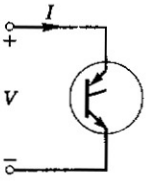
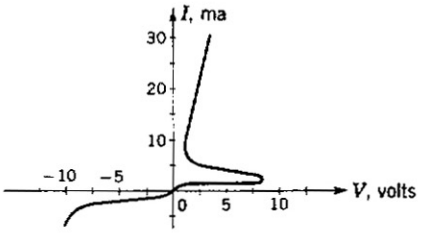
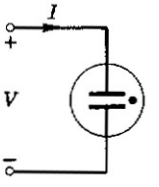
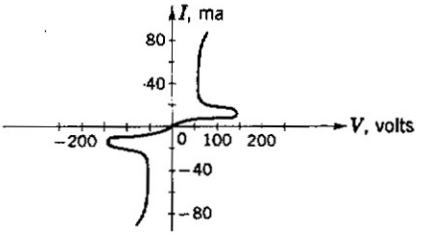
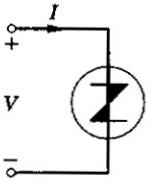
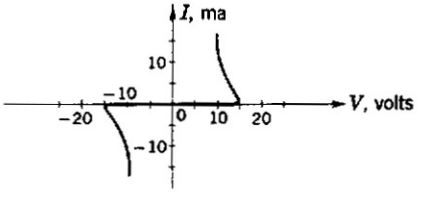
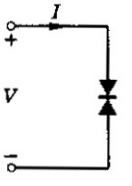
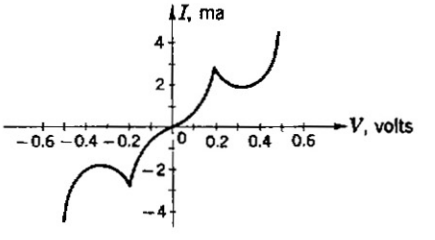
TABLE 1-1 (*Continued*)

Name	Symbol	v-i characteristic curve
Four-layer diode		
Glow tube		
Trigger diode		
Superconducting tunnel junction		

and is therefore called a *current-controlled resistor.*[1] For example, the tunnel diode is a voltage-controlled resistor but the glow tube is a current-controlled resistor. Observe that a strictly monotonically increasing resistor is both voltage-controlled and current-controlled and can therefore be described either in the form of Eqs. (1-16) or (1-17).

[1] More generally, any curve in the x-y plane is said to be an x-controlled curve if it is a single-valued function of x and a y-controlled curve if it is a single-valued function of y.

Another important property shared by some v-i curves is their symmetry with respect to the origin. Such elements are called *bilateral resistors* because in this case the two terminals may be interchanged without affecting the v-i curve (see Prob. 1-1). For the resistors listed in Table 1-1, the varistor, the glow tube, the trigger diode, and the superconducting tunnel junction are the only bilateral resistors. The rest are nonbilateral.

Exercise 1: It is sometimes convenient to describe a voltage-controlled resistor by an equation of the form $i = G(v)v$ and a current-controlled resistor in the form $v = R(i)i$. Find the functions $G(v)$ and $R(i)$ in terms of $i(v)$ in Eq. (1-16) and $v(i)$ in Eq. (1-17). Give a geometrical interpretation of $G(v)$ and $R(i)$.

Exercise 2: Is a monotonically (but not strictly) increasing resistor both current-controlled and voltage-controlled? If not, under what condition is it voltage-controlled? When is it current-controlled?

Exercise 3: A resistor which is neither voltage-controlled nor current-controlled is said to be a *multivalued resistor*. Give an example of a multivalued resistor. Can you describe a multivalued resistor in the form of Eq. (1-16) or (1-17)? Explain why. (See Appendix A.)

