Thermal Noise in Nonlinear Resistive Devices and its Circuit Representation

MADHU S. GUPTA, SENIOR MEMBER, IEEE

Abstract—The major developments, since the Nyquist theorem, in the study of thermal noise in linear and nonlinear dissipative devices are briefly summarized. Then the author's recently established theorem for calculating thermal noise in biased, near-equilibrium, nonlinear resistive devices is discussed, and examples of its use are presented. Based on the theorem, an equivalent circuit model for representing the noise in such devices is proposed, and is applied to analyze the behavior of nonlinear resistors in such applications as heat engines and refrigerators.

NOTATIONS

$\langle x \rangle$	Ensemble average of x.
$\bar{x}$	Time average of x.
$\dot{x}$	Time derivative of x.
v_A, i_A	Total instantaneous value of a voltage, current.
V_A, I_A	Ensemble average of v_A, i_A.
v_a, i_a	Fluctuation component of v_A, i_A.

I. INTRODUCTION

DISSIPATION of energy, fluctuations due to random noise, and nonlinearity of characteristics are three features of electrical and other systems with much in common: they occur in all systems of practical interest, they are abhorred and diligently avoided in many systems (such as in communication channels) while carefully exploited in others, and they are deeply interrelated through thermodynamic considerations. The connection between dissipation and fluctuations has now been known for several decades: fluctuations provide the mechanism for energy dissipation, and dissipative systems at finite temperatures are subject to thermal fluctuations. The Einstein relationship between diffusivity and mobility, and the Nyquist theorem of thermal noise, are a manifestation of this connection. The connection with nonlinearity, although less apparent and less well studied, has nevertheless been firmly established by some recent work [1], [2]. This paper brings out one aspect of this connection, by discussing the thermal fluctuations in nonlinear dissipative systems.

Specifically, this paper has three primary purposes (listed below under items 2 to 4); in the process it also serves an additional, secondary purpose (given in item 1):

1) To outline the major developments in the literature on thermal noise, both in linear and in nonlinear systems.

2) To describe, in electrical engineering language, a theorem, recently derived by the author, for calculating thermal noise in a class of nonlinear resistive devices.

3) To propose an approximate noise equivalent circuit for such nonlinear resistive devices based on the theorem.

4) To give examples of the use of the theorem and of the equivalent circuit.

The background and the need for each of these four objectives is briefly summarized in the next four paragraphs.

The study of fluctuations in the form of Brownian motion (1828) is almost as old as Ohm's law (1826). (For an early history of the subject, see some reprints in [3].) A rigorous, quantitative study of thermal noise may however be taken to begin with the publication of two papers by Nyquist [4] and Johnson [5] in 1928, a hundred years after Brown observed Brownian motion. These papers contain Nyquist's derivation of the fundamental theorem on thermal noise, and Johnson's experimental verification of that theorem. In the half century elapsed since, the understanding of the subject, as well as the generality of the basic results, has been enlarged by a number of investigators. Some of those advancements of engineering interest are summarized here in outline form (in Section II), as such a collection is not available elsewhere.

Perhaps the most restrictive among the assumptions required in the derivation of Nyquist's theorem is that of system linearity. As many devices of practical interest are nonlinear, attempts have long been made to extend or generalize the Nyquist theorem to nonlinear systems. Several difficulties were encountered in such early attempts, and are summarized by van Kampen [6]. It has now been known for over twenty years that the fluctuation spectrum for an arbitrary nonlinear device cannot be uniquely determined by the phenomenological characteristics of the device, and depends also upon the higher order correlation functions [7]. Despite this negative result, investigators have continued to look for a method of calculating thermal noise in nonlinear devices, because even though a rigorous general theorem cannot be found, an approximate result, or one with a restricted range of utility, will still be of some engineering interest. Several such *engineering attempts* are summarized in Section II-F. Recently, the author derived a thermal noise theorem, applicable to a limited class of purely resistive nonlinear systems close to thermal equilibrium, on which the present paper is based. This theorem, and its proof, are stated in general thermodynamic language in the original paper [1]. The theorem is restated and explained in Section III of this paper in the terminology of electrical engineering, with the hope of making it known to the engineering community and accessible to readers with engineering background.

Electrical engineers have long had a preference for expressing relationships in the form of equivalent circuits, because equivalent circuits are a convenient short hand for conceptualizing, memorizing, and using the relationships they represent, particularly for those accustomed to them. For this reason, an approximate equivalent circuit is proposed here for noisy, nonlinear, two-terminal, purely resistive devices close to equilibrium. This equivalent circuit model a) includes thermal

Manuscript received November 3, 1981; revised June 3, 1982. This work was supported in part by the National Science Foundation.

The author is with the Department of Electrical Engineering and Computer Science, University of Illinois at Chicago, Chicago, IL 60680.

0018-9219/82/0800-0788$00.75 © 1982 IEEE

fluctuations, b) approximates the electrical characteristics of the resistor at its terminals, and c) is thermodynamically consistent and correct. The need for such a model arises in a wide variety of problems. As one example of pragmatic interest, when nonlinear resistors are used as detectors of weak, high-frequency signals, their detection sensitivity and noise contribution are important specifications, and are most conveniently calculated from an equivalent circuit model which incorporates noise. As another example of conceptual interest, the paradox that a nonlinear resistor appears able to deliver a dc signal by rectifying the thermal noise from a linear resistor in thermal equilibrium (thus behaving as an electrical version of Maxwell's demon), should be resolved by a correct noise equivalent circuit.

Finally, this paper contains several applications of the thermal noise theorem, and of the noise equivalent circuit, to nonlinear resistive devices. The examples contain results of interest in themselves, provide specific test cases by which to judge the validity of the theorem, and serve to illustrate the use of the theorem and of the noise equivalent circuit.

II. A Survey of the Literature on Thermal Noise

A. What is Thermal Noise

Thermal noise is a consequence of the discrete (particle) nature of matter and energy. Most macroscopically observable physical variables, such as electric current, are only averages, over a large number of particles, of some parameter describing those particles. When observed more precisely, the statistical nature of the macroscopic variables become apparent from the fluctuations in their values around the average.

Thermal noise in a physical system is that part of the fluctuations which arises from the presence of thermal energy in the system. Systems composed of a large number of particles have a large number of degrees of freedom capable of storing energy. A macroscopic description of the system (e.g., specifying a few phenomenological variables) constraints the energy in only a few degrees of freedom; the energy in the large number of remaining degrees of freedom is lumped together under the heading of "thermal energy." If there is a coupling between the macroscopic and the thermal degrees of freedom it must be bidirectional, i.e., energy can be exchanged in either direction (this is a consequence of the quantum mechanical rule that every quantum transition has an inverse with equal probability). The flow of energy from macroscopic to thermal degrees of freedom is called "dissipation." The flow of energy from thermal to macroscopic degrees of freedom manifests itself as thermal fluctuations in the corresponding macroscopic variables.

B. Thermal Noise Theorem for Linear Systems (Nyquist, 1928)

The first and fundamental result in the theory of thermal noise is the Nyquist theorem. A highly restricted version of this theorem is stated in this section, with the several generalizations described in Section II-C.

Nyquist's theorem states that the spontaneous random fluctuations in the terminal voltage (or current) of an arbitrary linear resistor (any two-terminal electrical circuit with a purely resistive impedance), having a resistance R, maintained in thermal equilibrium at a temperature T (by means of a heat bath), are independent of such parameters as its conduction mechanism, composition, construction, or dimensions, and depend only upon the values of R and T in the following manner:

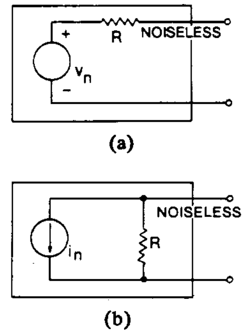

Fig. 1. (a) Thevenin, and (b) Norton noise equivalent circuits for a noisy linear one-port resistor in thermal equilibrium.

i) The spectral density (or the mean-square value per unit bandwidth B) of the open-circuit terminal noise voltage v_n is given by

$$S_v(f) = \frac{\overline{(v_n^2)}}{B} = 4kTR. \tag{1}$$

ii) The spectral density (or the mean-square value per unit bandwidth B) of the short-circuit terminal noise circuit i_n is given by

$$S_i(f) = \frac{\overline{(i_n^2)}}{B} = \frac{4kT}{R}. \tag{2}$$

iii) The available noise power (i.e., the noise power delivered to a matched load, which would be another resistor of the same value R) per unit bandwidth is given by

$$P_{av} = kT. \tag{3}$$

iv) The noisy resistor can be represented by a Thevenin equivalent, consisting of a noise voltage source of rms value $\sqrt{4kTRB}$ in a bandwidth B and a noiseless resistor R (as shown in Fig. 1(a)), or equivalently, by a Norton equivalent, consisting of a noise current source of rms value $\sqrt{4kTB/R}$ and a noiseless resistor R (as shown in Fig. 1(b)).

The four statements i) through iv) are equivalent and any one leads to the other three.

C. Derivation of Nyquist Theorem

Nyquist's theorem can be proved in a variety of ways; indeed a dozen "different" derivations appearing in the literature are cataloged in [8]. Fundamentally, however, the various classical derivations fall in one of three groups:

i) *Derivations based on kinetic theory [9]-[12]*, in which a detailed and specific microscopic model is postulated to describe a resistive system. For example, for a metallic resistor, the transport of electrons is described by a model of the scattering process which is responsible for the dissipation of energy. The conductance of the system is related to the average velocity of the electrons, while the terminal fluctuations are related to the random thermal velocities of the electrons, both of which are found in terms of the parameters of the scattering model. The requirement of thermal equilibrium is then introduced by assuming a Maxwellian distribution for electron velocities, or the law of equipartition of energy, or something equivalent to it. This leads to the Nyquist relationship between conductance and fluctuations, regardless of the assumed details of the scattering model.

ii) *Derivations based on Rayleigh-Jeans law [13], [14]*, in which the resistive system under consideration is placed in contact with another, known, convenient system such as a black-body cavity or a terminated transmission line, for which the frequency distribution of the available thermal power is governed by Rayleigh–Jeans law in one-dimensional form, or something equivalent to it (such as found by counting the number of modes on the transmission line connecting the two systems and determining the energy in each mode from the law of equipartition of energy). From the principle of detailed balance, there should be no transfer of thermal power between the two systems via fluctuations in thermal equilibrium. As a result, the noise power available from the resistive system under consideration is found.

iii) *Derivations based on Markovian transitions postulate [15], [16],* essentially lead to the fluctuation–dissipation theorem, which is a fundamental result in statistical mechanics. The dissipative system is described in terms of the set of transition probabilities between the various possible states that it can attain. The response of the system, as it relaxes back to equilibrium from some other initial state is found in terms of the transition probabilities. This response is taken to be independent of how the initial state was arrived at: whether as a result of an applied excitation or spontaneous fluctuations; this is the so-called Markovian or Onsager postulate. This postulate thus relates the autocovariance of the fluctuations to the impulse response of the system, and therefore the fluctuation spectrum to the conductance of the resistive system.

