MATHEMATICS FOR THE NONMATHEMATICIAN

MORRIS KLINE
Professor of Mathematics, Emeritus
Courant Institute of Mathematical Sciences
New York University

DOVER PUBLICATIONS, INC.
NEW YORK

Published in Canada by General Publishing Company, Ltd., 30 Lesmill Road, Don Mills, Toronto, Ontario.
Published in the United Kingdom by Constable and Company, Ltd., 10 Orange Street, London WC2H 7EG.

This Dover edition, first published in 1985, is an unabridged republication of the work first published by Addison-Wesley Publishing Company, Inc., Reading, Massachusetts, in 1967 under the title *Mathematics for Liberal Arts*. The Instructor's Manual published with the original edition, containing additional answers and solutions to the problems in the text, has been added to this edition.

Manufactured in the United States of America
Dover Publications, Inc., 31 East 2nd Street, Mineola, N.Y. 11501

Library of Congress Cataloging in Publication Data

Kline, Morris, 1908–
Mathematics for the nonmathematician.

Reprint. Originally published: Mathematics for liberal arts. Reading, Mass.: Addison-Wesley, © 1967. (Addison-Wesley series in introductory mathematics)
Includes bibliographies and index.
1. Mathematics—1961–   I. Title.
QA37.2.K6 1985   510   84-25923
ISBN 0-486-24823-2

TRIGONOMETRIC FUNCTIONS AND OSCILLATORY MOTION

All the effects of nature are only mathematical results of a small number of immutable laws.

P. S. LAPLACE

18-1 INTRODUCTION

In the seventeenth century one of the most pressing problems of the times was time itself. The increasing scientific activity, particularly in an age which had decided to measure and to seek quantitative laws, created the need for convenient, accurate methods of measuring time. Moreover, as we have already had occasion to mention in other connections, the seventeenth and eighteenth centuries were concerned with the very practical problem of improving the method by which ships determined their longitude at sea. Here a good clock is the simplest answer. Suppose that the longitude of a given place on land is known and that a ship has on board a clock set to agree with the time prevailing at that given locality. Since the earth turns through 360° of longitude in one day, it turns through 15° in each hour. Hence for each 15° that a ship is west, say, of the fixed locale, midday occurs one hour later compared to the time at the fixed position on land. If a ship's officer notes (by means of the sun's position) when midday occurs at his position at sea and finds, for example, that his clock reads 3 o'clock whereas it should, of course, read 12 o'clock, he knows that the longitude of his position is 45° west of the given reference locality on land. We can see then why scientists decided to search for a reliable and accurate clock.

The thought which suggested itself almost at once was to look for some physical phenomenon which repeated itself regularly. The day contains 24 hours; hence when the number of repetitions per day is known, the duration of each repetition is readily calculated. Where then could one find a repetitive or periodic physical phenomenon? Two prospects attracted the attention of seventeenth-century scientists. The first of these was the motion of a mass, called a bob, attached to a spring and oscillating up and down, and the second was the motion of a pendulum, that is, a bob attached to a string and swinging to and fro. Now first reactions to the possibility of using the motion of a bob on a spring or a pendulum as a measure of time are apt to be negative. The bob on a spring, for example, does go through each cycle, that is, each complete up and down motion, in the same time so far as the eye can judge, but the motion soon dies down. The same is true for the pendulum. But, if air resistance could be minimized or perhaps compensated for, then these motions might become truly periodic and should therefore merit investigation. The scientist or mathematician who expects to see at once the solution of a problem he sets out to study will never accomplish much. The best he can hope for at the outset is an idea or a clue to pursue.

In this chapter we shall examine first the physical problem of the motion of a bob on a spring, a prime example of oscillatory motion. To study such motions mathematicians created a new class of functions, the trigonometric functions. We shall then discuss these functions and see how they are used to derive some knowledge about the physical problem which motivated their introduction. Surprisingly, trigonometric functions proved to be admirably suited for the study of sound, electricity, radio, and a host of other oscillatory phenomena. Of these latter developments we shall learn more in the next chapter.

18-2 THE MOTION OF A BOB ON A SPRING

The problem of investigating the motion of a bob on a spring was undertaken by one of the greatest experimentalists in the history of physics, the Englishman Robert Hooke (1635–1703). Hooke was professor of mathematics and mechanics at Gresham College. His claim to fame also rests upon his success as an inventor. To his credit are a telescope moved by a clock mechanism and devices for measuring the moisture in the atmosphere, the force of the wind, and the amount of rainfall. He improved the microscope, the barometer, the air pump, and the telescope. One of his findings, namely, that white light passed through thin sheets of mica breaks into many colors, parallels Newton's work on light. He also discovered the cell structure of plants. Hooke was very much interested in designing a useful clock and thought that springs would furnish the essential device. While working on the action of springs, he discovered a basic law, still known as Hooke's law, which we shall discuss later.

Let us follow Hooke in studying the motion of a bob on a spring. The upper end of the spring is attached to a fixed support, and a bob is attached to the lower end. Because gravity pulls the bob downward, the spring will be extended until the tension in the spring offsets the force of gravity. The bob then comes to rest in some position

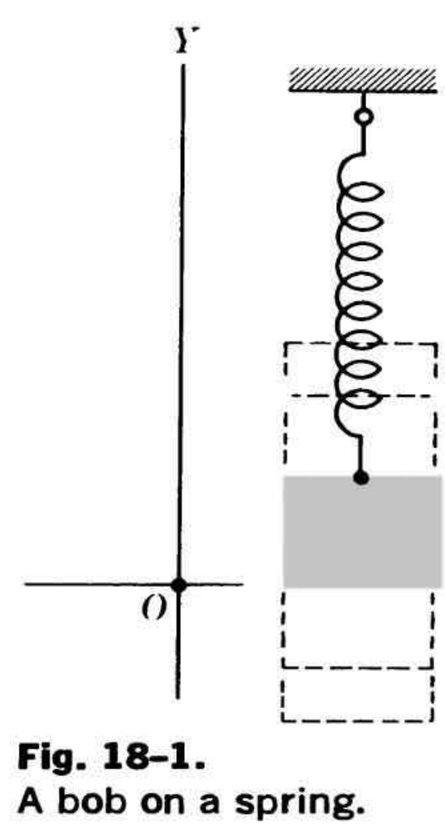

Fig. 18-1.
A bob on a spring.

which is called the *rest* or *equilibrium position* (Fig. 18–1). If one now pulls the bob downward some definite distance below the rest position and then releases it, the bob moves up to the rest position, continues past that point to some highest position, and then moves downward. When it reaches the point to which it had been pulled down, it starts upward and repeats its former motion. Following Galileo's plan of idealizing the physical situation, let us suppose that air resistance is negligible. (Strictly speaking, energy is also lost in the expansion and contraction of the spring, but this loss is negligible.) Then the bob will continue to move up and down endlessly.

To begin to get some mathematical description of this motion let us introduce a Y-axis alongside the bob (Fig. 18–1) and suppose that $y = 0$ corresponds to the rest position of the bob. When the bob is above or below the rest position, the bob is said to be displaced and the distance that it is above or below the rest position is called its *displacement*. To distinguish displacements above from those below the rest position, we shall call the former positive and the latter negative. Each displacement may then be described by a value of y. Thus $y = -\frac{1}{2}$ means that the bob is $\frac{1}{2}$ unit below the rest position.

To study the motion of the bob mathematically, it would be most helpful if we could find the formula which relates the displacement of the bob and the time it is in motion. Let us therefore seek such a formula.

18–3 THE SINUSOIDAL FUNCTIONS

No one of the formulas that we have considered thus far would be useful to represent the motion of the bob, for the peculiarity of the present phenomenon is that after each up and down motion, or oscillation, has been completed, the displacements go through their former sequence of values. Hence we apparently must seek a new type of formula which expresses the periodic character of the motion of the bob. We do not seem to have any clue, but a little imagination may supply one.

Suppose a point P moves around a circle of unit radius at a constant speed. Let us denote some of its positions by P_1, P_2, . . . (Fig. 18–2). We can, if we wish to, introduce a point Q on the vertical line through the center O such

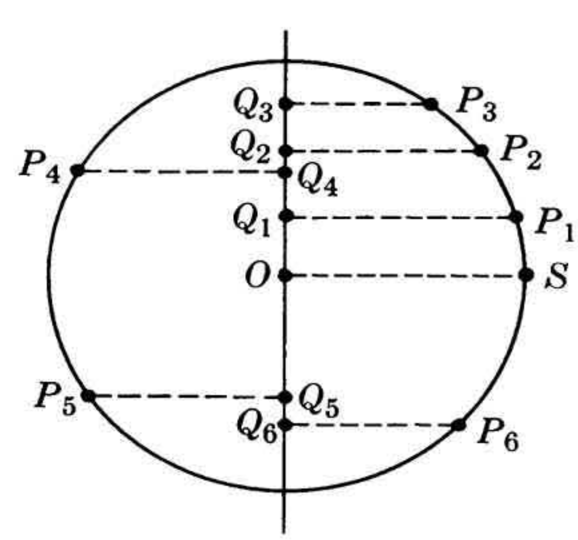

Fig. 18–2.
Successive positions of a point P which moves around a circle at a constant velocity, and the corresponding positions of Q.

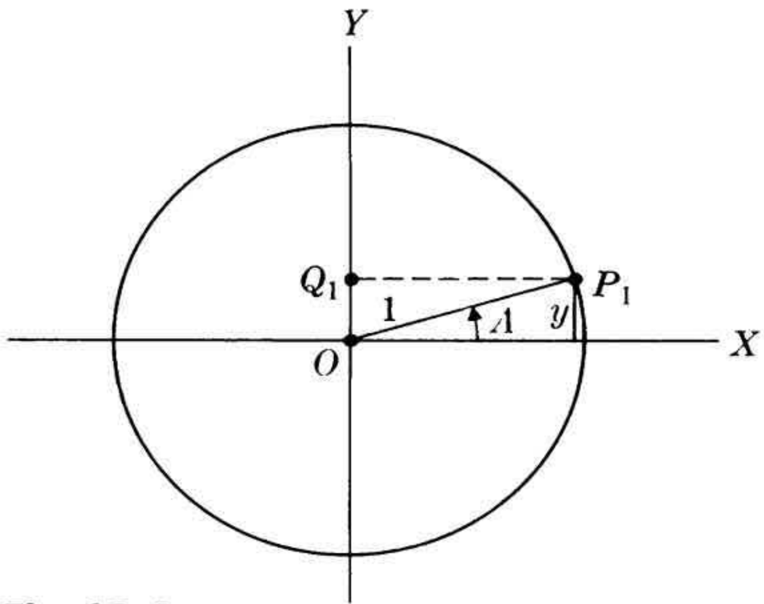

Fig. 18–3

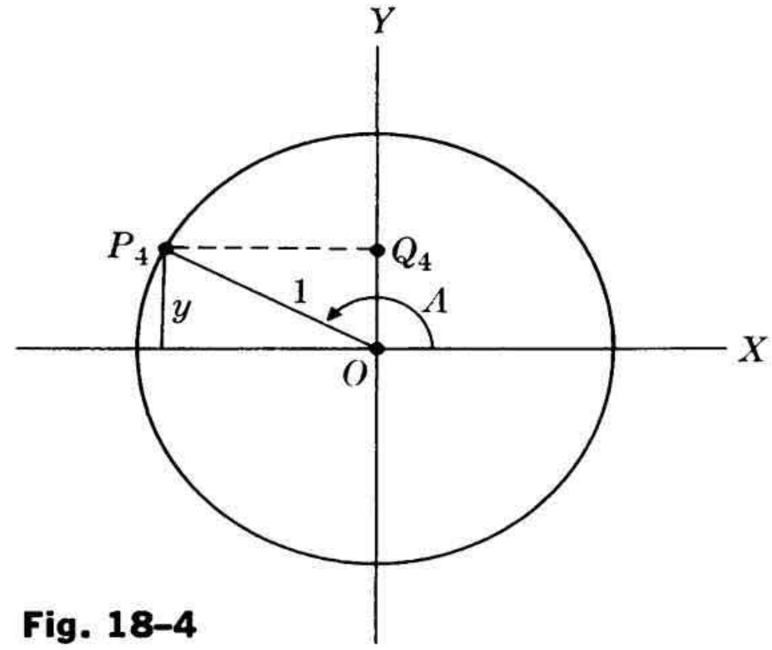

Fig. 18–4

that Q always has the same height that P has above or below the horizontal through O. The point Q is called the *projection* of P on the vertical line. Thus to the position P_1 of P there corresponds Q_1; to P_2 there corresponds Q_2; and so on. Why should we introduce the point Q? Well, let us imagine P moving around the circle through many revolutions starting from the position S at the right. What does its "shadow" Q do? It moves up from O to a highest position, moves down again to O, moves past O to a lowest position on the vertical line, moves up again to O, and then repeats this up and down motion. The motion of Q certainly seems to have the essential characteristics of the motion of the bob on the spring. Hence perhaps by pursuing further the motion of Q we may obtain the function we are seeking.