1-6-4 v-i CURVES OF DC SOURCES AND IDEAL DIODES

On many occasions we shall find it convenient to consider a dc voltage source and a dc current source as nonlinear resistors. This interpretation is valid because, by definition, a dc voltage source with terminal voltage E can be represented by a vertical line $v = E$ as shown in Fig. 1-10a. Similarly, a dc current source with terminal current I can be represented by a horizontal line $i = I$, as shown in Fig. 1-10b. In the special case where $E = 0$, the v-i curve of Fig. 1-10a becomes the $v = 0$ axis as shown in Fig. 1-10c. Since this coincides with the v-i curve of a short circuit, *a voltage source with zero terminal voltage is equivalent to a short circuit.* Similarly, when $I = 0$, the v-i curve of Fig. 1-10b becomes the $i = 0$ axis, as shown in Fig. 1-10d. Since this coincides with the v-i curve of an open circuit, *a current source with zero terminal current is equivalent to an open circuit.* Finally, a two-terminal resistor which does not exist in practice, but which is very useful conceptually, is the ideal diode whose symbol and v-i curve are shown in Fig. 1-11a and b, respectively. Analytically, an ideal diode is described by

$$\begin{aligned} i &= 0 && \text{for all } v < 0 \\ v &= 0 && \text{for all } i > 0 \\ p &= vi = 0 && \text{for all } v \text{ and } i \end{aligned} \tag{1-18}$$

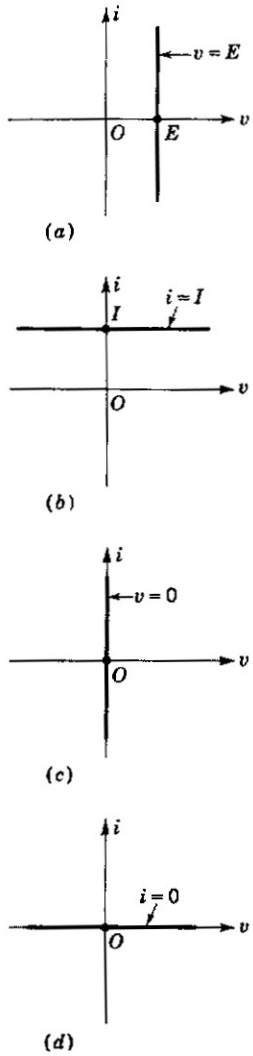


Fig. 1-10. The v-i curves of a dc-voltage source, a dc-current source, a short circuit, and an open circuit have one common property: they consist of either a vertical line or a horizontal line.

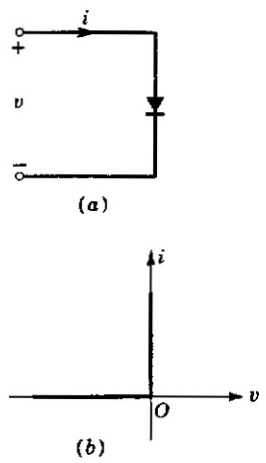


Fig. 1-11. The symbol and v-i curve of an ideal diode.

Observe that the last constraint is introduced to eliminate any point in the fourth quadrant from becoming a part of the v-i curve. It is also important to observe that an ideal diode becomes an open circuit for all $v < 0$ and a short circuit for all $i > 0$.

Before we leave this section, we wish to emphasize that the class of practical nonlinear resistors is not restricted to those listed in Table 1-1. In fact, when we reach Chap. 6, we shall be able to synthesize nonlinear resistors with almost any prescribed v-i curve of practical interest.

Exercise 1: Find the v-i curve of the ideal diode but with its terminals interchanged. Describe this curve analytically.

Exercise 2: A time-varying independent source may be represented by a family of v-i curves with the time t as a parameter. Sketch the v-i curves of a voltage source with terminal voltage $v_s(t) = 2t$ and a current source with terminal current $i_s(t) = 10 \sin \pi t$.

1-6-5 SOME PRACTICAL APPLICATIONS OF TWO-TERMINAL NONLINEAR RESISTORS

What are nonlinear resistors good for? How do we make use of their v-i curves to design practical electronic gadgets? Do certain types of v-i curves seem more appropriate for one application than another? These are some of the questions that will be answered in the latter part of this book, after we have built up enough theory to understand the basic principles involved in a practical design. However, to satisfy the impatient reader, we shall present in this section a qualitative description of some typical applications. Needless to say, this oversimplified treatment will become more quantitative and precise as the reader gains more ground in the subsequent chapters.

