theorems. Since we won't find it necessary to use these theorems in proving others, we will not list them formally as theorems. We consider them only as exercises.

Suppose in $\triangle ABC$, D divides $\overline{BC}$ in the ratio 1:2 from B to C, and E divides $\overline{AC}$ in the ratio 1:1. (See Figure A.8.) Let $\overline{AD}$ intersect $\overline{BE}$ in G. What are the numerical values of $DG:GA$ and $BG:GE$?

We can solve this problem as follows. In order that D may be the trisection point of $\overline{BC}$ nearer B, we assign the weights 2 to B and 1 to C. Then 2B + 1C = 3D. In order that E be the midpoint of $\overline{CA}$ we assign the same weight to A as to C. Having assigned 1 to C we assign 1 to A also. Then 1C + 1A = 2E. The point of (2B + 1C) + 1A is the same as the point 2B + (1C + 1A); that is, the point of 3D + 1A is the same as the point of 2B + 2E. This point is on $\overline{AD}$ and $\overline{BE}$; that is, this point is the intersection of $\overline{AD}$ and $\overline{BE}$, and it is named G. Therefore (2B + 1C) + 1A = 3D + 1A = 4G, and thus DG: GA = 1: 3. Also 2B + (1C + 1A) = 2B + 2E = 4G, and thus BG: GE = 1: 1.

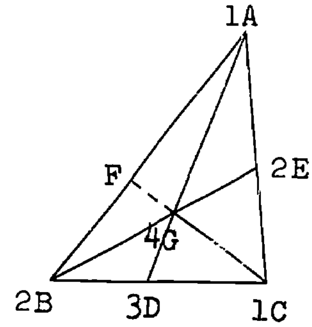


Figure A.8

We can extend our discoveries in this problem. Let CG $\cap$ $\overline{AB}$ = F. By Axioms 2 and 3, (2B + 1A) + 1C = 4G. Now 2B + 1A = 3H, where H is in $\overline{BA}$. But

$$4G = (2B + 1A) + 1C = 3H + 1C$$

implies G is in $\overline{HC}$, so that H is in CG. We thus have that H is in $\overline{BA}$ and in CG, and therefore H = F. Thus 2B + 1A = 3F and BF: FA = 1: 2. From 3F + 1C = 4G, it follows that FG: GC = 1: 3.

If we omit explanations, the solution of the above problem

can be written briefly as follows:

1. $2B + 1C = 3D$ and $3D + 1A = 4G$. Therefore $DG:GA = 1:3$.

2. $1C + 1A = 2E$ and $2B + 2E = 4G$. Therefore $BG:GE = 1:1$.

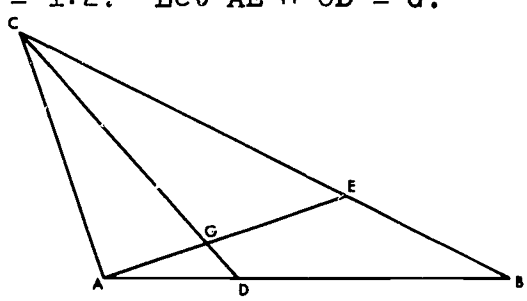
3. $2B + 1A = 3F$. Therefore $BF:FA = 1:2$.

4. $3F + 1C = 4G$. Therefore $FG:GC = 1:3$.

A.7 <u>Exercises</u>

1. Review the proof of the theorem about the medians of a triangle, then tell whether you think the proof applies only to the triangle represented in the diagram or to all triangles.
2. This is an experimental exercise. Draw any triangle, locate the midpoint of each side and draw the medians. In your diagram, do the medians meet at one point? Suppose they did not, or they did not in a drawing by your classmate. Try to find why the drawing does not agree with the theorem.
3. The lengths of the medians of a triangle are 15, 12, and 18 inches long. How long are the segments into which each median is divided by the point in which they meet?
4. Answer the question in Exercise 3 if the medians are 12, 13, and 14 inches long.

5. In $\triangle ABC$, D is in $\overline{AB}$ and AD:DB = 1:2. E is in $\overline{BC}$ and BE:EC = 1:2. Let $\overline{AE} \cap \overline{CD} = G$.



Prove that AG:GE = 3:4, and that CG:GD = 6:1.

(Hint: Assign weight 4 to A, 2 to B and 1 to C.)