The three classes of derivations of Nyquist's theorem mentioned above have an increasing order of generality. Kinetic derivations refer to a specific type of electrical conductor, radiative derivations prove the theorem for an arbitrary electrical conductor, while the statistical derivations prove the generalized theorem for all linear dissipative systems, electrical or otherwise. For the present purposes, the significant observation is that in each derivation, the two crucial assumptions on which the theorem is based are those of linearity and thermal equilibrium.

D. Generalizations of Nyquist Theorem

The Nyquist theorem of Section II-B is stated in a form even more restricted than the one first proved by Nyquist. Several generalizations and extensions of the theorem have been established, and are summarized here. The generalizations included in the following are those having some practical or engineering implications; many others which contribute primarily to the elegance or completeness of results are omitted here.

i) Generalization to Arbitrary Impedance: For a two-terminal network or device in thermal equilibrium, having a driving point impedance $Z(f) = R(f) + jX(f)$ at frequency f, Nyquist's theorem is applicable if the device is viewed as a series connection of a noiseless reactance $X(f)$ and a noisy resistor $R(f)$ with the same thermal noise as stated above. (Alternatively, the network admittance may be viewed as a parallel combination of a noiseless susceptance and a conductance with thermal noise.) Consequently, the theorem stated as (1) and (3) still holds, while (2) holds with the generalization

$$S_i(f) = \frac{4kTR}{R^2 + X^2}. \tag{4}$$

ii) Generalization to Interconnection of Impedances: When the linear, two-terminal network consists of a number of interconnected impedances, the terminal noise voltage or current can also be calculated from a network representation in which each impedance is replaced by its noise equivalent, and the noise sources are combined by Kirchhoff's current and voltage laws. This result was analytically as well as experimentally demonstrated by Williams [17].

iii) Generalization to Nonreciprocal Networks: For a linear, multiterminal network containing nonreciprocal elements, both the Thevenin theorem and the Nyquist theorem require a generalization. It was shown by Twiss [18] that the driving-point impedance appearing in the theorems must be replaced by a linear combination of the elements of an impedance matrix describing the nonreciprocal network.

iv) Generalization to High Frequencies and Low Temperatures (Quantum-Mechanical Correction): The Nyquist theorem, as stated in (1), cannot of course be valid at arbitrarily high frequencies, because when integrated over all frequencies, it leads to an infinite noise power, sometimes called the high-frequency catastrophe. A quantum-mechanical calculation [7] of the spectral density of thermal noise at a frequency f leads to the replacement of the factor kT in Nyquist theorem by the mean energy per oscillator $\langle E \rangle$ in an ensemble of quantum-mechanical harmonic oscillators having a natural frequency f and maintained at the temperature T given by

$$\langle E \rangle = \frac{hf}{2} \coth\left(\frac{hf}{2kT}\right) \tag{5}$$

$$= \frac{1}{2} hf + \frac{hf}{\exp(hf/kT) - 1} \tag{6}$$

where h is Planck's constant. In the limit of $hf/kT \to 0$, the energy $\langle E \rangle$ tends to kT, and the classical expression is recovered. This replacement can be carried out directly in (1) and (2) but not in (3). The first term in (6) is the zero-point energy of the harmonic oscillator, and since it cannot be extracted from the oscillator, it should not be included in the available noise power. Therefore, in a quantum-mechanical generalization of (3), kT should be replaced[1] by only the second term on the right

$$P_{av} = hf \Big/ \left[\exp\left(\frac{hf}{kT}\right) - 1\right] \tag{7}$$

as emphasized by Weber [19].

v) Generalization to Negative Temperature: A thermodynamical system can attain a negative temperature [21] provided it satisfies the following three conditions: a) The elements of the system are in thermodynamic equilibrium with each other, so that a temperature can be defined at all, b) there is an upper limit to the energy in the allowed states of the system so that negative temperature can be reached with a finite energy, and c) the system is thermally isolated from those systems which do not satisfy both a) and b) above (or that the time required to reach an internal quasi-equilibrium is small compared to the time in which appreciable energy is lost to or gained from the outside), so that the energy of the system is steady during an interval over which the temperature may be defined. A positive temperature system is "dissipative" in that, according to the second law of thermodynamics, it is possible to do work *on* it, resulting in an equivalent amount

[1] An alternative viewpoint, proposed by Siegman [20], is also possible in which the entire energy $\langle E \rangle$ may be formally treated as being available; then the noise power, delivered by a thermal noise source to a linear system, would have an additional part due to the zero-point energy, and this is held responsible for the minimum (quantum mechanical) noise of the linear system.

of heat being delivered *to* it, in a closed cycle without producing any other effect. In an analogous manner, the second law of thermodynamics allows that work may be performed *by* a system at a negative temperature by extracting an equivalent amount of heat *from* it, in a closed cycle without producing any other effect. The "dissipation" is therefore negative, or is actually "generation." The resistance (or dissipation constant) of the system is therefore also negative.[2] For a system having a negative temperature T and a negative resistance R, Nyquist's theorem is applicable in all of the four forms listed earlier [22]. Three of these are easy to accept because they involve only the ratio or the product of R and T, which is a positive quantity. That the fourth form is also applicable is evident from the fact that the "available" noise power is the power delivered to a *matched* load. For the negative resistor, the matched load is also a negative resistor, and the power delivered to this load can be negative (i.e., positive power can be extracted from the load). Haus and Adler [23] have introduced the term "exchangeable power" to replace the "available power" under such circumstances. This extention of Nyquist's theorem is not a mere mathematical exercise because negative-resistance, negative-temperature systems exist physically and have practial utility. The spin system in a paramagnetic crystal, an example of such a system, has applications in maser amplifiers and the extended Nyquist's theorem allows the calculation of the noise in such an amplifier [24].

vi) Generalization to Linear Dissipative Media (i.e., Distributed Systems): The calculation of thermal noise generated in linear dissipative circuits which are distributed in one dimension (such as lossy transmission lines, waveguides, and other transmission media) can be carried out by a direct and straightforward application of Nyquist theorem to an elementary length of the circuit. This leads to the simple result [25] that the noise power spectral density per unit length, generated in the circuit and *propagated in one direction* is $2\alpha kT$, where α is the voltage attenuation constant per unit length due to dissipation. A generalization of Nyquist theorem is required when the noise is generated in a linear dissipative medium distributed in three dimensions. Haus [26] has shown that the fluctuations in such media can be accounted for by adding a random-noise current source to one of the Maxwell's equations (Ampere's law for a dissipative electric medium), and has found the time and space correlations of this current source. The results can be further generalized [27] to anisotropic and nonuniform media, and permit the calculation of thermal noise radiated by media at uniform or nonuniform temperatures.

vii) Generalization to an Arbitrary Linear Dissipative System: Nyquist's theorem is not limited to electrical circuits having electrical fluctuations and electrical dissipation. It is applicable to any arbitrary thermodynamic linear dissipative system [281 The kinetics of the system are described by means of a generalizea velocity v and a generalized force F which are related to each other through the Langevin equation

$$m \frac{dv}{dt} + \alpha v = F \tag{8}$$

where m and α are the inertial and dissipational constants, respectively (for example, due to mass and friction). In the absence of an external force, F represents a net random force arising from the interaction of the thermal degrees of freedom with the generalized coordinate x (where $dx/dt = v$). Then it follows from the statistical definition of temperature T for the system that the dissipation constant

$$\alpha = \frac{1}{2kT} \int_{-\infty}^{\infty} \phi_F(\tau)\, d\tau \tag{9}$$

where $\phi_F(\tau)$ is the autocorrelation function of $F(t)$

$$\phi_F(\tau) = \langle F(0)\, F(\tau)\rangle_0. \tag{10}$$

The $\langle\ \rangle$ denotes ensemble average and the subscript $_0$ implies the absence of applied external forces. This theorem does not make any reference to the nature of the linear system considered or the microscopic origin of α. The result (9) is a generalization of Nyquist's theorem and is called the fluctuation dissipation theorem. It is applicable to any linear dissipative system, including electrical circuits.

viii) Extention to Nonequilibrium (or Driven) Systems: Most electronic devices of interest are operated under driven conditions, so that an extension of Nyquist theorem for nonequilibrium conditions is desirable. Landsberg and Cole [29] have shown by rigorous calculations that, for a system maintained in steady state, the noise current spectrum can still be written in terms of the equilibrium admittance of the system, along with a multiplicative correction factor. The correction factor is dependent both upon frequency and upon how strongly the system is driven.

E. Attempts at Generalization to Nonlinear Systems

Many attempts have been made in the past to derive a result similar to Nyquist theorem or fluctuation dissipation theorem for nonlinear systems. Several different techniques have been used in these attempts, each using a different set of axioms, including those based on a Langevin equation, a Fokker–Planck equation, an Onsager-like postulate for the regression of fluctuations, and the Master equation. The various methods have different implicit axioms and subtle conceptual problems, and their results are not always in agreement. A sizable amount of the literature on this subject has been critically reviewed by Lax [30] and by van Kampen [6]. The following list of some of the conclusions from this literature will point out the difficulties involved and will hopefully serve as a warning against expecting a general result of broad utility.

i) For a nonlinear function f of a random variable x, the quantities $\langle f(x)\rangle$ and $f(\langle x\rangle)$ are not necessarily identical. Therefore, a careful attention must be paid to the definition and interpretation of the phenomenological variables involved in the calculation of noise in nonlinear systems [6].

ii) The phenomenological characteristics of a system are expressed in terms of macroscopic parameters (such as voltage and current) which are themselves "coarse-grained" quantities in which fluctuations have been averaged out. There is no fundamental reason why the quantitative measures of fluctuations (such as the power spectrum) must always be expressible in terms of the phenomenological characteristics of the system [31].

iii) The assumption that either the fluctuations in a physical quantity, or the probability distribution of that fluctuation, follows the phenomenological laws is not always correct, and counterexamples are easily found. In particular, the regression

[2] While a negative temperature system is shown to result in a negative resistance, systems with negative incremental resistance, such as tunnel diodes, can have a positive temperature. Such devices are necessarily nonlinear, must be away from thermal equilibrium (i.e., biased) in order to reach the incremental negative resistance region, and their noise is then bias dependent. The linear Nyquist theorem, as stated in Section II-B, does not apply to these devices.