Let us introduce coordinate axes as shown in Fig. 18–3. If P starts from the X-axis and reaches, say the position P_1, then we may describe the position of P by the angle A shown in the figure. The height of Q above the X-axis *is the same* as the y-value of P. Now

$$\sin A = \frac{y}{1}.$$

Hence

$$y = \sin A. \tag{1}$$

Thus if the position of P is described by the angle A, then the position of the corresponding point Q on the vertical line is given by (1).

But now suppose P has moved to the position P_4 shown in Fig. 18–4. The angle A which describes the position of P_4 is the obtuse angle shown in the figure. This angle is no longer an acute angle of a right triangle, and we therefore have no right to speak of $\sin A$. However, let us extend the meaning of sine so that, by definition, $\sin A$ is the y-value of P_4. Since the height of Q_4 above O equals the y-value of P_4, we may continue to write $y = \sin A$ to describe the position of Q. That is, the distance of Q from O on the vertical line will be given by $y = \sin A$.

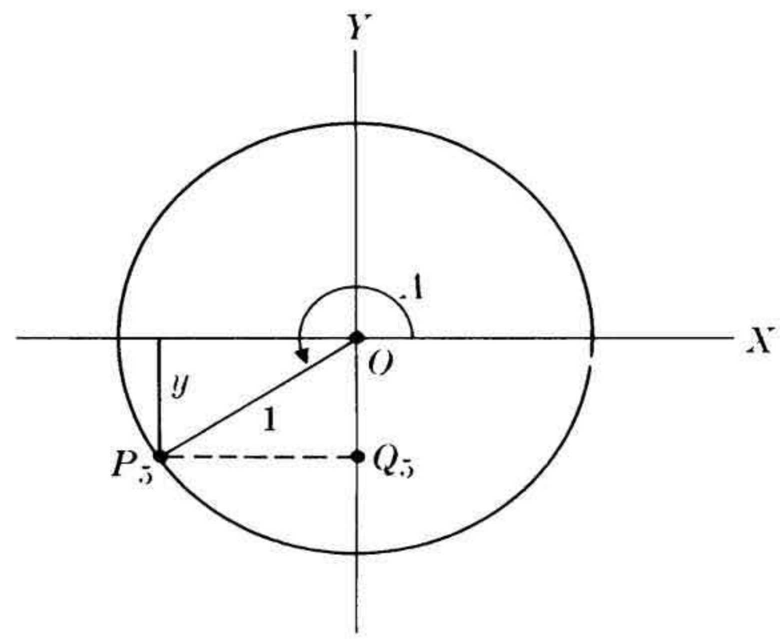

Fig. 18-5

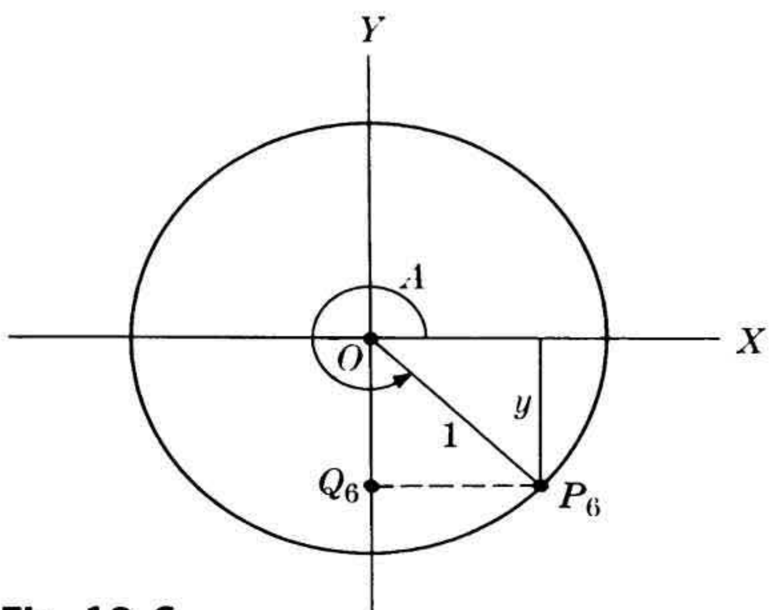

Fig. 18-6

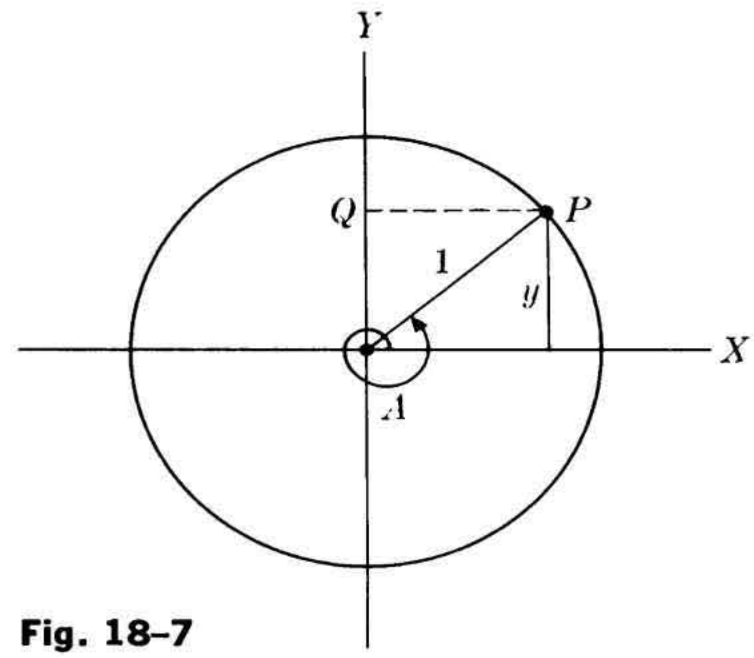

Fig. 18-7

Suppose next that P occupies the position P_5 shown in Fig. 18–5. The angle A which describes how far around the circle P has moved is now the angle shown. Let us agree again that by sin A we shall mean the y-value of P_5, which is also the distance below the X-axis of the point Q_5. Then we again may write $y = \sin A$ to describe the position of Q. Note that y is now a negative quantity.

If P reaches the position P_6 shown in Fig. 18–6, then its position is represented by the angle A shown, and if we again agree to mean by sin A the y-value of P_6, we shall be able to say here too that $y = \sin A$ describes the position of Q. In this instance also, y is a negative quantity.

As P returns to the X-axis and starts to repeat its revolution, the angle A which describes the position of P will now be 360° plus some additional angle (Fig. 18–7). It is only by including 360° for each revolution of P that we can keep track of the number of revolutions. However, let us note that the y-values of P will recur in precisely the same order in which they appeared on the first revolution. Despite the fact that on the second revolution the values of A are larger than 360°, we shall continue to mean by sin A the y-value of P. Thus sin 390° will be the same as sin 30°. As P goes through

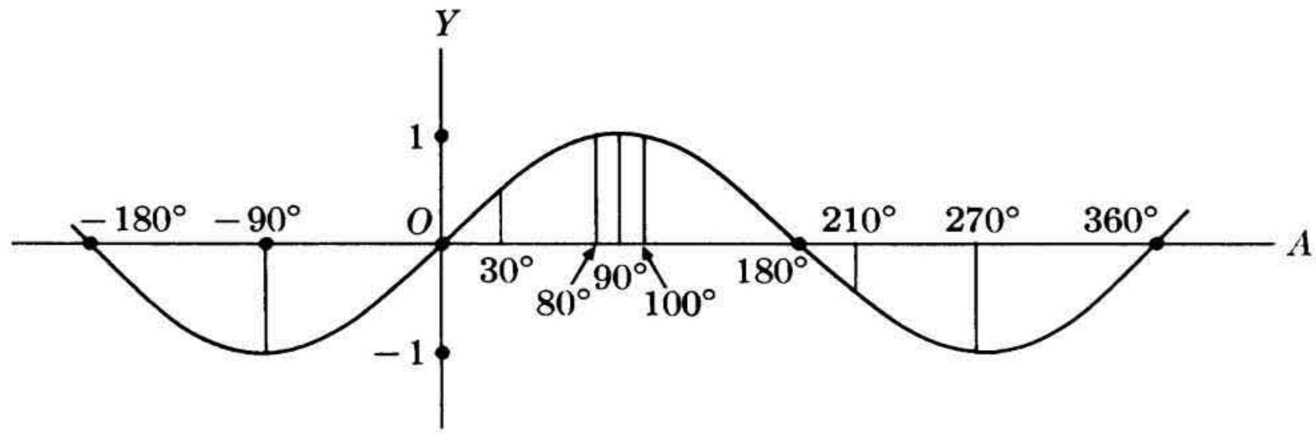

Fig. 18–8. The graph of $y = \sin A$.

its second revolution, Q repeats the motions of the first revolution. Hence it will still be true that $y = \sin A$ describes the position of Q on the vertical line. With each revolution of P, the y-values repeat, although the angle A increases by 360°. Since the motion of Q also repeats, its position on the vertical line will continue to be represented by $y = \sin A$.

If P were to revolve in the clockwise direction, then we would make only one change, i.e., call the values of A negative. The y-value of P, wherever it is, is still, by definition, $\sin A$ and this y-value would represent the position of Q.

Let us survey what we have done. To describe mathematically the position of the point Q, we have introduced a new function: $y = \sin A$. When A is an acute angle, as in Fig. 18–3, then $\sin A$ has the old meaning; that is, it is the ratio of the side opposite angle A to the hypotenuse of the right triangle in which A lies. (In the present case the hypotenuse is 1.) But when A is larger than 90°, then the equation $y = \sin A$ is a definition of what we mean by $\sin A$. Since there is a definite y-value for each value of A, positive or negative, we do indeed have a function.

To appreciate the nature of this function let us graph it. Figure 18–8 shows the graph. The values of A are plotted along the horizontal axis, and the corresponding y-values are plotted in the usual way.

Do we know the precise numerical value of y for each value of A? We do. For values of A which are between 0° and 90°, the y-values are the ordinary sine values which we find in our trigonometric table. In the interval from 90° to 180°, the values of $\sin A$ repeat, but in *reverse* order, the values which $\sin A$ has when A varies from 0° to 90°. This statement implies that $\sin 100° = \sin 80°$, $\sin 110° = \sin 70°$, and so forth. Stated in more general terms:

$$\sin A = \sin (180° - A). \tag{2}$$

In the interval from 180° to 360°, $\sin A$ has the same numerical values as when A varies from 0° to 180°. However, now $\sin A$ is negative. Thus sin

$210° = -\sin 30°$; $\sin 220° = -\sin 40°$; and in general:

$$\sin A = -\sin (A - 180°). \tag{3}$$

Since for each 360°-interval beyond the interval 0° to 360° sin A repeats the values that it has in the interval from 0° to 360°, $\sin 390° = \sin 30°$; $\sin 400° = \sin 40°$; and so on. In symbolic form,

$$\sin A = \sin (A - 360°). \tag{4}$$

The values of sin A for negative values of A are also shown in Fig. 18–8. If we look at the figure, we see that for any negative A-value, sin A is the negative of the sine of the corresponding positive A-value. That is, $\sin (-30°) = -\sin (30°)$; $\sin (-50°) = -\sin (50°)$, and in general:

$$\sin A = -\sin (-A). \tag{5}$$

Thus we have arrived at a definition of the function $y = \sin A$ for all values of A. Since we know quantitatively what sin A is for values of A between 0° and 90°, formulas (2) through (5) enable us to calculate sin A for all other values of A. The function we have just introduced is called a *periodic function* because the y-values repeat themselves in every 360°-interval of A-values. The interval of 360° is called the period of $y = \sin A$, and the entire set of y-values in one period is called the *cycle* of y-values.

EXERCISES

1. Using formulas (2) through (5) or Fig. 18–8, express the following sine values as sines of angles between 0° and 90°.

a) sin 120°	b) sin 150°	c) sin 210°	d) sin 260°
e) sin 270°	f) sin 300°	g) sin 350°	h) sin 370°
i) sin −50°	j) sin 750°		

2. What is the largest value of sin A? What is the smallest value of sin A?
3. At what value of A between 0° and 360° does the function $y = \sin A$ reach a maximum?
4. Why is $y = \sin A$ called a periodic function?
5. What purpose does the function $y = \sin A$ serve with respect to the location of Q, the projection of P?
6. What is the relationship between the function $y = \sin A$ and the trigonometric ratio sin A studied in Chapter 7?
7. Describe how sin A varies as A varies from 0° to 360°; from 360° to 720°.
8. For how many values of A between 0° and 360° is $\sin A = 0.5$?
9. Distinguish between the period and the cycle of $y = \sin A$.

Thus far we have described the size of angles in degrees. There is, however, no need to stick to this unit. Let us return to the motion of the point P in Figs. 18–3 through 18–7. The size of angle A can be specified by describing the *arc length* traversed by P from its starting point on the X-axis. This arc length is as much a measure of the size of angle A as the rather arbitrary agreement that a complete revolution of one side of A should be 360°.

Suppose that we agree to use the arc length traversed by P as a measure of A. How do we express in this new unit an angle of 90°, for example? When A is 90°, P has traversed one-quarter of the entire circumference. But the entire circumference of a circle of unit radius is 2π. Then the size of A in the new unit is $\pi/2$, that is about 1.57. We call this new unit *radians*. Thus an angle of 90° is also one of $\pi/2$ or 1.57 radians.

