Nonlinear Circuit Foundations for Nanodevices, Part I: The Four-Element Torus

LEON O. CHUA, FELLOW, IEEE

Invited Paper

Not all molecular and nanodevices are useful from an information technology perspective. Such devices are said to be inept in a precise technical sense that can be easily tested from an explicit mathematical criteria to be presented in this two-part tutorial review. Often an inept device can be redesigned into a smart device capable of computing and artificial intelligence by massaging the device's parameters, such as doping, concentration, geometrical profile, chemical moiety, etc., in accordance with the principle of local activity to be articulated in Part II. In particular, designing a smart nanodevice amounts to fine tuning the device parameters into a much smaller niche within the device's locally active parameter region called the edge of chaos where complexity abounds.

Molecular and nanodevices will remain novelty toys for nanodevice specialists unless they possess realistic nonlinear circuit models so that future nano circuit designers can simulate their exotic designs as easily and accurately as current CMOS circuit designers. A mathematically consistent theory for modeling nonlinear, high-frequency nanodevices, specially those which exploited exotic tunneling and entanglement quantum mechanical effects, such as Coulomb blockade, quasi-particle dynamics, Kondo resonance, Aharonov–Bohm nonlocal interactions, etc., will require the introduction of a complete family of fundamental circuit elements as model building blocks. They are presented via a doubly periodic table of circuit elements somewhat reminiscent of Mendeleev's periodic table of chemical elements. These fundamental circuit elements can be compactly represented by a loop of four generic species of circuit elements wrapped around the surface of a torus where any higher order element having an arbitrarily high order of frequency dependence can be generated from one of them, modulo the integer 4, ad infinitum. The significance of this four-element torus is that realistic circuit models of all current and future molecular and nanodevices must necessarily build upon an appropriate subset of nonlinear circuit elements begotten from this torus.

Keywords*—*Device modeling, four-element torus, local activity, nanodevices.

Manuscript received December 31, 2002; revised May 23, 2003. This work was supported in part by Office of Naval Research under Grants N00014-01-1-0741 and N00014-03-1-0698.

The author is with the Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, Berkeley, CA USA 94720.

Digital Object Identifier 10.1109/JPROC.2003.818319

I. INTRODUCTION

Many molecular and nanodevices reported so far in the literature are inept in the sense that they can never be used to design an amplifier, or an oscillator, let alone the garden variety of nonlinear circuit modules essential for digital signal processing and information technology. The main objective of this two-part tutorial review is to present the nonlinear circuit-theoretic foundations [1], [2] for identifying the basic character of a molecular or nanodevice and evaluating its information processing potentials. Since the results presented in this paper do not depend on the device's internal (physical, chemical, biological, etc.) operating mechanisms, they are applicable to all devices, including devices made from a single atom, as featured on the cover of a recent issue (13 June 2002) of *Nature* [3]–[5].

As soon as a researcher discovers that a device with two external conducting terminals (e.g., a single atom or a single molecule chemically attached to a pair of gold electrodes) can conduct electrical currents when a voltage is applied across the terminals, it is natural to measure the device's voltage–current $v-i$ characteristics by varying the applied voltage while measuring its corresponding current. Such a measurement process is usually mechanized by applying a low-frequency sinusoidal sweeping voltage across the device and recording the $v-i$ characteristics displayed on an oscilloscope.

One of the earliest $v-i$ characteristics measured from a two-terminal molecular device (made by sandwiching several layers of one type of molecule between gold electrodes [6], [7]) is shown in Fig. 1. Another $v-i$ characteristics measured from a single benzene-1, 4 dithiolate molecule (made by bonding dithiolated molecules into break junctions [8]) is shown in Fig. 2. Yet another $v-i$ characteristic measured from a single dithiol molecule (made from a more reliable method for making through-bond electrical contacts to molecules [9]) is shown in Fig. 3.

A common feature shared by the $v-i$ characteristics in Figs. 1–3, as well as by virtually all measured two-terminal

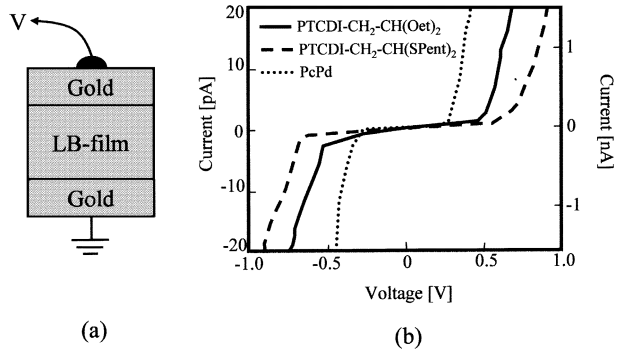

Fig. 1. (a) Two-terminal device made by sandwiching several layers of one type of molecule between gold electrodes. (b) $v - i$ characteristics measured from the device made from three different molecules. The current scale for the bold curve is in nanoamperes, as indicated on the right boundary (redrawn from [6, p. 272]).

nanodevice $v - i$ characteristics published so far in the literature is that the current i is a monotonically increasing function of the voltage v. One of our objectives in this paper (part II) is to understand why all such two-terminal devices are inept and therefore useless for information processing applications. In particular, we will present an explicit constructive criteria which allows a researcher to test whether an n-terminal device $(n \geq 2)$ is inept, and if so, the criteria will suggest how to massage the device parameters (e.g., doping, concentration, geometrical profile, chemical moiety, etc.) so as to obtain a smart device capable of information processing.

It is of course possible to design and fabricate smart molecular and nanodevices which exhibit more complex and exotic characteristics. One very recent promising example is based on the nanowire approach pioneered by Lieber's group from Harvard University, Cambridge [10]. Fig. 4(a) shows an exotic $v - i$ characteristic in the form of a pinched hysteresis loop measured from a two-terminal device made from an n-channel InP nanowire functionalized with CoPc redox molecules [10]. This $v - i$ characteristic, which is measured by applying a low-frequency periodic voltage waveform across the device, is a double-valued function, where the slope (i.e., the small-signal conductance [12]) changes from $G = 0$ to $G = 800$ nS at a bias voltage (i.e., dc operating point) of $V = 0.1$ V. This double-valued conductance property is depicted in Fig. 4(b) with the two-terminal nanowire device switching on and off via $\pm$5-V 1-s bias pulses. The authors [11] reported that these two distinct conductance states can be maintained even after the voltage pulse is set to zero, thereby suggesting possible applications as a nonvolatile memory.

Although the above two-terminal nanowire device is both exotic and promising, from a circuit-theoretic perspective, it is not correctly characterized because, even though not reported in [11], we will predict in Section IV-D that the pinched hysteresis loop $v - i$ characteristics in Fig. 4(b) is not invariant but changes with both the frequency and the waveform of the applied sweeping voltage. Notwithstanding the remarkable potentials of this device and the admirable ingenuity of the authors, their characterization is not a valid device model because it cannot be used to predict the outcome when their device is imbedded in any other circuit environment than the setup described by the authors [11]. A model must have some predictive ability when connected to an arbitrary external circuit environment, and the pinched hysteresis loop in Fig. 4(a) fails this critical test. This is only one among many other examples of "mistaken identity" scattered all over the literature (especially in biophysics journals) for over 50 years of devices improperly characterized by frequency-dependent and multivalued characteristics. In Section IV-D we will show that the two-terminal nanowire device of Fig. 4(a) can be modeled, at least qualitatively, as a two-terminal nonlinear circuit element, called a memristor, an acronym for *memory resistor* [13].

Our final example of a nanodevice in this section is a nanocoil made by rolling a graphite-like sheet of boron–carbon–nitrogen compound (BC_2N) into a cylinder a few micrometers in length with a very small diameter (1–2 nm) and a helical pitch close to those of armchair carbon nanotubes [14], as shown in Fig. 5(a). When the BC_2N nanocoil is doped with carriers, electrons are transported along the helical route of boron atoms, thereby simulating a classical inductor coil. Because the length of the nanocoil is several orders of magnitude larger than its diameter, any quantum mechanical effect due to coherent electron transport can be neglected at room temperature and a semiclassical transport analysis via Maxwell's equations should yield a realistic circuit model in typical experimental situations at room temperature. This analysis is given in [15], resulting in the *RL* circuit model shown in Fig. 5(c) consisting of a linear inductor (with an inductance equal to $L(\omega)$) in series with a linear resistor (with a resistance equal to $R(\omega)$) where both the inductance and the resistance are not constant, as would be the case had the helix been a conducting wire, but depends on the frequency ω of the sinusoidal current source $i(t) = I \sin \omega t$. In other words, if we measure the impedance $Z(j\omega) = R + j\omega L$ of this nanocoil in the laboratory, we will measure a different resistance *R* and inductance *L* for different choices of the frequency ω of the sinusoidal current source.

Although the frequency-dependent nanocoil circuit model derived in [15] fulfills successfully the original motivation of the authors, namely, to verify the chirality of the nanocoil, it cannot be used to predict its behavior when the nanocoil is connected to an external circuit containing any nonlinear device, such as a simple pn-junction diode. The reason is that the waveforms of the nanocoil in this external environment will no longer be sinusoidal even if the circuit is driven from a sinusoidal source. Since the frequency-dependent inductance $L(\omega)$ and the frequency-dependent resistance $R(\omega)$ in Fig. 5(c) no longer make sense in this situation, they have no predictive ability and is therefore not a valid circuit model.

The above nanocoil circuit model is another example of a mistaken identity. Indeed, as in the nanowire hysteresis models already alluded to above, the literature over the last 50 years abounds with examples of frequency-dependent circuit models [16], [17], none of which would pass the critical test of predictability.

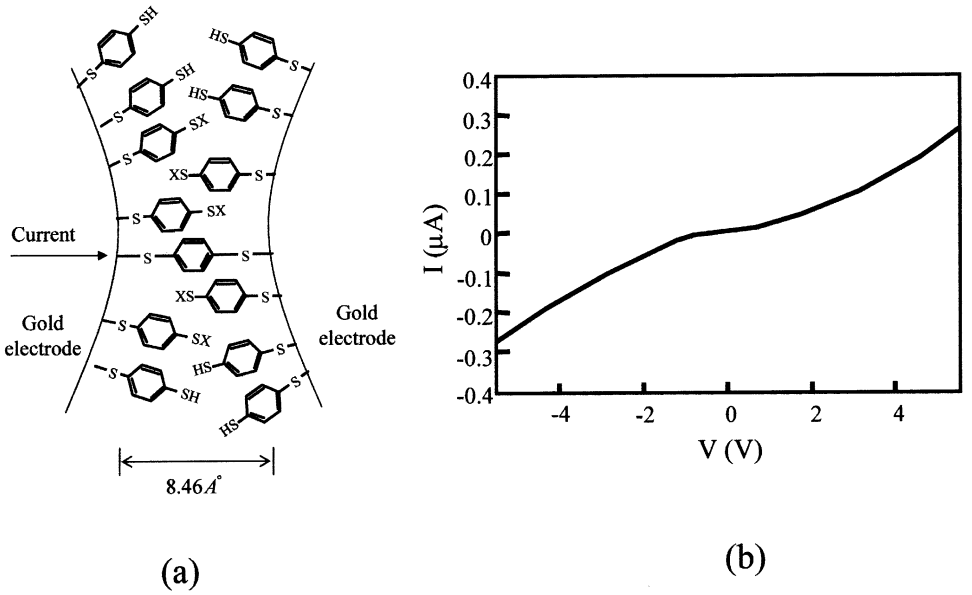

Fig. 2. (a) Schematic drawing of a benzene-1, 4-dithiolate self-assembled between the two facing gold electrodes formed after breaking the junction. (b) A typical measured $v - i$ characteristic (redrawn from [8]).

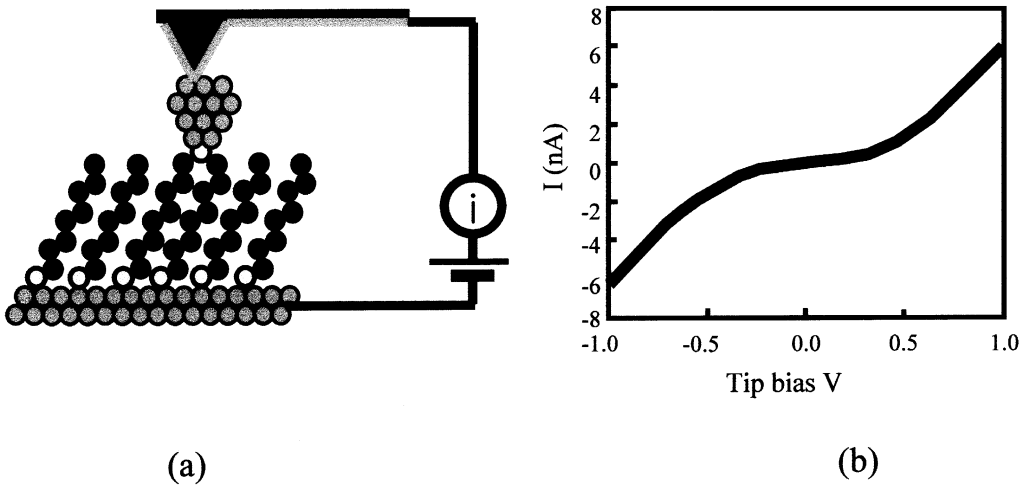

Fig. 3. (a) Schematic representation of the laboratory setup for chemically bonding metal contacts to either end of an isolated diothiol molecule. (b) A highly reproducible $v - i$ characteristic (redrawn from [9]).

To clear up this mess and to show how to make sense out of meaningless frequency-dependent circuit models, it is essential to define a family of fundamental nonlinear circuit elements from which frequency-dependent behaviors will emerge naturally from meaningful circuit models made from these building blocks. This will be presented in Section II.

II. Constitutive Relation and Device Modeling

This paper is concerned only with circuits made of lumped circuit elements where geometry, shape, and physical extent of the elements are irrelevant. A circuit element may have two or more external electrical terminals. Any interconnection via perfect conductors of two or more terminals from different circuit elements, or from the same element, results in a lumped circuit. We assume without loss of generality that the circuit is connected electrically in the sense that the voltage from any terminal of any element of the circuit is well defined with respect to a common datum node henceforth called the ground.

Our assumption of no spatial variables implies that all voltages $v_j(t)$ and all currents $i_k(t)$ are functions of only one independent variable, namely, the time t. This assumption is equivalent to the standard assumption in mechanics where each object is assumed to have zero spatial extent and can, therefore, be represented as a point in space located at the object's center of mass.

In addition, all circuit elements are assumed to be ideal with perfectly conducting terminals, and without any parasitic effects within the element proper, or with other neighboring elements. There is no loss of generality in this assumption, since any nonnegligible parasitic effects

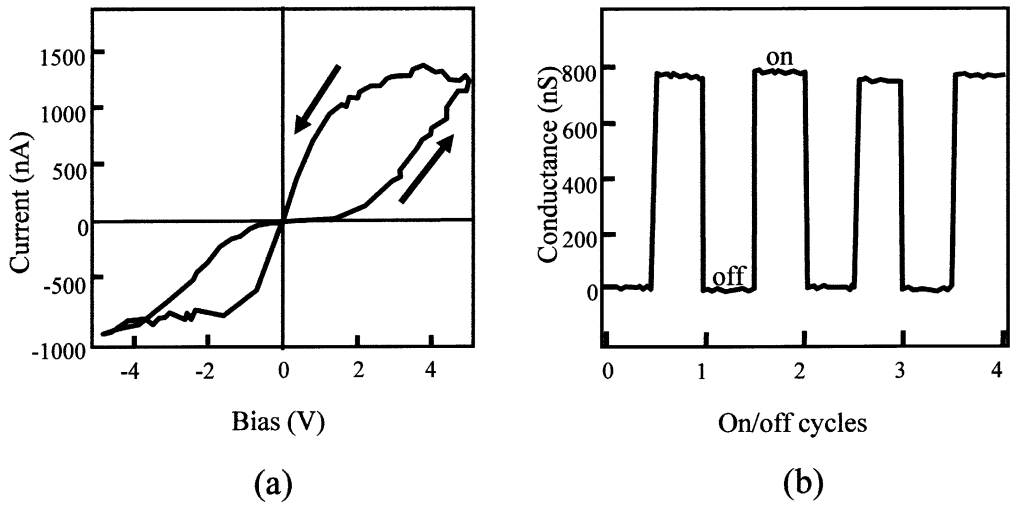

Fig. 4. Electrical characteristics exhibited by a two-terminal nanowire device (redrawn from [10]). (a) A "pinched" hysteresis loop in the voltage-current plane when driven by a low-frequency periodic voltage signal. (b) Its conductance switches from 0 to 800 nanosiemens when driven by a square-wave voltage signal.

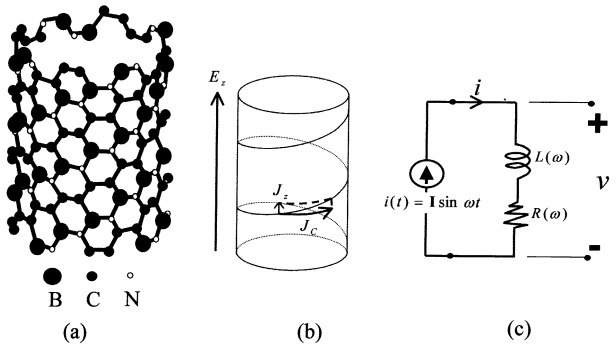

Fig. 5. (a) A nanocoil made from a BC_2N nanotube. The large circles are boron atoms which lie along a helix analogous to winding a conducting wire along a cylindrical core in making an inductor coil. (b) Idealized geometry of the nanocoil used to carry out a classical analysis of Maxwell's equations in terms of the electric field **E** and the chiral current density **J**, which have been decomposed into a vertical z-component (E_z and J_z) and a rotational component (E_c and J_c) along the circumference. (c) Circuit model of the nanocoil consists of a frequency-dependent inductance $L(\omega)$ in series with a frequency-dependent resistance $R(\omega)$.

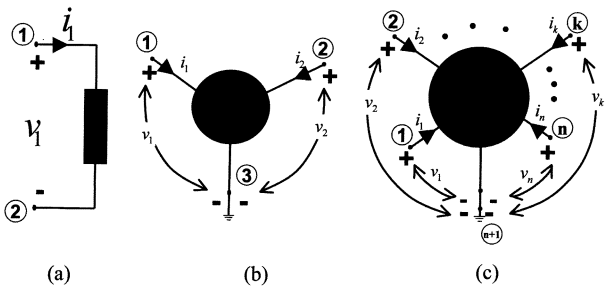

Fig. 6. Multiterminal black boxes and their associated reference convention. (a) Two-terminal black box with current i_1 entering the positive terminal ①. (b) Three-terminal black box with current i_j, entering the positive terminal ⓙ, $j = 1, 2$. (c) $(n+1)$-terminal black box with current i_j, entering the positive terminal ⓙ, $j = 1, 2, \ldots, n$.

in a real-world circuit (hardware) may be modeled by introducing additional circuit elements.

Finally, we assume that each circuit element is defined for all voltage and current waveforms for all frequencies. In other words, there are no time-rate dependent circuit parameters because such elements do not qualify as a model due to its lack of predictability under arbitrary external interconnections.

A. Definitions and Associated Reference Convention

Fig. 6 shows three black boxes with two, three, and $n+1$ electrical terminals attached to the respective boxes. From the external interconnections point of view, what is inside each box is irrelevant, hence the name black box. For example, the two-terminal black box in Fig. 6(a) may contain just a pn-junction diode, or it could contain a million arbitrary interconnected circuit elements. Since algebraic sum of all currents entering each black box is zero at all times in view of the Kirchoff current law (KCL) [12], only n terminal currents are independent in an $(n+1)$-terminal black box, $n = 1, 2, \ldots$. Similarly, in view of the Kirchoff voltage law (KVL) [12], only n terminal voltages are independent and we can choose any node as the datum node (ground) and then define the remaining terminal-to-datum voltages.

Although the reference direction for each terminal current and reference voltage polarity of each terminal voltage can be chosen arbitrarily, it is a standard procedure in circuit theory to define them as indicted in Fig. 6. This choice is called an associated reference convention [12], which will be the standing assumption in this paper. One convenient consequence of this choice is that whenever the total instantaneous power

$$p_T(t) \triangleq \sum_{j=1}^{n} v_j(t) i_j(t) \tag{1}$$

is positive at any time $t = t_k$ when the black box is embedded in an external circuit, then a net amount of power equal to

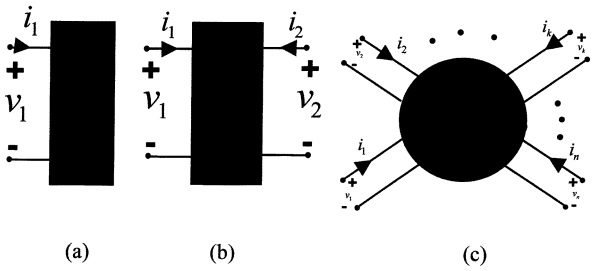

Fig. 7. Multiport black boxes and their associated reference convention. Each port has two wires but only one port current and one port voltage are needed for each port. (a) One-port black box. (b) Two-port black box. (c) n-port black box.

$p_T(t_k)$ watts enters the $(n+1)$-terminal black box at $t = t_k$. In other words, the black box behaves like a sink at $t = t_k$, since it accepts $p_T(t_k)$ watts of power from the external circuit at $t = t_k$. Conversely, if $p_T(t_k) < 0$, then the black box behaves like a source at $t = t_k$, since it supplies $|p_T(t_k)|$ watts of power to the external circuit at $t = t_k$. This physical interpretation is important when we define the fundamental concept of *local activity* in Part II of this paper.

There are many practical examples of black boxes having an even number $2n$ of terminals but can be uniquely characterized by only n terminal currents and n terminal voltages. This situation applies to the case where the currents in only one-half of the terminals are independent because by construction (e.g., n pairs of windings in an n-port transformer), the $2n$ terminals can be grouped into n pairs called ports so that the current entering one terminal of each port is always equal to the current leaving the other terminal of the same port at all times. An example of a one-port, two-port, and n-port black box is shown in Fig. 7, along with their respective associated reference convention. The total instantaneous power entering an n-port is given by the same equation (1), since both an n-port black box and an $(n+1)$-terminal black box have the same number n of current and voltage variables.