of fluctuations may not be identical with the relaxation response of a system for nonlinear systems [32].

iv) In general, the spectrum of fluctuations in a nonlinear system cannot be expressed in terms of the phenomenological variables alone, and requires a more detailed knowledge of the system than just the relationship between the phenomenological variables, which are only the first-order moments. The requisite detail depends on how the nonlinear system is characterized. For example, if the nonlinear system is described by a second-degree Volterra kernel, then, in the presence of a driving force, the first-order correction in the second moment (i.e., mean-square value) of an observed quantity is related to the third moment of that quantity in equilibrium [7]. Similarly, if the nonlinear system is described by the Kramer–Moyal expansion of the Fokker–Planck equation, the fluctuation spectrum is not uniquely determined by the linear phenomenological coefficients alone, and involves higher order coefficients [30].

v) The circuit representation of a noisy resistor, such as a Thevenin equivalent, tacitly assumes that the internal noise generation is independent of the load connected across the resistor. But the noise generation in a dissipative system will depend on the constraints under which the system is maintained, and the load is a form of constraint imposed on the system parameters. As a result, a load-independent description of the thermal noise cannot be developed for nonlinear systems in general [30].

vi) One encouraging result can also be added to the above gloomy picture: The Nyquist theorem is applicable to nonlinear systems [33] provided the system is at equilibrium, the fluctuations are small, and the impedance of the system is understood to be the small-signal (i.e., linearized part of the) impedance at equilibrium. However, van Kampen [6, p. 172] has pointed out that this result is not rigorously established, and is not valid universally.

It is clear that, unlike the Nyquist theorem, any calculation of thermal noise in terms of the phenomenological or "terminal" characteristics of a nonlinear device will have a limited range of validity among the family of nonlinear devices. For engineering applications, this is still acceptable, compared to the alternative in which the noise in *each* device must be separately calculated from first principles, provided the limits to the applicability of the results are known. Furthermore, even an approximate result will be useful, provided the approximation holds well in the typical circumstances. The remainder of this paper is devoted to such a pragmatic goal.

F. Engineering Calculations of Thermal Noise in Nonlinear Resistors

If the nonlinear system can be decomposed into constituent parts such that all the resistive parts are linear, and the nonlinearity is confined to the lossless, energy-storing (i.e., purely reactive) parts, thermal noise in the system is easily determined through the use of the linear Nyquist theorem and the constitutive equations (for example, see Anderson [34]). The present interest is therefore focussed on nonlinear resistive systems. From an engineering point of view, there is a need for a simplified, even if approximate, method of calculating thermal noise in nonlinear resistive systems under nonequilibrium conditions. Only a very small number of investigators have approached this problem with such pragmatic goals, and their attempts are summarized here. The limitations of each of their results are also pointed out. Finally, Gupta's result [1]

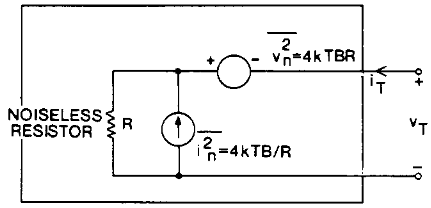

Fig. 2. Noise equivalent circuit for a one-port resistor in thermal equilibrium, as proposed by van Nie.

on the thermal noise of nonlinear resistors, on which this paper is based, is reviewed in Section III.

van der Ziel [35] approached the problem by asking how R, appearing in (1) and (2), should be defined if these equations are to remain applicable to nonlinear resistors. He considered a nonlinear resistor with a dc voltage V across it, a dc current I flowing through it, and having a small-signal admittance $y(f)$ at a signal frequency f, and considered the following possibilities for Nyquist's theorem:

i)

$$\overline{v_n^2(f)} = 4kTB\,\frac{V}{I} \qquad \overline{i_n^2(f)} = 4kTB\,\frac{V}{I}\,|y(f)|^2 \qquad (11a)$$

ii)

$$\overline{i_n^2(f)} = 4kTB\,\frac{I}{V} \qquad \overline{v_n^2(f)} = 4kTB\,\frac{I}{V}\Big/|y(f)|^2 \qquad (11b)$$

iii)

$$\overline{v_n^2(f)} = 4kTB\,\frac{dV}{dI} \qquad \overline{i_n^2(f)} = 4kTB\,\frac{dV}{dI}\,|y(f)|^2 \qquad (11c)$$

iv)

$$\overline{i_n^2(f)} = 4kTB\,\frac{dI}{dV} \qquad \overline{v_n^2(f)} = 4kTB\,\frac{dI}{dV}\Big/|y(f)|^2 \qquad (11d)$$

v) None of the above.

These expressions were tested on individual nonlinear resistive devices for which thermal noise has already been separately calculated from the first principles. van der Ziel thus came to the conclusion that while a particular expression may yield the correct result for an individual device, none of the expressions considered is universally valid, and therefore v) is the correct answer.

van Nie [36] observed that the terminal properties of a noisy linear resistor can be represented not only by the Thevenin model of Fig. 1(a), but also by any other two-terminal network for which the model of Fig. 1(a) is a Thevenin equivalent. Of the infinite number of such networks, van Nie selected one which satisfies an additional condition: that the rms value of the noise current flowing through the noiseless resistor inside the model be independent of the external circuit connected across the model terminals. This condition leads to the equivalent circuit shown in Fig. 2 containing two noise sources which are fully correlated (correlation coefficient = $1 + j0$). van Nie then postulated that the same model is applicable to nonlinear resistors having a terminal current–voltage characteristic given by

$$v_T = Ri_T + ai_T^2, \qquad R \neq 0 \qquad (12)$$

where R and a are constants (although possibly frequency dependent), with the noiseless linear resistor R in the model replaced by a noiseless nonlinear one. Notice that the noise

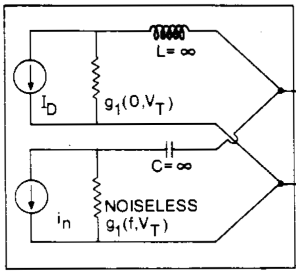

Fig. 3. Gunn's model for a nonlinear, isothermal, resistive, two-terminal device.

sources remain unchanged, with their magnitudes dependent on R but not a. Such a model for a noisy nonlinear resistor has only one distinguishing feature, pointed out by van Nie. If the noise current flowing through the noiseless resistor of the model is independent of the termination, then the rectified dc voltage developed across the resistor due to its nonlinearity is also independent of the impedance of the external circuit. The rectification of noise, and development of a dc voltage across the noiseless resistor, occurs even in thermal equilibrium; however, the second law of thermodynamics can now be rescued by postulating an internal dc source which does not depend upon the external circuit (as contrasted with Gunn's model, to be discussed next). van Nie's model suffers from a number of shortcomings.

i) The model assumes that the Nyquist theorem continues to hold when the resistor under consideration is nonlinear, and carries a bias current.

ii) The noise sources contained in the model are a function of R but not of a; i.e., the nonlinearity of the resistor is assumed not to influence the noise generation at all.

iii) The noise current, flowing through the noiseless resistor employed in van Nie's model, is independent of the terminating impedance only if the resistor, as well as the termination, are linear. The model therefore loses its only distinguishing feature in the presence of nonlinearity.

iv) If the nonlinear resistor, as well as its termination, are linearized in the neighborhood of their respective operating points, the noise current through the noiseless resistor in the model will indeed be independent of the termination. But such a linearized model does not possess any rectification property.

Gunn [37] proposed a model for an isothermal nonlinear resistive device which is somewhat similar to the model presented later in this paper. Gunn's model is shown in Fig. 3, and is described here with a little paraphrasing in the interest of clarity. The model employs four essential circuit elements:

i) $g_1(f, V_T)$ is the small-signal conductance of the device, which is a function of frequency f and of the dc component V_T of the terminal voltage v_T.

ii) $g_1(0, V_T)$ is the zero-frequency value of $g_1(f, V_T)$, and therefore equals the incremental dc resistance dI_T/dV_T, where I_T is the dc component of the terminal current i_T.

iii) i_n is a noise current source, having a mean-square value per unit bandwidth given by

$$\frac{\overline{i_n^2}}{B} = 4kT\left[g_1(f, 0) + \frac{1}{2}V_T g_2\right] \tag{13a}$$

consisting of two parts, the first of which is spontaneous and the second is dependent on the dc terminal voltage V_T, and g_2 is given by

$$g_2 = \frac{\partial g_1(f, V_T)}{\partial V_T}. \tag{13b}$$

iv) I_D is a dc current source, given by

$$I_D = I_S + \frac{1}{2}\int g_2\left(\frac{\overline{v_t^2}}{B}\right) df \tag{14a}$$

where $\overline{v_t^2}/B$ is the mean-square value of the noise voltage at the device terminals per unit bandwidth, and I_S is a "spontaneous dc current." Gunn gives neither a general expression for I_S nor an algorithm for finding it, although he does calculate it for a special case in which the device is connected across a conductance that is equal to $g_1(f, 0)$ in a narrow bandwidth B and infinite at all other frequencies, and for this special case

$$I_S = -\tfrac{1}{2} g_2 \overline{v_t^2}\big|_{V_T=0} = -kTB\, g_2/g_1(f, 0). \tag{14b}$$

As a generalization of his method of calculating I_S, this current source must be calculated for a given termination by requiring that the model satisfy the second law of thermodynamics.

In addition, the circuit model shown in Fig. 3 includes an infinitely large inductor for choking signals of all frequencies and an infinitely large capacitor for blocking dc, in the top and bottom halves of the model, respectively. Gunn's paper does not show or mention such low- and high-pass filters explicitly, but his discussion implicitly assumes the presence of these (or some other equivalent) ideal filter elements.