The advantage of radians over degrees is simply that it is a more convenient unit. Since an angle of 90° is of the same size as an angle of 1.57 radians, we now have to deal only with 1.57 instead of 90 units. The point involved here is no different from measuring a mile in yards instead of inches. If yards are just as good on other grounds, then it is far more convenient to speak of 1760 yards than 63,360 inches.

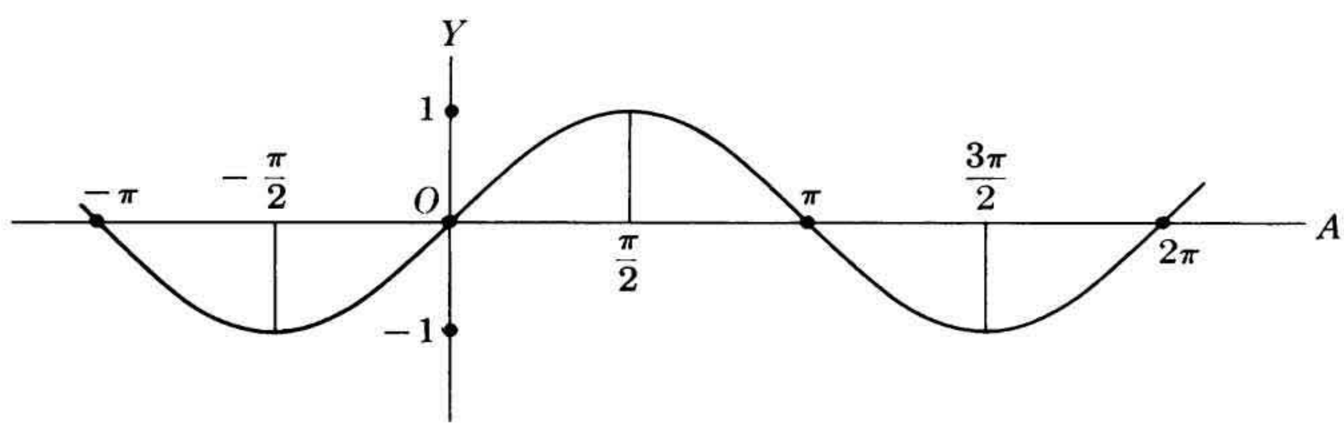

Fig. 18–9.
The graph of $y = \sin A$ when A is measured in radians.

The fact that we measure angles in radians does not disturb at all the meaning of the function $y = \sin A$. Instead of stating that $\sin 90° = 1$, we simply say that $\sin (\pi/2) = 1$. The same applies to any other value of A in the sinusoidal function we have introduced. Suppose, for example, that we wished to find the value of $y = \sin A$ when $A = \pi/6$. Because our table is set up in degrees, we note first that an angle of $\pi/6$ is of the same size as 30°, for $\pi/2$ radians is the same as 90°. Now from our tables we see that $\sin 30° = 0.5$, and so $\sin (\pi/6) = 0.5$.

Since we shall be using radians a good deal, we may as well become familiar with the function $y = \sin A$ when A is expressed in radians. Figure 18–9 shows the same function as Fig. 18–8 except that the units of A are now radians.

EXERCISES

1. Express the sizes of the following angles in radians: 90°, 30°, 180°, 270°, 360°, 420°.
2. The sizes of the following angles are in radians. Express the same angles in degrees.

$$\pi/2 \qquad 2\pi/3 \qquad 5\pi/2 \qquad 3\pi \qquad -\pi/2 \qquad 1$$

3. Find the value of

a) $\sin \pi$ b) $\sin (\pi/2)$ c) $\sin (\pi/3)$
d) $\sin (3\pi/2)$ e) $\sin 3\pi$ f) $\sin (5\pi/2)$.

4. Describe how $\sin A$ varies as A varies from 0 to 2π, as A varies from 2π to 4π.

The function $y = \sin A$ has a maximum value of $+1$ and a minimum value of -1. The maximum y-value, incidentally, is called the *amplitude* of the function. Such a function, even if it were suitable in all other respects, could not represent the motion of a bob whose maximum displacement is 2 or 3, say. This difficulty is easily obviated. Now that we have $y = \sin A$, we can readily manufacture hundreds of new functions whose amplitudes are whatever we choose to make them. Consider, for example, $y = 2 \sin A$. How does this function behave compared to $y = \sin A$? The answer is immediate. For any value of A, $y = 2 \sin A$ is twice as much as $y = \sin A$. Thus when $A = \pi/4$ or 45°, $\sin A = 0.71$, and $2 \sin A$ is 1.42. Figure 18–10 shows how $y = 2 \sin A$ looks compared to $y = \sin A$. If we want a sine function with amplitudes 3, $\frac{1}{2}$, or any other number, we can write one down immediately. As is evident from the nature of the function $y = 2 \sin A$, the function

$$y = D \sin A$$

has amplitude D.

Before we can use functions such as $y = \sin A$ or $y = 3 \sin A$ to represent the motion of a bob on a spring, we must clear one more hurdle. The func-

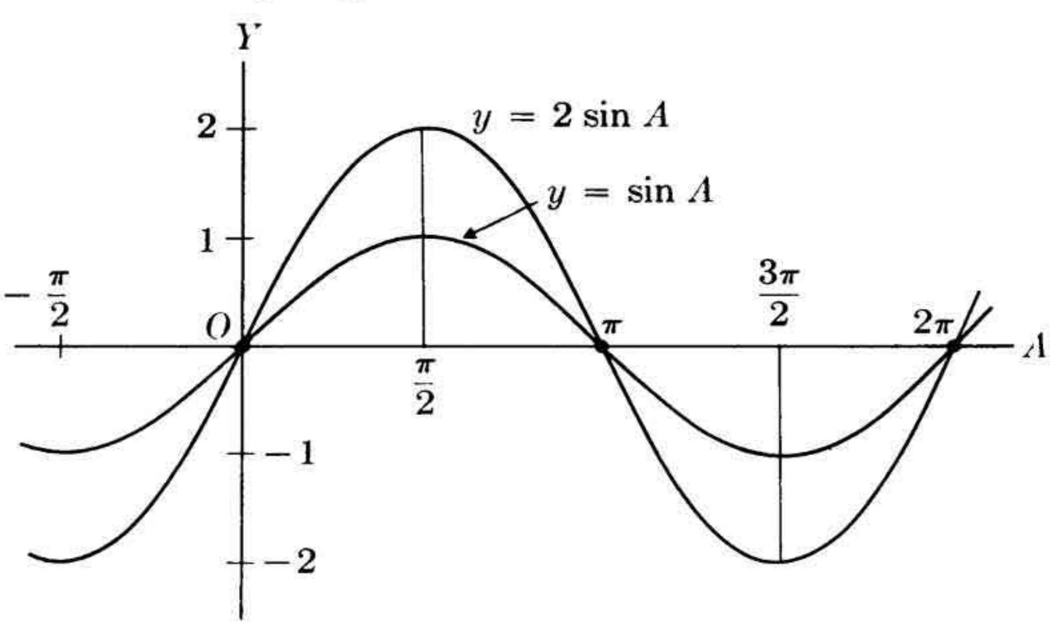

Fig. 18–10. Comparison of $y = \sin A$ and $y = 2 \sin A$.

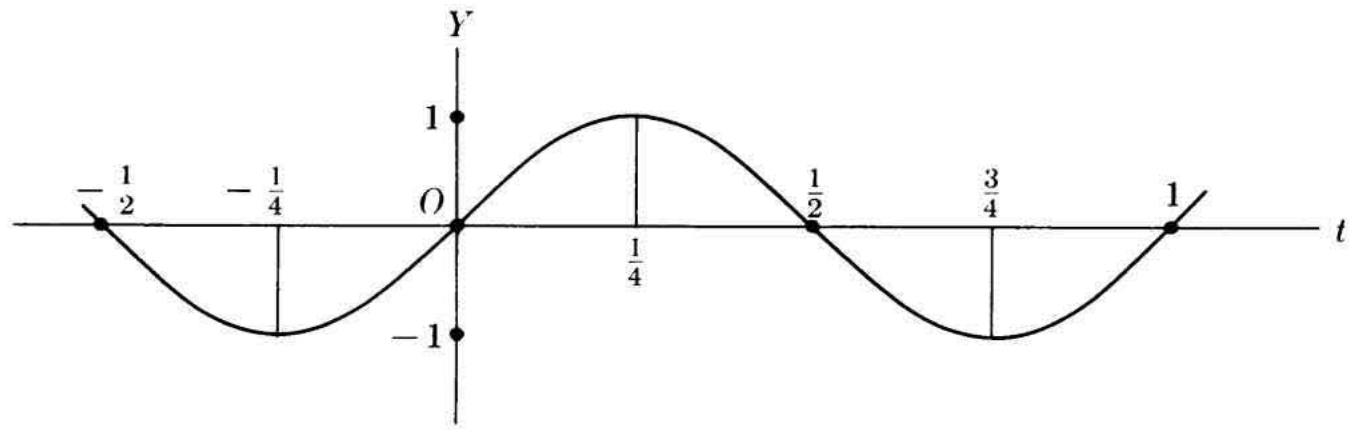

Fig. 18–11. The graph of $y = \sin 2\pi t$.

tion we seek should represent a relationship between displacement and time. The y-values of our functions do indeed represent the displacement of a point Q which moves up and down on a line, but our independent variable is an angle. Suppose, however, that the point P revolves around the circle f times in one second. Then, since for each revolution the angle A increases by 2π radians, the size of the angle which describes the amount of revolution of P in one second is $2\pi f$. If the point P revolves for t seconds and makes f revolutions per second, it will make ft revolutions in t seconds. The angle generated during these ft revolutions will be $2\pi ft$. Hence, the value of A in t seconds will be $2\pi ft$. Thus the function $y = \sin A$ becomes

$$y = \sin 2\pi ft. \tag{6}$$

This function requires some study. Suppose the point P makes one revolution per second. Then $f = 1$. The function (6) then is $y = \sin 2\pi t$. As t increases from 0 to 1, the quantity $2\pi t$ will increase from 0 to 2π. We must now ask, How will $\sin 2\pi t$ vary as $2\pi t$ varies from 0 to 2π? Since the *angle* which $2\pi t$ describes now varies from 0 to 2π, the function will go through the entire cycle of sine values. However, if we now label our horizontal axis with time values, we obtain the graph shown in Fig. 18–11.

Next let us consider a slightly more difficult case. Suppose the point P makes 2 revolutions per second so that $f = 2$. As t increases from 0 to $\frac{1}{2}$, $2\pi \cdot 2t$ will increase from 0 to 2π and $\sin 2\pi \cdot 2t$ will go through the entire cycle of sine values. As t increases from $\frac{1}{2}$ to 1, $2\pi \cdot 2t$ increases from 2π to 4π. Then $\sin 2\pi \cdot 2t$ takes on the values corresponding to angles from 2π to 4π. But in this range the sine function takes on the same values as it does in the range from 0 to 2π. Hence in the entire interval 0 to 1 for t, the graph will be as shown in Fig. 18–12. The conclusion, which emerges clearly from the graph, is that

$$y = \sin 2\pi \cdot 2t$$

goes through 2 complete cycles in one second or, as one says, it has a *frequency* of 2 cycles per second.

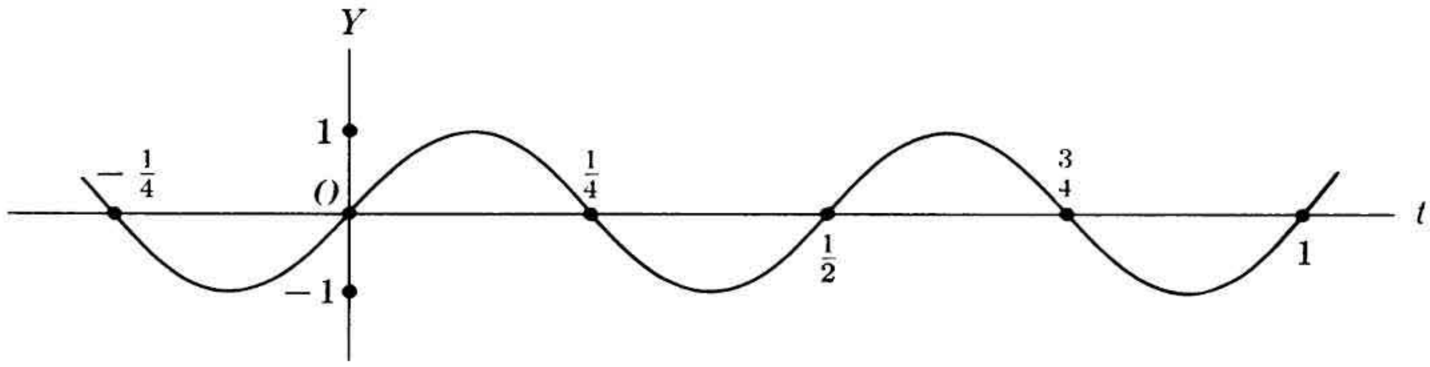

Fig. 18–12. The graph of $y = \sin 2\pi \cdot 2t$.