B. Concept of Device Modeling

In this paper each $(n + 1)$ terminal or n-port molecular or nanodevice will be denoted by a corresponding black box from Figs. 6 and 7. From the circuit designer's perspective, it is essential to have a realistic circuit model of the black box made of some appropriate circuit element building blocks [18], [19]. A circuit model for a molecular or nanodevice is said to be realistic if, and only if, when the device model is connected to any external circuit and driven by any voltage and/or current sources, the device terminal voltage and current waveforms calculated from this model agrees with the measured results to within a prescribed error bound. Depending on the applications, the acceptable error may be rather small (e.g., less than 5%) as in many analog signal processing applications, or it may be quite large if one is interested only in the device model's qualitative performance (e.g., whether it is capable of amplifying signals or performing logic operations, or whether its *I–V* relationship resembles that of a pinched hysteresis loop, etc.).

One can often improve the accuracy of a device circuit model by introducing additional circuit element building blocks. However, a complicated model would be more costly in terms of computation time; hence, a golden rule in device modeling is to choose the simplest model that will yield acceptable answers. In other words, there is no such thing as a best model for all applications, as beautifully articulated by Einstein in his famous quotation "A model must be as simple as possible, but not simpler."

The complexity of a device circuit model depends to a great extent on the choice of the circuit element building blocks. For example, if a device is inherently nonlinear, as is most certainly the case for useful molecular and nanodevices, then one must have a rich enough repertoire of nonlinear element building blocks [18], [19]. Clearly, trying to model a nonlinear device using only the classical circuit elements (linear R, L, C) is doomed to fail no matter how many circuit elements are used. Likewise, for a device that exhibits multivalued characteristics (such as the pinched hysteresis loop from a nanowire device), or frequency-dependent parameters (such as the frequency-dependent inductance $L(\omega)$ and frequency-dependent resistance $R(\omega)$ from a nanocoil), we would need circuit element building blocks that are capable of mimicking such exotic behaviors when properly assembled. Using a rigorous mathematical analysis via a device-independent axiomatic approach, it can be shown that an infinite family of such circuit elements must be introduced in order to provide a rigorous circuit-theoretic foundation for nonlinear device modeling [19]. Although in practice only a few (less than five) members of this infinite family are actually needed, it is easy to invent a *gedanken* device which cannot be correctly modeled if one retains only a finite subset from this family. Fortunately, we will provide a painless way of identifying any member of this infinite family of building blocks by tracing recursively around a closed loop of four fundamental nonlinear circuit elements drawn on the surface of a torus, henceforth referred to as the four-element torus.

C. Admissible Signal Pairs and Constitutive Relation

Let $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ and $\mathbf{i} = (i_1, i_2, \ldots, i_n)$ denote the voltages and currents associated with either an $(n+1)$-terminal black box (Fig. 6) or an n-port black box (Fig. 7). Let

$$\mathbf{v}(t) = \begin{bmatrix} v_1(t) \\ v_2(t) \\ \vdots \\ v_n(t) \end{bmatrix} \text{ and } \mathbf{i}(t) = \begin{bmatrix} i_1(t) \\ i_2(t) \\ \vdots \\ i_n(t) \end{bmatrix} \tag{2}$$

denote the voltage and current waveforms associated with the respective terminals and ports of Figs. 6 and 7 when the black box, henceforth denoted by $\mathcal{D}$, is connected to an external circuit $\mathcal{N}$ made of an interconnection of other black boxes, including voltage and/or current sources. Let $(v_k(t), i_k(t))$ denote the corresponding voltage and current waveforms measured at terminal k or port k of $\mathcal{D}$. Although in practice, any actual measurement must necessarily begin from some finite time t_0, for conceptual simplicity, it is convenient to choose $t_0 = -\infty$. Note that there is no loss of gen-

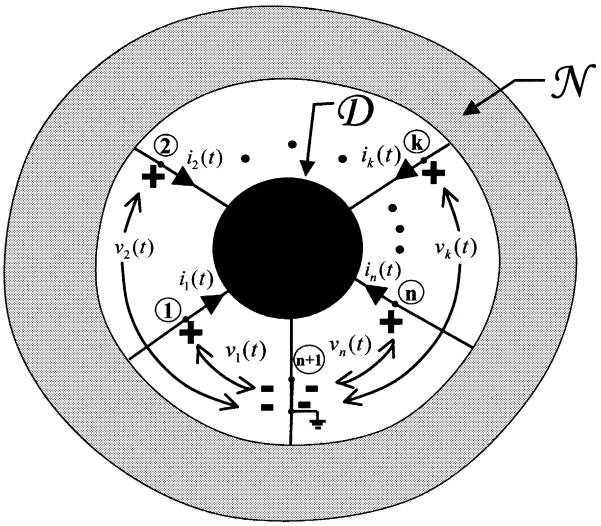

Fig. 8. $(n+1)$-terminal device $\mathcal{D}$ connected to an external circuit $\mathcal{N}$.

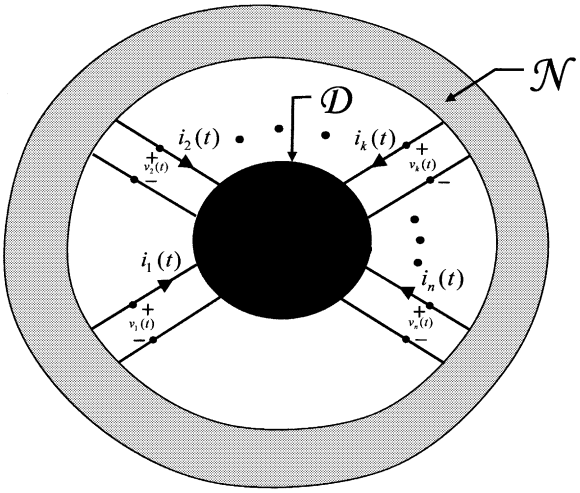

Fig. 9. n-port device $\mathcal{D}$ connected to an external circuit $\mathcal{N}$.

erality in this assumption, since if t_0 is finite, we can simply define $v_k(t)=0$ and $i_k(t)=0$ for $-\infty < t \le t_0$.

The n pairs of waveforms $(v_k(t), i_k(t))$, $k=1,2,\ldots,n$, henceforth denoted by the vector pair $(\mathbf{v}(t), \mathbf{i}(t))$, measured from the experimental setup in Fig. 8 or Fig. 9, is called an admissible signal pair of $\mathcal{D}$. Clearly, different choices of external circuit $\mathcal{N}$ would give rise to different admissible signal pairs. The collection of all admissible signal pairs (corresponding to all possible choices of external circuits) is called the constitutive relation of the $(n+1)$-terminal or n-ports device $\mathcal{D}$.

Observe that by definition, any device $\mathcal{D}$ is completely characterized by its constitutive relation, since the external circuit $\mathcal{N}$ constitutes the universe of all possible circuits. In other words, the constitutive relation of $\mathcal{D}$ is just an analog memory bank which stores all possible admissible signal pairs. Such an infinitely large analog database or lookup table of signal pairs is clearly impractical to implement in hardware. The only practical alternative is to devise a circuit model made from well-defined ideal circuit elements which would yield an acceptable approximation of the ideal constitutive relation of $\mathcal{D}$.

D. Complementary Signal Pairs $\{v_k^{(\alpha_k)}(t), i_k^{(\beta_k)}(t)\}$

Corresponding to each admissible *V–I* signal pair $(v_k(t), i_k(t))$ at each port k,[1] we can define an associated pair of waveforms $(v_k^{(\alpha_k)}(t), i_k^{(\beta_k)}(t))$, henceforth called a complementary signal pair, where

$$v_k^{(\alpha_k)}(t) \triangleq \begin{cases} \dfrac{d^{\alpha_k} v_k(t)}{dt^{\alpha_k}}, & \text{if } \alpha_k > 0 \\ v_k(t), & \text{if } \alpha_k = 0 \\ \int_{-\infty}^{t} \int_{-\infty}^{\tau_{k-1}} \cdots \int_{-\infty}^{\tau_2} v_k(\tau_1)\, d\tau_1\, d\tau_2 \cdots d\tau_{\alpha_k}, & \text{if } \alpha_k < 0 \end{cases} \tag{3}$$

[1]To reduce repetition, we will henceforth avoid writing "terminal k" and simply write "port k" in a generic sense which applies to both cases, i.e., $\mathcal{D}$ can be either an $(n+1)$-terminal black box, or an n-port black box.

and

$$i_k^{(\beta_k)}(t) \triangleq \begin{cases} \dfrac{d^{\beta_k} i_k(t)}{dt^{\beta_k}}, & \text{if } \beta_k > 0 \\ i_k(t), & \text{if } \beta_k = 0 \\ \int_{-\infty}^{t} \int_{-\infty}^{\tau_{k-1}} \cdots \int_{-\infty}^{\tau_2} i_k(\tau_1)\, d\tau_1\, d\tau_2 \cdots d\tau_{\beta_k}, & \text{if } \beta_k < 0 \end{cases} \tag{4}$$

and where α_k and β_k can be any positive integer, any negative integer, or zero. The adjective complementary is used to emphasize that the first waveform is derived from a voltage signal, whereas the second waveform is derived from a current signal, which live in complementary vector spaces [12]. Observe that given $v_k(t)$ or $i_k(t)$, we can generate (by hardware) the corresponding kth-order voltage signal $v_k^{(\alpha_k)}(t)$, or the kth-order current signal $i_k^{(\beta_k)}(t)$, respectively. In other words, the complementary signal pair $(v_k^{(\alpha_k)}(t),\ i_k^{(\beta_k)}(t))$ represents physical signals and can, therefore, be measured by appropriate instrumentations. Hence, just as we can apply a voltage signal $v_k(t)$ across port k of any black box $\mathcal{D}$ and measure the corresponding port current response signal $i_k(t)$, and vice versa, we can apply a kth-order voltage signal $v_k^{(\alpha_k)}(t)$ across port k of $\mathcal{D}$ and measure the corresponding kth-order port current response $i_k^{(\beta_k)}(t)$, and vice versa. Moreover, we can plot the loci traced out by the complementary signal pair $(v_k^{(\alpha_k)}(t),\ i_k^{(\beta_k)}(t))$ in the $i_k^{(\beta_k)}$-versus-$v_k^{(\alpha_k)}$ plane, for each port k, $k=1,2,\ldots,n$.

In general, for an arbitrary choice of (α_k, β_k), the plotted loci will depend on the choice of the testing signal $v_k^{(\alpha_k)}(t)$ (respectively, $i_k^{(\beta_k)}(t)$). For example, suppose the black box $\mathcal{D}$ consists of a single C-Farad linear two-terminal capacitor (i.e., $i(t)=C(dv(t)/dt)$) as depicted in Fig. 10(a) and suppose we apply a sinusoidal voltage testing signal $v(t)=A\sin\omega t$ of amplitude A and frequency ω across $\mathcal{D}$ and measure the current response, which in this case is given exactly by $i(t)=\omega CA\cos\omega t$, as shown in Fig. 10(b) for $A=1$ and $C=1$. Since we did not know what is inside of $\mathcal{D}$ (it is a *black* box) and are trying to identify experimentally its electrical characteristics by the only method possible,

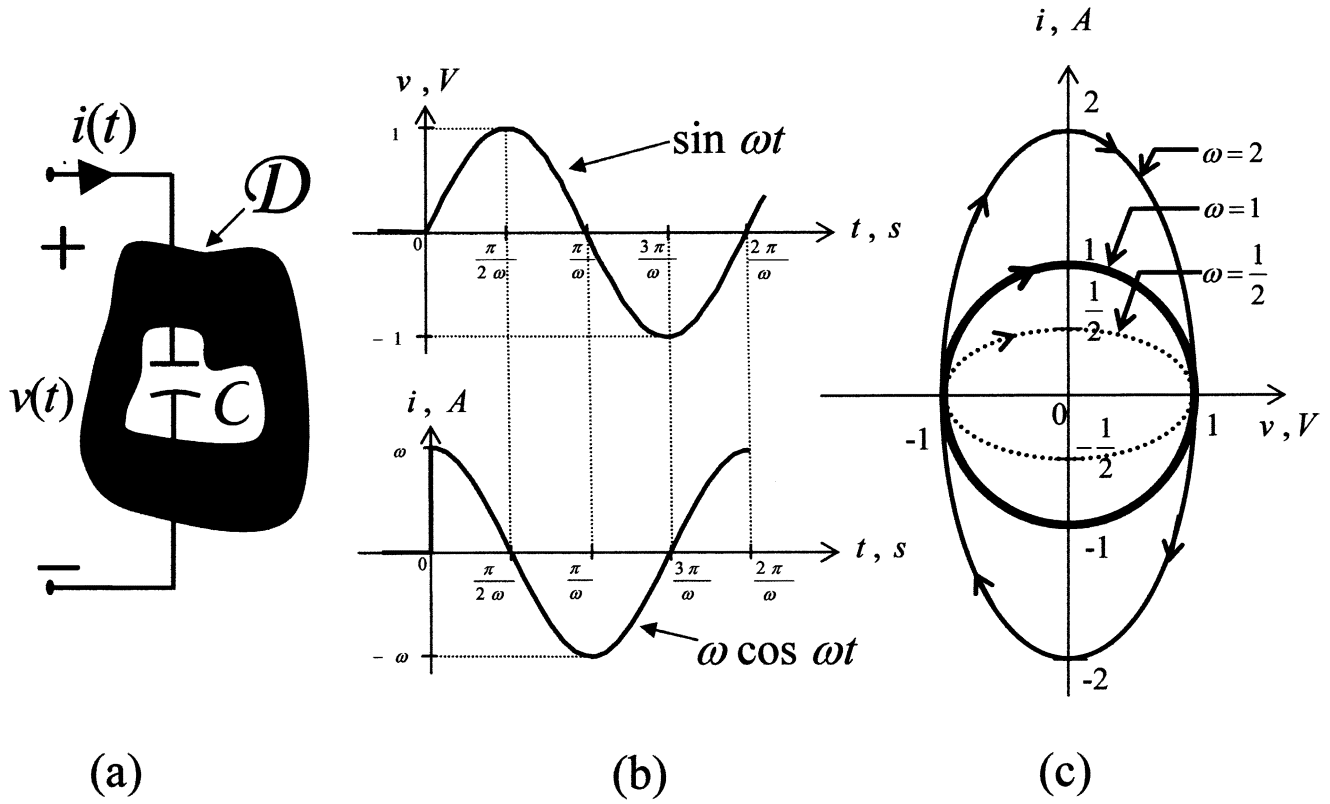

Fig. 10. (a) One-port black box $\mathcal{D}$ containing a C-Farad linear two-terminal capacitor. (b) Admissible signal pair $((v^{(0)}(t), i^{(0)}(t)) = (v(t), i(t)) = (\sin \omega t, \omega \cos \omega t)$, assuming $A = 1$ and $C = 1$. (c) Loci of $(v^{(0)}(t), i^{(0)}(t))$ plotted in the $i^{(0)}$-versus-$v^{(0)}$ plane for $\omega = 1/2, 1, 2$.

namely, apply different testing signals [in this case, $v(t)$] and record its corresponding response [in this case, $i(t)$]. Suppose we proceed with the traditional procedure of plotting the loci of $(v(t), i(t)) \triangleq (v^{(0)}(t), i^{(0)}(t))$ in the i-versus-v plane, anticipating some well-defined $v - i$ characteristics. Unfortunately, the actual outcome in this case is rather unexpected, as shown in Fig. 10(c) for three different choices of frequencies, $\omega = 1/2, 1, 2$, assuming $A = 1$ and $C = 1$. Observe that a different loci is measured for each of these three testing signals. Indeed, if we had repeated this experiment for all values of ω from $\omega = 0$ to $\omega = 2$, the set of corresponding loci would have covered the interior of the outer ellipse in Fig. 10(c) with points completely. Clearly, $\mathcal{D}$ cannot be characterized by a $v - i$ characteristic in this case because it is impossible to predict the current response $i(t)$ to an arbitrary testing voltage signal $v(t)$ from the loci in Fig. 10(c).

Now suppose we calculate $i^{(-1)}(t) \triangleq \int_0^t i(\tau) d\tau = q(t)$ using the current waveform $i(t)$ from Fig. 11(b), namely, $q(t) = (\omega A/\omega) \sin \omega t = A \sin \omega t$, as shown in the bottom part of Fig. 11(a) for $A = 1$. If we now plot the loci of $(v^{(0)}(t), i^{(-1)}(t))$ in the q-versus-v plane, instead of our earlier incorrectly chosen i-versus-v plane in Fig. 10(c), we would obtain a straight line with slope $C = 1$, as shown in Fig. 11(b). Note that this same loci would result for any choice of frequency ω. In fact, it is easy to see (given that $\mathcal{D}$ is just a C-Farad linear capacitor) that no matter what testing waveform $v = \hat{v}(t)$ we choose, where $\hat{v}(t)$ is no longer restricted to being sinusoidal, we would obtain the same straight-line loci in Fig. 11(b). Moreover, if we change the time rate of $\hat{v}(t)$, say changing the time scale from second to nanosecond, we would still obtain the same loci. In other words, we can predict the response $i(t)$ from any testing signal $v(t)$, by using the $v^{(0)} - i^{(-1)}$ characteristics in Fig. 11(b) to obtain $q(t) = C\hat{v}(t)$ first, and then differentiating $q(t)$ (once in this case) to obtain $i(t)$. We can at last conclude that the black box $\mathcal{D}$ in Fig. 10(a) is characterized by a loci (a straight line in this trivial example) in the $i^{(-1)} - v^{(0)}$ plane. Since every admissible signal pair $(v(t), i(t))$ measured from the black box $\mathcal{D}$ in Fig. 10(a) can be derived from the $i^{(-1)}$-versus-$v^{(0)}$ loci in Fig. 11(b), it follows that this loci is the correct constitutive relation of $\mathcal{D}$. Observe that in this case, instead of storing an infinitely large lookup table of admissible signal pairs of $\mathcal{D}$, the same information is contained compactly in a single loci in the $i^{(-1)}$-versus-$v^{(0)}$ plane.

E. Lessons From Aristotle's Law of Motion

The concept of predicting the motion (i.e, dynamical behaviors) of a physical object (e.g., stones, material bodies, planets, stars, etc.) by a law which relates two observable (i.e., measurable) attributes (i.e., physical variables) dates back at least 2300 years ago to Aristotle, who had postulated [20] that the force F acting on a body is proportional to its

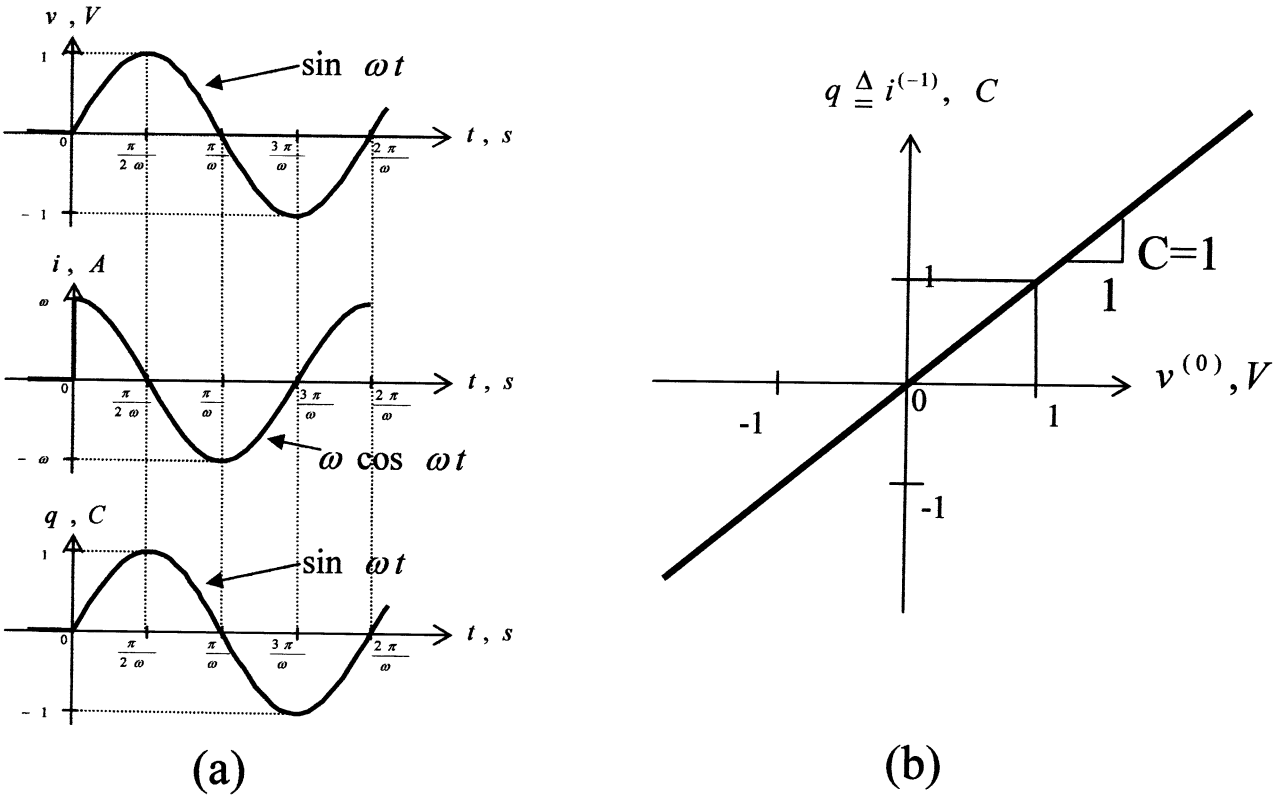

Fig. 11. (a) Three waveforms $v(t)$, $i(t)$, and $q(t) = \int_0^t i(\tau)d\tau$. (b) Loci of $(v^{(0)}(t), i^{(-1)}(t)) = (v(t), q(t))$ plotted in the $i^{(-1)}$-versus-$v^{(0)}$ plane for $\omega = 1/2, 1, 2$, assuming $A = 1$ and $C = 1$. Observe that the loci in this case is independent of ω.

velocity v of motion, namely, $F = mv$, where m is some constant of proportionality, as shown in Fig. 12(a).[2]

Although Aristotle's Law of Motion was a novel concept among the Greek philosophers of yore, no one has heard of it today because Aristotle's Law of Motion fails to predict the actual trajectory of a body in motion. We had to wait almost 2000 years to replace it with the Newton's Law of Motion $F = m(dv/dt)$, which is correct at least for mundane velocities $v \ll c$, where c is the velocity of light.