6. Using the data in Exercise 5, let $BG \cap \overline{CA} = F$ and find the numerical value of BG:GF and AF:FC.

*7. Add to the data in Exercise 5 that K is in $\overline{CA}$ and CK:KA = 1:2. Let $\overline{BK} \cap \overline{AE} = L$ and $\overline{BK} \cap \overline{CD} = M$. Prove: BL = LM = 3MK (This is a difficult exercise.)

A.8 **Another** Theorem

Our definition for addition of mass points applies to pairs of mass points. In other words, addition is a binary operation. To make it possible to add three mass points we introduce the Associativity Axiom, which says that aA + bB + cC can be found by either finding (aA + bB) first or (bB + cC) first. Either of these sums can be found and then a second addition completes the calculation by which aA + bB + cC is expressed as a mass point with one weight and one point. For our next theorem we need to know how to add four mass points. This can be done by

a repeated application of the Associativity Axiom as follows:

$$aA + bB + cC + dD = (aA + bB) + (cC + dD).$$

There are also other ways to associate. For instance, $aA + (bB + cC) + dD$. This reduces the addition from four to three mass points. We now prove a second theorem.

<u>Theorem</u> 2. The segments joining the midpoints of opposite sides of a quadrilateral bisect each other.

<u>Proof</u>. Let ABCD be the quadrilateral (Figure A.9) and let E be the midpoint of $\overline{AB}$, F the midpoint of $\overline{BC}$, G the midpoint of $\overline{CD}$, and H the midpoint of $\overline{DA}$. We have to prove that $\overline{EG}$ bisects $\overline{HF}$, and that $\overline{HF}$ bisects $\overline{EG}$.

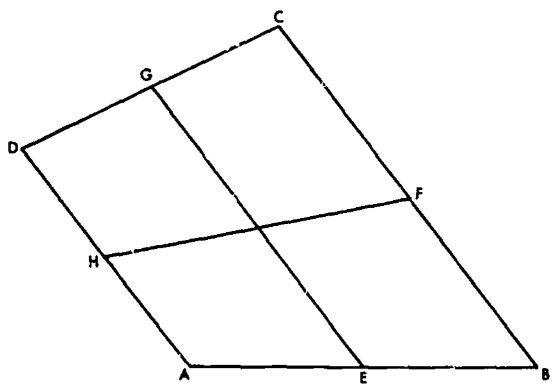


Figure A.9

We assign the weight 1 to each of A, B, C, D. Then we have the following equations.

(1) $1A + 1B = 2E$

(2) $1B + 1C = 2F$

(3) $1C + 1D = 2G$

(4) $1D + 1A = 2H$

By Axioms 3 and 2 we can show that

$(1A + 1B) + (1C + 1D) = (1D + 1A) + (1F + 1C)$.

Thus

$$2E + 2G = 2H + 2F$$

If K is the midpoint of $\overline{EG}$ then $2E + 2G = 4K$.

If L is the midpoint of $\overline{HF}$ then $2H + 2F = 4L$.

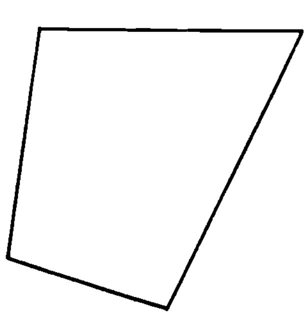
Thus $4K = 4L$

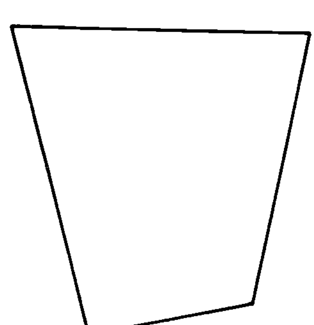
and $K = L$

Do you see that this completes the proof?

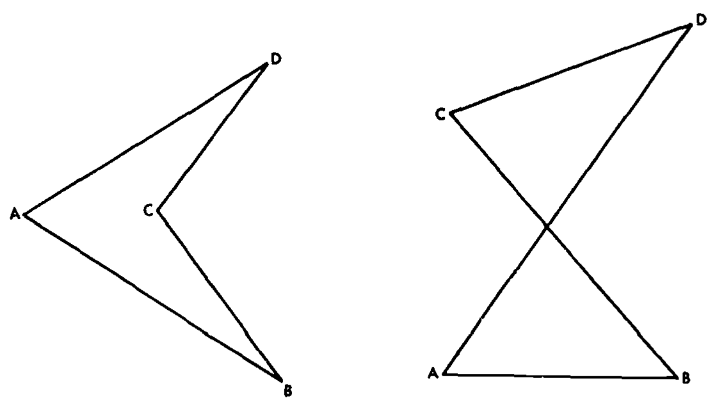
A.9 Exercises

1. The purpose of this exercise is to see if an experiment agrees with Theorem 2. In performing the experiment you should be careful to draw straight lines and to locate midpoints accurately. Perform the experiment on two different quadrilateral figures having shapes like these:

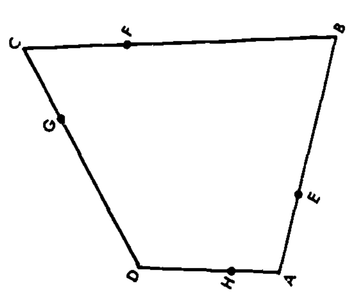




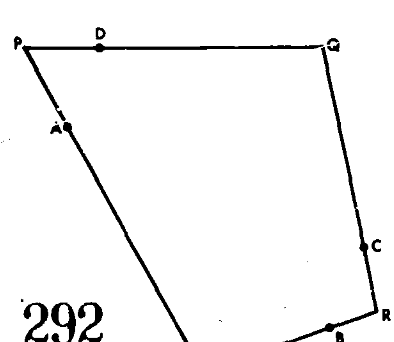
2. Verify whether or not the theorem is true for such figures like those below. They are named ABCD to tell you that the sides are $\overline{AB}$, $\overline{BC}$, $\overline{CD}$, $\overline{DA}$, in that order. This means that $\overline{AB}$ and $\overline{CD}$ are a pair of opposite sides and $\overline{BC}$ and $\overline{DA}$ are another pair of opposite sides.



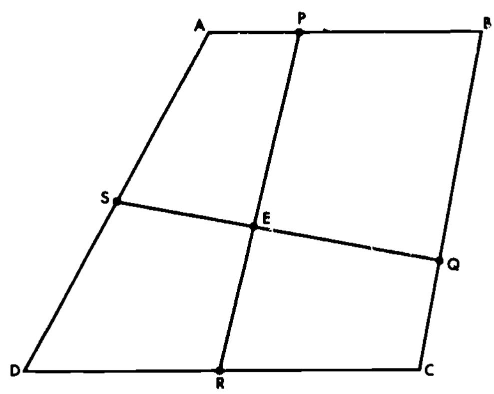
3. In the quadrilateral ABCD below, AE: EB = 1: 2, BF: FC = 2: 1, CG: GD = 1: 2, and DH: HA = 2: 1.
Prove: EG and FH bisect each other. (Hint: Assign weights 2 to A, 1 to B, 2 to C, and 1 to D.)



4. In the quadrilateral PQRS below, PA: AS = 1: 3, SB: BR = 3: 1, RC: CQ = 1: 3, QD: DP = 3: 1. Prove: $\overline{AC}$ and $\overline{BD}$ bisect each other



5. As shown for the quadrilateral ABCD below, AP: PB = 1: 2, BQ: QC = 2: 1, CR: RD = 1: 1, DS: SA = 1: 1. Let $\overline{SQ} \cap \overline{PR} = E$. Find the numerical values of RE: EP and SE: EQ.



A.10 Using a Definition

Consider the following problem:

> In $\triangle ABC$ (Figure A.10) D is the midpoint of $\overline{AB}$, E is the midpoint of $\overline{AC}$, and F is the trisection point of $\overline{BC}$ nearer B. Let $\overline{DE} \cap \overline{AF} = G$. We are required to show that G is the midpoint of $\overline{AF}$ and also the trisection point of $\overline{DE}$ nearer D.

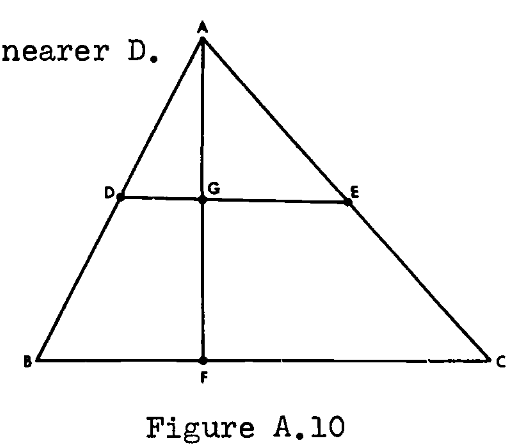


Figure A.10

We begin by assigning a weight of 1 to C. In order that

F be the trisection point of $\overline{BC}$ nearer B we assign 2 to B. Thus $2B + 1C = 3F$.