We can now anticipate what happens for any f. The function

$$y = \sin 2\pi ft$$

will go through f cycles in one second, or it has a frequency of f cycles per second.

To increase the amplitude of any of these functions, we have but to introduce the factor D. Thus the function

$$y = D \sin 2\pi ft \tag{7}$$

will have a frequency of f cycles per second and an amplitude of D. Let us note that while f is the number of revolutions per second of P, it is also the number of oscillations per second of Q.

The y-values of formula (7) oscillate above and below the zero value as t varies. We can make the number of oscillations per second what we please by merely inserting the proper value of f and we can do the same with respect to the amplitude by inserting the proper value of D. Of course, we do not know the proper values of f and D which fit the motion of a bob, but we shall see in the next section that it is not difficult to determine them.

Let us summarize what we have accomplished. We sought to represent the motion of a point which oscillates back and forth on a straight line. We were able to do so by introducing a point Q which is the projection of a point P moving around a circle at a constant velocity. Because the y-value of P equals the displacement of the oscillating point Q and because the y-value of P is expressible as a sinusoidal function, we can represent the motion of the oscillating point Q by such a function. That the approach to the oscillating point Q through the circle should be successful may be surprising, but, as Aristotle pointed out, "There is nothing strange in the circle being the origin of any and every marvel."

EXERCISES

1. Find the value of $2 \sin A$ when A is $30°$, $90°$, $\pi/2$, $\pi/3$.
2. What is the maximum value of $3 \sin A$? the minimum value?
3. What is the amplitude of $y = 4 \sin A$?

4. Find the value of

 a) $\sin 2t$ when $t = \pi/4, \pi/2, 3\pi/4, \pi$;

 b) $\sin 3t$ when $t = \pi/6, \pi/3, \pi/2, 2\pi/3$.

5. What is the shape of the graph of $y = \sin 2\pi \cdot 2t$ as t varies from 1 to 2?
6. Graph the function $y = \sin 2\pi \cdot 3t$ as t varies from 0 to 1.
7. Graph the function $y = 2 \sin 2\pi \cdot 2t$ as t varies from 0 to 1.
8. What is the frequency (in one second) of $y = \sin 2\pi \cdot 10t$?
9. Find the value of

 a) $y = \sin 2\pi \cdot 2t$ when $t = \frac{1}{8}, \frac{1}{4}, \frac{1}{3}$;

 b) $y = \sin 2\pi \cdot 4t$ when $t = \frac{1}{6}, \frac{1}{2}, 1$;

 c) $y = 2 \sin 2\pi \cdot 3t$ when $t = \frac{1}{6}, \frac{1}{4}, \frac{1}{12}$.

18-4 ACCELERATION IN SINUSOIDAL MOTION

What we have seen in the preceding article is that if a point P moves around a circle of unit radius at a constant speed and makes f revolutions per second, then the projection Q of P onto the vertical diameter moves up and down this diameter, and the displacement y of Q from its central position at O can be represented by formula (6), namely,

$$y = \sin 2\pi ft. \tag{8}$$

Moreover, if we wished to represent the same kind of oscillatory motion but with an amplitude D instead of 1, we have merely to modify (8) to read

$$y = D \sin 2\pi ft. \tag{9}$$

Our goal, however, is to represent the motion of the bob on the spring. Before we can do this we must learn one more fact about the motion of Q, namely, the acceleration of Q. The motion of Q, as we approached it, was determined by the motion of P which travels around the unit circle at a constant speed, say v. If an object moves along a circular path, then we know from our work in Chapter 15 that it must be subject to a centripetal acceleration, and by formula (24) of Chapter 15 this centripetal acceleration is v^2/r, where r is the radius of the circle. In our case, since P moves on a circle of unit radius, $r = 1$, and so the centripetal acceleration of P is v^2. This acceleration is directed toward the center of the circle.

We are, however, interested not in the motion of P, but in the motion of Q, which moves in the same way as the vertical motion of P. Hence we should seek the vertical acceleration of P. We learned in Chapter 14 that even though an object moves along a curve, we can study its motion by considering the horizontal motion and the vertical motion separately. By Galileo's principle these two motions are independent of each other. What we should like then is the vertical acceleration of P. When we considered the motion of a shell shot from a cannon inclined at an angle A to the ground, we found the horizontal and vertical velocities by dropping perpendiculars from the end point of the line segment representing velocity onto the horizontal and vertical axes (Section 14–4).

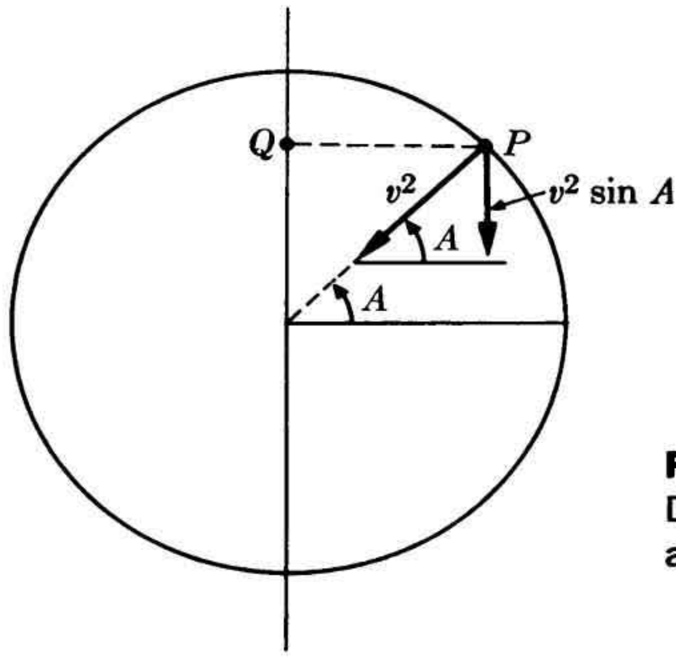

Fig. 18–13.
Determination of the vertical acceleration of P.

Now acceleration, like velocity, is a directed quantity or a vector. Moreover, the acceleration determines the velocity. Hence we should compute the vertical acceleration of the point P in the same manner as we computed the vertical component of the velocity of the shell. Figure 18–13 shows the centripetal acceleration v^2 of P as a line segment directed toward the center of the circle. If we drop a perpendicular from the end point of this line segment onto the vertical line through P, we obtain the vertical component of the acceleration. Angle A in the figure determines the position of P. We see, then, that the vertical component of the acceleration, which we shall denote by a, is

$$a = v^2 \sin A.$$

We know, however, from (1) that

$$\sin A = y,$$

where y is the ordinate of P. Hence

$$a = v^2 y.$$

However, since the acceleration is directed downward when y is positive, we must write

$$a = -v^2 y. \tag{10}$$

If now the moving point P makes f revolutions per second, then P covers f

circumferences per second; that is, $v = 2\pi f$ and $v^2 = 4\pi^2 f^2$. We substitute this result in (10) and obtain

$$a = -4\pi^2 f^2 y \tag{11}$$

as the acceleration of the vertical motion of P or of its shadow Q on the Y-axis.

What we have shown then is that the motion of Q, which is described by the formula

$$y = \sin 2\pi f t, \tag{12}$$

is subject to an acceleration of

$$a = -4\pi^2 f^2 y. \tag{13}$$

18-5 THE MATHEMATICAL ANALYSIS OF THE MOTION OF THE BOB

We may now undertake to represent mathematically the motion of the bob on the spring. We know that this motion is periodic and has a definite frequency and amplitude. However, we do not know that the motion is really sinusoidal; that is, as t varies, do the displacements of the bob follow precisely the variation of y in a function of the form

$$y = D \sin 2\pi f t? \tag{14}$$

If, for example, the motion of the bob should be faster on the upper half of its path than on the lower half, it could still have the same period for each complete oscillation and perform a fixed number of oscillations per second. Yet the motion would not be of the form (14). We need a little more insight into the motion of the bob than we now have.

This insight into the action of bobs on springs was supplied by Robert Hooke. The principle he discovered, still known as Hooke's law, is very simple. We all know that if we stretch or compress a spring, the spring seeks to restore itself to its normal length; that is, when stretched or compressed the spring exerts a force. Hooke's law says that the force is a constant times the amount of compression or extension. In symbols, if L is the increase or decrease in length of the spring and F is the force exerted by the spring, then $F = kL$, where k is a constant for a given spring. The quantity k is called the spring or stiffness constant and it represents the stiffness of the spring. If k is large, the spring exerts considerable force even for small L.

We shall now see what we can deduce from Hooke's law. Suppose that a bob of mass m is attached to a spring. Then we know that gravity pulls the bob downward some distance d where the bob comes to rest (Fig. 18–14). The rest position is reached when the force of gravity acting on the bob, or the weight of the bob, just offsets the upward force exerted by the spring. Now the force of gravity is $32m$, and, according to Hooke's law, the upward force exerted by a spring which is pulled downward a distance d is kd. Since

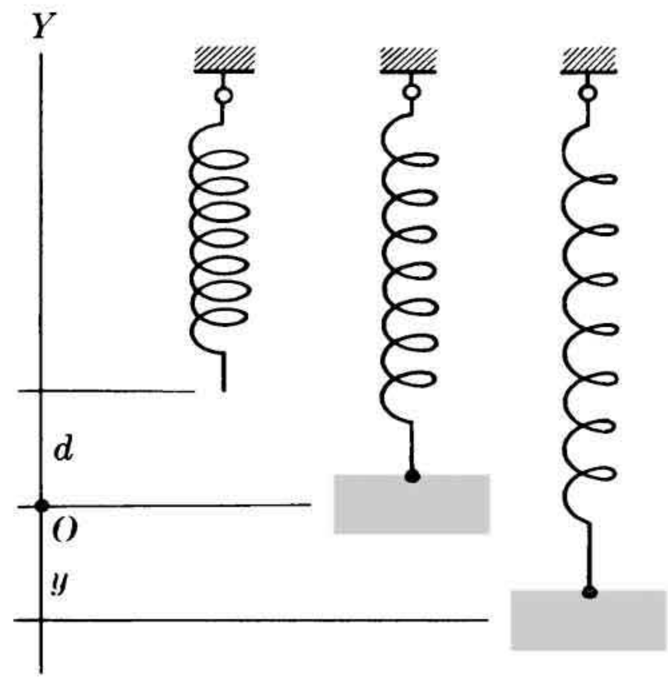

Fig. 18–14.
A bob on a spring in the rest position (center) and pulled down a distance *y* (right).

at the rest position these two forces just offset each other, we have

$$32m = kd. \tag{15}$$

Now suppose the spring is pulled downward an additional distance y. If we use the convention agreed upon in Section 18–2 that displacements above the rest position are to be positive and below the rest position negative, then the total extension of the spring is now $d - y$ because y itself is negative. The force that the spring exerts in an upward direction is, by Hooke's law,

$$k(d - y) \quad \text{or} \quad kd - ky. \tag{16}$$

However, the weight of the bob, or $32m$, exerts a constant downward force. Hence the net upward force is $kd - ky - 32m$. In view of equation (15) the *net* upward force is $-ky$. We now apply Newton's second law of motion, which says that when a force is applied to a mass, the force equals the mass times its acceleration. Thus we have

$$ma = -ky \tag{17}$$

or, by dividing both sides of this equation by m,

$$a = -\frac{k}{m}y. \tag{18}$$

Formula (18) is the basic law governing the motion of the bob on the spring. For, suppose the bob is pulled down some distance and then released. The spring exerts a force which pulls the bob back toward its rest position. The acceleration created by this force is precisely that given by (18). The acceleration now determines the velocity of the bob and the velocity determines the distance covered by the bob in any specified interval of time. The argument we are presenting here is in principle the same as the one we used in Chapter 13, where we discussed the motion of a body which is raised some distance from the surface of the earth and then dropped. In this situation, gravity immediately exerts an acceleration of 32 ft/sec², and thereafter the velocity and distance fallen by the object are determined. Of course, in the present case the acceleration is a more complicated expression, and the subsequent motion is not simply in one direction, but the argument is of the same nature.

We should now compare formula (13), namely,

$$a = -4\pi^2 f^2 y \tag{19}$$

and formula (18),

$$a = -\frac{k}{m}y. \tag{20}$$

In both cases the acceleration is a constant times the displacement. The constant is $4\pi^2 f^2$ in the former case and k/m in the latter. When the acceleration is given by (19), the motion itself [see (12) and (13)] is represented by

$$y = \sin 2\pi ft. \tag{21}$$

Since the acceleration (20) is precisely of the same form as (19), except for the label of the constant, and since the acceleration determines the motion, the motion of the bob must also be representable by a formula of the form (21).