Rectification In many practical applications such as electroplating, the power supply must be restricted to a *single-polarity* voltage or current source. Since the most economical power source is 60-Hz sinusoidal voltage, it is desirable to transform this alternating voltage into a single-polarity voltage. This conversion process is called *rectification,* and any network that carries out the desired transformation is called a *rectifier.* The simplest rectifier consists of an ideal diode in series with a linear resistor, as shown in Fig. 1-12. When the input voltage $v_i(t)$ is positive, the diode becomes a short circuit and $v_o(t) = v_i(t)$. However, when the input voltage $v_i(t)$ is negative, the diode becomes an open circuit and $v_o(t) = 0$.

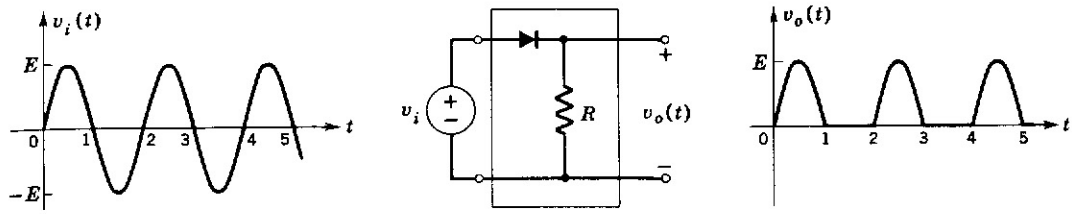


Fig. 1-12. An ideal rectifier converts a sinusoidal input voltage into a single-polarity output voltage.

The result is that the output voltage becomes zero during every other half cycle and is therefore a single-polarity voltage. Of course, this rectifier is an idealized circuit since it uses an ideal diode which does not exist in practice. However, an examination of Table 1-1 suggests that a practical rectifier may be constructed by replacing the ideal diode in Fig. 1-12 by a vacuum diode, a selenium diode, or a semiconductor junction diode.

The above procedure for arriving at a practical design by deriving first an idealized network (which is usually much easier to come by) and then approximating it by a practical circuit is a universal principle of creative design. This principle is based on the intuition that if two networks differ from each other only slightly (e.g., the v-i curves of corresponding resistors differ only slightly), then the corresponding voltage and current waveforms of the two networks must also differ only slightly. Mathematically, this is analogous to the variation of a *continuous* function; namely, a small variation in the value of the independent variable produces a correspondingly small variation in the value of the dependent variable. Because of its practical importance, we shall call the above assumption the small-variation postulate.

Frequency multiplication Another very common application of nonlinear resistors is to convert a low-frequency signal into a high-frequency signal. The ability to do this is instrumental in virtually all communication systems ranging from the simplest walkie-talkie to the most complex telemetry systems between artificial communication satellites. Amazingly, the principle for obtaining frequency multiplication is based on a simple observation from high school trigonometry; namely, the nth power of a sine or cosine function contains higher-harmonic components. For example, $\sin^3 x = \frac{3}{4} \sin x - \frac{1}{4} \sin 3x$, $\cos^4 x = \frac{3}{8} + \frac{1}{2} \cos 2x + \frac{1}{8} \cos 4x$, etc. Hence, if the v-i curve of a nonlinear resistor is described by a polynomial

$$i = a_0 + a_1 v + a_2 v^2 + a_3 v^3 + \cdots + a_n v^n \tag{1-19}$$

then upon applying a voltage signal $v = A \sin \omega t$, we obtain, with the help of various standard trigonometric identities, the expression

$$\begin{aligned} i(t) &= a_0 + a_1(A \sin \omega t) + a_2(A \sin \omega t)^2 + \\ &\qquad a_3(A \sin \omega t)^3 + \cdots + a_n(A \sin \omega t)^n \\ &= b_0 + b_1 \sin \omega t + b_2 \sin 2\omega t + b_3 \sin 3\omega t + \cdots + b_n \sin n\omega t \\ &\quad + c_1 \cos \omega t + c_2 \cos 2\omega t + c_3 \cos 3\omega t + \cdots + c_n \cos n\omega t \end{aligned} \tag{1-20}$$