Let us now consider what weight to assign to A. First, in order that D be the midpoint of $\overline{AB}$ we should assign to A the same weight that we assigned to B, that is, 2. In order that E be the midpoint of $\overline{AC}$ we should assign to A the same weight that we assigned to C, that is 1. Thus, we find ourselves assigning two weights to A, or to put it another way, at A we need two mass points: one is 2A, the other is 1A.

Suppose we add the mass points:

$$1A + 2A = 3A,$$

by our definition. [Recall: $aA + bA = (a + b)A$.] Then assigning weight 3 to A, we note that $2B + 1C + 3A$ can be calculated either as

$$(2B + 1C) + 3A \qquad (1)$$

or as

$$(2A + 2B) + (1A + 1C) \qquad (2)$$

Since $2B + 1C = 3F$, (1) becomes $3F + 3A$, which is equal to 6H, where H is in $\overline{FA}$, such that $FH:HA = 1:1$.

Since $2A + 2B = 4D$ and $1A + 1C = 2E$, (2) becomes $4D + 2E$ which is equal to 6K, where K is in $\overline{DE}$ such that $DK:KE = 1:2$. But whichever way we calculate $2B + 1C + 3A$, we get the same result. Thus $6H = 6K$ and $H = K$. Since H is on both $\overline{FA}$ and $\overline{DE}$, $H = \overline{FA} \cap \overline{DE} = G$.

The actual calculations are few and can be written briefly as follows.

$$2B + 1C + 3A \text{ is equal to:}$$

$$\begin{array}{lll} (2B + 1C) + 3A & \text{or} & (2A + 2B) + (1A + 1C) \\ = 3F + 3A & & = 4D + 2E \\ = 6H & & = 6K \end{array}$$

Therefore $H = K = G$.

Thus $FG: GA = 1:1$, and $DG: GE = 1:2$.

A.11 Exercises

1. (a) In the problem discussed in Section A.10 above we started by assigning a weight of 1 to C. Give the solution starting instead with the assignment of weight 2 to C.

   (b) In the solution of the above problem we calculated (1) $(2B + 1C) + 3A$ and (2) $(2A + 2B) + (1A + 1C)$. There are at least two other possibilities: (3) $(2B + 3A) + 1C$ and (4) $(1C + 3A) + 2B$. Perform the calculations suggested in (3) and (4) and interpret your results.

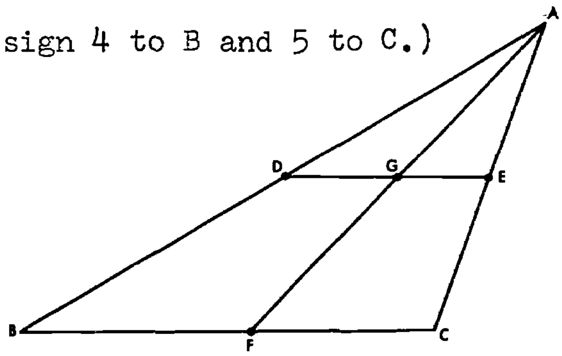
   (c) Perform the calculation $(3A + 1B) + (1B + C)$ and interpret the results.

   (d) There is still another calculation for locating G, in which the mass point at B is "split" into $1B + 1B$. What is it? What is its interpretation?

   (e) Still another calculation for locating G is sug-

gested by (2B + 1A) + (1C + 2A). Interpret this.

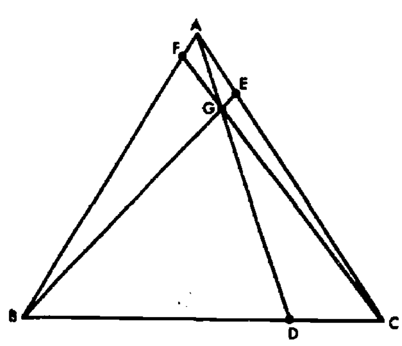
2. Suppose in $\triangle ABC$, D is the midpoint of $\overline{AB}$ and E is the midpoint of $\overline{AC}$, and F is in $\overline{BC}$ such that BF: FC = 5: 4 and $\overline{DE} \cap \overline{AF} = G$. (See the figure below.) Prove that G is the midpoint of $\overline{AF}$, and that DG: GE = 5: 4. (Hint: Assign 4 to B and 5 to C.)