We now realize that Aristotle's Law of Motion is not valid because he had chosen an incorrect pair of physical variables (force F and velocity v) to characterize a body in motion. Indeed, if instead of postulating his law of motion, Aristotle had the necessary experimental equipment to test his postulated hypothesis, he could have carried out a simple set of experiments by applying a fixed amount of force $F = F_0$ to a free body (initially at rest) and recording the corresponding velocity $v = v(F_0)$ of the body which had responded by moving in the same direction as the applied force F_0.

[2]Aristotle was the first philosopher to adopt the scientific approach that the laws of nature must be derived from experimental observations. Due to lack of scientific instrumentations. Aristotle's observations suggest to him that a nonliving body can move only as a result of a constant external cause of motion (which we now call force F); hence, the veolcity v of the object must be proportional to the applied force F. This qualitatively enunciated Aristotlelian dynamics has survived a thousand years until it was recast by the peripatetic physicists of the Middle Ages into the algebraic form $v \propto F/m$, and subsequently into an equation $F = mv$, henceforth called Aristotle's Law of Motion in this paper.

Aristotle would have then discovered to his amazement that the recorded velocity $v(F_0)$ is not constant, as predicted by his postulate, but increases linearly with time, namely, $v = kF_0t$. If he had then plotted his entire time recordings of $v(t)$, in the F-versus-v plane, Aristotle would have seen not a single point (F_0, v_0) as predicted by his law of motion in Fig. 12(a), but a continuous loci of points which traces out a horizontal line beginning from v_0 and moving to the right at a uniform rate, as depicted by the family of six time-parameterized loci (starting from v_0) in Fig. 12(b) for six different values of applied constant force $F_0 = 3, 2, 1, -1, -2, -3$. Repeating his experiments for all possible values of F_0, Aristotle would have eventually seen the entire first and third quadrants of the F-versus-v plane in Fig. 12(b) filled up completely with horizontal lines. Aristotle would have then concluded that his postulate has no predictive value and should never see the light of day.

Aristotle would be even more amazed to discover that without a continuously applied force (i.e., $F_0(t) = 0$, $t > 0$), a free body (in a frictionless environment) when given an initial velocity v_0 would continue to move at the same velocity $v = v_0$. When plotted in the F-versus-v plane in Fig. 12(b), the instrument would register a single stationary point (0, v_0). Since this behavior holds for any velocity $v_0 > 0$ when $F = 0$, the law of motion for the case $F = 0$ can be represented in the F-versus-v plane by the horizontal axis $F = 0$ in Fig. 12(b).

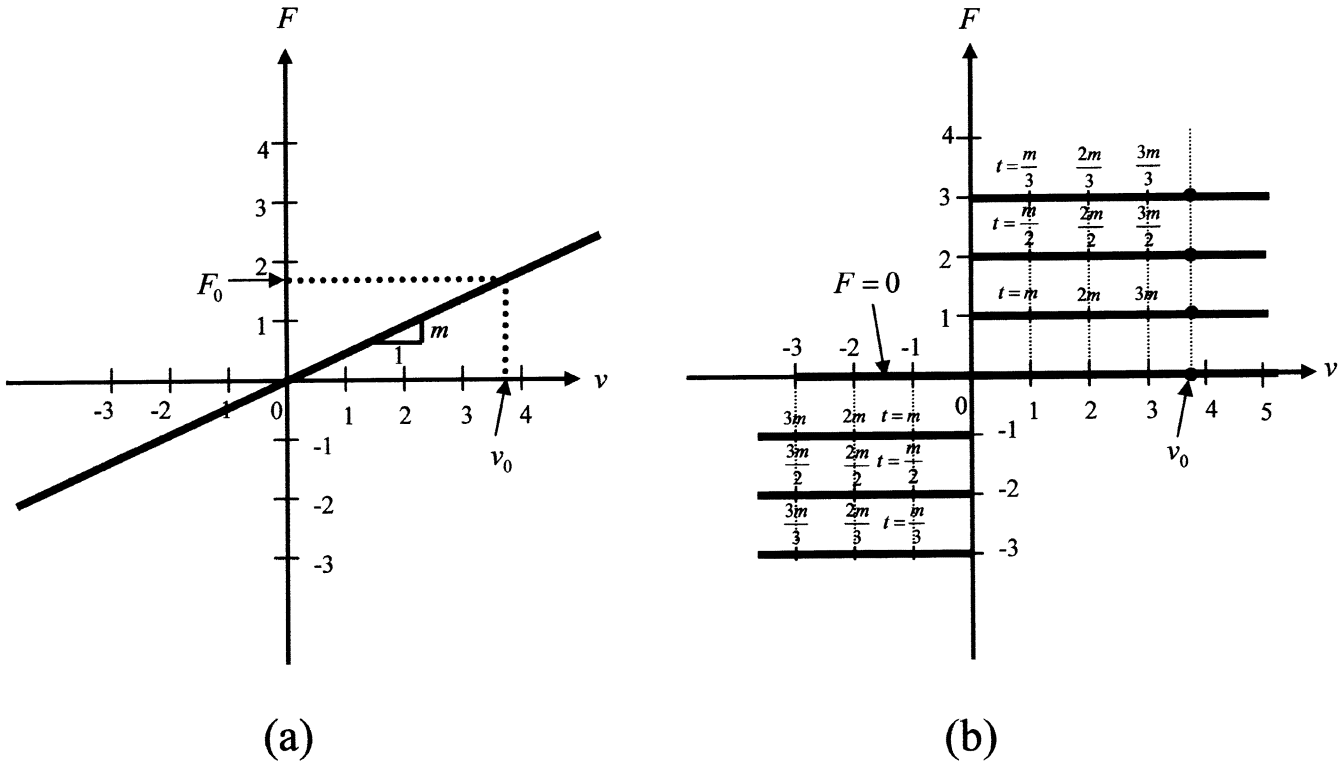

Fig. 12. (a) Aristotle's Law of Motion asserts that the velocity v of a body in motion is proportional to the applied force F, namely, $F = mv$, where $v = (dx/dt) \triangleq x^{(1)}$ is the velocity and m is a constant of proportionality. (b) In actual experiments, the relationship between the force F and its velocity v differs fundamentally from that of (a), as depicted in this figure showing six experimental results corresponding to six constant applied forces, namely, $F_0 = 3, 2, 1, -1, -2, -3$. The horizontal axis $F \equiv 0$ is not parameterized by time because every point (say $v = v_0$) on the horizontal axis is a stationary point.

This experimentally derived law of motion for zero applied force coincides with Galileo's Principle (or Newton's First Law) [21], at the bottom of the page.

A simple experimental setup which Aristotle could have used to arrive at the hypothetical scenarios described above is shown in Fig. 13 (a slightly modified version of [21, Fig. 3.3]). The body under investigation (the glider) is constrained to move along an almost frictionless horizontal track made with many tiny holes distributed uniformly along its entire length through which air is pumped to provide a cushion of air beneath the glider on which it moves. Since the glider literally floats as it moves there is almost no friction. A horizontal force is applied by hanging metal balls on the left or right pan, which is connected via a narrow tape to the glider via an almost frictionless pulley. A coordinate system is defined with the origin $(x = 0)$ at the center position of the track and positive (respectively, negative) distance is measured as the length of the tape measured from $x = 0$ to the position of the glider on the right (respectively, left). An electronic sensor system is distributed uniformly along the length of the track and its output signals are sent to an automatic microcomputer system which is programmed to calculate the distance, velocity, acceleration, etc., of the glider as a function of time, and to plot the relationship between any measured signal pairs, such as $F(t)$ versus $x(t)$, $F(t)$ versus (dx/dt), $F(t)$ versus (d^2x/dt^2), $F^{(-1)}(t)$ versus $x(t)$, $F^{(-1)}(t)$ versus (dx/dt), etc., where $F^{-1}(t) \triangleq \int_{-\infty}^{t} F(t_1)dt_1$.

If we use the experimental apparatus shown in Fig. 13 to carry out the above *gedanken* experiments which Aristotle would have carried out with the six constant forces $F = 3, 2, 1, -1, -2, -3$, the computer would have plotted the same $F(t)$-versus-(dx/dt) relationship shown in Fig. 12(b), where $v = (dx/dt)$. Such an output would have given Aristotle a definitive verdict that the object (respectively, black box) under investigation cannot be characterized by a relation (respectively, loci) in the F-versus-v plane. This conclusion is equivalent to saying that the two

Newton's First Law

If no net force acts on a body it will move in a straight line at constant velocity, or will stay at rest if initially at rest.

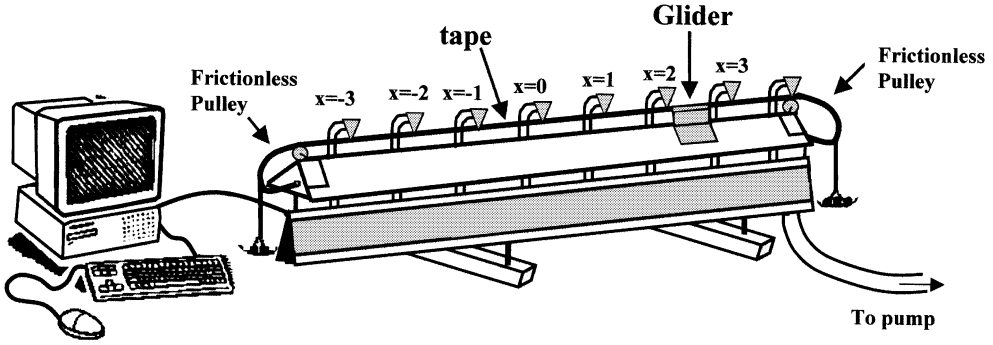

Fig. 13. Realistic laboratory set up showing an almost frictionless apparatus and the measurement system for gathering the relevant data to be processed by an automated software-driven microcomputing system. The software allows not only plotting the time evolution of the glider position $x(t)$, measured from the origin at the center, but it can also be programmed to automatically calculate $v = (dx/dt), v^{(1)} = (d^2x/dt^2), \ldots, v^{(n)} = (d^nx/dt^n)$, $\int_{-\infty}^{t} x(t_1)dt_1$, $\int_{-\infty}^{t}\int_{-\infty}^{t_2} x(t_1)dt_1dt_2$, etc., and to plot the loci resulting from any prescribed pair of waveforms $(F^{(m)}(t), x^{(n)}(t))$, for any positive and negative integers m and n.

variables consisting of the force F and veolcity v are not the correct choices for describing the dynamic characteristics of the object under investigation. Hence, Aristotle must choose another pair of variables and repeat the exercise, a trivial task for the microcomputer system in Fig. 13.

If Aristotle had chosen instead the force F and acceleration $a \triangleq (d^2x/dt^2) = (dv/dt) \triangleq v^{(1)}(t)$, and had repeated the exercise in the F-versus-$v^{(1)}$ plane for $F_0 = 3, 2, 1, -1, -2, -3$, he would have found that the loci of (F_0, $v^{(1)}(t)$) consists of six stationary points (corresponding to the six constant forces $F_0 = 3, 2, 1, -1, -2, -3$) lying on a straight line with some slope m. Repeating this exercise for all possible constant values of F, Aristotle would have obtained many more stationary points, all of them lying on the same straight line. All of these experimental results would suggest to Aristotle that the object under investigation can probably be modeled by the simple relationship $F = mv^{(1)}$. To verify his conjecture, Aristotle must apply many more testing signals, say $F = A\sin\omega t$, by modifying the experimental apparatus in Fig. 13 to allow an arbitrary Force $F = F(t)$ to be applied to the glider and then calculate the corresponding acceleration $v^{(1)}(t)$. Suppose Aristotle had carried out this experiment not only for sinusoidal signals over a large range of frequencies $\omega_1 \leq \omega \leq \omega_2$ and amplitudes, $A_1 \leq A \leq A_2$, but also for many other nonsinusoidal, as well as nonperiodic waveforms over different time scales. If in all cases, the loci of each experiment always lie on the same straight line $F = mv^{(1)}$, then Aristotle is entitled to claim that the object under investigation can be realistically modeled by the constitutive relation

$$F = mv^{(1)} = ma \tag{5}$$

over the frequency range and time scale of the testing signals. As long as the signal $F(t)$ is much slower than the velocity of light c, Aristotle would have found (5) to give an excellent prediction and he would have discovered Newton's Second Law 2000 years earlier. [3]

The preceding hypothetical scenario where Aristotle succeeded in choosing a suitable pair of physical variables (namely, force F and acceleration $v^{(1)}$) which give rise to a well-defined loci represents a lucky choice. If Aristotle had chosen some other combinations, say, force F and distance x, or force F and $x^{(-1)} \triangleq \int_{-\infty}^{t} x(t_1)dt_1$, he would not find a well-defined loci in either case. This does not imply, however, that no other pairs of physical variables would qualify. Indeed, if Aristotle had chosen instead the variable pair $p(t) \triangleq \int_{-\infty}^{t} F(t_1)dt_1$ and velocity v and had repeated the preceding exercise in the p-versus-v plane, he would have found that all loci fall neatly along the straight line $p = mv$ not only for constant force $F_0 = 3, 2, 1, -1, -2, -3$. Indeed, we can derive Fig. 15(a) from Fig. 14(a) as follows:

$$\text{Fig. 14(a)} \Rightarrow F = m\frac{d^2x}{dt^2}. \tag{6}$$

Integrating both sides of (6), we obtain

$$p(t) \triangleq \int_{-\infty}^{t} F(t_1)dt_1 = m\frac{dx}{dt}. \tag{7}$$

Consequently, both laws of motion $F = ma$ and $p = mv$ are equivalent. In fact, Newton's original law of motion is actually expressed in terms of $p(t)$, which he called the momentum and veolcity v, and which we now know is the correct version to use when the mass m is time-varying (e.g., mass of a rocket in motion) or when the velocity v is not

[3]Had Newton's Second Law been derived from this approach, the number m which we now call the mass of the object would represent simply the slope of a straight-line law of motion, which is a lot simpler to understand than to follow the long and tortuous history of the evolution of the concepts of mass [22].

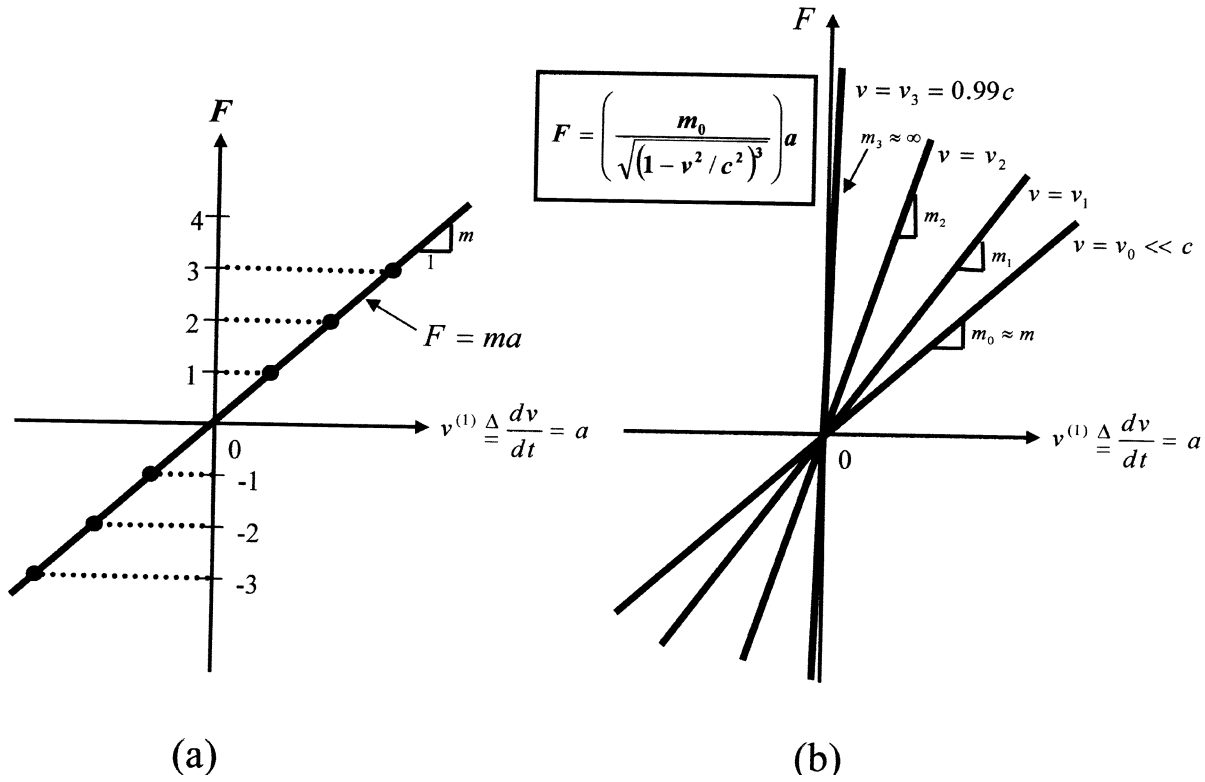

Fig. 14. (a) Popular version of Newton's Second Law, which asserts that the acceleration $a \triangleq v^{(1)} = x^{(2)}$ of a free body in motion is proportional to the applied force F. Newton's Second Law is usually expressed in the familiar form $F = ma$, where it is valid only for situations where the mass m is a constant and the velocity satisfies $v \ll c$. (b) When plotted in the F-versus-a plane, the relativistic version of Newton's Second Law is not a well-defined loci because F depends not only on the acceleration $a \triangleq x^{(2)}$, but also on the velocity $v \triangleq x^{(1)}$.

small compared to the velocity of light. Hence, we must write Newton's Second Law as follows:

$$\text{Newton's Second Law} \qquad \boxed{F = \frac{dp}{dt}, \quad p = mv} \tag{8}$$

Unfortunately, as is the case in all models, even the celebrated Newton's Second Law is an approximation of reality, and it loses its predictive ability (in the sense that it gives answers that do not agree with measurements) when the velocity $v = (dx/dt)$ of a body approaches the velocity of light c.

Thanks to Einstein's Special Theory of Relativity [21], we can show that the version of Newton's Second Law ($F = ma$) depicted in Fig. 14(a) must be replaced by the relativistic version

$$F = \left(\frac{m_0}{\sqrt{\left(1 - \frac{v^2}{c^2}\right)^3}} \right) a. \tag{9}$$

When we try to plot (9) in the F-versus-a plane, where $a \triangleq (dv/dt) = v^{(1)}$, we find the loci of $F(t)$ and $v^{(1)}(t)$ for velocity $v(t)$ close to c is no longer a single curve, let alone a straight line, but depends on the velocity $v(t)$ at each instant of time. In the F-versus-a plane, (9) represents a continuum of straight lines parameterized by the velocity v, as shown in Fig. 14(b). Observe that when $v \ll c$ (represented by the lowest straight line with slope m_0), (9) reduces to $F = m_0 a$, where $m_0 \approx m$, in agreement with Newton's Second Law in Fig. 14(a).

Observe, however, that (9) involves three variables, namely, F, v, and $v^{(1)}$. It follows that the earlier hypothetical choice by Aristotle of the variable pair F and a is in fact incorrect when the velocity $v(t)$ approaches the velocity of light.