Observe that although the voltage consists of a sinusoidal signal of angular frequency ω, the resulting current contains a constant term b_0, a component of the same frequency ω, and a number of higher-harmonic components $2\omega, 3\omega, \ldots, n\omega$. In practice, any of these harmonic components can be extracted by interposing a network known as a *filter* which essentially suppresses all other components except the desired one.[1] In fact, we could even avoid the use of filters if we could obtain nonlinear resistors with suitable v-i curves. For example, to generate a third-harmonic signal, we apply a well-known trigonometric identity

$$\cos 3x = 4 \cos^3 x - 3 \cos x \tag{1-21}$$

to obtain the desired v-i curve,

$$i = 4v^3 - 3v \tag{1-22}$$

Hence, if $v = \cos \omega t$, then

$$i(t) = 4 \cos^3 \omega t - 3 \cos \omega t = \cos 3\omega t \tag{1-23}$$

which is the desired third harmonic. The next step then consists of finding a practical nonlinear resistor with a v-i curve which approximates Eq. (1-22). Unfortunately, no commercially available resistor is close enough even as an approximation. Hence, it would be necessary to *synthesize* this v-i curve using commercially available resistors as building blocks. The principles and techniques for synthesizing arbitrary v-i curves will be given in Chap. 8.

[1] The design of filters is a very well-developed subject and is usually given in a senior-level course called network synthesis.

Exercise 1: Find the values of the coefficients $b_0, b_1, b_2, \ldots, b_n$ and $c_1, c_2, \ldots, c_n$ in Eq. (1-20) in terms of the constant A and the coefficients $a_0, a_1, a_2, \ldots, a_n$, where $n = 5$.

Exercise 2: Using Eq. (1-21), find the desired v-i curve for converting a 100-volt, 60-Hz sinusoidal voltage into a 10-amp, 180-Hz sinusoidal current.

Exercise 3: Verify the trigonometric identity $\cos^5 x = \frac{5}{8} \cos x + \frac{5}{16} \cos 3x + \frac{1}{16} \cos 5x$ and find the desired v-i curve for converting a 1-ma, 1-kHz sinusoidal current waveform into a 1-volt, 5-kHz sinusoidal voltage waveform.

Frequency mixing Given two sinusoidal waveforms with *commensurate* angular frequencies ω_1 and ω_2 (that is, the ratio ω_1/ω_2 is a rational number), we are frequently interested in generating a new sinusoidal waveform with a frequency given by $(m\omega_1 \pm n\omega_2)$, where m and n are any integers, including zero. Each new frequency corresponding to a given combination (m,n) is called a *beat frequency* and will be denoted by $\omega_{(m,n)}$. One of the most common requirements in signal processing (e.g., a radio receiver or an electronic organ) is the generation of appropriate beat frequencies.[1] We shall now demonstrate that in order to generate beat frequencies, it is necessary to perform a nonlinear operation. Again, the basis for doing this is given by the well-known trigonometric identities:

$$\sin x \sin y = \tfrac{1}{2} [\cos (x - y) - \cos (x + y)] \tag{1-24}$$

and

$$\sin x \cos y = \tfrac{1}{2} [\sin (x + y) + \sin (x - y)] \tag{1-25}$$

To demonstrate how we generate beat frequencies, consider applying two voltage sources $v_1 = A \sin \omega_1 t$ and $v_2 = B \sin \omega_2 t$ in series with a nonlinear resistor with a v-i curve given by $i = v^3$. The current $i(t)$ is given by

$$\begin{aligned} i(t) &= (A \sin \omega_1 t + B \sin \omega_2 t)^3 \\ &= A^3 \sin^3 \omega_1 t + 3A^2B \sin^2 \omega_1 t \sin \omega_2 t \\ &\qquad + 3AB^2 \sin \omega_1 t \sin^2 \omega_2 t + B^3 \sin^3 \omega_2 t \end{aligned}$$

If we now apply Eq. (1-25) and a number of standard trigonometric identities, we obtain, upon simplification, the expression

$$\begin{aligned} i(t) = (a_1 \sin \omega_1 t + b_1 \sin \omega_2 t) &+ (a_2 \sin 3\omega_1 t + b_2 \sin 3\omega_2 t) \\ &+ [a_3 \sin (\omega_2 - 2\omega_1)t + b_3 \sin (\omega_2 + 2\omega_1)t] \\ &+ [a_4 \sin (\omega_1 - 2\omega_2)t + b_4 \sin (\omega_1 + 2\omega_2)t] \end{aligned} \tag{1-26}$$

[1] The beat frequency is also called a *sideband frequency* and the collection of all beat frequencies is usually called *sidebands*. The definitions of beat frequency and sidebands are meaningful even if ω_1 and ω_2 are not commensurate with each other.