3. State a theorem which seems to be suggested by exercise 2 and the problem of section A.10.

4. Investigate the case in which we take D and E as trisection points of $\overline{AB}$ and $\overline{AC}$, both nearer A, instead of the midpoints. If BF: FC = 5: 4, what is the ratio AG: GF? Would the ratio AG: GF change if D and E are trisection points of $\overline{AB}$ and $\overline{AC}$ but we take BF: FC = 2: 3?

5. In $\triangle ABC$, D is in $\overline{BC}$ and $\frac{BD}{DC} = \frac{3}{1}$, E is in $\overline{CA}$ and $\frac{CE}{EA} = \frac{4}{1}$. $\overline{AD}$ and $\overline{BE}$ meet in G. CG meets AB at F.



(a) Find $\frac{AF}{FB}$.

(b) Prove: $\frac{AF}{FB} \cdot \frac{BD}{DC} \cdot \frac{CE}{EA} = 1.$

(Hint: Assign 1 to B. What should you assign then to C? Then to A?)

6. Solve Exercise 5 under the altered suppositions $\frac{BD}{DC} = \frac{3}{2}$ and $\frac{CE}{EA} = \frac{5}{3}$.

*7. Exercises 5 and 6 are special cases of a theorem called Ceva's theorem, named after an Italian who is said to have discovered it. Ceva's theorem says: In $\triangle ABC$, if F, D, E are interior points of $\overline{AB}$, $\overline{BC}$ and $\overline{CA}$ respectively and $\overline{AD}$, $\overline{BE}$, and $\overline{CF}$ meet in one point then

$$\frac{AF}{FB} \cdot \frac{BD}{DC} \cdot \frac{CE}{EA} = 1$$

Try to prove it. (Hint: Let BD = a, DC = b, CE = c, EA = d.) (Difficult)

*8. For the data in Ceva's Theorem prove $\frac{GD}{AD} + \frac{GE}{BE} + \frac{GF}{CF} = 1$, where G is the point in which $\overline{AD}$, $\overline{BE}$, and $\overline{CF}$ meet. (Difficult)

A.12 Mass Points in Space and a Theorem

At the beginning of this chapter we worked with mass points on a line. Then we worked with mass points in a plane. We end this chapter by discussing mass points in space.

We have considered the addition of four mass points in a plane. Suppose we have four points not all in the same plane as shown in Figure A.11.

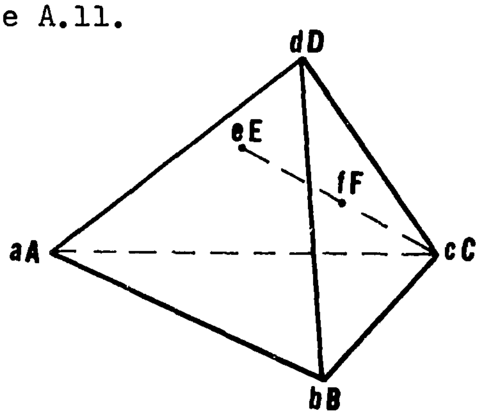


Figure A.11

Adding sets of three of these mass points, as we saw in Section A.4, determines points in the interiors of the triangles ABC, BCD, ABD, and CAD. Suppose eE in triangle ABD is the sum of the mass points aA, bB, and dD. Then the sum of eE and cC would be a mass point fF with F on segment $\overline{EC}$. Then fF is in the interior of the space figure ABCD (a triangular pyramid or tetrahedron).

We now prove a theorem about such a space figure which will remind you of the theorem about the medians of a triangle and its consequences.

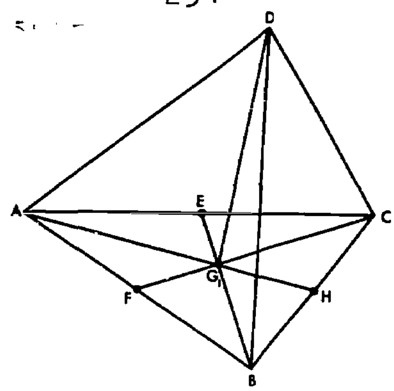


Figure A.12

We begin with four points A, B, C, and D not in a plane (see Figure A.12). Look at $\triangle ABC$ and its medians $\overline{AH}$, $\overline{BE}$, and $\overline{CF}$. We know from Theorem 1 that these medians meet in a point; name it G. The point in which the medians of a triangle meet is called the <u>centroid</u> of the triangle. In what ratio does the centroid G divide $\overline{AH}$, from A to H? Now, $\triangle BCD$, $\triangle ABD$, and $\triangle ADC$ also have centroids. Consider the segments joining the centroid of one of these triangles to the fourth point. One such segment is $\overline{GD}$ since it joins the centroid of $\triangle ABC$ to D. How many such segments are there? Do you think that these four segments meet at a point? Indeed they do and that is what our space theorem says.