Gunn showed that, to the first order (i.e., for small V_T), his model satisfies two thermodynamic principles: the second law of thermodynamics and the Onsager reciprocity principle; indeed, (13b) and (14b) were deduced by requiring that these two principles be satisfied. However, his model is unsatisfactory in a number of ways:

i) Gunn makes the distinction between ensemble averaged signals and random fluctuations by identifying them with dc and nonzero frequencies, respectively. This has the disadvantage that time-varying signals are choked by the inductor and passed by the capacitor in the model of Fig. 3, rather than being treated in a quasi-static manner.

ii) The model does not reproduce the nonlinear dc current-voltage characteristic at its terminals. Consider, for example, an ideal dc voltage source of magnitude V_T connected at the terminals of the model. The resulting dc terminal current I_T can be found by solving the circuit equations

$$\begin{aligned} V_T &= (I_T - I_D)\, g_1(0, V_T) \\ &= (I_T - I_S)\left.\left(\frac{\partial I_T}{\partial V_T}\right)\right|_{V_T} \end{aligned} \tag{15}$$

If I_S is a constant independent of V_T (as in (14b)), the terminal characteristic in (15) is linear, and if I_S is not a constant, it is not known.

iii) The spontaneous dc current source I_S is a rather unusual element: it flows in the reverse (high-resistance) direction in the nonlinear resistive device; it is present even in thermal equilibrium; its value depends not only on the temperature and the current-voltage characteristic of the nonlinear device but also on the temperature and characteristic of the termination across the device; there is no explicit expression given for it; and finally, its physical origin is obscure. All of these features arise because Gunn postulated that the current source I_D in his model is given by (14a). The second term in (14a) can be described to be due to the rectification of terminal noise by the nonlinear device. A careful examination of this term shows that the mere presence of voltage fluctuations at the terminals of the nonlinear device is sufficient to make it nonzero. Such fluctuations are present even in thermal equilibrium, so that Gunn's model has a rectified dc current term even in thermal equilibrium. The spontaneous current source I_S is merely a patch to repair that defect: it nullifies the rectified current present in thermal equilibrium.

III. Nonlinear Thermal Noise Theorem (Gupta 1978)

In 1978, Gupta [1] was able to calculate the thermal noise in a special class of nonlinear systems, and under some restrictive conditions, entirely in terms of phenomenological parameters. This theorem forms the basis of the noise model for nonlinear resistors described in the present paper. The original paper should be consulted for a derivation of the theorem or the precise statement of the axioms and approximations used in the derivation. The purpose of this section is to summarize only the assumptions and results of that thermodynamic theorem, and to state them in the language of electrical circuits to facilitate its use in electrical engineering.

A. Assumptions and Approximations

An explicit understanding of the constraints under which the present model applies is essential to guard against its incorrect use. The following is a list of the assumptions from which the theorem is derived.

Consider a thermodynamic system, enclosed by a boundary, with an opening (or "port") in that boundary through which the system is in contact with the remainder of the universe (which therefore serves as a "termination" at that port). This system is assumed to satisfy the following requirements.

i) The system is "classical," i.e., the quantum-mechanical effects are disregarded in the calculation of noise. However, quantum-mechanical processes with no classical analog and with phenomenologically observable consequences may go on in the system. Thus tunnel diodes are not excluded although tunneling is a quantum-mechanical phenomenon.

ii) The system must be large enough that a temperature and an entropy can be defined for the system.

iii) The system contains within its boundaries, in addition to energy, a physical quantity X, which can pass through the port of the system, and which is conserved, i.e., neither created nor destroyed within the system.

iv) The system has a uniform temperature T throughout.

v) The system also has a uniform value of Y, the force variable conjugate to X in energy representation (defined by the statement that YdX is the increase in system energy accompanying an increase of X by dX in the system).

vi) The system is purely resistive. This means that the system does not store any free energy in the quantity X, and any energy YdX given to the system is dissipated (i.e., converted into heat). This in turn implies that the excitation variable Y is an instantaneous function of $\dot{X}$, the time rate of increase of X, but not of X itself. Then the rate of dissipation of energy is given by $Y\dot{X}$.

vii) The system is Markovian, i.e., its future behavior depends upon its present state, and not upon the details of how that state was reached. In particular, if the system is perturbed from a steady state, its response is the same regardless of whether the perturbation was caused by spontaneous fluctuations or by an externally applied excitation.

viii) *Small-bias approximation.* When an excitation Y is applied, the system may be called a "driven" system. From assumption vi), the excitation Y is accompanied by a rate of increase of X, and a dissipation $Y\dot{X}$. It is assumed that the excitation is small and the system is not far from thermal equilibrium, so that the system can be described by the equilibrium distribution of states (the so-called canonical distribution).

ix) *Small-fluctuation approximation.* The fluctuations in the system variables (in particular $\dot{X}$) are small. This approximation excludes systems which are close to a critical point, where large fluctuations occur.

The above nine assumptions define the class of systems to which the thermal noise theorem of [1] is applicable, and the fluctuations of interest here are those in the rate of flow of X through the system port. It is evident that the fluctuations in $\dot{X}$ will be influenced by the nature of the termination presented at the system port. (As an obvious example, for a resistor in thermal equilibrium, an electrical short circuit at the terminals makes the fluctuations in terminal voltage vanish, while an open circuit makes the noise current disappear.) In addition, in a nonequilibrium system (such as an electrical resistance with a bias current), the termination also determines the driven state of the system (i.e., the amount of bias), and therefore further influences the fluctuations. In short, the fluctuations in the flow rate of X can be calculated only after the termination of the system is specified. The following additional requirement is therefore imposed to specify the termination.

x) The termination maintains the system in a fixed, steadily driven state, with a fixed value of the excitation Y. The requirement of steady state implies that the excitation Y is time-independent, while the requirement of fixed Y implies that Y is deterministic, or that $Y = \langle Y \rangle$.

B. Statement of the Theorem

With the above ten assumptions, the fluctuation spectrum of $\dot{X}$ can be calculated [1]. As the system is assumed to be purely resistive, the power spectral density is independent of frequency, and can be written as the mean-square value $\langle \dot{X}^2 \rangle$ normalized to the bandwidth B in which it is measured. Define a new variable $\dot{x} \equiv \dot{X} - \langle \dot{X} \rangle$ which is the fluctuation component of $\dot{X}$. The thermal noise theorem of [1] shows that

$$\langle \dot{x}^2 \rangle = 4kTB[\overline{P_{ex}}/\overline{y^2}] \tag{16}$$

where P_{ex} is the excess power dissipation in the system when a small periodic excitation $y(t)$, having a zero time average, is superimposed on the steady excitation Y_0 describing the driven state (i.e., $Y = Y_0 + y(t)$), and where overbar denotes time average.

If the system is nonlinear, the relationship between the excitation $\langle Y \rangle$ and the resulting phenomenological (i.e., ensemble-averaged) response $\langle \dot{X} \rangle$ is a nonlinear one, and can be expressed as a power series

$$\langle Y \rangle = c_1 \langle \dot{X} \rangle + c_2 \langle \dot{X} \rangle^2 + \cdots$$

where the constant c_1 describes the linearized response, c_2 describes the lowest (second) order measure of nonlinearity, and so on. The fluctuations can also be expressed in terms of these system constants.

C. Theorem for a Nonlinear Electrical Resistor

When the system under consideration is a purely resistive, nonlinear, two-terminal electrical device, the thermal noise theorem (16) can be restated in terms of terminal current and voltage. The choice of the excitation and response variables Y and $\dot{X}$ must, however, be made carefully to conform with the assumptions on which the theorem is based. For example, the choice of charge as the conserved variable X (and hence current as $\dot{X}$ and voltage as Y) is inappropriate because an electrical circuit element is by definition electrically neutral and does not store charge at all. A little thought shows that since current flow in a resistor is caused by the transport of charge carriers, the electromagnetic momentum of the current in the direction of transport is a conserved variable which can serve as X, while half the average (drift) velocity of the carriers is the corresponding excitation Y. The kinetic energy due to drift is then XY, and the power dissipated in the form of heat is $\dot{X}Y$. As the drift velocity is proportional to the current, the choice of Y as the terminal current, and the corresponding choice of $\dot{X}$ as the voltage, is still more direct. The thermal noise theorem then expresses $\langle v_n^2\rangle$, the mean-square thermal noise voltage, measured in a bandwidth B, at the terminals of the resistor at temperature T, carrying a steady bias current I_{dc}, as

$$\langle v_n^2\rangle = 4kTB\, \overline{P_{ex}}/(\overline{i_s^2}) \tag{17}$$

where P_{ex} is the excess power dissipation caused by the presence of a small, zero time average, periodic signal current i_s superimposed on I_{dc}. This can be rewritten in terms of the terminal current (I)-voltage (V) characteristic of the nonlinear resistor as

$$\langle v_n^2\rangle = 4kTB\left(\frac{dV}{dI} + \frac{1}{2} I \frac{d^2V}{dI^2}\right)\Bigg|_{I=I_{dc}}. \tag{18}$$

Having established a correspondence between the thermodynamic variables $\dot{X}$ and Y and the electrical circuit variables V and I, the implications of the ten assumptions and approximations, stated in Section III-A, can now be understood in circuit terms. These include the conditions that the resistor is passive, it has a small bias (so as to be near equilibrium), and its noise signal is small. In particular, the assumption x) implies the restriction that the current I_{dc} through the nonlinear resistor is a nonrandom dc current. As this terminal current contains no random noise in the bandwidth B of interest, while the terminal voltage does, it is clear that the resistor is terminated in an infinite impedance, or that the noise voltage appearing in (18) is the open-circuit noise voltage.

Strictly speaking then, the thermodynamic result (18) has been established in [1], and should be used, only under the open-circuit conditions with an ideal current generator as the source of the bias. This is a very severe limitation to the utility of the result. Therefore, two additional assumptions are now made, which extend the result to other termination conditions.

xi) The first assumption is that the noise internally generated in a nonlinear resistor depends only upon the resistor, and not upon the remainder of the universe. As a result, once a noise model for a nonlinear resistor has been found for one terminating condition, it applies for other terminations as well. This assumption cannot be established rigorously (indeed, exceptions will exist), and its utility can be judged only through a comparison between its consequences and known results.

xii) The second assumption is a relaxation of the conditions of assumption x). When terminations other than the open circuit are permitted, the terminal current i_T through the resistor will no longer be noiseless and deterministic (i.e., the same for all members of the ensemble), but will contain a random component. In view of the assumption of small fluctuations, stated in Section III-A, it appears reasonable to replace the dc current I_{dc} in (18) by the ensemble average of the terminal current

$$\langle v_n^2\rangle = 4kTB\left(\frac{dV}{dI} + \frac{1}{2} I \frac{d^2V}{dI^2}\right)\Bigg|_{I_T}, \qquad I_T = \langle i_T\rangle. \tag{19}$$

With these assumptions, the terminal noise voltage v_t and current i_t, for a given bias I_T, and for any arbitrary value of the terminating impedance connected to the nonlinear resistor in the frequency range of interest, can be found in terms of the open-circuit value of v_t at the same bias, as given in (19). To understand this procedure, assume that the noisy nonlinear resistor can be represented by some model. As the terminating impedance is varied at a fixed bias, the values of v_t and i_t change, but the model does not, because the model is independent of the terminating impedance (by assumption xi) and of the noise current i_t (by assumption xii). Therefore, v_t and i_t can be deduced entirely from circuit equations, involving the model and the termination. The model, however, can be linearized by virtue of assumption ix) that the noise is small. As a result, only the incremental characteristic of the nonlinear resistor need be known.