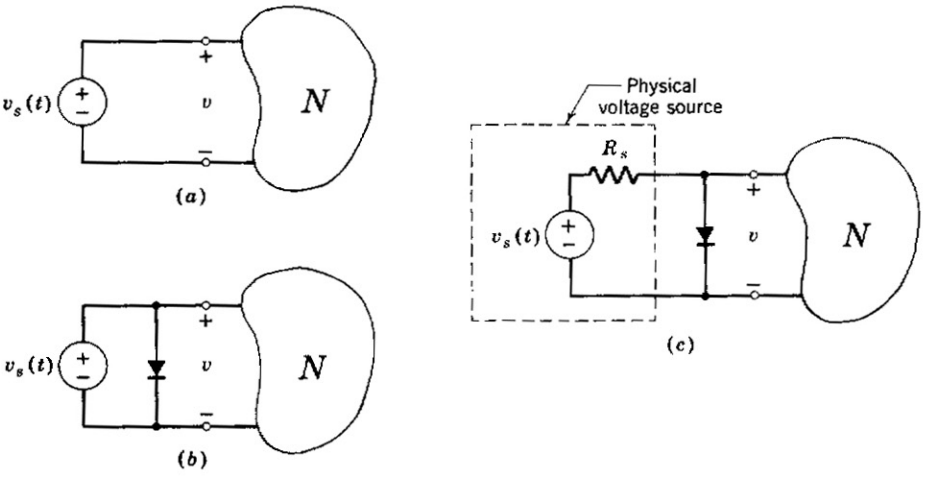
where the coefficients a_j, b_j are functions of A and B. Observe that in addition to sinusoidal terms having the same frequencies as the driving sources, the current $i(t)$ also contains the third-harmonic terms with frequencies $3\omega_1$ and $3\omega_2$ and the beat-frequency terms with frequencies $(\omega_2 \pm 2\omega_1)$ and $(\omega_1 \pm 2\omega_2)$. In the more general case where the v-i curve is described by a polynomial, we can expect, in general, sinusoidal terms with harmonic frequencies $m\omega_1$ and $n\omega_2$, as well as beat frequencies $m\omega_1 \pm n\omega_2$. In practice, any one of these beat frequencies may be extracted through a filter. This principle is widely used in telephone systems.

Exercise 1: Give an example of a pair of sinusoidal waveforms with incommensurate (i.e., not commensurate) frequencies. Is the sum of these two waveforms periodic?

Exercise 2: A speech synthesizer is an electronic system designed to simulate the human voice. An important component of this system is a mixer for generating as many beat frequencies as possible. Assuming that the v-i curve of the resistor is described by a polynomial, what must the general form of the polynomial be in order to generate beat-frequency terms with m and n equal to 0, ± 1, ± 2, and ± 3?

Limiting Any nonlinear resistor R with a v-i curve containing a (nearly) vertical segment can be used to *limit* the voltage across a two-terminal black box connected in parallel with R. For example, we can limit the terminal voltage across the black box N shown in Fig. 1-13*a* to nonpositive values by connecting an ideal diode across N as shown in Fig. 1-13*b*. This is because by definition, the voltage across an ideal diode is given by $v \leq 0$.

Fig. 1-13. The voltage across the black box N is constrained to nonpositive values by connecting an ideal diode in parallel with N.