<u>Theorem</u> 3. If A, B, C, D are points in space, not in a plane, and G_1 is the centroid of $\triangle ABC$, G_2 the centroid of $\triangle DAB$, G_3 the centroid of $\triangle DBC$ and G_4 the centroid of $\triangle DCA$, then $\overline{DG_1}$, $\overline{CG_2}$, $\overline{AG_3}$, and $\overline{BG_4}$ meet in a point which divides each

of these segments in the ratio 1:3 from
centroid to the point.

To prove this theorem we assign weight 1 to each of A, B, C, D. Then we consider $1A + 1B + 1C + 1D$.

One way to calculate this is to associate $(1A + 1B + 1C)$ which is $3G_1$. Then $3G_1 + D = 4H$, where H is a point in $\overline{G_1D}$ such that $G_1H:Hd = 1:3$. Thus $1A + 1B + 1C + 1D = 4H$, and whether we calculate it as $(1A + 1B + 1D) + 1C$, or $(1B + 1C + 1D) + 1A$, or $(1A + 1C + 1D) + 1B$, we continue to get $4H$. Do you see that this completes the proof?

A.13 Summary

In this chapter we studied some properties of mass points deductively. We started by defining mass points and addition of mass points. The first axiom assured us that we had an operational system. The second and third provided the properties of commutativity and associativity. We deduced three statements which you may find useful to remember. We labeled them theorems. One asserts that the medians of a triangle meet in a point. Another asserts that the segments joining midpoints of opposite sides of a quadrilateral bisect each other. The third is about four points in space, not in a plane, and the centroids of the four triangles determined by each triple of four points. It asserts that the segments joining the centroid of each triangle to the fourth point meet in a point that [illegible]ides each segment in the ratio 1:3 from the centroid to

the point.

We also solved many exercises by deductions and thus proved many other statements which however we did not call theorems, even though they are theorems.

A.14 Review Exercises

1. Draw $\overline{AB}$ making it 3 inches long. Let C be its mid-point. Locate the centers of mass for the following mass points:

   (a) $2A + 1B$ (d) $1A + 1B + 1C$

   (b) $1A + 2B$ (e) $1A + 2C + 3B$

   (c) $2A + 1C$ (f) $2A + 4B + 3C$

2. Solve for x and locate X in a drawing of AB where $\overline{AB}$ is a one inch segment.

   (a) $3A + xX = 4B$ (c) $xX + 2A = 4B$

   (b) $2A + xX = 3B$ (d) $xX + 3A = 5B$

3. Let A have weight 8 and let $\overline{AB}$ be a given segment. Let C be the center of mass for mass points at A and B. What weight should you assign B for each of the following descriptions of C?

   (a) C is the midpoint of $\overline{AB}$.

   (b) C is the trisection point of $\overline{AB}$ nearer A.

   (c) C is the trisection point of $\overline{AB}$ nearer B.

   (d) C is the point of $\overline{AB}$ such that AC:CB = 2:3.

4. In $\triangle ABC$, D is the midpoint of $\overline{BC}$ and E is the point in $\overline{CA}$ such that $CE:EA = 4:1$.

   (a) If weight 1 is assigned to B, what weights should you assign to C and A so that D is the center of mass of the mass points at B and C, and E is the center of mass of the mass points at C and A?

   (b) If $\overline{AD} \cap \overline{BE} = G$, compute the values of $AG:GD$ and $BG:GE$.

   (c) If $CG \cap \overline{AB} = F$, compute $AF:FB$.

5. In $\triangle ABC$, D is in $\overline{AB}$ and $AD:DB = 1:2$, E is in $\overline{BC}$ and $BE:EC = 2:1$. F is in CA and $CF:FA = 1:2$. Prove that $\overline{DF}$ and $\overline{AE}$ bisect each other.

6. In quadrilateral ABCD, E, F, G, H, are respectively in $\overline{AB}$, $\overline{BC}$, $\overline{CD}$, $\overline{DA}$. Each of $AE:EB$, BF,FC, and $CG:GD$ is equal to $2:1$, $DH:HA = 1:8$, and $\overline{EG} \cap \overline{FH} = K$. Prove $EK:KG = 4:1$ and $FK:KH = 3:2$.

Index