To illustrate the above procedure, consider finding i_t for the case where the terminating impedance in the frequency range of interest is zero (i.e., the short-circuit case). The short-circuit noise current i_n can be related to the open-circuit noise voltage v_n through the incremental resistance of the device

$$\langle i_n^2\rangle = \langle v_n^2\rangle \left(\frac{dI}{dV}\right)^2\Bigg|_{I_T} \tag{20}$$

$$= 4kTB\left[\frac{dI}{dV} - \frac{1}{2} I \left(\frac{d^2I}{dV^2}\Big/\frac{dI}{dV}\right)\right]_{I_T}. \tag{21}$$

For later convenience, this result will now be rewritten in terms of a different parameter characterizing the lowest (second) order nonlinearity. The quantity β, defined as

$$\beta \equiv \frac{1}{2}\left(\frac{d^2I}{dV^2}\Big/\frac{dI}{dV}\right)\Bigg|_{I_{dc}} \tag{22}$$

has the units of $(\text{volts})^{-1}$, and serves as a convenient measure of the nonlinearity of a device with quadratic nonlinearity. It is identical with the "current sensitivity" defined for nonlinear resistors when they are used as low-level detectors [38]. In terms of β, the mean-square short-circuit noise current can be written as

$$\langle i_n^2\rangle = 4kTB[g(I_T) - \beta I_T] \tag{23}$$

where $g(I_T)$ is the incremental conductance of the nonlinear resistor at its "operating point." In analogy with the linear resistors, a "noise temperature" may be defined for a nonlinear resistor as

$$t_n \equiv \langle i_n^2\rangle/4kBg = T[1 - \beta I_T/g(I_T)]. \tag{24}$$

D. Implications of the Theorem

The following five observations on the nonlinear thermal noise theorem will help in understanding its use and limitations:

i) Limiting Values: When applied to linear resistors, for which $\beta = 0$, the result in (23) correctly reduces to the usual statement of Nyquist's theorem, contained in (2). It also reduces to Nyquist's theorem when the resistor is *strictly* at thermal equilibrium, i.e., the bias current I_T is zero. This limiting value is in agreement with the result of Bernard and Callen [33] mentioned in Section II-E. It may also appear from (24) that the noise temperature of the device can be made zero (or even negative!) merely by a proper choice of the bias current I_T. Such is not the case, because assumption viii), requiring small departure from equilibrium (i.e., small I_T), invalidates the theorem at such large I_T values; that assumption is equivalent to the condition that the noise temperature t_n be close to the physical temperature T of the nonlinear device.

ii) Bias Dependence of Noise: The noise spectral density in (21) consists of a "linear" part, and an "excess" part which contains the factor I_T. As the nonlinear thermal noise theorem has been deduced thermodynamically, the microscopic mechanism responsible for the "excess" noise is not prescribed. A survey of the literature on fluctuation phenomena shows that numerous papers have been published on the theory of fluctuations in nonequilibrium systems at steady state, in which "excess" noise is proportional to the square of the excitation applied to the system. Most of these theories were constructed in the context of searching an explanation for the $1/f$ noise and are therefore based on models which do not lead to "thermal" noise. These models typically require the system to be driven sufficiently away from thermal equilibrium so that the "excess" noise mechanism (e.g., the heating of the charge carriers in a resistor due to the bias [39]) may become effective. There are also some theories which attempt to explain $1/f$ noise as a quasi-equilibrium phenomenon [40]. The important point to remember is that the excess noise due to these various mechanisms, if it is observable, would be in addition to the thermal noise calculated here.

iii) Noise Cooling by Biasing: The noise temperature t_n of the nonlinear resistor, given by (24), differs from the physical temperature T by an amount dependent upon the bias current. In particular, t_n is lower than T if $\beta I_T/g$ is positive,[3] which is the case for forward-biased Schottky-barrier diodes. This result is not surprising [41], and has been known to microwave mixer designers for decades [38], [42]. Physically, the condition $t_n < T$ implies that thermal energy in the form of electrical noise can flow from another resistor at temperature T to the nonlinear resistor also at the same temperature T; this is the refrigerator action (discussed in greater detail in Section VI-B), where the work required for the process is derived from the source supplying the bias I_T.

iv) Lack of Additivity: It is apparent from (21) that, if two arbitrary nonlinear devices at the same temperature are connected in parallel, the mean-square short-circuit noise current of the combination cannot be calculated by applying the theorem to the I–V characteristic of the combination. The reason for this failure of the theorem can be traced back to the fact that, in general, the two nonlinear resistors will be at different noise temperatures, and there will be a net flow of thermal energy from one to the other. The theorem does apply provided the nonlinear resistors are so selected that, when connected in parallel, the current divides between them in a manner that results in equal noise temperatures for the two resistors. An obvious example of such a choice of nonlinear resistors is one where the resistors are alike to within a scale factor (for example, devices having different areas of cross section, but otherwise identical with each other). That the theorem is indeed applicable to this case follows from the scale-factor linearity of (21) with respect to I. The absence of additive linearity requires that the theorem should be used cautiously, for example, when applied to electron devices in which there are two distinct streams of charge carriers with a net energy transfer between them.

v) Absence of Duality: The current–voltage duality in circuit theory [43] is a consequence of the duality of Kirchhoff's voltage and current laws, and is a topological result. By contrast, the thermal noise theorem is a physical result, not merely a topological consequence. (Thus Nyquist's theorem cannot be derived from Kirchhoff's laws alone; its derivation requires additional results from thermal physics, such as the Boltzmann distribution for the canonical ensemble). From a thermodynamic viewpoint, the current and voltage are two very different variables. For example, in a system at thermal equilibrium, electrical currents must be absent, but voltages (i.e., electrostatic potential differences) need not be; an unbiased semiconductor p-n junction is a case in point. Therefore, there is no reason for expecting the circuit theoretical principle of duality to apply to the expressions for thermal noise. A comparison of (19) and (21) shows that duality indeed does not apply.

IV. Applications of the Theorem

The purpose of this section is to illustrate the use of the theorem stated in the previous section. The theorem will be applied to some nonlinear electron devices for which the noise results have already been determined from theoretical models based upon first principles and from experimental measurements. A comparison of known results with the results of the theorem will help us recognize the range of validity of the assumptions on which the theorem is based.

A. p-n and Schottky-Barrier Junctions

At sufficiently low frequencies and low injection levels, the ideal semiconductor junction may be treated as a purely resistive device, described by the instantaneous current–voltage relationship [44]

$$I = I_0 \left[\exp\left(\frac{qV}{kT}\right) - 1\right]. \tag{25}$$

The conditions under which this result holds are well documented. From (21), the mean-square short-circuit noise current at the diode terminals at temperature T is

$$\langle i_n^2 \rangle = 2qI_0 B \left[\exp\left(\frac{qV}{kT}\right) + 1\right]. \tag{26a}$$

Equation (26a) expresses the result as a function of V; with the help of (25) it can be rewritten in terms of I as

$$\langle i_n^2 \rangle = 2q(I + I_0)B + 2qI_0 B. \tag{26b}$$

This last equation is also the result derived by applying the shot theorem to the two streams of carriers of magnitude $I + I_0$ and I_0 crossing the junction in opposite directions, and is experimentally well established [45].

The two values of $\langle i_n^2 \rangle$, as calculated by the thermal noise theorem and by the shot-noise model, agree for arbitrary bias current I. (Of course, the diode characteristics are given by

[3] Notice that the sign of $\beta I_T/g$ is independent of the choice of reference directions for v_T and i_T.

(25) only under the so-called "low-injection" assumption.) This situation is fortuitous and should not be expected in every case, because the theorem contained in (21) has been derived under the assumption of quasi-equilibrium. In general, the theorem will hold only in the limit of small bias currents.

B. Tunnel Diodes

At low frequencies, the tunnel diode can also be represented by means of a nonlinear resistor with instantaneous I–V relationship. Most physical models of the tunnel diode yield a vary complex I–V relationship despite numerous simplifying assumptions. For the present purposes, a sufficiently accurate and experimentally verified empirical I–V relationship is given by the following approximation:

$$\begin{aligned} I &= A V e^{-\alpha V} + B[e^{\beta V} - 1] \\ &\equiv I_{\text{tun}} + I_{\text{jn}} \end{aligned} \tag{27}$$

where A, B, α, and β may be treated as empirical coefficients [46], determined by curve-fitting of this equation to the experimental data. The first term I_{tun} on the right-hand side represents the tunneling component of the total current, while the second term I_{jn} is the usual junction current, similar to that in (25). The tunnel diode may, therefore, be viewed as a parallel combination of a usual p-n junction and a tunnel junction carrying only the tunneling component of current I_{tun}. For small biases, I_{tun} exceeds I_{jn}, while for large biases I_{jn} is much larger than I_{tun}. In the tunneling regime of the diode characteristic, the noise resulting from the tunneling part of the current can be found from the theorem discussed. For a purely tunneling junction, described by

$$I_{\text{tun}} = A V \exp(-\alpha V) \tag{28}$$

the mean-square short-circuit noise current at the terminals is found from (21) to be

$$\begin{aligned} \langle i_n^2 \rangle &= \frac{2kTBI_{\text{tun}}}{V} [(1 - \alpha V) + (1 - \alpha V)^{-1}] \\ &\approx \frac{2kTBI_{\text{tun}}}{V} [2 + a^2 V^2] \end{aligned} \tag{29}$$

in the limit of low bias values.

The short-circuit mean-square noise current $\langle i_n^2 \rangle$ in the tunneling component I_{tun} of diode current can be assumed to be pure shot noise and is known to be

$$\langle i_n^2 \rangle = 2qBI_{\text{tun}} \coth\left(\frac{qV}{2kT}\right) \tag{30a}$$

both on theoretical grounds [47] and from experimental measurements [48]. In the limit of low bias, this can be approximated as

$$\langle i_n^2 \rangle \approx 2qBI_{\text{tun}} \left[\left(\frac{2kT}{qV}\right) + \frac{1}{3}\left(\frac{qV}{2kT}\right)\right] \tag{30b}$$

where the two lowest order terms in the power series expansion of coth function are retained. A direct comparison of (29) and (30b) shows that they are identical, provided

$$\begin{aligned} \alpha &= \frac{1}{\sqrt{6}} \cdot \frac{q}{kT} \\ &= 15.7\ (\text{V})^{-1} \text{ at room temperature.} \end{aligned} \tag{31}$$

This compares very well with the value 16.8 per volt for α determined by fitting (27) to the experimentally measured I–V characteristics of tunnel diodes [46]. The results are valid only for small V. As αV approaches unity, (28) shows that the first derivative vanishes (this is the onset of negative differential resistance region), and (21) cannot be used.