Fortunately, it turns out that (9) is an exact differential in the sense that if we integrate both sides of (9), we would obtain

$$\begin{aligned} \int_{-\infty}^{t} F(t_1)dt_1 &= \int_{-\infty}^{t} \left[\left(\frac{m_0}{\sqrt{\left(1 - \frac{v^2}{c^2}\right)^3}} \right) \frac{dv}{dt_1} \right] dt_1 \\ &= \int_{0}^{v} \frac{d}{dv} \left[\frac{m_0 v}{\sqrt{\left(1 - \frac{v^2}{c^2}\right)}} \right] dv \\ &= \frac{m_0 v}{\sqrt{1 - \frac{v^2}{c^2}}} \end{aligned} \tag{10}$$

which is a function involving only the two variables $F^{(-1)} \triangleq p(t)$ and $v = x^{(1)}$. In other words, $F^{(-1)}$ and $x^{(1)}$ are the correct choice of variables for characterizing the law of motion of a body traveling close to the velocity of light, resulting in

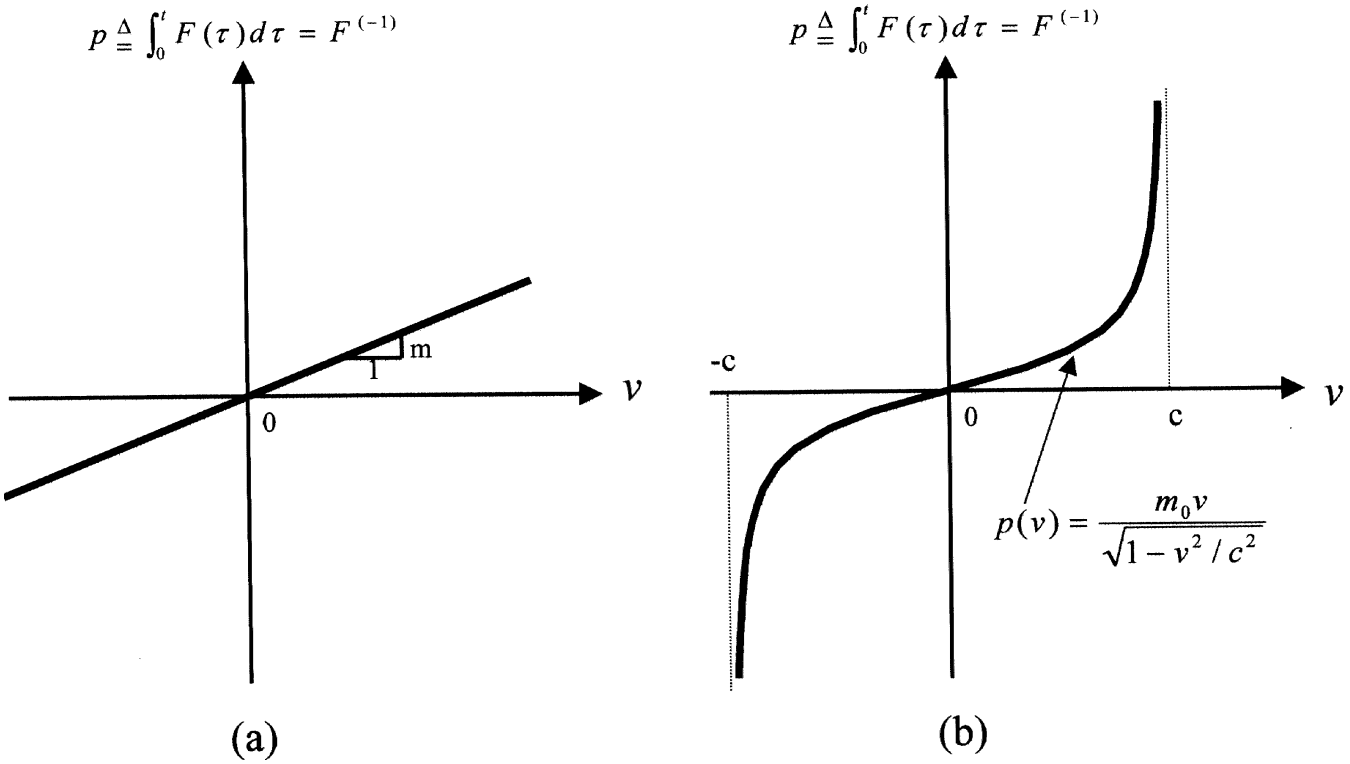

Fig. 15. (a) Less common but original version of Newton's Second Law is plotted in the momentum $p \triangleq F^{(-1)}$ versus the velocity $v \triangleq x^{(1)}$ plane. In this version, Newton's Second Law asserts that $F = (dp/dt)$, where $p \triangleq mv$. Note that for constant m, this version coincides with the popular version $F = ma$. The advantage of this version of Newton's Second Law is that it is valid even when m varies with time, a situation where the first version ceases to hold. (b) When plotted in the momentum p-versus-velocity plane, Newton's Second Law can be easily extended to Einstein's Relativistic Law of motion by merely changing the linear momentum function $p = mv$ from (a) into a nonlinear momentum function $p = p(v)$. Since $p(v)$ does not depend on any additional variable, as in Fig. 14(b), it follows that the variable pair $p \triangleq F^{(-1)}$ and $v \triangleq x^{(1)}$ is the correct choice for characterizing the constitutive relation of a free body in motion.

the following law of motion that has so far predicted all measurement outcomes correctly[4]

$$\text{Relativistic Law of motion} \qquad \boxed{F = \frac{dp}{dt}, \quad p = \frac{m_0 v}{\sqrt{1-\frac{v^2}{c^2}}}} \qquad (11)$$

When plotted in the p-versus-v plane, (11) gives a well-defined nonlinear $p - v$ characteristic, as shown in Fig. 15(b).

The most important lesson from the above result is that a body (respectively, black box) may be characterized by several distinct but equivalent models which cannot be distinguished from any measurements, when the models are linearized (i.e., approximated by a linear model). In other words, it is impossible to tell which pair of variables are the correct choice when the corresponding models are linear, as demonstrated by the two equivalent versions of Newton's Second Law of motion corresponding to Fig. 14(a) and Fig. 15(a), respectively. However, when the more realistic nonlinear models are invoked, then each model is unique in the sense that a measurement can be devised to identify which pair of variables must be chosen to correctly predict all possible measurement outcomes. This important observation is particularly relevant in Section V.

[4]Equation (11) is sometimes written in the more familiar form $p = m(v)v$ where $m(v) \triangleq (m_0/(\sqrt{1 - v^2/c^2}))$ is called the relativistic mass and m_0 is called the rest mass. Observe that when $v \ll c$, $m(v) \approx m_0$ and (11) reduce to (8).

III. Axiomatic Definition of Elementary Circuit Elements

There are three standard two-terminal circuit elements from electrical engineering: resistor, inductor, and capacitor. They are often referred to loosely as resistance R, inductance L, and capacitance C, respectively. There is little harm in abusing terminologies as long as the device is linear. However, there is a fundamental conceptual difference between the two words "resistor" and "resistance" (or respectively, "inductor" and "inductance," or "capacitor" and "capacitance") when the device is nonlinear. The difference between these two terms is best understood by comparing them to the mechanical analog between a free body and its mass m where it is clear that a free body is just the name of an object whereas the mass refers to an intrinsic attribute of the object [22], namely, the slope of the straight-line F-versus-$v^{(1)}$ characteristic in Fig. 14(a). When a device is nonlinear, which is the case in most useful nanodevices, it is essential

that we understand the three words "resistor," "inductor," and "capacitor" as names of three different circuit element species, whereas the three terms "resistance," "inductance," and "capacitance" refer to an intrinsic local attribute of each element species.

The above three standard circuit elements were originally defined to model three distinct device prototypes, henceforth called circuit element species, or simply circuit elements in this paper. For example, when a voltage is applied across a two-terminal device made by winding a very fine and insulated metallic wire around a nonmetallic core, heat is generated and dissipated into the environment due to the random collisions among the free electrons propagating through the wire. Since these incessant collisions effectively resist the flow of electrons, the device is aptly called a resistor and defined via the constitutive relation (known universally as Ohm's Law) [21]

$$v = Ri \tag{12}$$

where R is the resistance.

If we replace the nonmetallic core in the above device by a ferromagnetic core and increase the diameter of the wire sufficiently so that the heat generated is negligible, we will find that the net energy entering the device is not lost but rather stored in the device's magnetic field, and can be fully recovered, as demonstrated by switching the energized device momentarily across a light bulb and observing a brief flash of light in the absence of any external source. Since the energy in this case is stored via magnetic induction, the device is therefore called an inductor and defined via the constitutive relation [21]

$$v = L\frac{di}{dt} \tag{13}$$

where L is the inductance.

The third device prototype was made originally of a pair of parallel metallic plates separated by a di-electric, such as air. Although no dc current can flow through this device, energy can be stored in the electric field generated by applying a battery across the device. Like the inductor, this device is also lossless, as can be demonstrated by the same light bulb experiment. This device was originally called a condenser in view of an analogy from thermodynamics, but was subsequently renamed a capacitor and defined via the constitutive relation [21]

$$i = C\frac{dv}{dt} \tag{14}$$

where C is the capacitance.

While the above brief history of the three standard circuit elements is both instructive and motivating, it has fostered a rigid mindset among generations of engineers educated with a background in classical circuit theory which makes it rather difficult for them to cope with the modern world of nanoelectronic devices, where the device behaviors and operating mechanisms are completely different from those made from wires, coils, and parallel metal plates. Yet future nanodevice engineers will continue to model their exotic devices by finding some "magical" but nonexisting circuit combinations of such standard circuit building blocks as linear resistors, inductors, capacitors, and their nonlinear and multiport versions.

To avoid repeating the above scenario from recurring in future generations, we will use an axiomatic approach in this paper to define a new family of circuit elements which are essential for modeling unconventional and exotic devices, including biological, molecular, and nanodevices. Our axiomatic approach is completely rigorous and device independent, since our definitions of the elementary circuit elements do not depend on the operating mechanisms inside the device but are couched in terms of externally measurable constitutive relations. In other words, our elementary circuit elements are black boxes.

In order to model the many unconventional and exotic quantum mechanical phenomena [6] (e.g., tunneling, entanglement, Coulomb blockade, single-electron dynamics, kondo effect, Aharonov–Bohm effect, electron–photon interactions, polaron and bipolaron propagation, ion–membrane dynamics, ultrafast chemical reactions, redox chemical activations, etc.) which are essential to the nonlinear dynamics of many biological, molecular, and nanodevices, it is necessary to enlarge our repertoire of model building blocks by introducing a sufficiently rich family of elementary circuit elements which play the same role as the set of basis vectors used to define a vector space. Just as a function space requires an infinite number of basis vectors in order to guarantee that every function belonging to the function space can be approximated via the basis vectors to any prescribed accuracy [23], an infinite set of elementary circuit elements will be needed to guarantee that every real or hypothetical nonlinear device can be realistically modeled to within any prescribed tolerance. Fortunately, just as most commonly encountered nonlinear functions can be realistically approximated by using only a small number of basis functions, we can expect that most nanodevices can be adequately modeled by using only a small number (less than five in most practical cases) of elementary circuit elements to be introduced below.

A. (α, β) Element

Definition of (α, β) Element

A two-terminal or one-port black box characterized by a constitutive relation in the $v^{(\alpha)}$-versus-$i^{(\beta)}$ plane is called an (α, β) element, where $v^{(\alpha)}$ and $i^{(\beta)}$ are complementary variables[5] derived from the voltage $v(t)$ and current $i(t)$ via (3) and (4), respectively.

[5]A pair of variables ($v^{(\alpha)}$, $i^{(\beta)}$) are said to be complementary if $v^{(\alpha)}(t)$ is derived from a voltage signal $v(t)$, while $i^{(\beta)}(t)$ is derived from a current signal $i(t)$. Here, α and β are any integer (negative, positive, or zero) which need not be identical. It follows from Tellegen's Theorem [12] that any set of b signals $v_k^{(\alpha)}(t)$, $k = 1, 2, \ldots, b$ which satisfy the Kirchhoff Voltage Law (KVL) for a circuit made of b arbitrary two-terminal elements is orthogonal at all times to any set of b signals $i_k^{(\beta)}(t)$, $k = 1, 2, \ldots, b$ which satisfy the Kirchhoff Current Law (KCL) in the same circuit. In particular, the two vector functions of time $\mathbf{v}^{(\alpha)} \triangleq [v_1^{(\alpha)}(t), v_2^{(\alpha)}(t), \cdots, v_b^{(\alpha)}(t)]^T$ and $\mathbf{i}^{(\beta)} \triangleq [i_1^{(\beta)}(t), i_2^{(\beta)}(t), \cdots, i_b^{(\beta)}(t)]^T$ lie in complementary orthogonal vector spaces.

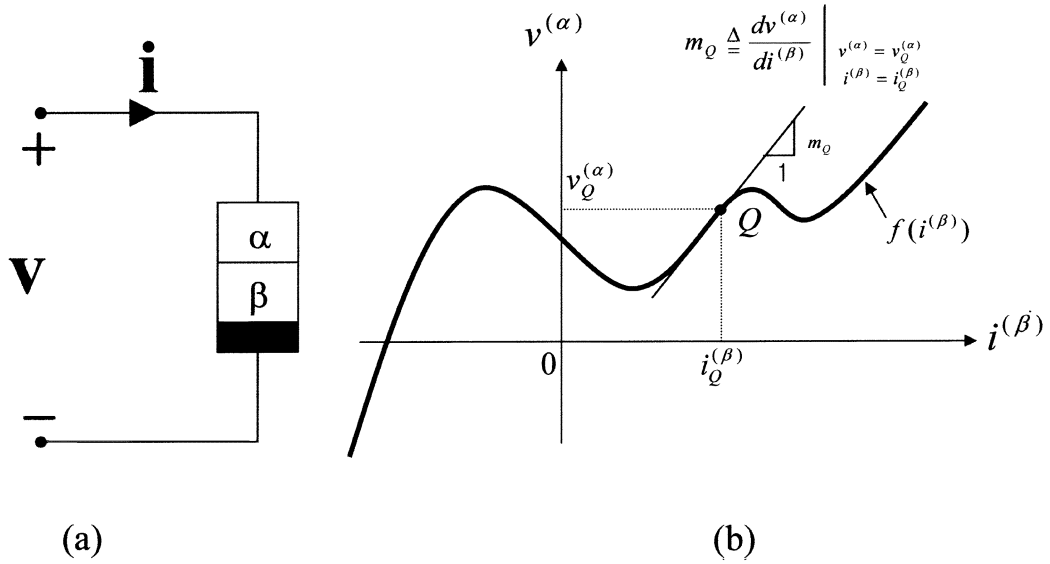

Fig. 16. (a) Symbol of a two-terminal or one-port (α, β) element defined by a loci of points, called its constitutive relation, in the $v^{(\alpha)}$-versus-$i^{(\beta)}$ plane. (b) A hypothetical loci (sometimes called a $v^{(\alpha)} - i^{(\beta)}$ curve or $v^{(\alpha)} - i^{(\beta)}$ characteristic in this paper) plotted on the $v^{(\alpha)} - i^{(\beta)}$ plane. Unless otherwise stated, in this paper, we will consistently assume $v^{(\alpha)}$ as the vertical axis, and assume $i^{(\beta)}$ as the horizontal axis.

The symbol we have chosen for an (α, β) element is shown in Fig. 16(a), where the upper integer α is always associated with the voltage-based signal $v^{(\alpha)}(t)$, and the lower integer β is always associated with the current-based signal $i^{(\beta)}(t)$. Throughout this paper, the Greek letter α is reserved exclusively for $v^{(\alpha)}(t)$, henceforth referred to as the voltage exponent. Similarly, the Greek letter β is reserved exclusively for $i^{(\beta)}(t)$, henceforth referred to as the current exponent. A thick black band is needed at the bottom of the symbol in Fig. 16(a) in order to distinguish the voltage index from the current index when the (α, β) element is drawn on a horizontal or vertical position.

To avoid confusion, the constitutive relation for an (α, β) element will always be drawn in this paper with $v^{(\alpha)}$ chosen as the vertical axis, and with $i^{(\beta)}$ chosen as the horizontal axis. In general, the constitutive relation can be represented as a set of points, called a loci, a curve, or a characteristic in this paper, depending on the context. A hypothetical $v^{(\alpha)}$-versus-$i^{(\beta)}$ characteristic is shown in Fig. 16(b). Any point Q on the $v^{(\alpha)} - i^{(\beta)}$ characteristic is called an operating point and we will use the generic symbol m_Q to denote the slope of the $v^{(\alpha)} - i^{(\beta)}$ characteristic at an operating point Q located at $(v_Q^{(\alpha)}, i_Q^{(\beta)})$ assuming the $v^{(\alpha)} - i^{(\beta)}$ curve defined by $v^{(\alpha)} = f(i^{(\beta)})$ is differentiable at Q.

Taking the Taylor-Series expansion of $f(\bullet)$ in a small neighborhood around Q, we obtain

$$v^{(\alpha)} = f\left(i^{(\beta)}\right) = i_Q^{(\beta)} + m_Q\left(i^{(\beta)} - i_Q^{(\beta)}\right) + \text{h.o.t.} \quad (15)$$

where "h.o.t." denotes higher order derivatives of $f(i^{(\beta)})$ at Q.

Notwithstanding the fact that the $v^{(\alpha)} - i^{(\beta)}$ curve shown in Fig. 16(b) is highly nonlinear, we will see in Part II of this paper that whether a two-terminal device is useful or inept depends only on the $v^{(\alpha)} - i^{(\beta)}$ characteristic in an infinitesimal neighborhood of an operating point Q, namely, it suffices to examine the small-signal characteristic at Q, defined by

$$\boxed{\delta v^{(\alpha)}(t) = m_Q \delta i^{(\beta)}(t)} \quad (16)$$

where

$$\boxed{m_Q \triangleq \frac{dv^{(\alpha)}}{di^{(\beta)}}\Big|_{\substack{v^{(\alpha)}=v_Q^{(\alpha)} \\ i^{(\beta)}=i_Q^{(\beta)}}} = f'\left(i^{(\beta)}\right)|_{i^{(\beta)}=i_Q^{(\beta)}}} \quad (17)$$

and

$$\delta v^{(\alpha)} \triangleq v^{(\alpha)} - v_Q^{(\alpha)} \text{ and } \delta i^{(\beta)} \triangleq i^{(\beta)} - i_Q^{(\beta)}$$

B. Small-Signal Impedance $Z_Q(j\omega)$

Taking the single-sided Laplace transform $\mathcal{L}\{g(t)\} = \int_0^\infty g(t)e^{-st}dt \triangleq G(s)$ for each of the time functions (henceforth called signals) in (16), we obtain

$$\mathcal{L}\left\{\delta v^{(\alpha)}(t)\right\} = m_Q \mathcal{L}\left\{i^{(\beta)}(t)\right\}. \quad (18)$$

Invoking the standard Laplace transform operation on time derivatives and time integrals (assuming zero initial conditions) of the signals $v^{(\alpha)}(t)$ and $i^{(\beta)}(t)$ in (18), we obtain

$$\mathcal{L}\{v(t)\} = Z_Q(s)\mathcal{L}\{i(t)\} \quad (19)$$

where

$$\boxed{Z_Q(s) = s^{(\beta-\alpha)}m_Q} \quad (20)$$

is called the *small-signal impedance* of the $(\alpha - \beta)$ element at the operating point Q. Observe that the small-signal

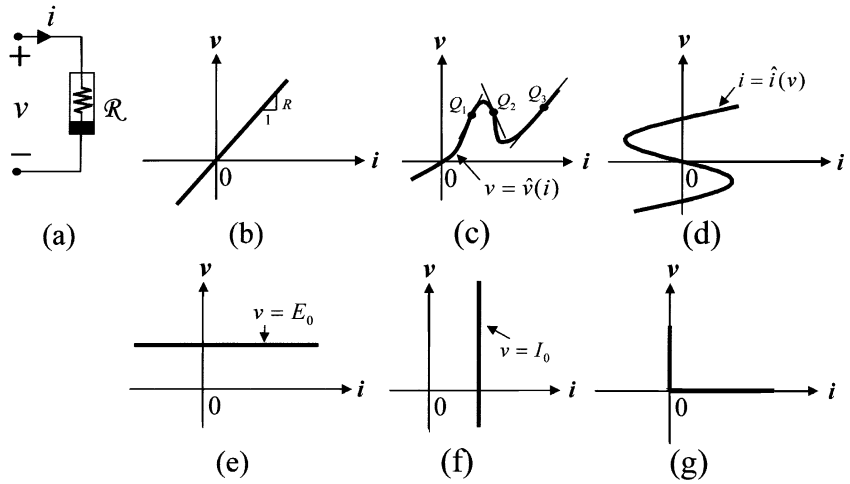

Fig. 17. (a) Symbol of a two-terminal resistor defined by a constitutive relation in the $v - i$ plane. (b) When the constitutive relation is described by a straight line $v = Ri$, we recover Ohm's Law as the simplest special case. In this case, the generic nonlinear resistor reverts to the classical circuit element called the resistor, a term often used synonymously with its resistance R. In the linear case, no ambiguities result if we delete the box in (a) enclosing the conventional resistor symbol. The SI unit for the resistance is called the ohm (Ω). (c) A hypothetical current-controlled $v - i$ characteristic defined by a single-valued function $v = \hat{v}(i)$. (d) A hypothetical voltage-controlled $v - i$ characteristic defined by a single-valued function $i = \hat{i}(v)$. (e) A resistor characterized by a constant voltage $v = E_0$ constitutive relation is called a $v^{(0)}$ source; following tradition, we call this a dc voltage source, or even simply a battery, by an abuse of terminology which confuses a device for its model. (f) A resistor characterized by a constant current $i = I_0$ constitutive relation is called an $i^{(0)}$ source; following tradition, we call this a dc current source. (g) A resistor characterized by the constitutive relation $v \geq 0$ when $i = 0$, $i \geq 0$ when $v = 0$, or equivalently, $v\, i = 0$ is called an ideal diode.

impedance $Z_Q(s)$ is a function only of the complex variables $s \triangleq \sigma + j\omega$ and the slope m_Q of the $v^{(\alpha)} - i^{(\beta)}$ curve, and is well-defined at every operating point Q where the curve $v^{(\alpha)} = f(i^{(\beta)})$ is differentiable.

Since it is often convenient to apply sinusoidal signals as testing signals, it suffices to set $\sigma = 0$ in such situations and the resulting small-signal impedance simplifies to

$$\boxed{Z_Q(j\omega) = (j\omega)^{\beta-\alpha} m_Q} \tag{21}$$

in the frequency domain $s = j\omega$.

When the exponent $(\beta - \alpha)$ in (21) is a positive or negative odd integer, the small-signal impedance $Z_Q(j\omega)$ is a pure imaginary number at each frequency ω and will henceforth be called a *conservative* impedance for reasons that will be clear later.

When the exponent $(\beta - \alpha)$ in (21) is a positive, zero, or negative even integer, the small-signal impedance $Z_Q(j\omega)$ is a pure real number at each frequency ω and will henceforth be called a *nonconservative* impedance.

IV. Frequency-Independent Elements

An (α, β) element is said to be frequency independent iff its associated small-signal parameter at all operating points does not depend on the frequency ω of the applied sinusoidal excitations. In this section, we will show that there are only four frequency-independent (α, β) elements, three of which coincides with the three classical circuit elements described earlier, namely, the Resistor, the Inductor, and the Capacitor. The fourth element, called the Memristor [13], though unconventional, is particularly relevant to nanodevices involving ion transports or other unconventional charge transport mechanisms [10], [11]

A. Every (0,0) Element Is Called a Resistor $\mathcal{R}$

A two-terminal (α, β) element with $\alpha = 0$, $\beta = 0$ and characterized by any constitutive relation $\mathcal{R}(v, i) = 0$ (where $v \triangleq v^{(0)}, i \triangleq i^{(0)}$) in the $v-i$ plane is called a resistor [2]. We will use the symbol shown in Fig. 17(a) to denote a two-terminal resistor. A resistor is said to be linear iff its constitutive relation is a straight line through the origin in the $v - i$ plane with a constant slope equal to R, as shown in Fig. 17(b). In this case, we recover $v = Ri$, i.e., Ohm's Law from (12) where the small-signal resistance R is simply called the resistance, since it does not depend on the operating point Q.

The constitutive relation $\mathcal{R}$ shown in Fig. 17(c) defines the voltage v as a single-valued function of the current i, namely, $v = \hat{v}(i)$; hence, $\mathcal{R}$ is called a current-controlled resistor. In this case, the slope R at any operating point Q on $\mathcal{R}$ is called the small-signal resistance R at Q. Note that $R(Q_1) > 0$, $R(Q_2) < 0$ and $R(Q_3) > 0$ in Fig. 17(b). When $R(Q) < 0$, we called it a small-signal *negative resistance* at Q.