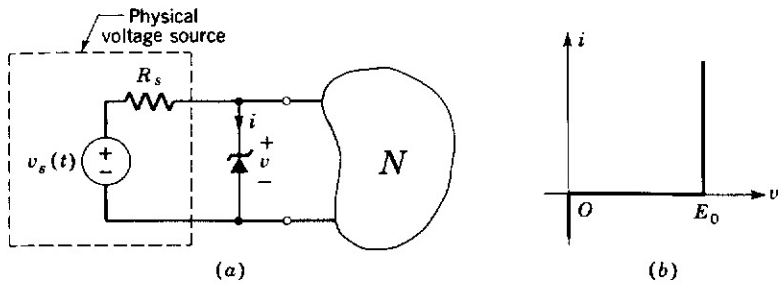


Fig. 1-14. The voltage across N is limited to a maximum value equal to the constant voltage E_0 of the zener diode.

A paradoxical situation arises when one questions what happens if a voltage source with a positive terminal voltage (say $v_s = 2$ volts) is connected across the network shown in Fig. 1-13*b*. By definition, the terminal voltage of this voltage source must remain constant regardless of the external network connected across it. But also by definition, the voltage across an ideal diode cannot be positive. The basic problem here is that we are connecting two incompatible ideal elements in parallel, thereby rendering the definitions inconsistent. In other words, this paradox arises because of an overidealization. It is no different from many paradoxes of a similar nature, most notably among which is the paradox: "What happens if one connects a short circuit across a voltage source with a nonzero terminal voltage?" The best way to resolve this type of paradox is to exclude all such incompatible connections. But how can we forbid anyone from making an incompatible connection in practice? The answer is that there is no such thing as an incompatible connection *in practice* because there are no such things as an ideal voltage source and an ideal diode. Any physical voltage source has a small internal resistance R_s in series with it, as shown in Fig. 1-13*c*. Once we introduce R_s, the paradox disappears because whenever $v_s(t)$ becomes positive, the diode becomes a short circuit and the entire voltage appears across R_s. Hence, the voltage across N can never be positive.

The same principle can be applied to limit the voltage across N from exceeding a prescribed value E_0. For example, if we connect a zener diode with a constant voltage $E_z = E_0$ across N as shown in Fig. 1-14*a*, then from the v-i curve of the zener diode shown in Fig. 1-14*b* (observe that the reference polarity and directions are opposite to those shown in Table 1-1) it is clear that $0 \leq v \leq E_0$. This circuit is commonly used for overload protection. For example, in a typical application, the black box N consists of a sensitive instrument (such as a voltmeter) whose maximum permissible voltage is equal to E_0.

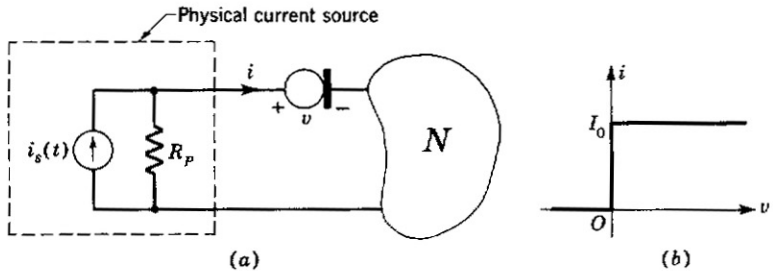


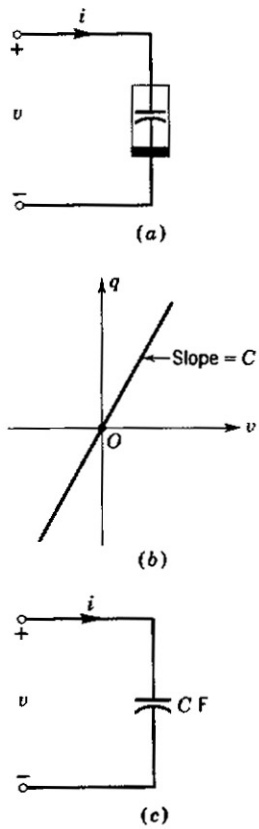
Fig. 1-15. The current entering N is limited to a maximum value equal to the constant current I_0 of the constant-current diode.

By analogous reasoning, any resistor R with a v-i curve containing a (nearly) horizontal segment can be used to *limit* the current entering a black box connected in series with R. For example, if we connect a constant-current diode with constant current I_0 in series with N as shown in Fig. 1-15*a*, then from the v-i curve of the constant-current diode shown in Fig. 1-15*b* it is clear that $0 \leq i \leq I_0$ (the resistor R_p is introduced to avoid a similar paradox).