C. Applicability of the Theorem

Two related issues will be discussed in this section: i) Why does the nonlinear thermal noise theorem lead to the noise in semiconductor p-n junction and tunnel diodes, which has traditionally been called shot noise in the engineering literature, and ii) for what class of devices can the nonlinear thermal noise theorem be expected to be useful.

The first question requires the definition of the terms "thermal noise" and "shot noise." Thermal noise has been formally defined [49] as the "random noise associated with the thermodynamic interchange of energy necessary to maintain thermal equilibrium between a system and its surroundings." Therefore, even if a part of the device is far from thermal equilibrium, the device noise can be called thermal noise if it is determined primarily by that part of the device which is at or near equilibrium. The term shot noise was initially used to describe the noise in an electron tube arising from the random process of thermionic emission. Gradually it became established in the literature of probability theory and statistics as the name of any random process which can be expressed as

$$s(t) = \sum_n h_n(t - t_n) \tag{32}$$

where t_n are random points in time, having some specified distribution, and $h_n(t)$ are real functions of time, having some specified dependence on t_n.

For the present purposes, narrower definitions of thermal and shot noise are required which apply to the fluctuations in the currents and voltages at the terminals of a free-electron device. In order to develop definitions which apply to a wide variety of such devices, it is necessary to focus on the two essential features of all such devices. First, the terminal currents in the device are caused by charge carriers which are drawn from a source or pool of carriers in some region of the device. Second, the voltages at the terminals of the device produce fields with which the above carriers interact in some interaction region of the device, thereby giving rise to the particular terminal behavior characteristic of that device. The fluctuations in the terminal voltages and currents of the device arise from certain stochastic properties of the carriers, such as velocity, which are randomly distributed among the carriers. In the most general case, these random distributions will be determined by the source region as well as the interaction region of the device.

The terms "shot noise" and "thermal noise" can now be defined in terms of the above device description [10]. The noise present at the terminals of the device can be called shot noise when the randomness in the stochastic properties of the carriers is determined by the source region of the device, and not by the interaction region. The terminal currents can then be expressed in the form of (32), with the distribution of t_n governed by the source region, and the shape of $h_n(t)$ governed by the interaction region. The noise at the terminals of the device can be called thermal noise provided the random properties of the carriers are established in a region where the carriers are in an approximate thermal equilibrium. It is clear from these definitions that the terms thermal noise and shot noise are not mutually exclusive; when the random distribution of the stochastic properties among carriers in the interaction region is determined by the source region (the two regions may even be coincident), and the carriers remain near thermal equilibrium throughout, either term may be used.

Finally, it is well to remember that the above definitions are based only on an intuitively appealing description of the device. To quote van Kampen [31], "the distinction often made between shot noise and thermal noise is vague and of limited applicability."

The limitations to the applicability and utility of the nonlinear thermal noise theorem will now be pointed out. In thermal equilibrium, the only noise that a classical system can exhibit is thermal noise. Therefore, for a device in thermal equilibrium, the Nyquist theorem determines the total noise of the device. By contrast, when a device is maintained in a nonequilibrium state through biasing, it may (and usually does) have other sources of noise. Therefore, it is convenient to treat the noise in a biased device as being composed of a thermal part and a nonthermal part; the thermal part is that part which arises due to thermal fluctuations in a near-equilibrium region of the device. The nonlinear thermal noise theorem accounts for only this part, and the device noise calculated with its help is only a part of the total device noise; it may be large compared to the nonthermal noise, or it may be entirely masked by the nonthermal noise, depending on the device and its biasing.

In the light of the above discussion, the principal limitation in the use of the nonlinear thermal noise theorem for calculating the noise in a particular device can now be understood: the theorem calculates only thermal noise, and only in near-equilibrium devices. Thus the theorem cannot be used to determine the noise in a semiconductor p-n junction diode reverse biased well into avalanche, because this noise is predominantly due to the random, nonequilibrium (i.e., nonthermal) avalanche process; the thermal fluctuations in carrier velocities outside the avalanche region have little effect on the total device noise [50]. By contrast, the noise in a space-charge limited (SCL) diode, which is ultimately thermal in origin [51], [52], cannot be calculated by applying the nonlinear thermal noise theorem directly to the terminal current–voltage characteristic of the device, because the device must be biased far away from equilibrium in order to reach the SCL regime, while the theorem applies only to near-equilibrium systems.

V. Equivalent Circuit for Nonlinear Resistors with Thermal Noise

For a linear system, the knowledge of the mean-square noise voltage under open-circuit conditions (or with some other known termination) is sufficient for calculating the noise voltage or current for all other terminating conditions, and for the construction of an equivalent circuit, through the use of Thevenin or Norton theorems. As these theorems do not apply to nonlinear networks, a noise model for nonlinear resistors cannot be deduced from the results of the previous section. Instead, the equivalent-circuit model proposed here is found inductively, by first making some reasonable postulates without *a priori* justification, and then examining their consequences for validity. The model thus found is approximate, and has a limited range of utility. Furthermore, the model is not unique, and simplicity is an important criterion for selecting it from among the alternatives.

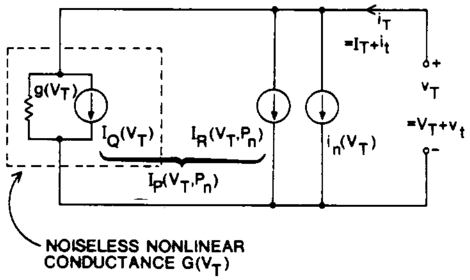

Fig. 4. Proposed noise equivalent circuit model for a nonlinear resistor, employing a noise current source, a noiseless conductance, and nonrandom, dissipation-controlled, current sources.

A. Proposed Noise Equivalent Circuit

The noise equivalent circuit for a nonlinear resistor must reduce to that for linear resistors, shown in Fig. 1, in the limit of zero nonlinearity. It can, therefore, be found by a suitable extension of the linear circuit of Fig. 1. The extension proposed here is shown in Fig. 4, which contains two more circuit elements than the linear circuit. The four circuit elements appearing in the proposed model of Fig. 4 are defined as follows.

Let the voltage and current at the terminals of the nonlinear resistor (and the model) be v_T and i_T, respectively. Their ensemble averages

$$V_T \equiv \langle v_T \rangle \quad \text{and} \quad I_T \equiv \langle i_T \rangle \tag{33}$$

are then the "bias voltage" and "bias current," while the noise voltage and current, defined by

$$v_t \equiv v_T - V_T \quad \text{and} \quad i_t \equiv i_T - I_T \tag{34}$$

have zero ensemble averages. The bias current and voltage are functionally related through the nonlinear terminal characteristic of the nonlinear resistor.

As an aid in understanding how the noise model is constructed, consider first a hypothetical case in which there is no net flow of noise power. The expressions for bias current and voltage should reduce to some simpler form in this special case, and will be denoted by I_C and V_C, respectively. The nonlinear functional relationship between I_C and V_C is sometimes called the "cold device characteristic" in the jargon of microwave detector engineering, to distinguish it from the I_T–V_T relationship which applies in the presence of noise power transport. This cold characteristic can be expressed as a Taylor series around the origin

$$I_C = \left.\frac{dI_C}{dV_C}\right|_0 V_C + \frac{1}{2}\left.\frac{d^2 I_C}{dV_C^2}\right|_0 V_C^2 + \cdots . \tag{35}$$

As in Section III-B, and in writing the nonlinear thermal noise theorem as (23), the interest here will be confined to small values of V_C, or to the nonlinear resistors with the lowest (second) order nonlinearity, so that the terminal characteristic can be written as a quadratic. This quadratic may be expressed in terms of two parameters describing the nonlinear resistor: the incremental conductance $g(V_C)$, defined as dI_C/dV_C, and the current sensitivity $\beta(V_C)$, defined by (22). The terminal characteristic can, therefore, be written as

$$I_C = g(0)\, V_C [1 + \beta(0)\, V_C] \tag{36a}$$

or, alternatively, as

$$I_C = g(V_C)\, V_C [1 - \beta(V_C)\, V_C] \tag{36b}$$

where the higher powers of the dimensionless quantity (βV_C) have been neglected.

Although the temperature is not explicitly included as an argument, the parameters $g(V_C)$ and $\beta(V_C)$ are obviously dependent on the temperature of the nonlinear resistor. The nature of this dependence is not of immediate concern here, because the model developed here is for use at a single temperature; the parameter values are those applicable at that temperature.

The proposed model of Fig. 4 consists of the following circuit elements, each controlled by the value of the terminal voltage or current as follows:

i) The noiseless nonlinear conductance $G(V_T)$ represents a resistive circuit element having the nonlinear terminal characteristic of (35). As (36b) suggests, this circuit element can be thought of as being composed of two separate elements connected in parallel. The first is a noiseless linear, but controlled, conductance $g(V_T)$, equal to the incremental conductance of the nonlinear resistor at its quiescent operating point (defined by the bias)

$$g(V_T) \equiv \left. \frac{dI_T}{dV_T} \right|_{V_T}. \tag{37}$$

The second is a noiseless controlled current source $I_Q(V_T)$, intended to account for the remainder of the quiescent current through $G(V_T)$, and is thus given by

$$\begin{aligned} I_Q &\equiv I_T(V_T) - V_T g(V_T) \\ &\approx -\beta(V_T) I_T V_T \end{aligned} \tag{38}$$

with the help of (36b), where higher order terms in βV_T have been neglected. The source I_Q is, therefore, a controlled current source, controlled by the power dissipation $I_T V_T$ due to the phenomenological parts of the terminal current and voltage.

ii) The deterministic current source I_R represents the rectification of noise power by the nonlinear resistor. If the average power dissipation due to the noise parts of the terminal voltage and current is denoted by P_n

$$P_n \equiv \langle i_t v_t \rangle \tag{39}$$

the current source I_R is defined by the equation

$$I_R \equiv \beta P_n \tag{40}$$

in accordance with the detector current sensitivity interpretation of β. Clearly, I_R is also a controlled current source, controlled by the noise power dissipation.

iii) The random current source i_n is the only noisy element in the model, and therefore accounts for the thermal noise generated in the nonlinear resistor. Its mean-square value depends upon the bias, and is given by (23). If necessary, it can be written in terms of V_T, after substitution from (36), as

$$\langle i_n^2 \rangle \approx 4kTB\, g(0)[1 + \beta(0)\, V_T] \tag{41a}$$

or alternatively, as

$$\langle i_n^2 \rangle \approx 4kTB\, g(V_T)[1 - \beta(V_T)\, V_T] \tag{41b}$$

where again the higher powers of (βV_T) have been neglected.[4]

For the sake of brevity in subsequent work, the two dissipation-controlled current sources I_Q and I_R would be combined into a single phenomenological current source I_P, defined as

[4] Note that the nonlinear thermal noise theorem written in the form (41a) with the help of approximations, is identical with (13a).