The constitutive relation $\mathcal{R}$ shown in Fig. 17(d) defines the current i as a single-valued function of the voltage v, namely, $i = \hat{i}(v)$; hence, $\mathcal{R}$ is called a voltage-controlled resistor.

When the constitutive relation $\mathcal{R}$ is constrained to lie on a horizontal line $(v = E_0)$, as in Fig. 17(e), or a vertical line $(i = I_0)$, as in Fig. 17(f), we call the resistor $\mathcal{R}$ a voltage

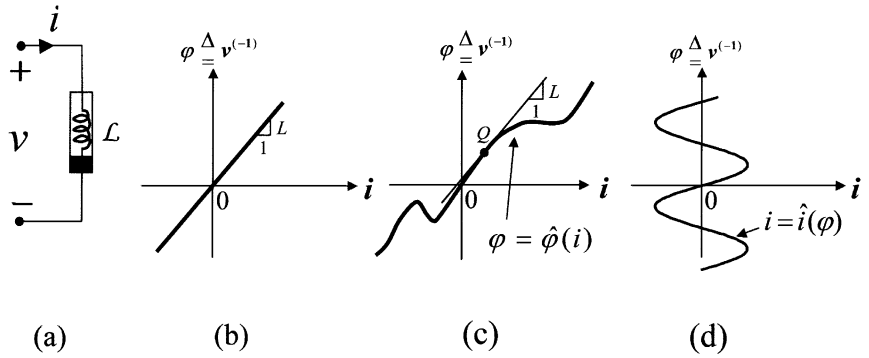

Fig. 18. (a) Symbol of a two-terminal inductor defined by a constitutive relation in the $\varphi - i$ plane, where $\varphi(t) \triangleq \int_{-\infty}^{t} v(\tau)d\tau \triangleq v^{(-1)}$ is called the flux, regardless of whether or not one can associate it with a physical magnetic flux linking a current-carrying coil. (b) When the constitutive relation is described by a straight line $\varphi = Li$, the generic nonlinear inductor reverts to the classical circuit element called the inductor, a term often used synonymously with its inductance L. In the linear case, no ambiguities result if we delete the box in (a) enclosing the conventional inductor symbol. The SI unit for the inductance is called the Henry (H). (c) A hypothetical current-controlled $\varphi - i$ characteristic defined by a single-valued function $\varphi = \hat{\varphi}(i)$. (d) A hypothetical flux-controlled $\varphi - i$ characteristic defined by a single-valued function $i = \hat{i}(\varphi)$.

source or a current source, respectively.[6] Note that in these two special cases, one of the two coordinates (voltage in Fig. 17(e) [respectively, current in Fig. 17(f)] is constrained to assume a fixed value $v = E_0$ (respectively, $i = I_0$) independent of its associated current (respectively, voltage). In particular, a voltage source with $v = E_0 = 0$ is called a short circuit, and a current source with $i = I_0 = 0$ is called an open circuit.

A resistor which behaves like a short circuit when $i \geq 0$, and an open circuit when $v \leq 0$, as shown in Fig. 17(g) is called an ideal diode.

In general, an (α, β) element is called a $v^{(\alpha)}$ source (respectively, $i^{(\beta)}$ source) iff $v^{(\alpha)}$ (respectively, $i^{(\beta)}$) is constrained to assume a constant value, independent of the value of its associated complementary variable. Since only one of the two port variables is controllable via the external circuit, a $v^{(\alpha)}$ source and an $i^{(\beta)}$ source have only one degree of freedom and are called constraint elements.

We note that most commercially available two-terminal devices described by a $V - I$ curve (e.g., pn-junction diodes, zener diodes, varistors, tunnel diodes, etc.) can be realistically modeled (at least in relatively low-frequency applications) as nonlinear two-terminal resistors [1].

B. Every (−1, 0) Element Is Called an Inductor $\mathcal{L}$

A two-terminal (α, β) element with $\alpha = -1$, $\beta = 0$ and characterized by any constitutive relation $\mathcal{L}(\varphi, i) = 0$ (where $\varphi \triangleq v^{(-1)}$, $i \triangleq i^{(0)}$) in the $\varphi - i$ plane is called an inductor [1]. We will use the symbol shown in Fig. 18(a) to denote a two-terminal inductor. An inductor is said to be linear iff its constitutive relation is a straight line through the origin in the $\varphi - i$ plane with a constant slope equal to L, as shown in Fig. 18(b). In this case, we obtain $\varphi = Li$, where the slope L is called the small-signal inductance, or simply the inductance, since it does not depend on the operating point Q. Differentiating both sides of $\varphi = Li$, we recover the familiar constitutive relation ($v = L(di/dt)$) of the classical linear inductor given earlier in (13).

The constitutive relation $\mathcal{L}$ shown in Fig. 18(c) depicts the flux $\varphi \triangleq v^{(-1)}$, where

$$\varphi \triangleq \int_{-\infty}^{t} v(t_1)dt_1 \tag{22}$$

as a single-valued function of the current i, namely

$$\boxed{\varphi = \hat{\varphi}(i)} \tag{23}$$

and, hence, $\mathcal{L}$ is called a current-controlled inductor. In terms of the port voltage and the port current, the constitutive relation (23) assumes the equivalent form

$$\boxed{v(t) = L(i)\frac{di}{dt}} \tag{24}$$

where

$$L(i) \triangleq \frac{d\hat{\varphi}(i)}{di} \tag{25}$$

is called the small-signal inductance $L(i)$. When the slope $L < 0$ at some operating point Q, we call it a *negative inductance*.

The constitutive relation $\mathcal{L}$ shown in Fig. 18(d) depicts the current i as a single-valued function of the flux φ; hence, $\mathcal{L}$ is called a flux-controlled inductor.

The flux-controlled sinusoidal $\varphi - i$ relationship ($i = I_0 \sin k\varphi$, $k \triangleq ((2e)/\hbar)$, where $e =$ electron charge and $\hbar =$ Planck's constant) shown in Fig. 18(d) can be used to model

[6]Some readers may feel uncomfortable in calling a voltage source or a current source a resistor. There is in fact a sound theoretical reason for this axiomatic method of classification because it can be proved [2] that an n-port black box made of arbitrary interconnections of nonlinear resistors, linear resistors, voltage sources, and current sources always give rise to an n-port resistor. Since logic and any axiomatic theory dictates that mixtures of elements from one species (in this case, the species are resistors) must yield an element of the same specie (called the closure principle in mathematics), voltage sources and current sources must be classified as resistors in order for this fundamental closure principle to hold.

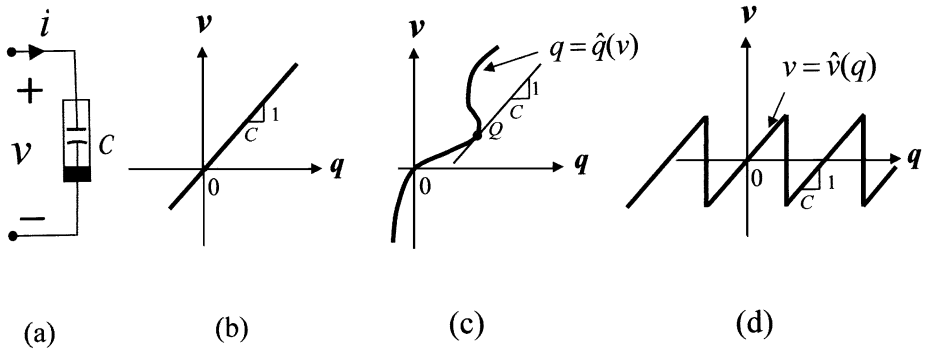

Fig. 19. (a) Symbol of a two-terminal capacitor defined by a constitutive relation in the $v-q$ plane, where $q(t) \triangleq \int_{-\infty}^{t} i(\tau)\, d\tau \triangleq i^{(-1)}$ is called the charge, regardless of whether or not one can associate it with a physical charge residing on the plate of a capacitor with an applied dc voltage. (b) When the constitutive relation is described by a straight line $q = C\, v$, the generic nonlinear capacitor reverts to the classical circuit element called the capacitor, a term often used synonymously with its capacitance C. In the linear case, no ambiguity results if we delete the box in (a) enclosing the conventional capacitor symbol. The SI unit for the capacitance is called the farad (F). (c) A hypothetical voltage-controlled $v-q$ characteristic defined by a single-valued function $q = \hat{q}(v)$. (d) A hypothetical charge-controlled $v-q$ characteristic defined by a single-valued function $v = \hat{v}(q)$.

the superconducting component due to the Cooper pairs in a Josephson junction [24]–[26]. It is important to observe that the physical variable φ is equal to $((2e)/\hbar)$ times the difference between two quantum mechanical phases, and not the familiar magnetic flux in a current-carrying coil from electromagnetic field theory.

C. Every (0, −1) Element Is Called a Capacitor $\mathcal{C}$

A two-terminal (α, β) element with $\alpha = 0$, $\beta = -1$ and characterized by any constitutive relation $\mathcal{C}(q, v) = 0$ (where $q \triangleq i^{(-1)}$, $v \triangleq v^{(0)}$) in the $q - v$ plane is called a capacitor [1]. We will use the symbol shown in Fig. 19(a) to denote a two-terminal capacitor. A capacitor is said to be linear iff its constitutive relation is a straight line through the origin in the plane, as shown in Fig. 19(b). In this case, we obtain $q = Cv$, where the reciprocal[7] slope $C \triangleq ((dq)/(dv))$ is called the small-signal capacitance, or simply the capacitance, since it does not depend on the operating point Q. Differentiating both sides of $q = Cv$, we recover the familiar constitutive relation $(i = C((dv)/(dt)))$ of the classical linear capacitor given earlier in (13).

The constitutive relation $\mathcal{C}$ shown in Fig. 19(c) depicts the charge q as a single-valued function of the voltage v

$$\boxed{q = \hat{q}(v)} \tag{26}$$

and, hence, $\mathcal{C}$ is called a voltage-controlled capacitor. In terms of the port voltage v and the port current i, the constitutive relation (26) assumes the equivalent form

$$\boxed{i(t) = C(v)\frac{dv}{dt}} \tag{27}$$

[7]The reader may find it more convenient to interchange the v and q axes in Fig. 19 so that the vertical axis denotes the charge q and the horizontal axis denotes the voltage v. Our current choice is in keeping, however, with our standing assumptions where $v^{(\alpha)}$ is plotted as the vertical axis and where $i^{(\beta)}$ is plotted as the horizontal axis.

where

$$C(v) \triangleq \frac{d\hat{q}(v)}{dv} \tag{28}$$

In this case, the reciprocal slope $C(v)$ at any operating point Q on $\mathcal{C}$ is called the small-signal capacitance $C(v_Q)$. When the reciprocal slope $C < 0$ at some operating point Q, we call it a *negative capacitance*.

The constitutive relation $\mathcal{C}$ shown in Fig. 19(d) depicts the voltage v as a single-valued function of the charge q; hence, $\mathcal{C}$ is called a charge-controlled capacitor. This $q-v$ curve can be used to model the charge-versus-voltage characteristic of a two-terminal single-electron device [27].

D. Every (−1,−1) Element Is Called a Memristor $\mathcal{M}$

A two-terminal (α, β) element with $\alpha = -1$, $\beta = -1$ and characterized by any constitutive relation $\mathcal{M}(\varphi, q) = 0$ (where $\varphi \triangleq v^{(-1)}$, $q \triangleq i^{(-1)}$) in the $\varphi - q$ plane is called a memristor, a contraction of *memory resistor* [13]. We will use the symbol shown in Fig. 20(a) to denote a two-terminal memristor. A memristor is said to be linear iff its constitutive relation is a straight line through the origin in the $\varphi - q$ plane with a slope equal to M, as shown in Fig. 20(b). In this case, we obtain $\varphi = Mq$, where the constant slope M is called the small-signal memristance, or simply the memristance, since it does not depend on the operating point Q. Differentiating both sides of $\varphi = Mq$, we obtain $v = Mi$, which is Ohm's Law once again. Since it is not possible to distinguish a two-terminal linear memristor from a two-terminal linear resistor, its existence and relevance as a new species cannot be predicted from classical linear circuit theory, and had to await the advent of nonlinear circuit theory to see the light of day [13].

The constitutive relation $\mathcal{M}$ shown in Fig. 20(c) depicts the flux $\varphi \triangleq v^{(-1)}$ as a single-valued function of the charge $q \triangleq i^{(-1)}$, namely

$$\boxed{\varphi = \hat{\varphi}(q)} \tag{29}$$

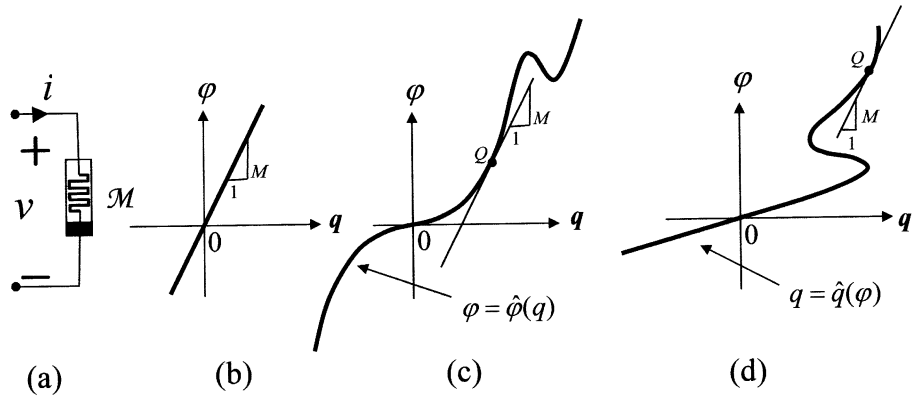

Fig. 20. (a) Symbol of a two-terminal memristor defined by a constitutive relation in the $\varphi - q$ plane, where $\varphi(t) \triangleq \int_{-\infty}^{t} v(\tau)\, d\tau \triangleq v^{(-1)}$ is the flux and $q(t) \triangleq \int_{-\infty}^{t} i(\tau)\, d\tau \triangleq i^{(-1)}$ is the charge. (b) When the constitutive relation is described by a straight line $\varphi = Mq$, the nonlinear memristor reduces to a linear memristor which is indistinguishable from that of a linear resistor. (c) A hypothetical charge-controlled $\varphi - q$ characteristic defined by a single-valued function $\varphi = \hat{\varphi}(q)$. (d) A hypothetical flux-controlled $\varphi - q$ characteristic defined by a single-valued function $q = \hat{q}(\varphi)$.

and, hence, $\mathcal{M}$ is called a charge-controlled memristor. In terms of the port voltage and the port current, the constitutive relation (29) assumes the equivalent form

$$\boxed{v = M(q)i} \tag{30}$$

where

$$M(q) \triangleq \frac{d\hat{\varphi}(q)}{dq}. \tag{31}$$

is called the small-signal memristance $M(q)$.

An examination of (30) shows that a two-terminal charge-controlled memristor behaves like a linear resistor described by Ohm's Law, except that its small-signal resistance $M(q)$ defined by (31) is not a constant, but varies with the instantaneous value of the charge $q(t) \triangleq \int_{-\infty}^{t} i(t_1)dt_1$ which bookkeeps the past events of the input current $i(t)$. In other words, the resistance M has a "memory," and the name "memory resistor," or *memristor* for short, was assigned to this heretofore missing fourth basic nonlinear circuit element. It follows from Ohm's Law that the SI unit of memristance is the ohm (Ω).

The constitutive relation $\mathcal{M}$ shown in Fig. 20(d) depicts the charge $q \triangleq i^{(-1)}$ as a single-valued function of the flux $\varphi \triangleq v^{(-1)}$, namely

$$q = \hat{q}(\varphi) \tag{32}$$

and, hence, $\mathcal{M}$ is called a flux-controlled memristor. In terms of the port voltage and the port current, the constitutive relation in (32) assumes the equivalent form

$$\boxed{i = G(\varphi)v} \tag{33}$$

where

$$\boxed{G(\varphi) \triangleq \frac{d\hat{q}(\varphi)}{d\varphi}} \tag{34}$$

is called the small-signal memductance $G(\varphi)$. It follows that $G(\varphi)$ has the same SI unit [siemen (S)] as that of conductance.

Indeed, if we apply a time-varying voltage $v(t)$ across a two-terminal flux-controlled memristor and calculate its associated flux waveform $\varphi(t) \triangleq \int_0^t v(t_1)dt_1$ (assuming $\varphi(0) = 0$), the corresponding memductance $G(\varphi(t))$ will also be a function of time. This observation has misled numerous researchers [16], [17], including celebrated Nobel laureates A. L. Hodgkin and A. F. Huxley [28], into calling any element described by (33) a time-varying conductance, where $G = G(x_1, x_2, \cdots, x_n)$ may depend on one (as in a memristor) or more state variables. Such erroneous interpretation of the intrinsic characteristic properties of an unconventional element had unfortunately led to endless paradoxes and confusions. One well-known researcher had been so utterly perplexed that he had added the revealing adjective "anomalous" to the title of one of his insightful papers [17].

Observe that when the correct complementary port variables ($v^{(-1)}$ and $i^{(-1)}$ in this case) are chosen, the resulting constitutive relation will never involve the time variable t, as is clear in (30) and (33).

Readers interested in a more detailed exposition on the memristor are referred to [13]. Here, we simply present a few examples in order to further stress the overriding importance of identifying the relevant complementary port variables ($v^{(\alpha)}$, $i^{(\beta)}$) in the modeling of any new nano or molecular device. In view of their unconventional physical, chemical, and/or biological operating mechanisms (e.g., tunneling, ion-membrane transports, quasi-particle excitations, etc.), their relevant port variables may consist of some unconventional complementary port variables $v^{(\alpha)}$ and $i^{(\beta)}$, where $\alpha \neq 0$ and $\beta \neq 0$.

Example 1: Consider a flux-controlled memristor $\mathcal{M}$ characterized by the constitutive relation

$$q = 8 \sin \varphi. \tag{35}$$

Let us apply a sinusoidal voltage signal $v(t) = 10 \sin \omega t$ across $\mathcal{M}$ and derive the exact equation of the corresponding

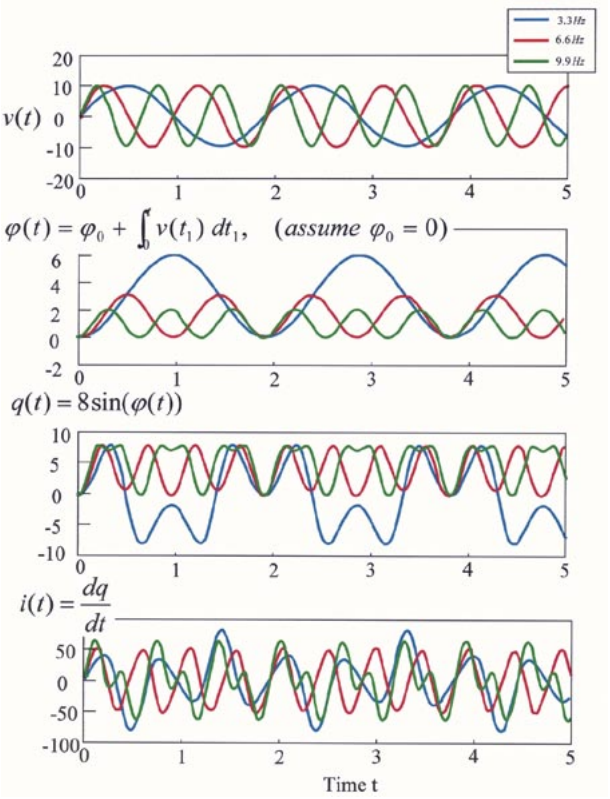

Fig. 21. Waveforms of the time functions $v(t)$, $\varphi(t)$, $q(t)$ and $i(t)$ plotted from exactly derived equations.

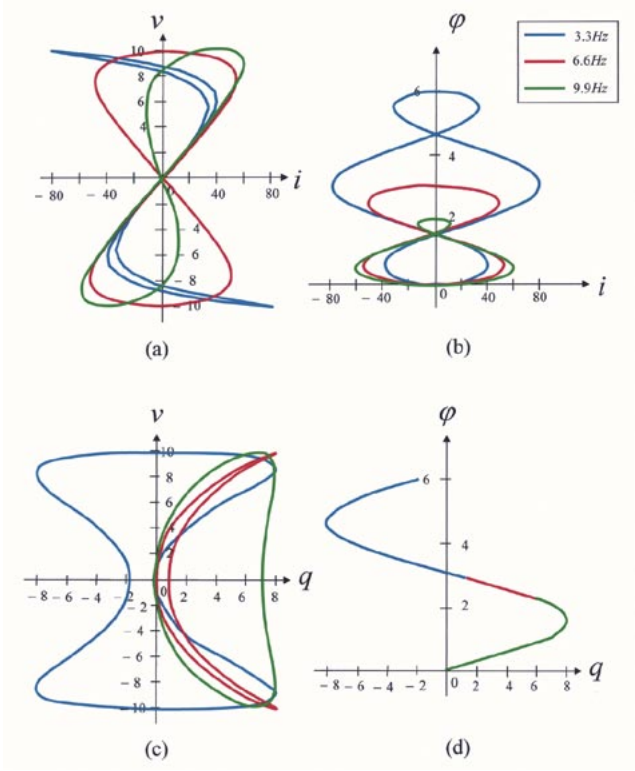

Fig. 22. Loci of projections of pairs of signals taken from Fig. 21 corresponding to the three different frequencies $\omega = 3.3$ Hz, 6.6 Hz, and 9.9 Hz, respectively.

flux $\varphi(t)$, charge $q(t)$, and current $i(t)$ for three different values of the frequency ω, namely, $\omega = 3.3$ Hz, 6.6 Hz, and 9.9 Hz. These exact equations are listed as follows:

Case 1) $\omega = 3.3$ Hz

$$\left.\begin{aligned} v(t) &= 10\sin(3.3t) \\ \varphi(t) &= \int_0^t v(t_1)dt_1 = -3.03\left[\cos(3.3t) - 1\right] \\ q(t) &= 8\sin\left(\varphi(t)\right) \\ i(t) &= \frac{dq}{dt} = 80\cos\left(3.03\left[\cos(3.3t) - 1\right]\right)\sin(3.3t) \end{aligned}\right\} \tag{36}$$

Case 2) $\omega = 6.6$ Hz

$$\left.\begin{aligned} v(t) &= 10\sin(6.6t) \\ \varphi(t) &= \int_0^t v(t_1)dt_1 = -1.52\left[\cos(6.6t) - 1\right] \\ q(t) &= 8\sin\left(\varphi(t)\right) \\ i(t) &= \frac{dq}{dt} = 80\cos\left(1.52\left[\cos(6.6t) - 1\right]\right)\sin(6.6t) \end{aligned}\right\} \tag{37}$$

Case 3) $\omega = 9.9$ Hz

$$\left.\begin{aligned} v(t) &= 10\sin(9.9t) \\ \varphi(t) &= \int_0^t v(t_1)dt_1 = -1.01\left[\cos(9.9t) - 1\right] \\ q(t) &= 8\sin\left(\varphi(t)\right) \\ i(t) &= \frac{dq}{dt} = 80\cos\left(1.01\left[\cos(9.9t) - 1\right]\right)\sin(9.9t) \end{aligned}\right\} \tag{38}$$

The waveforms of $v(t)$, $\varphi(t)$, $q(t)$, and $i(t)$ corresponding to the three frequencies $\omega = 3.3, 6.6,$ and 9.9 are plotted in Fig. 21.