However we do not know what f is in the case of the bob. But k/m in the case of the bob plays the role of $4\pi^2 f^2$ in the case of the motion of the point Q. That is

$$4\pi^2 f^2 = \frac{k}{m},$$

so that

$$f^2 = \frac{1}{4\pi^2}\cdot\frac{k}{m}$$

and

$$f = \sqrt{\frac{1}{4\pi^2}\cdot\frac{k}{m}} = \sqrt{\frac{1}{4\pi^2}}\,\sqrt{\frac{k}{m}}$$

or

$$f = \frac{1}{2\pi}\sqrt{\frac{k}{m}}. \tag{22}$$

In other words, if we let f in (21) be the value given by (22), we can write the formula for the bob's motion in terms of the quantities k and m. Thus if we substitute this value of f in (21), we obtain

$$y = \sin 2\pi \left(\frac{1}{2\pi} \sqrt{\frac{k}{m}} \right) t$$

or

$$y = \sin \sqrt{\frac{k}{m}}\, t. \tag{23}$$

We made one misleading statement in the preceding discussion. We said that the acceleration of the bob determines the motion of the bob, and so the formula of the bob's motion must be of the form (21). The acceleration does determine the essential characteristics of the motion, but the initial velocity and initial displacement do have some effect. This point may become clearer if we compare the present case with the vertical motion of objects. All bodies rising or falling near the surface of the earth are subject to an acceleration of 32 ft/sec^2, and this fact determines the essential nature of the motion. But if a body is thrown up into the air, the final formula depends also on the initial velocity given to the object and on the position from which it is thrown up. In the case of the bob, if it is pulled down to a distance D below the rest position and then released, this initial position must enter into the formula for the motion. To complete our determination of the formula we are obliged, with the limited mathematics at our disposal, to call upon observation, which tells us that in each oscillation the bob will rise to a height of D above the rest position and then descend a distance D below it. That is, the amplitude of the motion, D, is determined by the initial displacement. Thus the final formula for the motion of the bob is

$$y = D \sin \sqrt{\frac{k}{m}}\, t. \tag{24}$$

We can now draw several conclusions about the motion of the bob. Formula (22) gives the frequency of the bob's motion in terms of k and m. We see that the spring constant k, which represents the stiffness of the spring, and the mass m of the bob determine the frequency of the motion. If we wished to have the bob make, for example, two complete oscillations per second, we could pick values of k and m so that f in (22) should be 2. The period of the bob's motion, that is, the time required to make one oscillation, is

$$T = \frac{1}{f} = \frac{2\pi}{\sqrt{k/m}} = \frac{2\pi}{\sqrt{k/m}} \frac{\sqrt{m/k}}{\sqrt{m/k}} = \frac{2\pi\sqrt{m/k}}{\sqrt{(k/m)\cdot(m/k)}}$$

or

$$T = 2\pi \sqrt{\frac{m}{k}}\cdot \tag{25}$$

As Hooke observed, this formula for the period is immensely significant. The period is independent of the amplitude of the motion; that is, whether one pulls the bob down a great distance or a short distance and then releases it, the time for the bob to go through each complete oscillation will be the same.

This fact is immensely useful. At the very outset of our treatment of the bob's motion we pointed out that the resistance of the air and internal energy losses in the spring will cause the motion to die down. At the time we decided to ignore this fact and to suppose that there was no loss of energy. But there is. Suppose, however, that we were to give the bob a little upward push every time it reached its lowest position, i.e., add energy to the motion and keep the bob moving. Such an action might alter the amplitude, but would *not* affect the period, and each successive oscillation of the bob would therefore continue to take the same amount of time. Hence the oscillations of the bob on the spring can be used to measure time or to regulate the motion of some hands on a dial which would show time elapsed.

Of course, the motion of a bob on a spring is not quite the practical device for a clock. The device actually used can be found in every modern pocket or wrist watch. There a spring coiled in a spiral and carrying a weight called the balance wheel expands and contracts regularly. Each second the wheel is given a little "kick" which restores the energy the spring loses on each oscillation. (The energy comes from a mainspring which is wound up by hand usually once a day.) The spiral spring regulator was invented and patented by Christian Huygens in 1675. The first chronometer which was sufficiently accurate to be used by ships to determine longitude was invented by John Harrison, who in 1772 won a prize of £20,000 offered by the British government for such a device.

EXERCISES

1. If a mass of 2 lb pulls a spring down 6 in., what is the spring constant? [*Suggestion:* Use (15)]
2. Suppose that one attaches a mass of 2 lb to a spring whose stiffness constant is 50. Calculate the number of oscillations per second which the mass would make if set into vibration.
3. What is the period of a mass vibrating at a rate of 100 oscillations per second?
4. Suppose a mass is set to vibrating on a spring at the rate of 50 oscillations per second. If the mass has a maximum displacement of 3 in., what formula describes the motion?
5. Suppose a mass of 3 lb is attached to a spring whose stiffness constant is 75. The mass is pulled down 3 in. below the rest position and then released. Write a formula relating displacement and time.
6. Suppose that you are given a spring with a stiffness constant of 50. Calculate the mass that you would have to place on the spring to produce a period of oscillation of one second.

7. Suppose that a mass oscillates on a spring so that the relation between displacement and time is

   a) $y = 4 \sin 2\pi \cdot 5t$ b) $y = 4 \sin 10t$.

   Describe the motion of the mass for (a) and (b).
8. Suppose that you wished to decrease the number of oscillations per second which a mass makes on a given spring. How would you alter the mass?
9. Suppose that a tunnel is dug through the earth and a man of mass m steps into the tunnel. Inside the earth the force of gravity on a mass m at a distance r from the center is $F = GmMr/R^3$, where M is the mass and R is the radius of the earth. This force is directed toward the center. To distinguish distances above and below the center, let r be positive above and negative below the center. Then the acceleration acting on the mass m is $a = -GMr/R^3$. Discuss the subsequent motion of the man (Fig. 18–15).

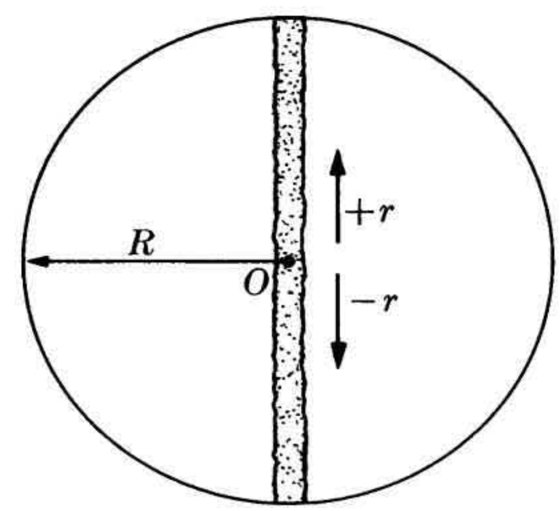

Fig. 18–15

18–6 SUMMARY

The mathematical objective of this chapter was to introduce a new type of mathematical function, the sinusoidal function. There is not just one sinusoidal function, for all functions of the form $y = D \sin 2\pi ft$, no matter what D and f may be, are sinusoidal. The sinusoidal functions are also called trigonometric functions because they are obtained by extending the concept of the sine of an angle, a concept which was first created and studied in trigonometry. Other trigonometric functions can be derived from an extension of the concepts of cosine and tangent of an angle and other trigonometric ratios which we did not study. All trigonometric functions are highly useful in scientific work.

The creation of trigonometric functions was motivated by the study of vibratory or oscillatory motion. We have used the motion of a bob on a spring to illustrate such a motion, and we have shown how the mathematical description of this motion can be used to deduce information about it. We have yet to see some of the major uses of sinusoidal functions.

Topics for Further Investigation

1. The mathematics of pendulum motion.
2. The trigonometric function $y = \cos A$.

Recommended Reading

BROWN, LLOYD A.: "The Longitude," in James R. Newman: *The World of Mathematics,* Vol. II, pp. 780–819, Simon and Schuster, Inc., New York, 1956.

KLINE, MORRIS: *Mathematics and the Physical World,* Chap. 18, T. Y. Crowell Co., New York, 1959. Also in paperback, Doubleday and Co., New York, 1963.

RIPLEY, JULIEN A., JR.: *The Elements and Structure of the Physical Sciences,* Chap. 15, John Wiley and Sons, Inc., New York, 1964.

TAYLOR, LLOYD WM.: *Physics, The Pioneer Science,* Chap. 15, Dover Publications, Inc., New York, 1959.

WHITEHEAD, ALFRED N.: *An Introduction to Mathematics,* Chaps. 12 and 13, Holt, Rinehart and Winston, Inc., New York, 1939.

THE TRIGONOMETRIC ANALYSIS OF MUSICAL SOUNDS

Motion appears in many aspects—but there are two obvious kinds, one which appears in astronomy and another which is the echo of that. As the eyes are made for astronomy so are the ears made for the motion which produces harmony: and thus we have two sister sciences, as the Pythagoreans teach, and we assent.

PLATO

19-1 INTRODUCTION

In this chapter we intend to show how trigonometric functions have given man his first real insight into the nature of musical sounds, and how this knowledge is utilized in the design of such devices as the telephone, the phonograph, the radio, and sound films.

The mathematical study of musical sounds did not start with the application of trigonometric functions. Indeed, it goes back to the very first emergence of any real mathematics and science, namely the beginning of the classical Greek period. For example, the Pythagoreans discovered that the lengths of two equally taut plucked strings whose sounds harmonize are related by simple numerical ratios such as 2 to 1, 4 to 3, and 3 to 2. The lower note in each case originates with the longer string. They also designed musical scales whose notes, as measured quantitatively by the lengths of the vibrating strings, possessed precise numerical values. From Pythagorean times onward, mathematicians and scientists were convinced that musical sounds had important mathematical properties, and music, along with arithmetic, geometry, and astronomy, became part of the quadrivium. These four subjects were studied together right through the medieval period. Although Greek, Arab, and medieval mathematicians continued to investigate musical sounds and wrote books on music, their work was essentially limited to the construction of new systems of scales for instrumental and vocal music.

It was the mathematicians and scientists of the seventeenth century who initiated other investigations and made the next series of important discoveries. Familiar names, such as Galileo, his French pupil and colleague Father Marin Mersenne (1588–1648), Hooke, Halley, Huygens, and Newton, obtained significant new results. Whereas the Pythagoreans had studied strings of different length but equal tension, Mersenne studied the effect of changing tension and mass of a string and found that an increase in mass and a decrease in tension produce lower notes in a string of given length. This discovery was very important for stringed instruments such as the violin and the piano; to secure the range of pitch which these instruments possess by variations in length only would require exceedingly long strings. Galileo and Hooke demonstrated experimentally that each musical sound is characterized by a definite number of air vibrations per second, a statement which will mean more to us in a few moments. The determination of the velocity of sound (about 1100 feet per second in air) was another achievement. It is of interest that the clocks which some of these men designed and constructed were essential to the progress made in the study of sound because, as we can see from the results cited, the ability to measure small intervals of time was an indispensable condition for any work in this field.

The best mathematicians of the eighteenth century, Leonhard Euler, Daniel Bernoulli (1700–1782), Jean le Rond d'Alembert (1717–1783), and Joseph Louis Lagrange, studied vibrating strings, such as the violin string, and vigorously disputed whether trigonometric functions were adequate to represent the vibrations. The mathematical analysis of sound waves soon followed and proved to be the chief tool in the theoretical mastery of musical sounds. We can readily see why mathematics was invaluable in these investigations, for observation of the air, even of air in the process of propagating sound, reveals nothing.

Before we undertake to study just what the nineteenth-century mathematicians and scientists learned, we must make some distinctions. The first is a matter of terminology. We shall be interested in the analysis of musical sounds as opposed to noise. However, in the present context, the term "musical sound" is used in a technical sense and includes not only those sounds commonly understood to be music, but also the sounds of ordinary speech. As a matter of fact, the physicist's meaning might be more appropriately represented by the term *intelligible sound.* Just what is meant by either phrase will be clear in a few moments.

The second distinction one must make is between sound as a motion of air and sound as a sensation which human beings experience. The former is a physical phenomenon which takes place in space and whose physical and mathematical properties are fixed. On the other hand, the sensations which human beings may receive because moving air strikes their ears and stimulates certain nerves depend upon their auditory mechanism and may vary from one person to another. There are, for example, physical sounds which humans cannot hear at all. Though we shall have something to say about the perception of sound, our first and main concern will be to understand the physical phenomenon.

19-2 THE NATURE OF SIMPLE SOUNDS

The variety of sounds given off by musical instruments, the human voice, phonographs, radios, and whirring machinery, for example, is so great that one cannot hope to study all of them in one swoop. Hence it would seem wise to start one's investigation with simple sounds. But which sounds are simple? If we rely upon our ears to decide this question, then the sounds given off by tuning forks seem to be simple. The ear may, indeed, be deceived here, but let us follow up this suggestion.

If either prong of a tuning fork is struck, both prongs will move inward and then outward very rapidly and will repeat this motion for a long time. Let us consider one prong, say the right one shown in Fig. 19–1. Before the prong is struck, it occupies what we might call the rest position. After being struck, the tip is displaced some distance to the right. It then moves to the left, to a position somewhat to the left of the rest position, and then moves to the right. The sequence then repeats itself many times. The displacement of the tip varies with time, and the first question one might raise is, What is the relationship between displacement and time? There are two considerations which suggest that the formula is sinusoidal: First of all, the prong resembles a spring-and-bob arrangement. The spring is the prong itself, though the motion is a sidewise oscillation rather than an expansion and contraction. The mass which corresponds to the bob is the mass of the prong itself, though admittedly this mass is not concentrated in one place as it is in the case of a bob on a spring. The second consideration is that as the tip of the prong moves farther and farther out from the rest position, the force which the prong exerts to return to the rest position may be expected to increase with the displacement. The simplest assumption one might make in this case is that the force increases directly with the displacement. From formula (17) of the preceding chapter we can see that this is indeed the mathematical law which underlies and determines the sinusoidal motion of the bob. Hence it seems reasonable to expect that the relation between displacement of the tip of either prong and time is sinusoidal. The amplitude of this relation is the maximum displacement of the tip, and the frequency is the frequency per second with which the prong oscillates.