Exercise 1: The maximum permissible range of voltages of a hypersensitive instrument is given by $-10 \leq v \leq 5$. Design an overload protection circuit and specify the v-i curve of any nonlinear resistor used in the circuit.

Exercise 2: Explain what happens if $i_s(t) > I_0$ in the circuit shown in Fig. 1-15. Replace the constant-current diode with an appropriate nonlinear resistor so as to limit the terminal current entering N to $|i| < 20$ ma.

Fig. 1-16. Symbols for a two-terminal capacitor.



1-7 TWO-TERMINAL CAPACITORS

A two-terminal black box which can be characterized by a curve in the v-q plane is called a *two-terminal capacitor* and will be denoted by the symbol shown in Fig. 1-16*a*. Observe that one edge of this symbol is darkened for the same reason as it was for the resistor.

1-7-1 LINEAR CAPACITORS

An important subclass of capacitors can be characterized by a straight line through the origin of the v-q plane, as shown in Fig. 1-16*b*. This subclass is called *linear capacitors* and will be denoted by the conventional symbol shown in Fig. 1-16*c*. A linear capacitor can be described analytically by

$$q = Cv \tag{1-27}$$

where the constant C represents the slope of the straight line and is called the *capacitance* associated with the capacitor. The unit of capacitance is the *farad* and will be denoted by F. To find the current entering a linear capacitor, we substitute Eq. (1-27) for q in Eq. (1-7) and obtain

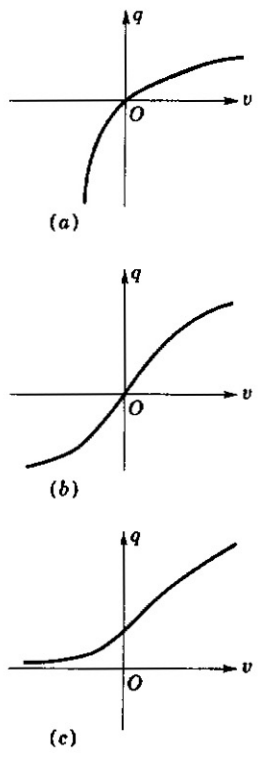
$$i(t) = C\frac{dv(t)}{dt} \tag{1-28}$$

A linear capacitor is therefore completely characterized by one number, namely, its capacitance. Again, we would differentiate between the terms *capacitor* and *capacitance.*

1-7-2 NONLINEAR CAPACITORS

If a capacitor is characterized by a v-q curve other than a straight line through the origin, it is called a *nonlinear capacitor.* In this case, the capacitor can no longer be described by a single number, and hence the entire v-q curve must be given. An example of a practical, nonlinear capacitor is the metal-oxide-semiconductor (MOS) capacitor whose v-q curve is shown in Fig. 1-17*a*. This nonlinear capacitor is used quite extensively in integrated circuits, where the conventional linear capacitor becomes impractical to fabricate. Although there are at present only a few practical nonlinear capacitors available commercially, it is expected that more will become available in the near future. In fact, as will be shown in Chap. 3, it is possible to synthesize a capacitor with any prescribed v-q curve with the help of a new network component called the *mutator.*

Fig. 1-17. The v-q curves of three typical nonlinear capacitors.



There are other reasons for studying nonlinear capacitors. One reason is that components of many physical and biological systems behave in a manner analogous to that of a nonlinear capacitor. Hence the study of such systems can often be achieved by constructing an electrical network model to simulate the behavior of these systems. A simple example is the displacement-vs.-force curve of the nonlinear spring shown in Fig. 1-7*a*. This mechanical element is usually modeled by a nonlinear capacitor with a similar v-q curve, as shown in Fig. 1-17*b*. Another example is given by the volume-vs.-pressure curve of the ventilatory part of the human respiratory system. This biological component can be modeled by an analogous v-q curve as shown in Fig. 1-17*c*.

We shall denote the v-q curve of a nonlinear capacitor by

$$q = q(v) \tag{1-29}$$