Suppose we did not know the collection of signals depicted in Fig. 21 are derived, or measured, from a memristor characterized by the constitutive relation defined by (35), and were asked to develop a circuit model which would reproduce all of the waveforms in Fig. 21, within some specified tolerance. Without given any additional information, the best one can do is to group these waveforms for each frequency ω into the four complementary port variable pairs $(v,i)\triangleq(v^{(0)},i^{(0)})$, $(\varphi,i)\triangleq(v^{(-1)},i^{(0)})$, $(v,q)\triangleq(v^{(0)},i^{(-1)})$, and $(\varphi,q)\triangleq(v^{(-1)},i^{(-1)})$, and project each pair of signals into the corresponding $v-i$ plane, $\varphi-i$ plane, $v-q$ plane, and $\varphi-q$ plane, respectively, thereby suppressing the time variable. Since the signals in Fig. 21 are all periodic, the loci in each case will be either a closed loop, or an open-ended curve where the signals simply retrace the curve twice. Repeating this exercise for the three sets of signals corresponding to

$\omega = 3.3$ Hz, 6.6 Hz and 9.9 Hz, we would in general expect to obtain for each pair of signals three different loci, one for each frequency, as depicted in Fig. 22(a)– (d), respectively. If we find all loci form closed loops in all four cases, we can conclude that the collection of signals in Fig. 21 cannot be modeled by one of the four basic circuit elements defined so far. In this case, we must try to fit these signals to one of the higher order (α, β) elements to be defined in Section V. If this attempt fails as well, then we will need to invoke the more advanced modeling techniques to be presented in Part II of this paper.

An examination of Fig. 22 shows that Fig. 22(a)– (c) contain three different closed loops for each frequency ω. In particular, observe that each loop in Fig. 22(a) always goes through the origin in the $v - i$ plane, henceforth called a "pinched" loop. These observations imply immediately that the collection of signals from Fig. 21 cannot be modeled by a nonlinear resistor, inductor, or capacitor. In sharp contrast, we find Fig. 22(d) contains only one open-ended loci. In particular, we notice that the signal pair ($\varphi(t)$, $q(t)$) for each of the three frequencies falls into a part of the same curve, albeit in different regions, as indicated by the three respective colors highlighting the $\varphi - q$ curve. This observation not only shows that the collection of signals in Fig. 21 can be reproduced by a memristor characterized by (35), but it also suggests a very strong possibility that this memristor can be used as a circuit model for predicting the response due to any other applied signals. To validate this model, it is necessary to apply many more testing signals over different amplitude and frequency ranges, a task that can be automated via software.

Example 2. A More Realistic Josephson Junction Circuit Model: The classical circuit model for a Josephson junction consists of a parallel connection of a linear capacitor $\mathcal{C}$, a linear resistor $\mathcal{R}$, and a nonlinear inductor $\mathcal{L}$ with the constitutive relation $i = I_0 \sin k\varphi$, $k \triangleq ((2e)/\hbar)$. A more rigorous quantum mechanical analysis of the Josephson junction dynamics reveals the presence of an additional small current component due to interference among quasi-particle pairs [24]–[26]. This heretofore neglected current component can be represented approximately by

$$i = G(\cos k_0\varphi)v \tag{39}$$

where G and k_0 are device constants which need not concern us in this paper. One of the reasons why this component had been ignored in the past is due to the lack of a familiar circuit element for its representation. Observe, however, that if we define a flux-controlled memristor with the constitutive relation[8]

$$\boxed{q = G_0 \sin(k_0\varphi)} \tag{40}$$

where $G_0 \triangleq G/k_0$, then differentiating both sides of (40) would give us exactly (39). In other words, the heretofore neglected current component defined in (39) is nothing more than a memristor and can be simply added in parallel to the classical model to obtain the more realistic Josephson junction circuit model shown in Fig. 23. It is instructive to note that the Josephson junction provides us with the simplest nonlinear circuit model of a device made from all four basic circuit elements introduced so far.

[8]We have already seen some signals generated by this element in Fig. 21 for $G_0 = 1$ and $k_0 = 1$.

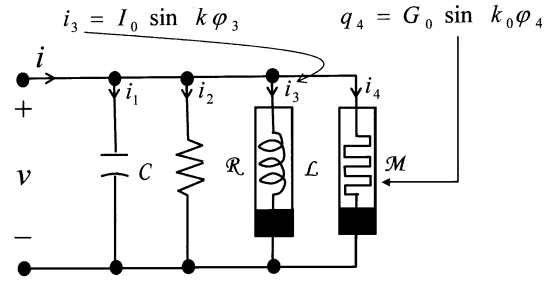

Fig. 23. More realistic circuit model of a Josephson junction circuit consists of a parallel connection of four basic circuit elements, a linear capacitor C, a linear resistor $\mathcal{R}$, a nonlinear inductor $\mathcal{L}$ with constitutive relation $i_3 = I_0 \sin k\varphi_3$, and a nonlinear memristor with constitutive relation $q_4 = G_0 \sin k_0\varphi_4$ which incorporates the small, hitherto neglected current component due to interference among quasi-particle pairs.

Example 3. Modeling a Two-Terminal Nanowire Device: We are now ready to demystify the strange electrical phenomena observed from the two-terminal nanowire device [10], [11] described earlier in Section I, whose conductance switches from almost zero to approximately 800 nS upon the application of a square-wave voltage signal [Fig. 4(b)], and whose $v - i$ characteristics formed a pinched hysteresis loop in the $v - i$ plane when driven by a low-frequency periodic voltage signal. In particular, we will show that these exotic behaviors can be modeled by a memristor.

Consider first the piecewise-linear $q - \varphi$ constitutive relation[9] of a two-terminal flux-controlled memristor shown in Fig. 24(a). Observe that by virtue of the piecewise-linear nature of the $q - \varphi$ curve, the slope $G(\varphi) \triangleq (d\hat{q}(\varphi))/(d\varphi)$, henceforth called the small-signal memductance (contraction for "memory conductance") is equal to 0 when $|\varphi| \leq 2.5$ Wb, and $G(\varphi) = 800$ nS when $|\varphi| > 2.5$ Wb. Consequently, if we apply the square-wave voltage signal $v(t)$ shown in Fig. 24(b), or equivalently, the triangular-wave flux signal $\varphi(t)$ shown in Fig. 24(c), across this memristor, we would obtain the square-wave conductance $G(t)$ shown in Fig. 24(d). Note that we have chosen the parameters of the memristor constitutive relation [i.e., the $q - \varphi$ curve in Fig. 24(a)] so that the memductance $G(\varphi)$ of this memristor switches from $G = 0$ to $G = 800$ nS, thereby reproducing approximately the square-wave conductance measured from a nanowire device, shown earlier in Fig. 4(b). Observe that the conductance G [or, more precisely, the memductance $G(\varphi)$] in Fig. 24(d) is equal to zero (equivalent to an open circuit) from $t = 0$ to $t = 0.25$ because it takes this amount of time for the flux signal $\varphi(t)$ in Fig. 24(c) to increase from $\varphi = 0$ to $\varphi = 2.5$ Wb, while the applied voltage

[9]For ease of comparison with the measured results in Fig. 4, we have chosen the current i and the charge q as the vertical axis throughout this example. Hence, the slope $G(\varphi)$ of the $q - \varphi$ curve in Fig. 24(a) has the unit of a conductance, namely, nanosiemens (nS).

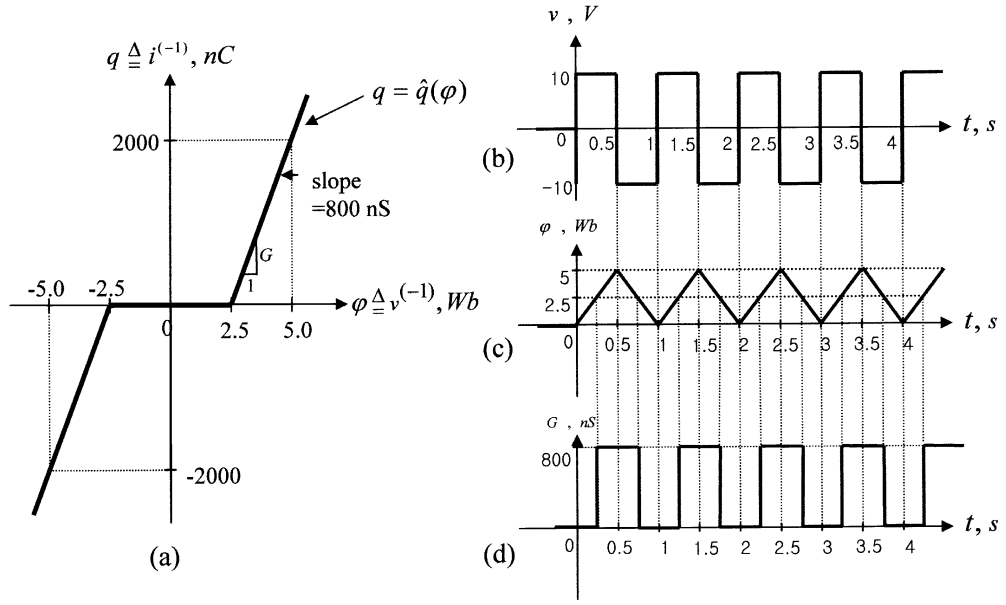

Fig. 24. (a) $q-\varphi$ characteristic of memristor. (b) Periodic voltage waveform $v(t)$. (c) Corresponding waveform of the flux $\varphi(t)=\int_0^t v(\tau)\,d\tau$, assuming $\varphi(0)=0$. (d) Corresponding conductance waveform $G(t)$.

$v(t)$ remains constant at $v = 10$ V in Fig. 24(b). Observe also that the conductance G in Fig. 24(d) does not drop abruptly to zero at $t = 0.5$ s when the applied voltage switches abruptly to $v =$ -10 V, but instead it had to wait until $t = 0.75$ s when $\varphi(t)$ decreases from its maximum of $\varphi = 5$ Wb at $t = 0.5$ s to $\varphi = 2.5$ Wb at $t = 0.75$ s. In other words, unlike in a resistor, the conductance $G(t)$ at $t = t_0$ of a memristor depends not only on the value of the applied voltage $v(t)$ at $t = t_0$, but also on the entire past history of $v(t)$, $-\infty < t \leq t_0$ of the applied voltage signal. The memductance $G(\varphi)$ (respectively, memristance $M(q) \triangleq G^{-1}(\varphi)$) of a memristor, therefore, behaves like the conductance (resp, resistance) of a resistor with a memory of what had transpired.

Instead of applying a square-wave voltage input signal as in Fig. 24(b), let us apply next a sinusoidal memristor voltage signal $v(t) = 5\sin\omega t$ across the memristor [with the same $q-\varphi$ curve shown in Fig. 24(a)] for four different frequencies, $\omega = 2$ Hz, 1 Hz, (1/2) Hz, and 5 Hz, respectively, as shown in Figs. 25(a), 26(a), 27(a), and 28(a), respectively. By integrating $v(t)$ with $\varphi(0) = 0$, we obtain the corresponding memristor flux signal $\varphi(t) = (5/\omega)(1-\cos\omega t)$, as plotted in Figs. 25(b), 26(b), 27(b), and 28(b), respectively. Substituting each of these four flux signals to the constitutive relation

$$\boxed{\begin{aligned} q = \hat{q}(\varphi) &= 0, \quad |\varphi| \leq 2.5 \\ &= 800(\varphi - 2.5), \quad |\varphi| \geq 2.5 \end{aligned}} \tag{41}$$

[shown graphically in Fig. 24(a)], we obtain the corresponding memristor charge signal $q(t)$ and memristor current signal $i(t)$ shown in Figs. 25(c)–(d), 26(c)–(d), 27(c)–(d), and 28(c)–(d), respectively.[10]

Now suppose we did not know that these signals were derived [or measured from a real device having the constitutive relation defined by (41)] from a memristor and had followed our ingrained habit by mapping the current signal $i(t)$ shown in Figs. 25(d), 26(d), 27(d), and 28(d), respectively, and its corresponding voltage signal $v(t)$ shown in Figs. 25(a), 26(a), 27(a), and 28(a) respectively, in the current-versus-voltage $i-v$ plane, we would then obtain the four pinched hysteresis loops shown in Figs. 25(e), 26(e), 27(e), and 28(e)[11] respectively, reminiscent of the pinched hysteresis loop of the nanowire device seen earlier in Fig. 4(a). Since the $v-i$ curves (i.e., the pinched hysteresis loops) in Figs. 25(e), 26(e), 27(e), and (28e) are all different from one another even though they were measured from the same black box [in this case, the memristor with the $q-\varphi$ curve in Fig. 24(a)] using a sinusoidal testing signal $v(t) = 5\sin\omega t$ of the same amplitude (only one parameter, namely, the frequency ω, is varied in these experiments) it follows that the constitutive relation of this black box is not a resistor. If we had repeated the exercise by projecting the signals from Figs. 25–28 onto the (φ, i) and (q, v) plane, we would find the loci would also form closed loops, though not pinched as in Figs. 25(e), 26(e), and 27(e).[12]

[10]We have deliberately chosen a piecewise-linear constitutive relation in Fig. 24(a) in order that all of the signals shown in Figs. 25–28 can be derived analytically; hence, these waveforms are exact and not calculated numerically, as would usually be the case for memristor models of real devices.

[11]In Fig. 28(e), the horizontal segment $-5 \leq v \leq 5$ along the v axis can be interpreted as the limiting form of a pinched hysteresis loop whose associated current had collapsed to zero, since the flux signal did not reach its critical value of $\varphi = 2.5$ Wb.

[12]The formation of pinched hysteresis loops, as in Figs. 25(e), 26(e), and 27(e) in the $i-v$ plane, is the hallmark of a memristor.

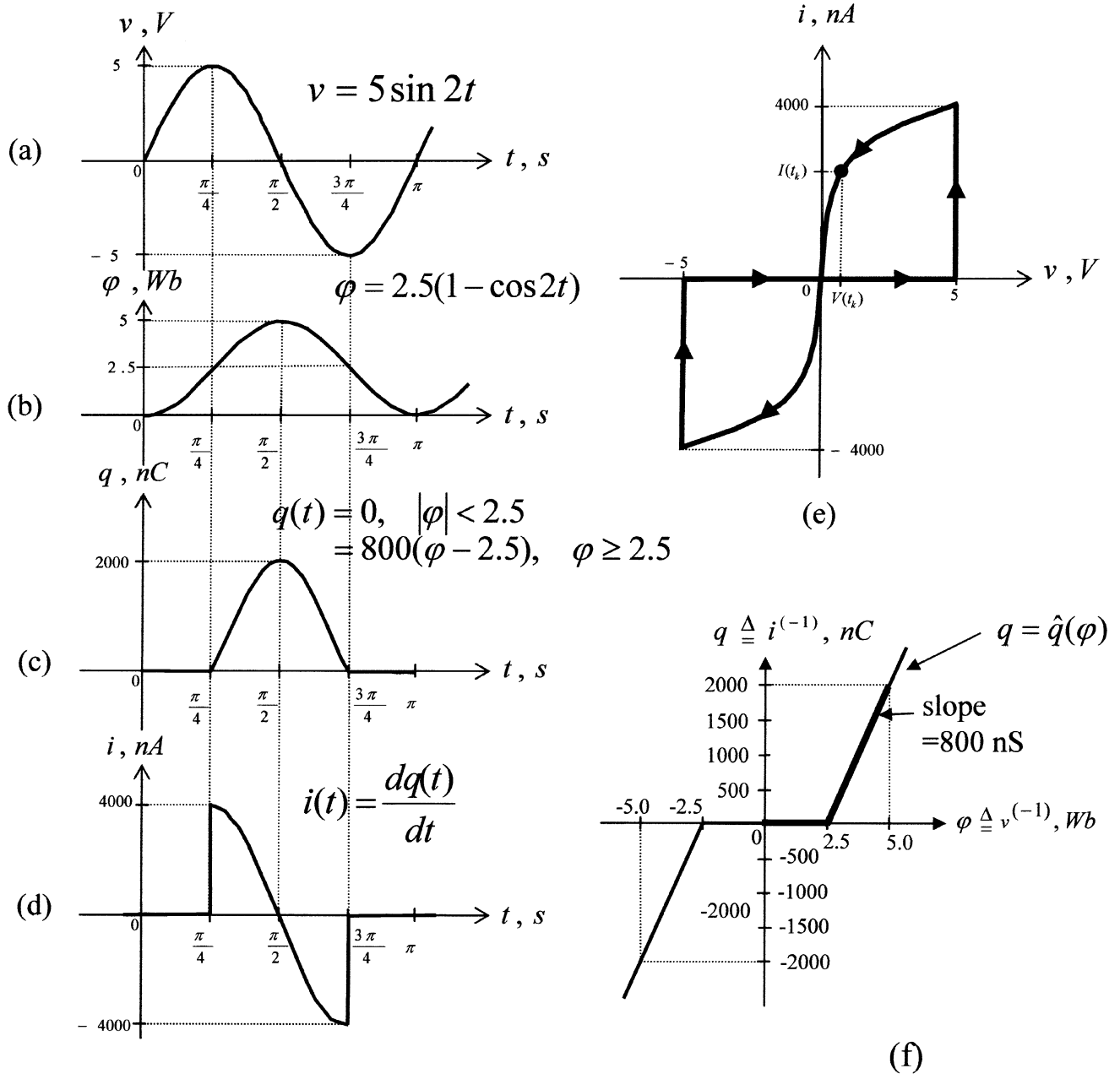

Fig. 25. Waveforms resulting from applying the voltage signal $v(t) = 5 \sin 2t$ across a memristor whose constitutive relation $q = \hat{q}(\varphi)$ is given by Fig. 24(a).

Moreover, we will again find these loops to vary with the frequency ω. On the basis of the above observations, we can unequivocally conclude that the black box cannot be modeled as a nonlinear resistor, inductor, or capacitor.

Before giving up, however, let us try the only remaining complementary port variable pair which has not yet been examined, namely, the $q - \varphi$ plane. The results are depicted in the bold portions of the $q - \varphi$ curve shown in Figs. 25(f), 26(f), 27(f), and 28(f), respectively, which incidentally is identical to the $q-\varphi$ curve in Fig. 24(a), plotted with different scales. Observe that even though the $(q(t), \varphi(t))$ signal pairs, corresponding to the four different frequencies $\omega = 2$ Hz, 1 Hz, (1/2) Hz, and 5 Hz, fall on different regions of the $q-\varphi$ curve, they always follow the $q-\varphi$ curve of Fig. 24(a). If we had performed additional tests, using different testing signals and over a wide range of frequencies, we would have found that the corresponding $q-\varphi$ loci always trace out parts of the same $q - \varphi$ curve. In other words, we can model the black box by a flux-controlled memristor with the $q - \varphi$ constitutive relation defined by (41), as depicted in Fig. 24(a).