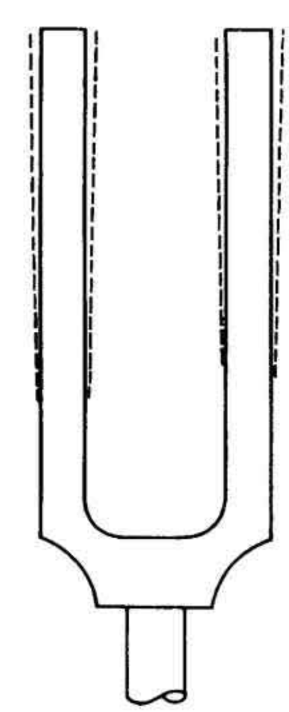

Fig. 19–1. A vibrating tuning fork.

Of course, we are not so much interested in the motion of the tuning fork as we are in the sound it creates. Hence what matters next is, How does the air respond to the vibration of the tuning fork? The fundamental fact about the behavior of air which is of importance in this connection is that air pressure seeks to become uniform everywhere. This means that if the air

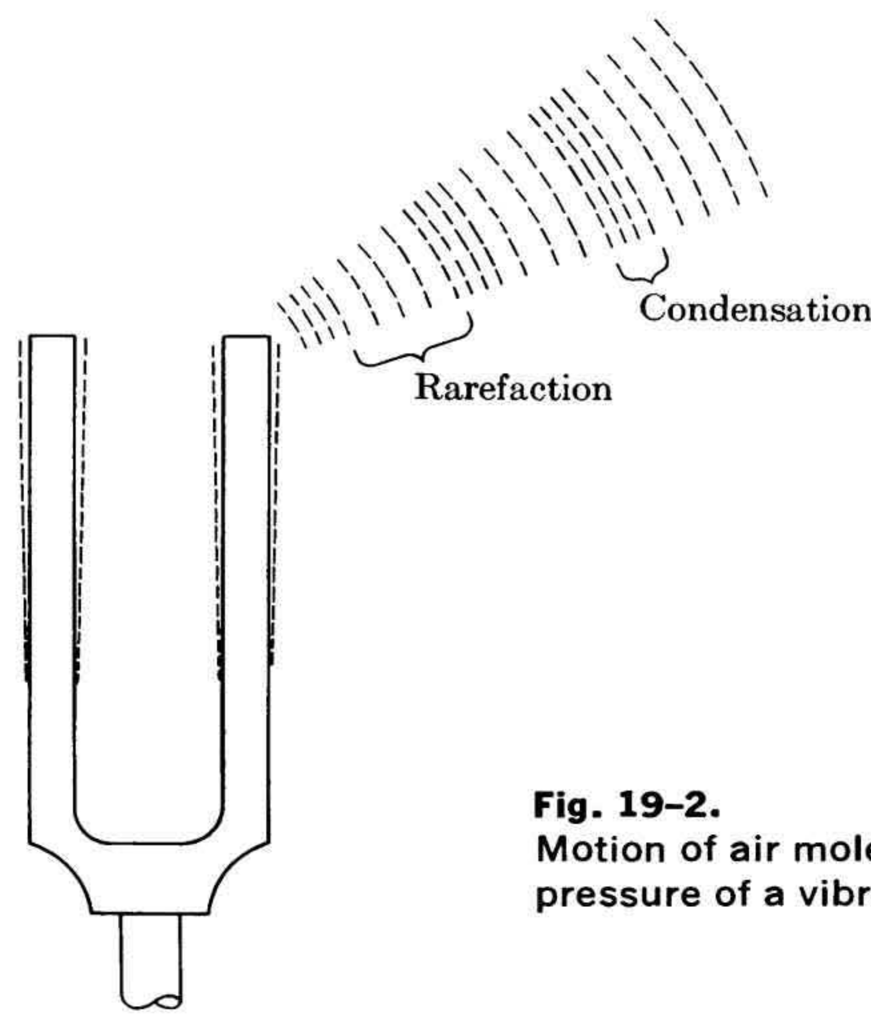

Fig. 19–2.
Motion of air molecules under pressure of a vibrating tuning fork.

pressure for any reason should become high in one place, the air will spread out from that place into neighboring regions where the pressure is lower and so try to equalize the pressure in the entire region under consideration. With this physical fact in mind, let us see what happens when the right prong of the tuning fork moves, say to the right. The prong pushes the molecules of air near it to the right and thus crowds them into a place occupied by other molecules. The pressure becomes high in this place, and since the molecules of air cannot move to the left because the prong is there, they will move off farther to the right (and in other directions) in order to equalize the pressure. But this motion means that the crowding now occurs a little farther away from the tuning fork and again, to equalize the pressure, the molecules move farther to the right. The process continues, and the crowding, or *condensation*, as it is usually called, moves off to the right.

The prong, having moved as far to the right as it can, will now move back not only to its rest position but farther to the left. This motion leaves an empty region—the place that the prong had occupied—and so the molecules of air on the right rush into this empty space. Molecules still farther to the right also move to the left because the pressure has become less to their left. Thus a state of low pressure, or *rarefaction*, as it is called, moves to the *right* as molecules move to the left to equalize the pressure in their neighborhood. With each successive vibration of the prong, a condensation and rarefaction move off to the *right* (Fig. 19–2). The successive condensations and rarefactions also move out in other directions, but it is sufficient for our purposes to follow what happens in one direction.

The action of the air is somewhat complicated because it consists of billions of molecules, and they do not all behave in exactly the same way. However, there is an average effect. It is convenient to speak of a series of typical molecules to the right of the prong which represent the average behavior of the entire collection. If we consider the action of any one typical molecule, say one near the prong, then what it does is to move to the right when the prong moves to the right. When the prong moves to the left, the typical molecule will also move to the left because the air pressure has been lowered. Like the prong, it will move past its rest position, and continue to the left. Then, as the prong moves to the right, the molecule will be pushed to the right, will pass its rest position, go farther to the right, and, from this time on, it will continue to oscillate.

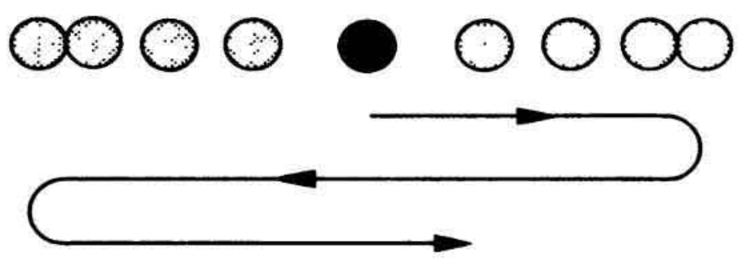

Fig. 19–3.
Motion of a typical molecule.

Typical molecules farther to the right will behave like the typical molecule near the prong; however, their reaction will be slightly delayed since condensations and rarefactions reach them a little later. Figure 19–3 illustrates the motion of a typical molecule reacting to a series of prong oscillations.

Two important facts emerge from the above discussion. The first is that the average, or typical, molecule follows, in effect, the motion of the prong. Any one molecule acts as if it were attached to the prong by a spring. When the prong moves to the right, it contracts the spring. The latter seeks to restore its length and so pushes the molecule to the right. While the molecule moves to the right, the prong moves to the left, and hence the spring is extended. It now seeks to contract and so pulls the molecule to the left. The molecule moves to the left, and the spring contracts. But now the prong is ready to move to the right, and consequently the motion of prong and molecule repeats itself. The action of air pressure is indeed like the action of the spring. In fact, Hooke used the phrase, "the spring of the air," to describe the effect of air pressure.

The second fact is that the sound wave which moves from the prong to some person's ear, say, consists of the series of condensations and rarefactions induced by the prong's motion. Each molecule merely oscillates about its rest position, but in doing so it produces the increase and reduction of pressure which cause the neighboring molecules to oscillate.

The nature of the sound wave may perhaps be made clearer by comparing it with a water wave. If the end of a stick is quickly moved back and forth in still water, a series of waves will spread out from the end of the stick.

However, the individual water molecules do not move out. Each oscillates about its original position, but the increase and decrease in pressure which the stick creates cause the molecules farther away to duplicate the motion of the molecules near the stick.

Since the motion of any typical molecule whether near or far from the prong is the same, let us study the motion of any one of these molecules. Specifically, let us seek the relationship between the displacement from rest position and the time that the molecule is in motion. What formula relates displacement and time? We have already produced two crude physical arguments suggesting that for the prong, displacement and time are related by a sinusoidal formula. Since the motion of any typical air molecule duplicates the action of the prong, the formula which relates displacement and time for the typical air molecule should also be sinusoidal. Actually these physical arguments do not really prove that the formula is sinusoidal. However, this fact can be established either by a rather complex mathematical analysis of air motion or, experimentally, by converting the air pressure to electric current (by means of a microphone, for example) and by then displaying the current on a cathode-ray tube (television tube).

We shall take for granted that for a typical air molecule, the formula relating displacement and time is sinusoidal. Hence, if y is the displacement and t is the time, then, in view of what we have learned in the preceding chapter, the formula is

$$y = D \sin 2\pi f t, \tag{1}$$

where D is the amplitude, or maximum displacement, and f is the number of oscillations, or the frequency of cycles per second. We wish to emphasize that the formula applies to the sounds produced by tuning forks or to what we have reason to believe are simple sounds.

To use formula (1), we must know D and f. The value of f is the frequency with which the tuning fork oscillates. A frequency commonly used to standardize the pitch of sounds is 440 per second. This then is a typical value of f. The value of D, the amplitude of the motion of a typical air molecule, is *not* the amplitude of the prong's motion, but depends upon the medium in which sound spreads out or is propagated. It depends, so to speak, on the "springiness" of the medium. In air, 0.001 inch can be considered to be a reasonable value for D. Hence a typical formula for a simple sound is

$$y = 0.001 \sin 2\pi \cdot 440t. \tag{2}$$

Thus, a typical air molecule oscillating in accordance with formula (2) shuttles back and forth about its mean, or rest, position 440 times per second or, as we say, it goes through 440 complete cycles in one second. The farthest distance from the mean position that it reaches, that is, the amplitude of its motion, is 0.001 inch.

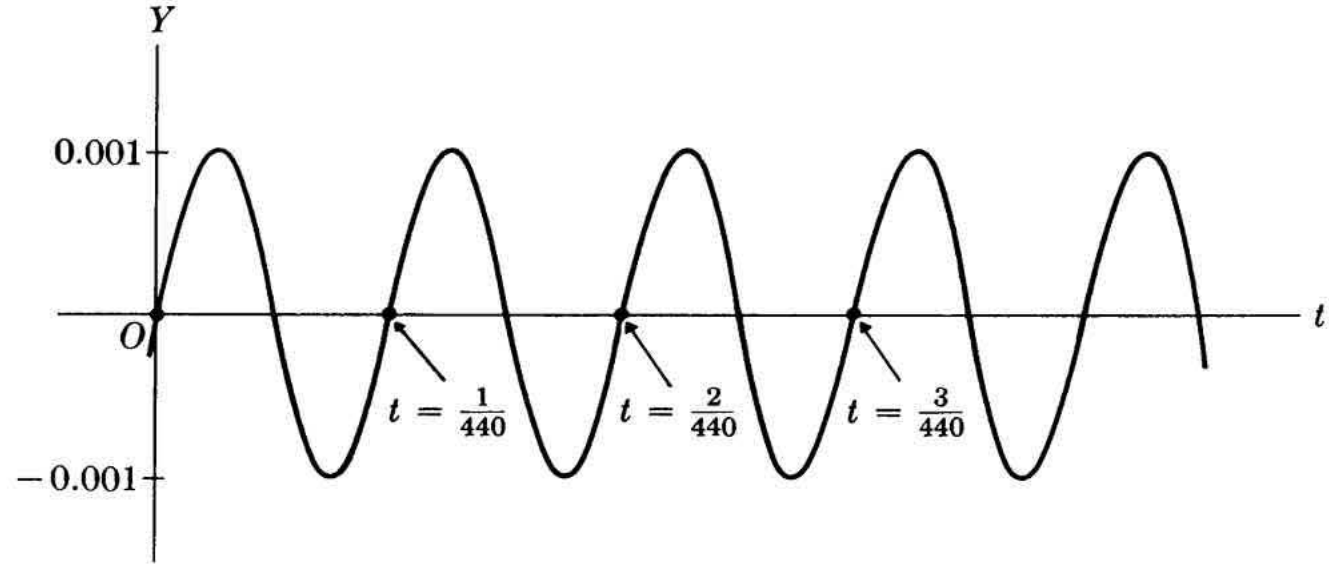

Fig. 19–4.
Graph of displacement versus time of a typical molecule executing 400 oscillations per second.