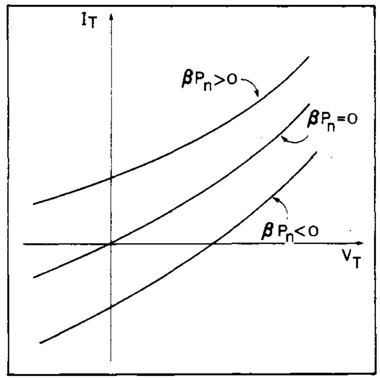

Fig. 5. Terminal current-voltage characteristic for a nonlinear resistor in the presence of noise power transport.

$$\begin{aligned} I_P &\equiv I_Q + I_R \\ &= \beta(V_T)\, [P_n - I_T V_T]. \end{aligned} \tag{42}$$

In order to develop some understanding of the above equations, and thus of the model, consider the determination of the terminal current-voltage $(I_T\text{-}V_T)$ characteristic for the nonlinear resistor as predicted by the model. The characteristic given by (35) or (36) applies only in the absence of noise power transport, i.e., for $P_n = 0$. The effect of P_n on the characteristic follows from the model and (42)

$$I_T = gV_T + I_P = gV_T + \beta(P_n - I_T V_T). \tag{43}$$

This can be solved for I_T to yield

$$I_T = (gV_T + \beta P_n)/(1 + \beta V_T) \approx (gV_T + \beta P_n)(1 - \beta V_T) \tag{44}$$

after neglecting the higher powers of βV_T. In the absence of noise power transport, this reduces to the cold characteristic (36b) as expected, and is shown schematically in Fig. 5. With a net noise power delivered to (or by) the nonlinear resistor, the characteristic shifts as indicated in Fig. 5. Such a shift of characteristic due to detection is an experimentally observed fact for diodes [38, fig. 2.9]).

B. Condition of Validity

The proposed model for nonlinear resistors is clearly approximate, having a limited range of applicability. It applies only for "small" terminal voltages, such that the second-order terms in the dimensionless quantity βv_T can be neglected. The condition of validity of the model can, therefore, be explicitly written as

$$(\beta V_T)^2 \ll 1 \tag{45a}$$

and

$$\beta^2 \langle v_t^2 \rangle \ll 1. \tag{45b}$$

This condition has already been invoked in (36b), (38), (41b), and (44). The need and significance of this condition is best understood by examining its several consequences.

The first consequence of the condition is that the model is valid only if the transported noise power P_n is "small." This condition can be expressed quantitatively, by recognizing that the noise source i_n is uncorrelated with all external noise sources which may be connected at the terminals of the nonlinear resistor. As a result, the average noise power delivered by this source must be positive, or

$$\langle i_n v_t \rangle < 0. \tag{46}$$

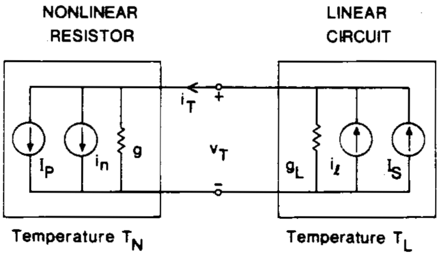

Fig. 6. A nonlinear resistor at temperature T_N, connected across an isothermal, linear, two-terminal resistive circuit at temperature T_L.

Replacing i_n by $i_t - gv_t$, and using (39) and (45b) leads to the result

$$P_n \ll g/\beta^2. \tag{47}$$

A second consequence of the condition (45) is that the terminal current–voltage characteristic of the proposed model does not display an incremental negative resistance within the range of validity of the model. In the absence of noise power transport, the terminal characteristic of the model is given by (36b), and describes a passive element if g is positive. The incremental conductance $\partial I/\partial V$ for a fixed P_n, found from (44), is also positive when (47) holds.

A third consequence of the condition of validity is that the incremental conductance g can be defined in (37) without the need to account for the P_n dependence of current–voltage characteristic, as in (44). It follows from (44) that

$$\left(\frac{\partial I_T}{\partial V_T}\right)_{P_n} = \left(\frac{dI_T}{dV_T}\right)\Bigg|_{P_n=0} + \text{terms with second and higher powers of } \beta V_T. \tag{48}$$

Expressed alternatively, the incremental conductance g can be treated as a function of either I_T alone, or V_T alone, with an error which is only of second order in βV_T.

C. Thermal Noise in Circuits Containing Nonlinear Resistors

The equivalent circuit of a noisy nonlinear resistor is useful for calculating noise voltages and currents in circuits containing such a resistor. An illustrative example of this calculation is presented here; its purpose is to illustrate the use of the equivalent circuit, to explore the validity of its consequences, and to obtain some results needed in the next section.

Consider a circuit in which the nonlinear resistor is connected across an arbitrary but isothermal, near-equilibrium, linear, resistive network. (The isothermal requirement is introduced so that a single noise temperature can be ascribed to the linear circuit.) This circuit is represented in Fig. 6, where the nonlinear resistor has been replaced by its noise equivalent circuit, while the remainder of the linear circuit has been replaced by its Norton equivalent. The nonlinear resistor is characterized by three parameters: the incremental conductance at the operating point, g (in mhos); the detection sensitivity β (in volts^{-1}); and the temperature T_N (in Kelvins). The linear circuit, shown here in Norton equivalent form, is also characterized by three parameters: the short-circuit signal current I_S (in amperes), the linear Norton conductance g_L (in mhos), and the noise temperature of the isothermal linear circuit T_L (in Kelvins). In order to keep the following analysis tractable, two further restrictions are imposed on the circuit under consideration. First, it is assumed that the linear circuit and the nonlinear resistor are connected electrically but not thermally. Such a condition can indeed be approached in practice by a suitable coupling. This assumption allows the heat flow and the thermoelectric effects to be ignored; to include those processes would require a considerably more detailed model than the purely electrical description at the terminals, as used here. Second, it is assumed that the entire circuit is close to equilibrium. This condition holds when the power flow between the resistors is small. This in turn implies that *each* of the two excitations which cause the power flow, namely the temperature differential $(T_L - T_N)$ and the current source I_S, is small in magnitude. This assumption permits the thermal noise current in the linear resistor to be determined by the Nyquist relationship (2).

The circuit described above will now be analyzed to determine the values of the terminal voltage and current, and the noise power transported between the linear circuit and the nonlinear resistor. Kirchhoff's current law yields

$$i_T = gv_T + i_n + I_P = -g_L v_T + i_l + I_S. \tag{49}$$

Taking the ensemble average on each side of the equation, and using (42), gives

$$I_T = gV_T + \beta(P_n - V_T I_T) = -g_L V_T + I_S \tag{50}$$

and subtraction of (50) from (49) gives

$$i_t = gv_t + i_n = -g_L v_t + i_l. \tag{51}$$

The terminal noise voltage and current are found from (51) as

$$v_t = (i_l - i_n)/(g + g_L) \quad \text{and} \quad i_t = (g_L i_n + g i_l)/(g + g_L) \tag{52}$$

so that

$$P_n \equiv \langle v_t i_t \rangle = [g\langle i_l^2 \rangle - g_L \langle i_n^2 \rangle]/(g + g_L)^2 \tag{53}$$

where i_l and i_n have been assumed uncorrelated. From (2) and (23)

$$\langle i_l^2 \rangle = 4kBT_L g_L \tag{54}$$

and

$$\langle i_n^2 \rangle = 4kBT_N[g - i_T\beta] \tag{55}$$

where B is the effective bandwidth over which noise power transport occurs. Upon substitution in (53), the net noise power delivered to the nonlinear resistor is given by

$$\begin{aligned} P_n &= 4kBg_L[g(T_L - T_N) + \beta I_T T_N]/(g + g_L)^2 \\ &= [g(T_L - T_N) + \beta T_N(I_S - g_L V_T)]/\beta^2\theta \end{aligned} \tag{56}$$

where θ is a quantity with the units of temperature, defined as

$$\theta \equiv (g + g_L)^2/(4kBg_L\beta^2). \tag{57}$$

Equations (50) and (56) provide three relationships among the three unknowns I_T, V_T, and P_n, and can be solved for them. Thus V_T is the solution of the quadratic equation[5]

[5] Although the model validity is limited by the assumption that higher order terms in βV_T are negligible, there is no inconsistency if the quadratic term is retained during the circuit analysis stage. The purpose of this retention will become apparent in (72). None of the intervening discussions are materially affected by this retention, and the quadratic terms in (58) and (64) may be ignored, if desired, to yield results correct to the first order in βV_T.

$$g_L(\beta V_T)^2 + (g + g_L - \beta I_S - g_L T_N/\theta)(\beta V_T) + [g(T_L - T_N)/\theta - \beta I_S(1 - T_N/\theta)] = 0. \quad (58)$$

Of the two roots of the quadratic equation (58), only one lies in the region of validity of the model, defined by the inequality (45a), for small values of the two excitations: the current source I_S, and the temperature differential $(T_L - T_N)$. For example, if $T_N < \theta$, the acceptable root is given by

$$\beta V_T = -\frac{1}{2}\left(1 + \frac{g}{g_L} - \beta\frac{I_S}{g_L} - \frac{T_N}{\theta}\right) + \left[\frac{1}{4}\left(1 + \frac{g}{g_L} - \beta\frac{I_S}{g_L} - \frac{T_N}{\theta}\right)^2 - \frac{g}{g_L}\left(\frac{T_L - T_N}{\theta}\right) + \beta\frac{I_S}{g_L}\left(1 - \frac{T_N}{\theta}\right)\right]^{1/2}. \quad (59)$$

Substitution of this root in (50) and (56) determines the values of I_T and P_n, respectively. The solutions found are obviously implicit, because g, β, and θ are functions of V_T.