Finally, we remark that the samples chosen in this example were specially designed for pedagogical reasons—namely, to allow the results be easily verified and reproduced via analytical formulas. In modeling practical devices, it is highly unlikely that analytical formulas can be derived and we must rely on numerical on graphical techniques, to be implemented via computers. For example, by massaging the shape of the $q - \varphi$ curve, we can tailor the pinched

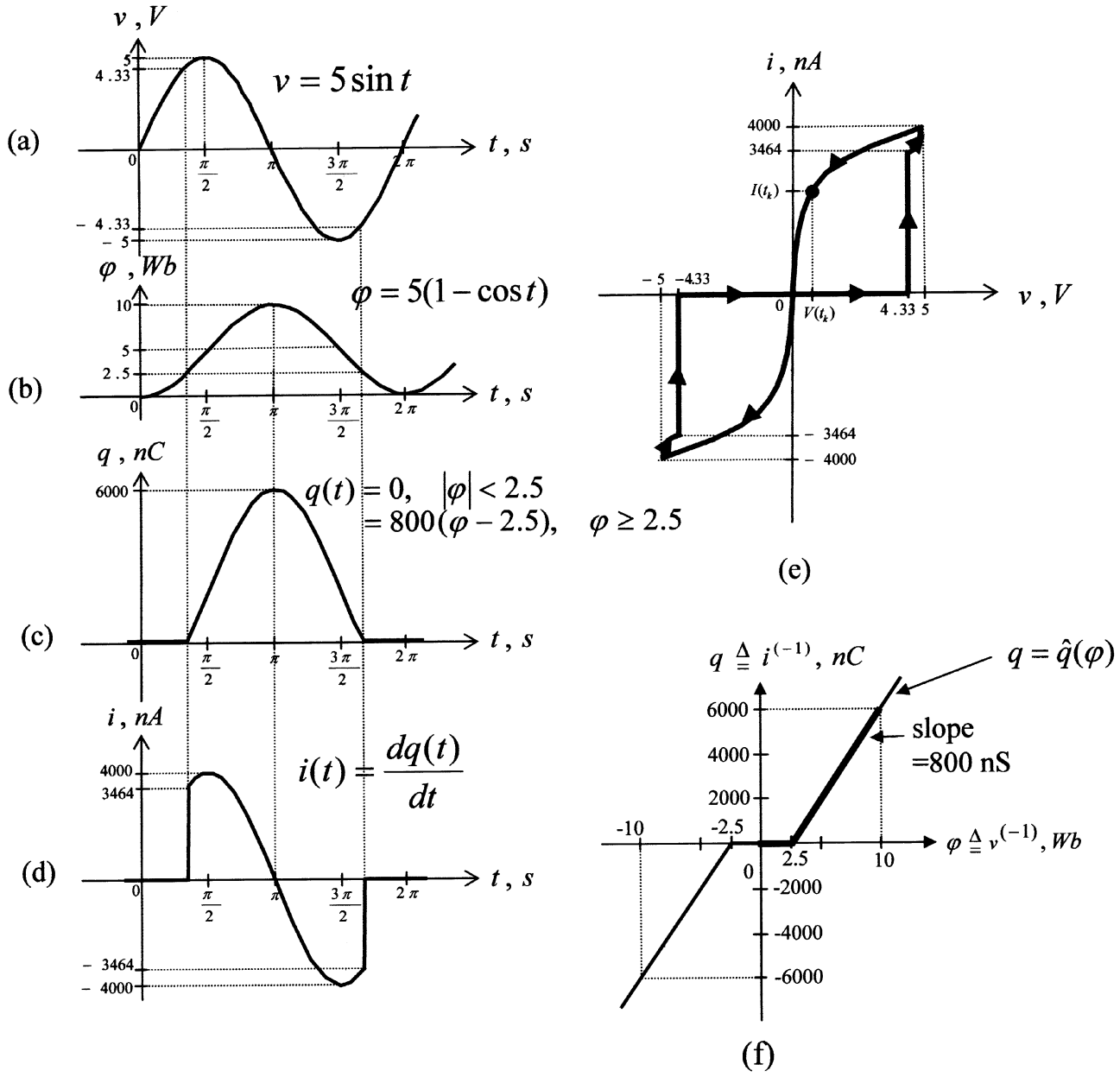

Fig. 26. Waveforms resulting from applying the voltage signal $v(t) = 5 \sin t$ across a memristor whose constitutive relation $q = \hat{q}(\varphi)$ is given by Fig. 24(a).

hysteresis loop to match the shape of measured loops. For example, by choosing the $q - \varphi$ curve shown in Fig. 29(f), and choosing the testing signal $v(t) = 5\sin(4/3)t$, as shown in Fig. 29(a), we can derive numerically the corresponding signals $\varphi(t)$, $q(t)$, and $i(t)$ shown in Fig. 29(b), (c), and (d). The corresponding loci shown in Fig. 29(f) now resembles the measured pinched hysteresis loop for the nanowire device shown in Fig. 4(a). By destroying the symmetry of the $q - \varphi$ curve, we can easily obtain an even better match of this measured loop.

E. The Basic Circuit Element Quadrangle

The circuit element species defined in Sections IV-A to IV-D—resistor $\mathcal{R}$, inductor $\mathcal{L}$, capacitor $\mathcal{C}$, and memristor $\mathcal{M}$—are truly basic not only because they include the three classical circuit elements $\mathcal{R}$, $\mathcal{L}$, $\mathcal{C}$ as special cases, but also because the value of their associated small-signal resistance $R(i)$, inductance $L(i)$, capacitance $C(i)$, and memristance $M(q)$, respectively, do not change with the frequency ω of an infinitesimally small sinusoidal testing signal about any fixed operating point Q. They are frequency independent in this sense. The relationships defining these four basic circuit elements are summarized in the diagram shown in Fig. 30, henceforth called the basic circuit element quadrangle.

V. Frequency-Dependent Elements

Each of the four (α, β) elements defined above via $(\alpha, \beta) = (0, 0), (-1, 0), (0, -1)$, and $(-1, -1)$ and dubbed a resistor $\mathcal{R}$, an inductor $\mathcal{L}$, a capacitor $\mathcal{C}$, and a memristor $\mathcal{M}$, respectively, was given a comprehensive presentation in

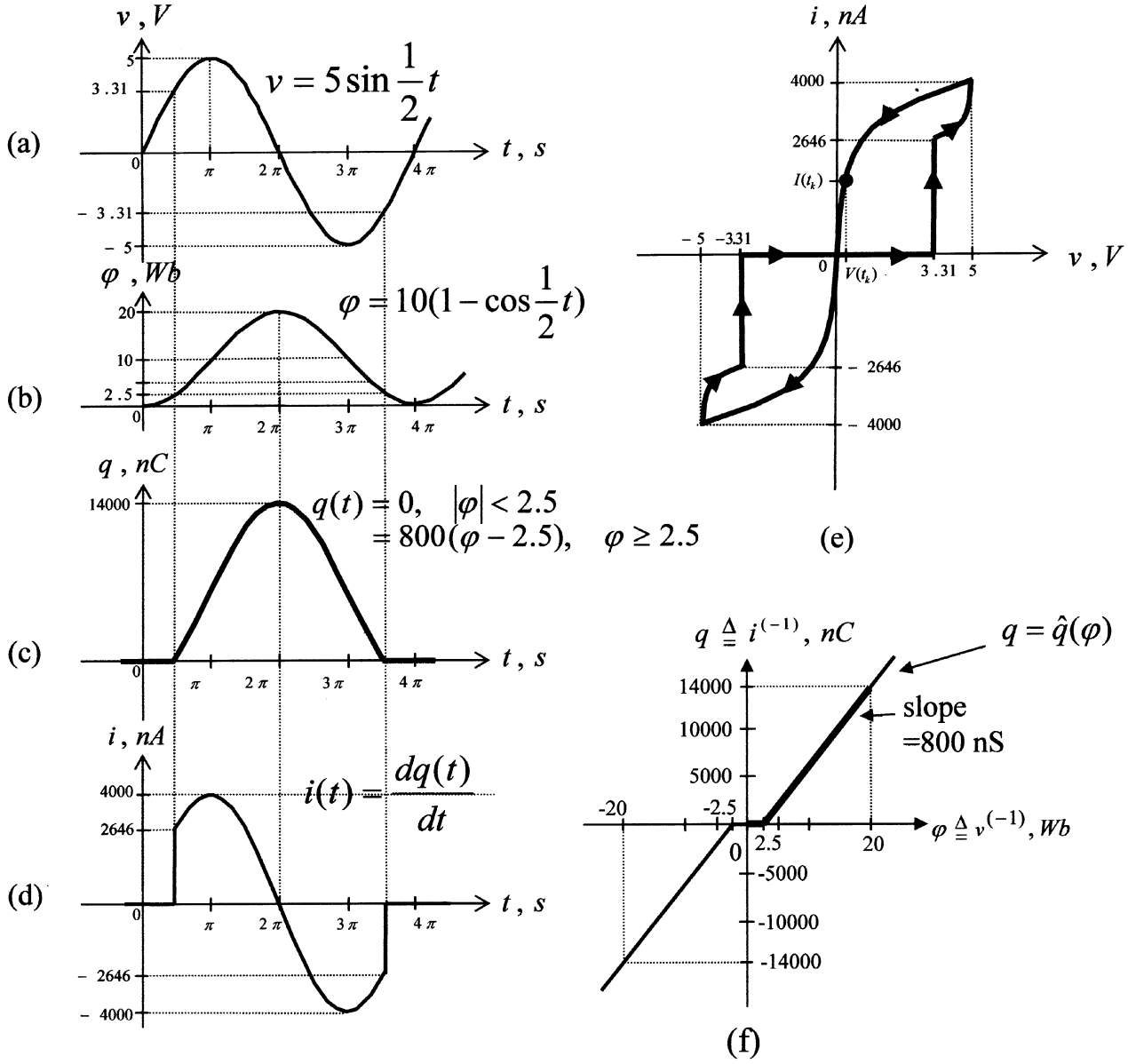

Fig. 27. Waveforms resulting from applying the voltage signal $v(t) = 5\ \sin(1/2)t$ across a memristor whose constitutive relation $q = \hat{q}(\varphi)$ is given by Fig. 24(a).

Section IV not only because each one represents a fundamental and frequency-independent element but also because it provides an intuitive proof that each element is independent in the sense that no element described by a nonlinear constitutive relation from this basic set can be derived from the other three elements. In fact, it follows from the same reasoning that no two-terminal (α, β) elements characterized by a nonlinear constitutive relation for arbitrary α and β can be derived from the other elements.[13] Consequently, from an axiomatic perspective, a complete set of two-terminal circuit element building blocks must include all (α, β) elements where α and β can assume any positive, zero, or negative integers. A part of this infinite (α, β) element array—henceforth called the complete circuit element set—is tabulated in Fig. 31, where the horizontal axis α and the vertical axis β include all positive, zero, and negative integers. Each grid point (α, β) represents an (α, β) element, as depicted by the element symbol defined in Fig. 16(a). Recall that the upper and lower integers within each element symbol in this table specifies the voltage exponent α and the current exponent β, respectively. For example, the four basic elements—resistor, inductor, capacitor, and memristor—are located at the 4 point (0,0), (-1), (0, -1), and $(-1, -1)$ in the $\alpha - \beta$ plane, respectively. It is proved in [18] that for every (α, β) element we delete from this table, we can conceive of a hypothetical nonlinear device which cannot be adequately modeled.

In the special case where the constitutive relations are linear, however, only a subset of this table is necessary. In particular, only the elements on any one row, or any one column, are sufficient. This observation follows from (21) where all linear (α, β) elements with an identical

[13]This element independence property is not true when the element's constitutive relation is linear, as evidenced from our earlier observation that a linear memristor is indistinguishable from a linear resistor.

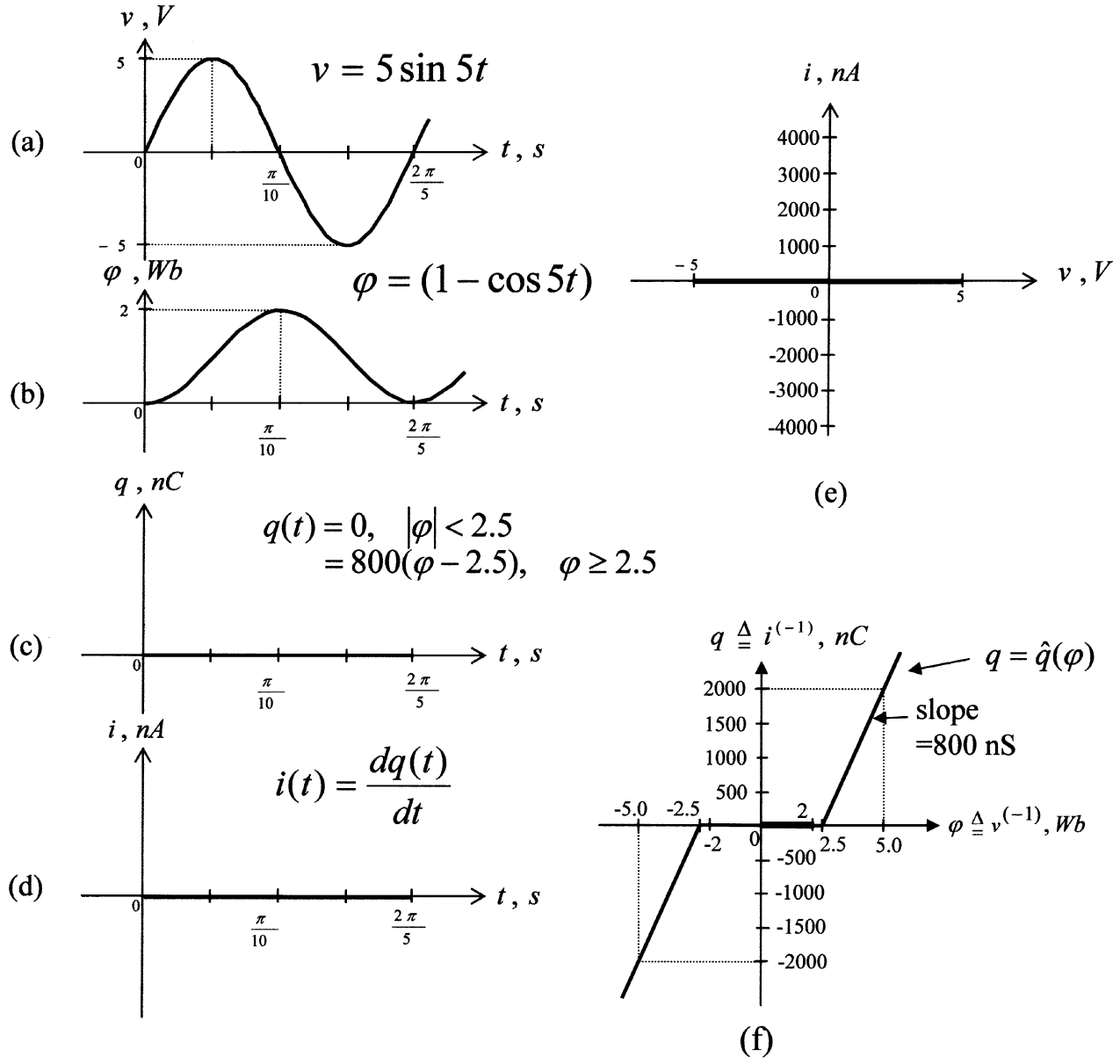

Fig. 28. Waveforms resulting from applying the voltage signal $v(t) = 5\ \sin\ 5t$ across a memristor whose constitutive relation $q = \hat{q}(\varphi)$ is given by Fig. 24(a).

voltage–current exponent difference $\gamma \triangleq \beta - \alpha$ have an identical small-signal impedance $Z_Q(j\omega)$ at all operating points Q and, hence, are indistinguishable from each other. It follows that all linear elements in Fig. 31 which lie on a diagonal line (with unit slope) $\beta - \alpha = n$, $n = \cdots -3, -2, -1, 0, 1, 2, 3, \cdots$ are equivalent; they are shown in Fig. 31 in the same color.

Although no nonlinear (α, β) element from the complete element set in Fig. 31 can be synthesized from other elements, we will show in the next section that they can be partitioned into four mutually exclusive groups whose members share certain important circuit-theoretic properties.

A. Four Classes of Frequency-Dependent Elements

We will now show that even though the global (i.e., large-signal) dynamic behaviors of the nonlinear α, β elements in Fig. 31 are unique and distinct from each other, their local (i.e., small-signal) behaviors resemble that of a frequency-dependent resistance function $R(\omega)$ or a frequency-dependent reactance function $X(\omega)$, where $R(\omega)$ and $X(\omega)$ denote the real and imaginary part, respectively, of the impedance[14]

$$Z_Q(j\omega) = R(\omega) + jX(\omega) = (j\omega)^{\beta-\alpha} m_Q. \qquad (42)$$

In particular, depending on the value of $\gamma \triangleq \beta - \alpha$, we can classify an (α, β) element into one of the following mutually exclusive categories.

[14]In classic linear circuit theory [29], the real part $R(\omega)$ and the imaginary part $X(\omega)$ of an impedance $Z(j\omega)$ is called the resistance function and reactance function, respectively. It follows from the definition of "Fourier transform" that $R(\omega)$ is an even function of ω (i.e., $R(-\omega) = R(\omega)$) whereas $X(\omega)$ is an odd function of ω (i.e., $X(-\omega) = -X(\omega)$)

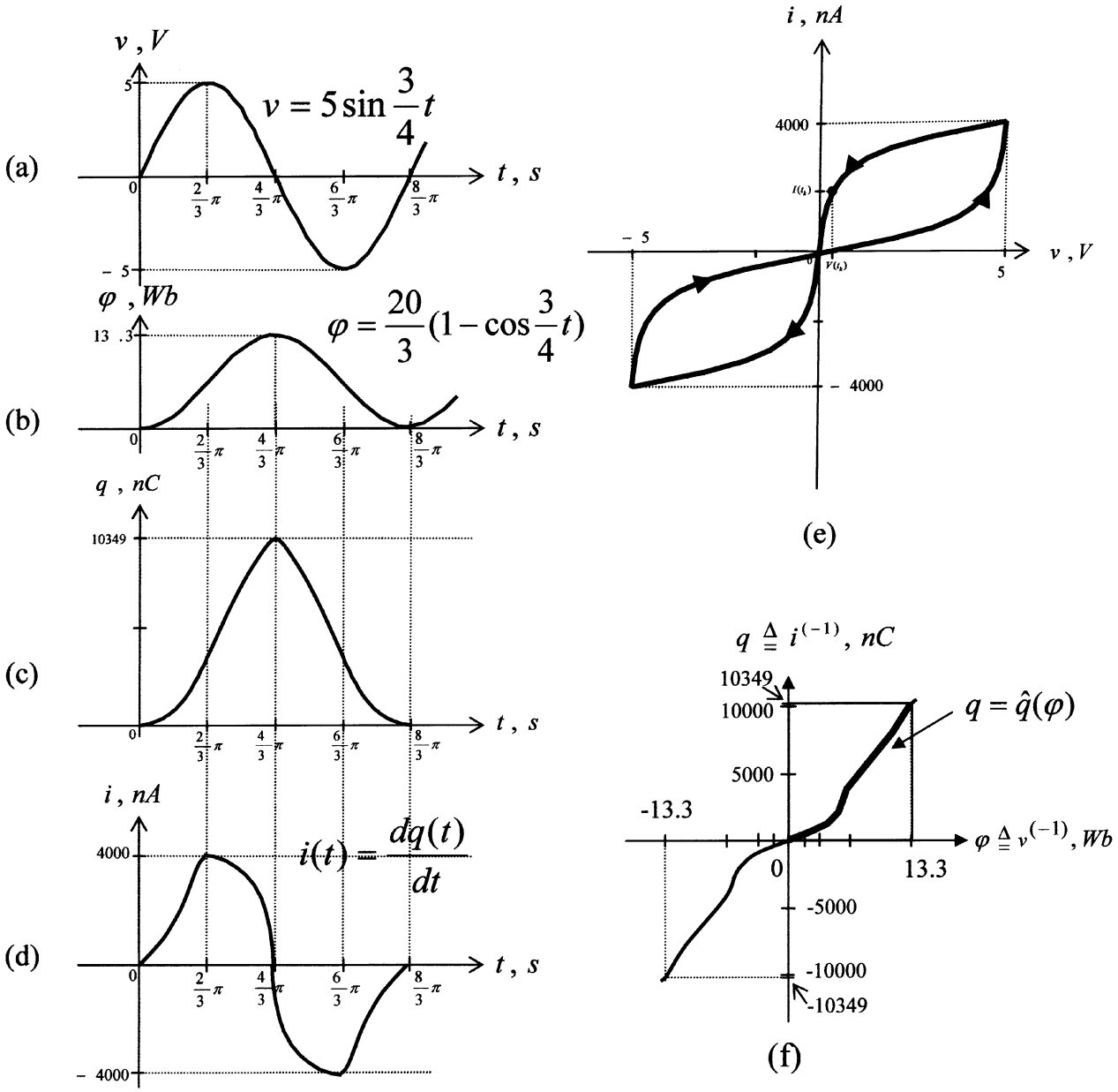

Fig. 29. Waveforms resulting from applying the voltage signal $v(t) = 5\ \sin(3/4)t$ across a memristor whose constitutive relation $q = \hat{q}(\varphi)$ is given in (f).

1) *Frequency-Dependent Resistors* ($\beta - \alpha = \pm 2n$, $n =$ even integer)

Under this category, $Z_Q(j\omega)$ from (42) reduces to a real positive function of the frequency ω (assuming $m_Q > 0$), namely

$$Z_Q(j\omega) = \omega^{\pm 2n} m_Q \triangleq R(\omega). \tag{43}$$

The frequency-dependent resistors in Fig. 31 are colored in red and lie along red diagonal lines labeled with the element category name "resistance" $R(\omega)$.

2) *Frequency-Dependent Negative Resistors* ($\beta - \alpha = \pm 2n$, $n =$ odd integer)

Under this category, $Z_Q(j\omega)$ from (42) reduces to a real negative function of the frequency ω (assuming $m_Q > 0$, namely

$$Z_Q(j\omega) = -\omega^{\pm 2n} m_Q \triangleq R^*(\omega). \tag{44}$$

The frequency-dependent negative resistors in Fig. 31 are colored in orange and lie along orange diagonal lines labeled with the element category name "negative resistance" $R^*(\omega)$.

We note in passing that the two elements (2,0) and (0,2) under this category have found useful applications in active circuit design [30] and are known under the names FDNR D element and FDNR E element, respectively, where FDNR is an acronym for frequency-dependent negative resistor.

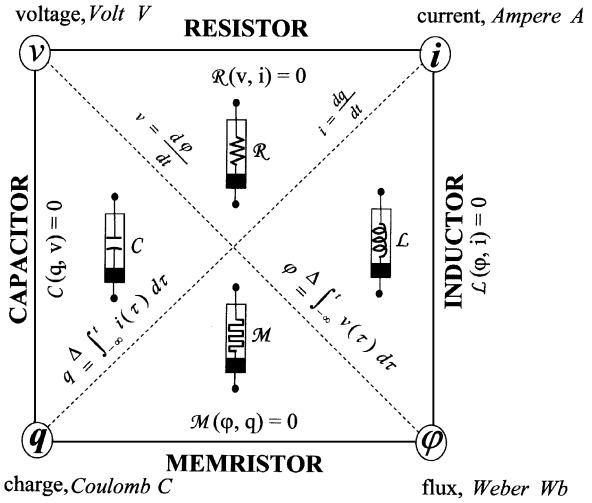

Fig. 30. Basic circuit element quadrangle.