Figure 19–4 illustrates the relationship between displacement and time for a simple sound such as formula (2) represents. Of course the typical molecule shuttles backward and forward, but on the graph its displacements are plotted as ordinates and the time elapsed is shown by the corresponding abscissas.

Although formula (2) represents only simple sounds—we have yet to discuss the formulas describing more complicated sounds—it enables us to understand what we meant earlier by the phrase "intelligible sounds." We see that a simple sound has a regularity or periodicity. The motion of the air molecules repeats itself a number of times a second. When the ear receives many cycles of this motion, it can identify the sound. If, on the other hand, the motion of the air molecule is not regular but varies irregularly with time, the ear still hears sound, but sound that does not convey any meaning, i.e., noise.

EXERCISES

1. What is the basic mathematical formula which represents simple sounds? State the physical meaning of the various letters in the formula.
2. State the formula which describes the relationship between displacement and time for a simple sound whose frequency is 300/sec and whose amplitude is 0.0005 in.
3. If $y = 0.002 \sin 2\pi \cdot 540t$ is the mathematical description of a sound, what are the frequency and amplitude of this sound?
4. If a sound has a frequency of 400 cycles/sec, how many cycles would the ear receive in 1/20 sec?

19–3 THE METHOD OF ADDITION OF ORDINATES

We have now a good mathematical representation of simple sounds. But interesting musical sounds, whether vocal or instrumental are, as a rule, not simple, and the really significant contribution of mathematics to the understanding of musical sounds lies in the analysis of more complex sounds. To comprehend this contribution, we must first examine a relevant mathematical idea. Instead of considering simple sinusoidal functions such as (2), let us take the function

$$y = \sin 2\pi t + \sin 4\pi t. \tag{3}$$

What sort of relationship between y and t does formula (3) represent?

A good way to investigate this question is to draw a graph of the above function. Since we wish to obtain merely some general idea of how y varies with t, we shall seek only a sketch rather than a very accurate graph. We could proceed by selecting values of t, calculating the corresponding values of y, and then plotting the points whose coordinates have thus been determined. However, there is a quicker method which is also more perspicuous. Let us consider the two functions:

$$y_1 = \sin 2\pi t \tag{4}$$

and

$$y_2 = \sin 4\pi t. \tag{5}$$

We have used the notation y_1 and y_2 to distinguish the dependent variables in (4) and (5) from the y in formula (3). Formulas (4) and (5) are easily graphed. Formula (4) is the ordinary sine function which goes through the regular cycle of sine values in each unit of t. Formula (5) has a frequency of 2 in each unit of t; that is, the y-values go through the complete cycle of sine values twice in each unit of t. Let us sketch both functions on the same set of axes (Fig. 19–5).

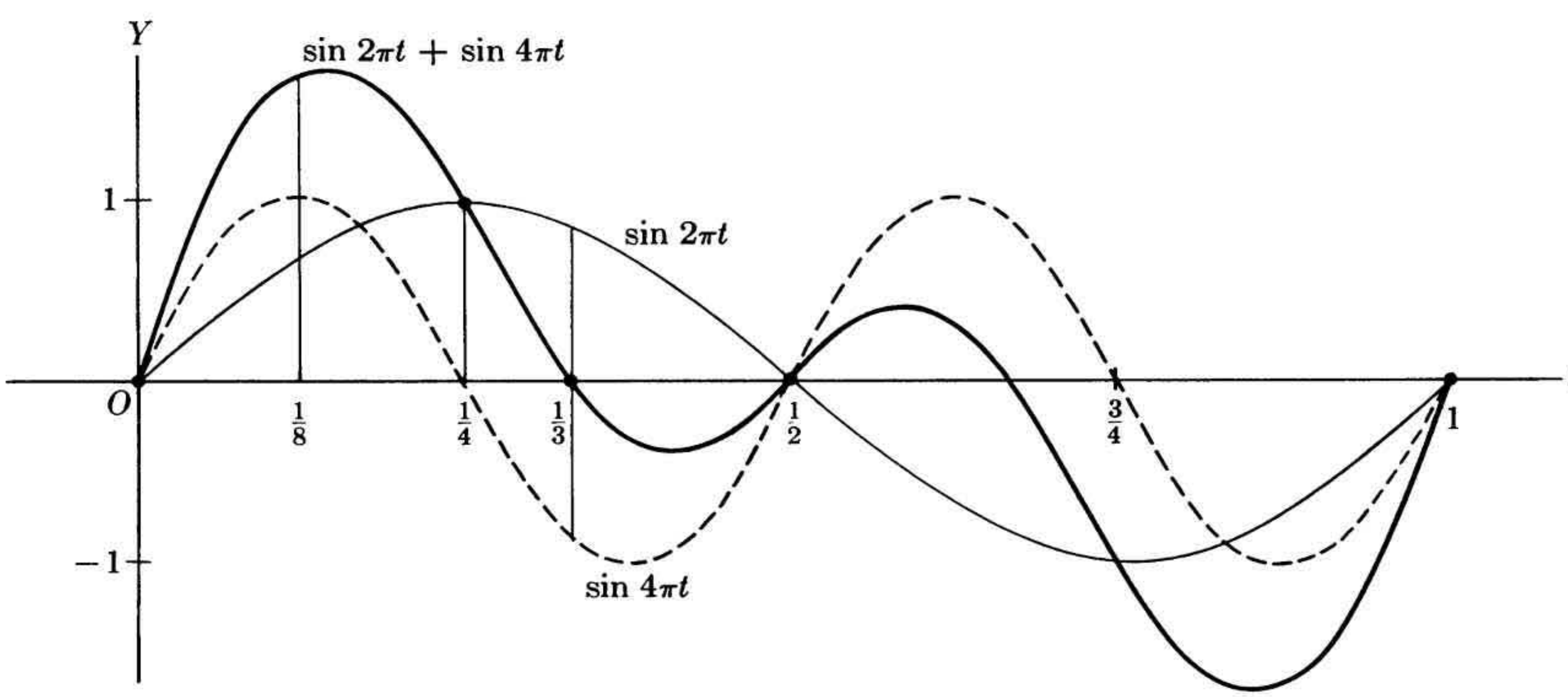

Fig. 19–5.
The graph of $y = \sin 2\pi t + \sin 4\pi t$ obtained by addition of ordinates.

Now the y of formula (3) is clearly the sum of y_1 and y_2. Hence adding the values of y_1 and y_2 at various values of t will yield y. Since we are interested only in a sketch, let us perform the addition by using Fig. 19–5 to obtain the values of y_1 and y_2. Thus for $t=0$, the graphs show that y_1 and y_2 are both zero. Hence y, the sum of y_1 and y_2, also is zero. At $t=\frac{1}{8}$, we see from the graph that y_1 is about 0.7 and y_2 is 1. Hence $y=1.7$ when $t=\frac{1}{8}$. At $t=\frac{1}{4}$, we find that $y_1=1$ and $y_2=0$. Hence $y=1$ when $t=\frac{1}{4}$. At $t=\frac{1}{3}$, y_1 is about 0.85 and y_2 is about -0.85. In adding the last two values for y_1 and y_2, we must take into account that one is positive, the other negative, and their sum zero. Hence at $t=\frac{1}{3}$, $y=0$. By selecting a few more values of t and estimating the corresponding y_1- and y_2-values, we can obtain more y-values. Finally, we join the various points which belong to the graph of formula (3) by a smooth curve. The result is the heavy-lined curve shown in Fig. 19–5. The method just described for graphing y as a function of t provides a rough sketch. If one wishes to obtain a more accurate graph, he can calculate the value of y for each value of t.

How far need we carry this process of determining y-values corresponding to various t-values? We note that the function $y_1=\sin 2\pi t$ repeats itself when t becomes larger than 1. The function $y_2=\sin 4\pi t$ goes through two full cycles in the interval from $t=0$ to $t=1$ and begins its third cycle of sine values as soon as t increases beyond 1. Thus at $t=1$ both functions begin to repeat the values which they had taken on at $t=0$ and, in the interval from $t=1$ to $t=2$, both functions will repeat the behavior exhibited in the interval from $t=0$ to $t=1$. Since y_1 and y_2 repeat their former behavior, it follows that y, which is the sum of y_1 and y_2, will also repeat its former behavior. In other words, in the interval from $t=1$ to $t=2$, y will behave precisely as it did in the interval from $\mathrm{t}=0$ to $\mathrm{t}=1$. And in each succeeding unit interval of t-values, the function will repeat the behavior exhibited in the interval from $t=0$ to $t=1$. If we therefore determine the behavior of y in the interval from 0 to 1, we know how it behaves for all larger values of t.

There are several major facts to be learned from this example. First of all, since the function (3) repeats its behavior in every unit of t-values, it is periodic. Moreover, because the term $\sin 4\pi t$ goes through two cycles of sine values in exactly the t-interval in which $\sin 2\pi t$ goes through one cycle, the entire function repeats itself with the frequency with which $y=\sin 2\pi t$ repeats itself. Hence the frequency of formula (3) is one cycle per unit of t. Thirdly, the shape of the graph of formula (3) shows that the formula is *not* sinusoidal even though it is periodic. In other words, the sum of two sine functions can yield a function whose shape is quite different from that of a sine function, but the sum can nevertheless repeat itself.

We might expect that functions built up of three or more sine functions could have quite strange shapes and yet be periodic if the summands *all* began to repeat at some value of t, say $t=1$, the values they had taken on at $t=0$.

EXERCISES

1. By following the method described in the text sketch the graph of

   a) $y = \sin 2\pi t + \sin 6\pi t$, b) $y = \sin 2\pi t + \frac{1}{2}\sin 4\pi t$, c) $y = \sin 2\pi t + \sin 3\pi t$.

2. What is the frequency, in one unit of t, of the function

   a) $y = \sin 2\pi t + \sin 8\pi t$,

   b) $y = 2 \sin 2\pi t + \sin 4\pi t$,

   c) $y = \sin 2\pi t + \sin 4\pi t + \sin 6\pi t$,

   d) $y = \sin 2\pi \cdot 100t + \sin 2\pi \cdot 200t + \sin 2\pi \cdot 300t$?

19-4 THE ANALYSIS OF COMPLEX SOUNDS

We have already mentioned that the sounds given off by almost all musical instruments and by the human voice are not simple sounds; that is, they are not representable by functions of the form (1). Yet these sounds are intelligible, which means that they must be periodic or that the pattern of displacement versus time must repeat itself. The shapes of the curves which represent such sounds are, however, quite varied. In fact, to each sound there corresponds a characteristic shape. For example, Fig. 19-6 shows the shape corresponding to the sound of a piano note C. To obtain this graph, the sound is converted to electric current and the vibration of the current is made visible by means of a cathode-ray tube. In view of the variety of musical sounds it may seem that we have reached an impasse in our attempt to analyse all such sounds mathematically. But by a stroke of good luck mathematics provided the very theorem which gives us remarkable insight into all complex sounds. The stroke of good luck was the mathematician Joseph Fourier (1768–1830).

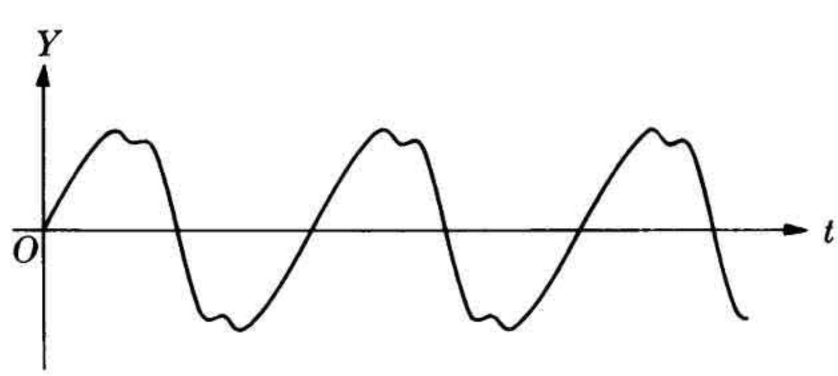

Fig. 19-6. Displacement versus time of a typical molecule for the note C on a piano.

Fourier was the son of a French tailor. While attending a military school he became intrigued with mathematics. Since he realized that his low birth would not permit him to become an army officer, he let himself be persuaded by members of the Church to study for the priesthood. However, he abandoned the priesthood to accept a professorship of mathematics at the military school that he had attended. Later he became a professor at the École Normale and at the École Polytechnique, universities founded by Napoleon.

Fourier's main interest was mathematical physics, and his most important work in that domain concerned the conduction of heat; for example, he studied how heat travels along metals. His chief contribution, a book entitled *The Analytical Theory of Heat* (1822), is one of the great classics of mathematics. In the development of the theory of heat Fourier established a mathematical theorem whose value extends far beyond the physical application for which it was intended. Our interest in the theorem lies in what it does to analyze complex musical sounds.

Fourier's celebrated theorem says that any periodic function is a sum of simple sine functions of the form $D \sin 2\pi ft$. Moreover, the frequencies of these component functions are all integral multiples of one frequency. To illustrate the significance of this theorem, let us suppose that y is a periodic function of t. Then the formula which relates y and t must be of the form

$$y = \sin 2\pi \cdot 100t + 0.5 \sin 2\pi \cdot 200t + 0.3 \sin 2\pi \cdot 300t + \cdots \qquad (6)$$

The numbers in this formula depend, of course, on the choice of the initial periodic function, but let us suppose that they are correct and see what they stand for. The numbers 1, 0.5, 0.3 are the amplitudes of the respective sinusoidal components of the entire periodic function. The lowest frequency per second, that of the first term, is 100. The second term has frequency 200, or twice the lowest frequency. The third term has frequency 300, or three times the lowest frequency, and so on. The dots at the end of formula (6) imply that we might need additional terms like the ones shown, to represent any given periodic function. In accordance with the theorem, all frequencies occurring in such additional terms must be multiples of 100.

Before we consider the significance of Fourier's theorem for the study of musical sounds, we should satisfy ourselves that formulas such as (6) do represent periodic functions. In this connection, two results of our work in Section 19–3 should be helpful. We learned there that the sum of two sine functions can produce a rather peculiarly shaped but nevertheless periodic graph. Moreover, because the second term in formula (3) had twice the frequency of the first one, the frequency of the *entire* function was the lower of the two frequencies. The situation in (6) is very much the same. It is a sum of sine terms, and the graph of this sum may indeed have a peculiar or irregular shape. But the shape will repeat itself because during the time that the first term goes through one cycle, namely the interval $t = 0$ to $t = \frac{1}{100}$, the second term will go through two cycles, and the third term through three, so that the entire function will repeat itself as soon as the first term does. Since the frequency of the first term is 100 in one unit of t, the entire function has the frequency of the first term.

And now what does Fourier's theorem have to do with the analysis of musical sounds? The application of this theorem to music was made by a German, Georg S. Ohm, a teacher of mathematics and physics, who lived in the first half of the nineteenth century. As pointed out earlier in this section, every musical sound is a periodic function; that is, the relation between displacement and time of a typical air molecule oscillating under the pressure exerted originally by the source of the sound is a periodic function of t. But Fourier's theorem says that every such function is a sum of simple sine functions of the type illustrated in (6). Each simple sine function corresponds to a simple sound such as is given off by a tuning fork. Hence one arrives at the important conclusion that *every* musical sound is a sum of simple sounds. Moreover, the frequencies per second of these simple sounds are all multiples of one lowest frequency. To put the matter differently, every musical sound can be duplicated by a combination of tuning forks, each vibrating with the proper frequency and amplitude.

The musical sound whose graph is shown in Fig. 19–6, for example, is a sum of five simple sounds. The frequencies of these sounds and their respective amplitudes are tabulated below. The amplitudes are expressed in terms of the first one entered, which is chosen to be 1.

Frequency	512	1024	1536	2048	2560
Amplitude	1	0.2	0.25	0.1	0.1

We should note that the frequencies are all multiples of the lowest one, which is 512. The formula representing this sound is then

$$y = \sin 2\pi \cdot 512t + 0.2 \sin 2\pi \cdot 1024t + 0.25 \sin 2\pi \cdot 1536t + 0.1 \sin 2\pi \cdot 2048t + 0.1 \sin 2\pi \cdot 2560t.$$

The assertion that every musical sound is no more than a combination of simple sounds is so surprising that, although it is backed by unassailable mathematics, one wishes to see it confirmed by experimental evidence. Such evidence is available. First of all, a trained ear can recognize the simple sounds present in a complex sound. Secondly, if one releases the dampers on the strings of a piano and then strikes a note, a number of other strings will also begin to vibrate, namely those whose basic frequencies are the same as the component frequencies present in the note struck. The physical explanation is that the note struck gives off several frequencies—the frequencies of its component simple sounds. Each of these frequencies sets off air vibrations which in turn force into vibration all other strings whose basic frequencies are the same as those of the simple sounds.

Perhaps the best experimental evidence is furnished by some specially designed instruments. The distinguished nineteenth-century physician, physicist, and mathematician Hermann von Helmholtz (1821–1894) gave two kinds of demonstrations. In the first one he designed special pipes, called resonators, each of which selected and rendered audible only that frequency which was suited to the dimensions of the pipe. A resonator in the neighborhood of a complex sound will pick up and render audible any component of the sound whose frequency excites the resonator. By using resonators of different sizes Helmholtz was able to show that the frequencies present in the complex sound were just those called for by Fourier's theorem. Then Helmholtz demonstrated the reverse. He set up electrically driven tuning forks of the proper frequency and amplitude such that the combination of simple sounds duplicated a given complex sound. A modern version of this latter device is the electronic music synthesizer.

There is no question, then, that any musical sound is no more than a sum of simple or sinusoidal sounds. The simple sound of lowest frequency is called the fundamental, or first partial, or first harmonic. The simple sound whose frequency is twice that of the lowest one is called the second partial or second harmonic; and so on. The frequency of the entire complex sound is the frequency of the first harmonic for the reason already given in our discussion of Fourier's theorem. The amplitudes of the individual sine terms are the amplitudes or strengths of the harmonics present.

EXERCISES

1. State Fourier's theorem.
2. Suppose that a complex sound is representable by the function
$$y = 0.001 \sin 2\pi \cdot 240t + 0.003 \sin 2\pi \cdot 480t + 0.01 \sin 2\pi \cdot 720t.$$
What is the frequency of the complex sound? What is the amplitude of the third harmonic?
3. Write the formula for a musical sound whose frequency is 500/sec and whose first, second, and third harmonics have amplitudes of 0.01, 0.002, and 0.005, respectively.
4. If the relationship between displacement and time for the fundamental of a musical sound is $y = 3 \sin 2\pi \cdot 720t$, what is the frequency of the third harmonic?
5. Explain why the frequency of a complex musical sound is always that of the first harmonic.

19-5 SUBJECTIVE PROPERTIES OF MUSICAL SOUNDS

Musical sounds as received by the ear seem to possess three essential properties; that is, the ear recognizes what are commonly called the pitch, the loudness, and the quality of a sound. One of the major values of the mathematical analysis of musical sounds is that it clarifies and makes precise just what we mean by these properties. We shall consider them in turn.

In our subjective judgment, sounds vary from low or deep tones to high or piercing ones. Verbal descriptions of the pitch of sounds are, of course, qualitative and vague. If one experiments with tuning forks of different pitch, he readily discovers that high pitch means high frequency of fork vibration and therefore high frequency of oscillation of the air molecules. Correspondingly, low pitch means that the fork and the air molecules vibrate with low frequencies. Prior to the availability of the analysis examined in the preceding section, the notion of pitch was not clear for complex sounds. But we now know that all musical sounds have a definite frequency, namely the frequency of the fundamental. Thus, although complex sounds contain other frequencies, that is, the frequencies of the higher harmonics, it is the over-all frequency of the composite sound which determines whether it appears high- or low-pitched to the ear. For example, as one strikes the notes on a piano going from left to right, the fundamental frequency steadily rises.

The loudness of a musical sound is determined by the amplitude of the corresponding molecular motion, but the relationship between loudness and amplitude is not quite so simple as that between pitch and frequency. Let us note, first of all, that amplitude means the maximum displacement of the typical air molecule or the largest y-value of the corresponding graph. Physicists call the square of this amplitude the intensity of the sound. Thus intensity is still a physical or objective property of a musical sound. Among sounds of a given frequency, the more intense sound will seem louder to the ear. However, this is no longer true if the frequencies of the sounds differ. The average ear is most sensitive to a frequency of about 3500 per second and less so to frequencies above and below this value. Hence a very intense sound at a frequency of 1000 per second may sound softer to the ear than a less intense one at a frequency of 3500 per second. As a matter of fact the average human ear does not hear at all sounds above about 16,000 vibrations per second, no matter how intense they are. Loudness depends not only on the intensity and the frequency of the sound but also on the shape of the graph within any one period. Two sounds may possess the same frequency and the same amplitude, but may have differently shaped graphs. Such sounds will, in general, not sound equally loud to the ear.

The most interesting and from an aesthetic standpoint the most important aspect of musical sounds is their quality. It is this property which determines whether or not a sound is pleasing. The quality of a sound depends upon which harmonics are present in the sound and the amplitudes of these harmonics. Thus a sound emitted by a piano and a sound of the same frequency emitted by a violin create different effects on the ear because they differ in the harmonics present and in the amplitudes of these harmonics. Since the harmonics and their amplitudes determine the shape of the graph, it follows that, mathematically, the quality of a sound is the shape of the graph within any one period.

Sounds or tones vary greatly with respect to harmonics and their amplitudes. Some sounds, for example, the sounds of tuning forks, some notes on the flute, and sounds produced by wide-stopped organ pipes, possess only a few harmonics or, in effect, merely the first. On the other hand, most instruments give off sounds containing many harmonics, but some of these may have small or almost zero amplitude. For example, the sounds of organ pipes are, in general, weak in the higher harmonics. The sounds of a violin possess a great number of harmonics and are usually strong in the first six harmonics. The relative amplitudes of the harmonics present in violin sounds are about the same for all notes; however, there are enough differences for the ear to distinguish, say the A- from the D-string, even though both are sounded at the same frequency. The uniformness of quality may explain why the sounds of a violin are so pleasing. The sounds of a piano also contain many harmonics, but the relative amplitudes of the harmonics in any one sound depend upon the velocity with which the hammer strikes the string.

The vowel sounds of the human voice are rich in harmonics. For example, the sound of "oo" as in tool, expressed at a fundamental frequency of 125 vibrations per second, has as many as 30 detectable harmonics. The relative amplitudes of the first six are 0.4, 0.7, 1, 0.2, 0.2, and 0.2, respectively. The higher harmonics, though present, have lower amplitudes. However, not only do the number of harmonics present and their relative amplitudes vary considerably from one vocal sound to another, but even the same sound issued at two different pitches will have different harmonics and amplitudes.

The physical reason for the differences in quality among the many types of musical instruments is, of course, the nature of the device itself. The piano and violin both use vibrating strings, but piano strings are struck whereas violin strings are bowed. The clarinet, oboe, and bassoon are operated by forcing air against vibrating reeds. Air is forced past the edge of an opening in the organ pipe also, but here the edge or lip is rigid. In addition, each instrument possesses a resonance device which emphasizes certain harmonics. The sounding board of the piano, the hollow box of the violin, and the pipes of an organ are resonance devices.

Although two people may not quite agree about their reactions to sounds, it is on the whole true that the qualities of sounds which we describe by such words as soft, piercing, rich, dull, braying, hollow, bright, and the like, are due to the harmonics and their relative amplitudes. Sounds which contain only the first harmonic are soft but dull. For brightness and acuteness the higher harmonics are essential. Sounds which possess the first six harmonics are grand and sonorous. If harmonics beyond the sixth or seventh are present and have appreciable amplitudes, the tones are piercing and rough. In general, the amplitudes of harmonics decrease as the frequencies increase. However, if the amplitudes of higher harmonics are too large compared with that of the fundamental, the tone is described as poor rather than rich.

EXERCISES

1. Suppose that two sounds are represented respectively by $y = 0.06 \sin 2\pi \cdot 200t$ and $y = 0.03 \sin 2\pi \cdot 250t$. Which one is louder? Which one is higher pitched?
2. Explain in mathematical terms the meaning of a simple sound.
3. What is the mathematical criterion of a musical sound as opposed to noise?
4. Which mathematical properties of the formula for a complex musical sound represent the pitch of the sound and the quality of the sound?
5. Discuss the assertion that music is basically just mathematics.

Topics for Further Investigation

1. The construction of musical scales. Include, in particular, the work of J. S. Bach on the equal-tempered scale.
2. The human voice as a source of musical sounds.
3. The functioning of the human ear.

Recommended Reading

BENADE, ARTHUR H.: *Horns, Strings and Harmony*, Doubleday and Co., New York, 1960.

FLETCHER, HARVEY: *Speech and Hearing*, D. Van Nostrand Co., Princeton, 1929.

HELMHOLTZ, HERMANN VON: *On the Sensations of Tone*, Dover Publications, Inc., New York, 1954.

JEANS, SIR JAMES H.: *Science and Music*, Cambridge University Press, London, 1937.

MILLER, DAYTON C.: *The Science of Musical Sounds*, 2nd ed., The Macmillan Co., New York, 1926.

OLSON, HARRY F.: *Musical Engineering*, McGraw-Hill Book Co., Inc., New York, 1952.

REDFIELD, JOHN: *Music, A Science and an Art*, A. A. Knopf, Inc., New York, 1926.

SEARS, FRANCIS W. and MARK W. ZEMANSKY: *University Physics*, 3rd ed., Chaps. 21 to 23, Addison-Wesley Publishing Co., Inc., Reading, Mass., 1964.

TAYLOR, LLOYD WM.: *Physics, The Pioneer Science*, Chaps. 24 to 28, Dover Publications, Inc., New York, 1959.

VON BERGEIJK, WILLEM A., JOHN R. PIERCE and EDWARD E. DAVID, JR.: *Waves and the Ear*, Doubleday and Co., New York, 1960.

WOOD, ALEXANDER: *The Physics of Music*, 6th ed., Dover Publications, Inc., New York, 1961.