3) *Frequency-Dependent Inductor* ($\beta-\alpha=(-1)^n(2n+1)$, $n=$ any integer)

Under this category, $Z_Q(j\omega)$ from (42) reduces to a positive imaginary function of the frequency ω (assuming $m_Q>0$), namely

$$Z_Q(j\omega)=j\omega\,[L(\omega)] \tag{45}$$

where

$$L(\omega)\triangleq\begin{cases}\omega^{2n}m_Q, & \text{when n=even integer}\\ \omega^{-2(n+1)}m_Q, & \text{when n=odd integer.}\end{cases} \tag{46}$$

Observe that when $n=0$, $L(\omega)=m_Q$ is just the frequency-independent small-signal inductance defined in (24) and (25). Hence, it is logical to call $L(\omega)$ in (46) a frequency-dependent inductance at the operating point Q. The frequency-dependent inductors in Fig. 31 are colored in blue and lie along blue diagonal lines labeled with the element category name "inductance" $L(\omega)$

4) *Frequency-Dependent Capacitor* ($\beta-\alpha=(-1)^{n+1}(2n+1)$, $n=$ any integer)

Under this category, $Z_Q(j\omega)$ from (42) reduces to a negative imaginary function of the frequency ω (assuming $m_Q>0$), namely

$$Z_Q(j\omega)=\frac{1}{j\omega\,[C(\omega)]}=-j\left(\frac{1}{\omega\,[C(\omega)]}\right) \tag{47}$$

where

$$C(\omega)\triangleq\begin{cases}\dfrac{\omega^{2n}}{m_Q}, & \text{when } n=\text{even integer}\\ \dfrac{\omega^{-2(n+1)}}{m_Q}, & \text{when } n=\text{odd integer}\end{cases} \tag{48}$$

Again, observe that when $n=0$, $C(\omega)=(1/m_Q)$ is just the frequency-independent small-signal capacitance defined in (24) and (25). Hence, it is logical to call $C(\omega)$ in (46) a frequency-dependent capacitance at the operating point Q. The frequency-dependent capacitors in Fig. 31 are colored in green and lie along green diagonal lines labeled with the element category name "capacitance" $C(\omega)$

B. Modeling the Frequency-Dependent Nanocoil

We will now show that the frequency-dependent nanocoil described in Section I [14], [15] can be modeled by a series circuit made up of frequency-dependent linear inductors and capacitors from element categories 3 and 4. For simplicity, suppose we connect the three elements (−1,0), (−3,0), and (−5,0) (with parameters $m_Q=m_1$, m_3, and m_5, respectively) in series. The impedance of the resulting linear one-port is given by[15]

$$\begin{aligned}Z(j\omega)&=j\omega\,m_1+\frac{1}{j\omega\left(\frac{1}{\omega^4m_3}\right)}+j\omega\left(\omega^4m_5\right)\\&=j\omega\left[m_1-m_3\omega^2+m_5\omega^4\right]\\&=j\omega L(\omega)\end{aligned} \tag{49}$$

where

$$L(\omega)\triangleq m_1-m_3\omega^2+m_5\omega^4 \tag{50}$$

is equivalent to a frequency-dependent inductance. By connecting a sufficiently large number of frequency-dependent linear inductors and capacitors in series, and by choosing appropriate values for their respective parameters m_Q, where $-\infty<m_Q<\infty$, we can approximate any given smooth even function $L(\omega)$ of ω to any desired accuracy.[16]

C. The Four-Element Torus

A careful examination of the (α, β) elements in Fig. 31 reveals some remarkable symmetries. In particular, the four colors (red, blue, orange, and green) corresponding to the four circuit element categories (frequency-dependent resistors, inductors, negative resistors, and capacitors) along diagonal lines keep repeating themselves, ad infinitum, Observe also that if we add or subtract any multiple of four to either α or β, or to both α and β of any (α, β) element, we would obtain a higher or lower order element belonging to the same element category. Moreover, adding or subtracting unity to both α and β of any (α, β) element simply moves along the diagonal line (of the same color) where the (α, β) element originated, thereby preserving the element category of all higher or lower order elements generated via this simultaneous incrementing process. In addition, both the local and the global[17] circuit-theoretic properties of all

[15]Note that the linear impedance $Z(j\omega)$ from (49) gives a frequency-dependent reactance function $X(\omega=m_0\omega-m_2\omega^3+m_4\omega^5$, which is an odd function of ω, as expected. By factoring out ω, the resulting inductance function $L(\omega)$ becomes an even function of ω.

[16]We can, of course, also model any given smooth even function $R(\omega)$ to any desired accuracy by connecting a sufficiently large number of frequency-dependent linear resistors and negative resistors in series.

[17]The preservation of the global circuit-theoretic properties among all elements along each diagonal line in Fig. 31 will be proved rigorously in Part II of this paper.

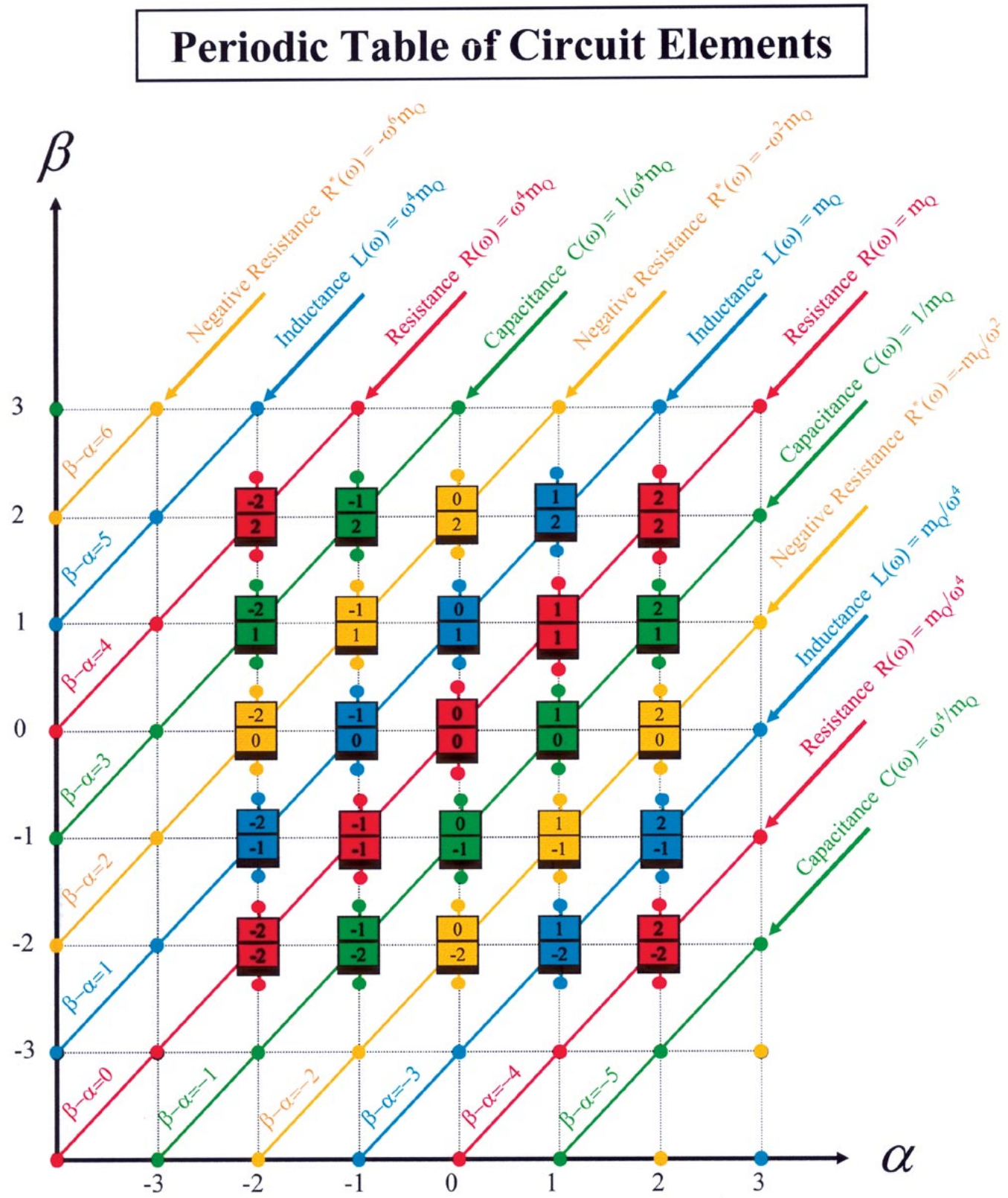

Fig. 31. The periodic table of all two-terminal (α, β) circuit elements. All elements printed in the same color belong to the same circuit element species, namely, frequency-dependent resistors (red), inductors (blue), negative resistors (orange), and capacitors (green).

circuit elements situated on any diagonal line in Fig. 31 are also preserved. The remarkable resemblance between the periodicity of the circuit element properties and that of the chemical element properties in the periodic table of chemical elements inspires us to henceforth call Fig. 31 the *periodic table of circuit elements*.

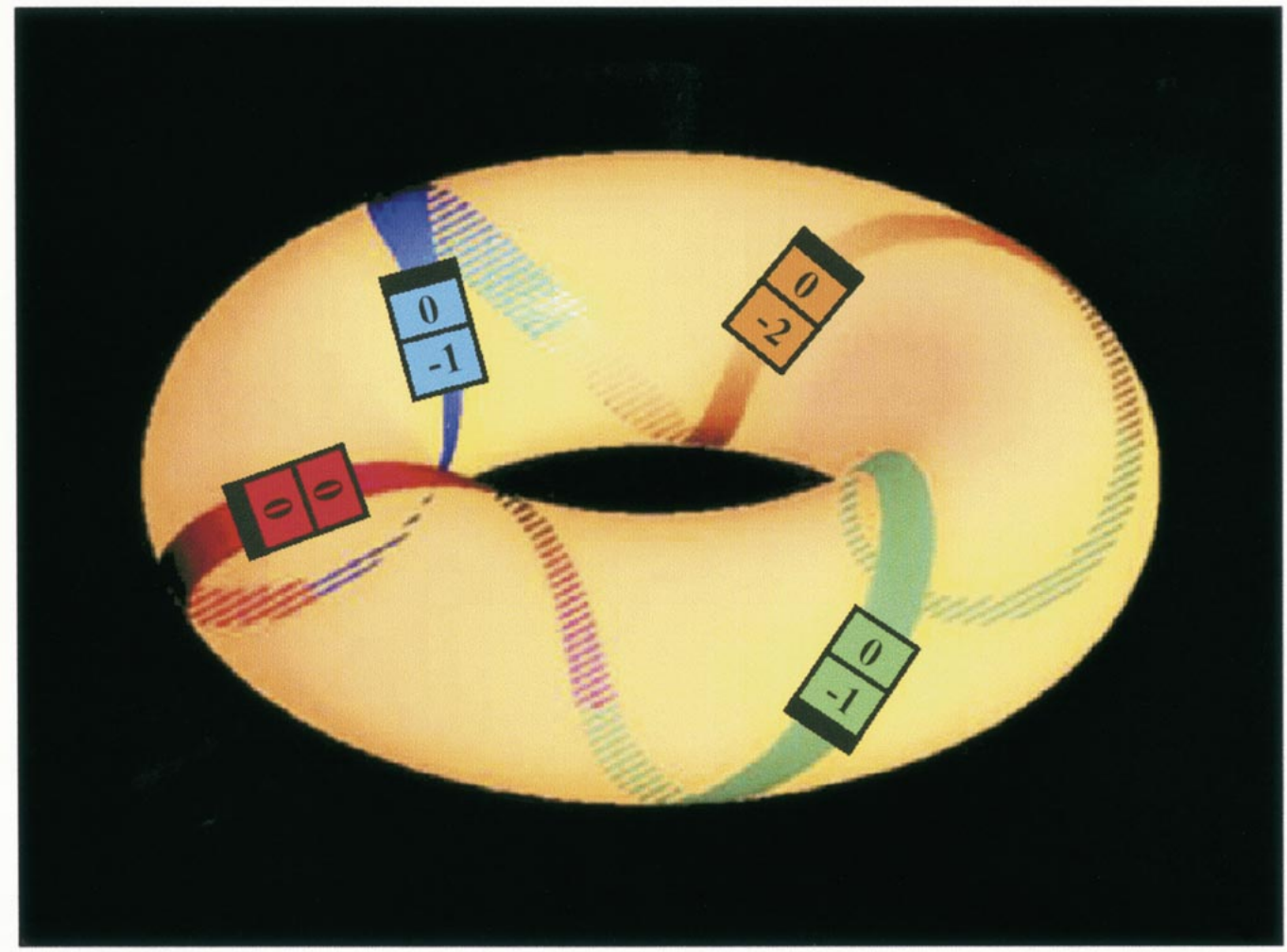

Fig. 32. Four-element torus. The four seed elements wrapped around the torus surface is periodic along the α-circumference (longitude) modulo ± 4, as well as along any β-cross section (latitude) modulo ± 4. It is also doubly periodic with respect to both α and β, modulo ± 1. The four-element torus contains the same information as the periodic element table by exploiting these threefold periodicity properties.

In view of the doubly periodic structure of Fig. 31, modulo ± 4, it suffices for us to map all (α, β) elements from the periodic table of circuit elements onto the surface of a torus, as shown in Fig. 32. Although we could have picked any four elements, one from each element category, as a seed and used them to generate all other (α, β) elements, we have chosen the red element (0,0) corresponding to the resistor $\mathcal{R}$, the blue element (-1,0) corresponding to the inductor $\mathcal{L}$, the green element (0, -1) corresponding to the capacitor $\mathcal{C}$, and the orange element (-2,0) corresponding to the lowest order frequency-dependent negative resistor $\mathcal{R}^*$. These four elements are wound around the torus via interconnections (in color) which preserve the same neighboring element category as in Fig. 31. This simple four-element torus contains the same information as that of the periodic table of circuit elements (Fig. 31) via the following encoding algorithm.

1) Each (α, β) element is periodic along the α axis (longitude) modulo ± 4. For example, starting from the seed element (0,0), we can generate the entire family of frequency-dependent resistors along the horizontal line $\beta = 0$ in Fig. 31. e.g., (4,0), (-4,0), (8,0), (-8,0), etc.
2) Each (α, β) element is periodic along the β axis (lattitude) modulo ± 4. For example, starting from the seed element (0,0), we can generate the entire family of frequency-dependent resistors along the vertical line $\alpha = 0$.
3) Each (α, β) element is doubly periodic (along both α and β axis modulo ± 4. For example, starting from the seed element (0,0), we can generate the entire family of frequency-dependent resistors along the diagonal line going through (0,0).

VI. Conclusion

Following exactly the same axiomatic approach from Section III-A, we can define an (α, β) $(n+1)$-terminal element (Fig. 6) or an (α, β) n-port (Fig. 7) by simply substituting the scalars $v^{(\alpha)}$ and $i^{(\beta)}$ by n-dimensional vectors $\mathbf{v}^{(\alpha)}$ and $\mathbf{i}^{(\beta)}$. The four-element torus in Fig. 32 follows by similar substitutions.

We note in closing that a realistic circuit model of a nanodevice will in general require an appropriate interconnection of building blocks chosen from the four-element torus. The circuit-theoretic properties (local activity, losslessness, etc.) of a nanodevice can only be derived from its circuit model. In Part II (to appear elsewhere) of this paper, we will develop the fundamental concepts of local activity and edge of chaos. These concepts will then be used to construct a systematic procedure for tuning the physical parameters of a nanodevice to endow it with computation and information processing potentials.

ACKNOWLEDGMENT

The author would like to thank Dr. S. Yoon for her invaluable assistance on the preparation of this manuscript, and J. Flak for drawing the four-element torus.

REFERENCES

[1] L. O. Chua, *Introduction to Nonlinear Network Theory*. New York: McGraw-Hill, 1969.
[2] ——, "Nonlinear circuit theory," in *Proc. 1978 Eur. Conf. Circuit Theory and Design*, R. Brayton *et al.*, Eds., 1978, pp. 65–173.
[3] J. Park, A. N. Pasupathy, J. I. Goldsmith, C. Chang, Y. Yaish, J. R. Petta, M. Rinkoski, J. P. Sethna, H. D. Abruna, P. L. McEuen, and D. Raph, "Coulomb blockade and the Kondo effect in single-atom transistors," *Nature*, vol. 417, pp. 722–725, June 13, 2002.
[4] W. Liang, M. P. Shores, M. Bockrath, J. R. Long, and H. Park, "Kondo resonance in a single-molecule transistor," *Nature*, vol. 417, pp. 725–729, June 13, 2002.
[5] R. H. M. Smit, Y. Noat, C. Untiedt, N. D. Lang, M. C. van Hemert, and J. M. van Ruitenbeek, "Measurement of the conductance of a hydrogen molecule," *Nature*, vol. 419, pp. 906–909, Oct. 31, 2002.
[6] J. Jortner and M. Ratner, *Molecular Electronics*. Oxford, U.K.: Blackwell Sci., 1997.
[7] S. Roth, M. Burghard, and C. M. Fischer, "Resonant tunneling and molecular rectification in Langmuir-Blodgett films," in *Molecular Electronics*, J. Jortner and M. Ratner, Eds. Oxford, U.K.: Blackwell Sci., 1997, pp. 255–280.
[8] M. A. Reed, C. Zhou, C. J. Muller, T. P. Burgin, and J. M. Tour, "Conductance of a molecular junction," *Science*, vol. 278, pp. 252–254, Oct. 10, 1997.
[9] X. D. Cui, A. Primak, X. Zarate, J. Tomfohr, O. F. Sankey, A. L. Moore, T. A. Moor, D. Gust, G. Harris, and S. M. Lindsay, "Reproducible measurements of single-molecular conductivity," *Science*, vol. 294, pp. 571–574, Oct. 19, 2001.
[10] Y. Huang, X. Duan, Y. Cui, and C. M. Lieber, "Gallium nitride nanowire nano devices," *Nano Lett.*, vol. 2, no. 2, pp. 101–104, 2002.
[11] X. Duan, Y. Huang, and C. M. Lieber, "Nonvolatile memory and programmable logic from molecule-gated nanowires," *Nano Letters*, vol. 2, no. 5, pp. 487–490, 2002.
[12] L. O. Chua, C. A. Desoor, and E. S. Kuh, *Linear and Nonlinear Circuits*. New York: McGraw-Hill, 1987.
[13] L. O. Chua, "Memristor—the missing circuit element," *IEEE Trans. Circuit Theory*, vol. CT-18, pp. 507–519, Sept. 1971.
[14] Y. Miyamoto, S. G. Louie, and M. L. Cohen, "Chiral conductivities of nanotubes," *Phys. Rev. Lett.*, vol. 76, pp. 212–2124, 1996.
[15] Y. Miyamoto, A. Rubio, S. G. Louie, and M. L. Cohen, "Self-inductance of chiral conducting nanotubes," *Phys. Rev. B*, vol. 60, no. 19, pp. 13 885–13 889, Nov. 15, 1999.
[16] K. S. Cole, *Membranes, Ions, and Impulses*. Berkeley, CA: Univ. of California Press, 1972.
[17] K. S. Mauro, "Anomalous impedance, a phenomenological property of time-variant resistance," *Biophys. J.*, vol. 1, pp. 353–372, 1961.
[18] L. O. Chua, "Device modeling via basic nonlinear circuit elements," *IEEE Trans. Circuits Syst.*, vol. CAS-27, no. 11, pp. 1014–1044, Nov. 1980.
[19] L. O. Chua and P. M. Lin, *Computer-Aided Analysis of Electronic Circuits: Algorithms & Computational Techniques*. Englewood Cliffs, NJ: Prentice-Hall, 1975.
[20] K. Mainzer, *Thinking in Complexity. The Complexity Dynamics of Matter, Mind, and Mankind*, 4th ed. Berlin, Germany: Springer-Verlag, 2004.
[21] M. Mansfield and C. O'Sullivan, *Understanding Physics*. Chichester, U.K.: Wiley, 1998.
[22] M. Jammer, *Concepts of Mass in Contemporary Physics and Philosophy*. Princeton, NJ: Princeton Univ. Press, 2000.
[23] K.Klaus Deimling, *Nonlinear Functional Analysis*. Berlin, Germany: Springer-Verlag, 1985.
[24] B. D. Josephson, "Possible new effects in superconductive tunneling," *Phys. Lett.*, vol. 1, pp. 251–253, 1962.
[25] A. M. Kadin, *Introduction to Superconducting Circuits*. New York: Wiley, 1999.
[26] T. P. Orlando and K. A. Delin, *Foundation of Superconductivity*. Reading, MA: Addison-Wesley, 1991.
[27] H. Koch and H. Lübig, *Single-Electron Tunneling and Mesoscopic Devices*. Berlin, Germany: Springer-Verlag, 1992.
[28] A. L. Hodgkin and A. F. Huxley, "A quantitative description of membrane current and its application to conduction in nerve," *J. Phys.*, vol. 117, pp. 500–544, 1952.
[29] E. A. Guillemin, *Synthesis of Passive Network*. New York: Wiley, 1957.
[30] L. T. Bruton, *RC-Active Circuits: Theory and Design*. Englewood Cliffs, NJ: Prentice-Hall, 1980.

Leon O. Chua (Fellow, IEEE) received the M.S. degree from the Massachusetts Institute of Technology, Cambridge, and the Ph.D. degree from the University of Illinois, Urbana-Champaign, in 1962 and 1964, respectively.

Since 1970, he has been with the University of California, Berkeley, where he is currently a Professor of Electrical Engineering and Computer Sciences. For developing an axiomatic foundation for nonlinear circuit theory, he is widely recognized as the foremost pioneer of this area. He is responsible for introducing chaos theory and nonlinear dynamics to engineering and is recognized widely for inventing Chua's circuit, which has become a standard reference, in the literature on chaos. His seminal work on cellular neural networks (CNNs) has become a very active research area worldwide. Among all neural network architectures, the CNN is the only one that has been successfully implemented into a practical fully programmable chip capable of solving ultrahigh-speed mission-critical image processing problems. He has been awarded seven U.S. patents

Dr. Chua was elected a Foreign Member of the European Academy of Sciences (Academia Europea) in 1997. He received the IEEE Browder J. Thompson Memorial Prize in 1972, the IEEE W. R. G. Baker Prize in 1978, the Frederick Emmons Award in 1974, the M. E. Van Valkenhurg Award in 1995 and in 1998, and the IEEE Neural Networks Pioneer Award in 2000. He has been awarded nine honorary doctorates (Doctor Honoris Causa) from major European universities and Japan. In 2002, he received the Top 15 Cited Authors in Engineering Award chosen from the Current Contents (ISI) database of all cited papers in the engineering disciplines in the citation index from 1991 to October 31, 2001, from all branches of engineering.