The values of V_T, I_T, and P_n determined above can now be used to show that the model does describe the expected properties of a nonlinear resistor and does not violate any thermodynamic principles. To demonstrate this, consider the following special cases:

i) If the nonlinear resistor is biased by an ideal current source I_S, both g_L and i_l vanish. From (52), the terminal noise voltage and current are

$$i_t = 0 \qquad v_t = -i_n/g(V_T). \quad (60)$$

The bias current I_T is equal to the source current I_S, and the bias voltage is related to it by the relationship

$$I_T = g(V_T)\, V_T/(1 + \beta V_T) \approx g(V_T)\, V_T(1 - \beta V_T) \quad (61)$$

which is identical with (36b), the terminal characteristic of the nonlinear resistor.

ii) If the nonlinear resistor is biased by an ideal voltage source, the noise quantities are

$$i_t = i_n \qquad v_t = P_n = 0 \quad (62)$$

while the bias current and voltage are still related to each other by (61).

iii) If the nonlinear resistor is connected across a linear one at the same temperature, both I_S and $(T_L - T_N)$ vanish. Then

$$V_T = 0 \qquad I_T = 0 \qquad \text{and} \qquad P_n = 0 \quad (63)$$

from (59), (50), and (56), respectively.

iv) If the nonlinear resistor is connected across a linear one at a different temperature, only I_S vanishes. Under this condition, elimination of I_T between the two equations in (50) shows that the terminal voltage V_T is the solution of the quadratic equation

$$g_L(\beta V_T)^2 + (g + g_L)\,\beta V_T + \beta^2 P_n = 0. \quad (64)$$

For reasons explained in connection with (59), only one of the two solutions of the quadratic is acceptable, and is given by

$$\beta V_T = -\frac{1}{2}\left(1 + \frac{g}{g_L}\right) + \left[\frac{1}{4}\left(1 + \frac{g}{g_L}\right)^2 - \frac{\beta^2 P_n}{g_L}\right]^{1/2}. \quad (65a)$$

This shows that

$$\beta V_T \gtrless 0 \quad \text{according as} \quad P_n \lessgtr 0 \quad (65b)$$

or that βV_T and P_n have opposite signs. The sign of βV_T, in turn, depends upon the temperature differential. For $I_S = 0$ and $T_N < \theta$, it is clear from (59) that βV_T and $(T_L - T_N)$ also have opposite signs. These results agree with the observed [37] and expected behavior of nonlinear resistors. For example, if a diode is connected across a resistor that is hotter than the diode, P_n and βV_T will be positive and negative, respectively; i.e., the noise power will flow from the hotter resistor to the colder diode, and the diode will develop a dc bias voltage in the reverse direction and carry a current in the forward (lower resistance) direction. This result can also be deduced by drawing a load line in Fig. 5.

v) If the two temperatures T_N and T_L are unequal, the nonlinear resistor may function as a heat engine. If, in addition, the source I_S is also present, the nonlinear resistor can serve as a refrigerator. These cases are examined in detail in Sections VI-A and -B.

vi) At the terminals of the nonlinear resistor, there exists a simultaneous flow of two separate conserved quantities: the extensive variable X (discussed in Section III-C) and the thermal energy. Their flow rates are the terminal voltage V_T and the noise power P_n, respectively. In the limiting case of very small values, these two fluxes can be expressed as linear functions of two forces, which can be identified by first determining $\dot{S}$, the time rate of increase of entropy in the circuit. Entropy increases in part because the signal power $I_T V_T$ is dissipated in the nonlinear resistor at temperature T_N, and in part because the noise power P_n is transported across a small temperature differential $(T_L - T_N)$, so that

$$\dot{S} \approx V_T[I_T/T_N] + P_n[(T_L - T_N)/T_N^2]. \quad (66a)$$

The quantities in $[\cdot]$ in (66a) are intensive variables, and are, therefore, the thermodynamic forces conjugate to the fluxes V_T and P_n, respectively. From (50) and (56), the fluxes can be linearized and expressed as a function of the forces

$$V_T = [I_T/T_N]\,[(\theta + T_N)/\theta T_N g] - [(T_L - T_N)/T_N^2]\,[T_N^2/\beta\theta] \quad (66b)$$

$$P_n = -[I_T/T_N]\,[T_N^2/\beta\theta] + [(T_L - T_N)/T_N^2]\,[T_N^2 g/\beta^2\theta]. \quad (66c)$$

The coefficients of the forces in the cross-terms are equal, demonstrating that the Onsager reciprocity relations are obeyed.

vii) If the linear conductance g_L is frequency dependent, an analysis similar to that presented in this section can be carried out, with the total noise power P_n replaced by the integral, over all frequencies, of the noise power per unit bandwidth delivered to the nonlinear resistor.

It is apparent from the above that the model of the nonlinear resistor behaves in the manner expected of a nonlinear resistor embedded in a linear circuit. Similarly consistent results are also found for a circuit consisting of two nonlinear resistors and a bias source. Finally, the model has another desirable feature: it does not make an artificial distinction between small signals and noise. Consider, for example, a nonlinear resistor connected across a series connection of two ideal voltage sources having a dc voltage V_{dc} and a small-signal voltage v_{ac}. This situation can be viewed in two different ways. If v_{ac} is treated as noise, then $V_T = V_{dc}$, and $v_t = v_{ac}$. On the other hand, if v_{ac} is treated as a signal, then $V_T = V_{dc} + v_{ac}$, and

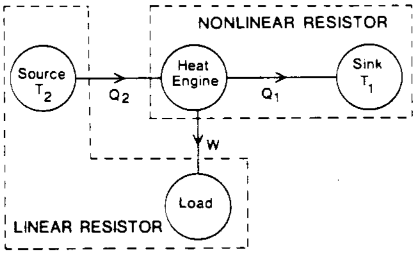

Fig. 7. One possible arrangement for using a nonlinear resistor as a heat engine. $T_2 > T_1$.

$v_t = 0$ with condition (45), either treatment results in the same expressions for the current and power flow in the nonlinear resistor. The choice of the model of Fig. 4 for a nonlinear resistor was based on these positive results.

VI. Applications of the Noise Equivalent Circuit

The purpose of this section is to present two examples to illustrate the use of the proposed noise equivalent circuit for nonlinear resistors.

A. Nonlinear Resistors as Heat Engines

By definition, a heat engine is a mechanism for extracting work by allowing the flow of heat energy from a higher to a lower temperature body. If Q_2 is the heat energy lost by a source at a temperature T_2, Q_1 is that received by a sink at the lower temperature T_1, and W is the work done by the heat engine (see Fig. 7), then the first law of thermodynamics states that

$$Q_2 = Q_1 + W \tag{67}$$

and the second law of thermodynamics in Carnot's theorem form states that

$$\frac{W}{Q_2} \leqslant \frac{T_2 - T_1}{T_2}. \tag{68}$$

A net transport of thermal energy from one body to another occurs provided the two are at different temperatures. This transfer of thermal energy between them takes place via heat if the two bodies are in thermal contact, and via noise power if they are in electrical contact. If there is a net noise power transfer, a nonlinear resistor can deliver dc power by noise rectification and therefore serves as the electrical analog of a heat engine [53].

A number of different arrangements can be envisioned for connecting a nonlinear resistor as a heat engine. Perhaps the simplest is the one in which the dc power generated is delivered to a load consisting of a linear resistor, the load resistor is placed in thermal contact with the heat source, at the higher temperature T_L, and the nonlinear resistor is placed in contact with the heat sink at the lower temperature T_N as indicated in Fig. 7. This arrangement is described by the circuit of Fig. 6, which has already been analyzed in Section V-C, provided I_S is equated to zero.

It is clear that the work done by the heat engine per unit time is

$$W = -I_T V_T = g_L V_T^2 \tag{69}$$

which is necessarily nonnegative for a passive load ($g_L > 0$). The dc voltage V_T developed across the load has already been solved for in Section V-C, and is given by (59) with I_S equated to zero

$$\beta V_T = -\frac{1}{2}\left(1+\frac{g}{g_L}-\frac{T_N}{\theta}\right)+\left[\frac{1}{4}\left(1+\frac{g}{g_L}-\frac{T_N}{\theta}\right)^2 - \frac{g}{g_L}\left(\frac{T_L-T_N}{\theta}\right)\right]^{1/2}$$
$$= -\frac{g}{g_L}\cdot\frac{(T_L-T_N)/\theta}{1+(g/g_L)-(T_N/\theta)} - \left(\frac{g}{g_L}\right)^2 \cdot \left(1+\frac{g}{g_L}-\frac{T_N}{\theta}\right)^{-1}\left(\frac{T_L-T_N}{\theta}\right)^2+\cdots \tag{70}$$

where the series converges for small temperature differentials $(T_L - T_N)$. The net noise power delivered by the linear resistor, given by (56), is

$$Q_2 = P_n = \frac{g}{\beta^2}\left(\frac{T_L - T_N}{\theta}\right) - \frac{g_L T_N}{\beta^2 \theta}\cdot \beta V_T. \tag{71}$$

The efficiency of the heat engine is thus given by

$$\eta \equiv \frac{W}{Q_2} = \frac{g_L (\beta V_T)^2 \theta}{g(T_L - T_N) - (\beta V_T) g_L T_N}. \tag{72}$$

Notice that although the numerator of the expression for η in (72) is a second-order quantity in (βV_T), the entire expression is not, and therefore the efficiency η can be calculated correctly to first-order terms. To emphasize this, (72) can be rewritten as follows, with the aid of (58) and after the substitution of $I_S = 0$:

$$\eta = \frac{-\beta V_T}{1 + (g/g_L) + \beta V_T}. \tag{73}$$

Since the heat engine is constructed so as to make $P_n > 0$, it follows from (65b) that βV_T is negative, and η in (73) is indeed positive.

It will now be demonstrated that, within the range of validity of the model, the efficiency η calculated from (73) does not exceed the Carnot efficiency, in accordance with the second law of thermodynamics. From (73), η depends on βV_T, which in turn depends on $(T_L - T_N)$ as in (70). While the efficiency, expressed as a function of βV_T in (73), appears to be unbounded, the validity of the model is limited by (45a) to small values of $(-\beta V_T)$, and therefore to small $(T_L - T_N)$. The assumption that $T_N < \theta$ is implicit in the choice of the sign in (59). If, in addition, T_L is also restricted such that $(T_L + T_N) < 2\theta$, it follows that

$$T_L - \theta < \theta - T_N < \sqrt{\theta(\theta - T_N)}. \tag{74}$$

Further, since the geometric mean of two positive numbers is bounded from above by the arithmetic mean

$$2(T_L - \theta) < \frac{g}{g_L}\theta + \frac{g_L}{g}(\theta - T_N). \tag{75}$$

After addition of $2\theta - T_N$, and then division by $(T_L - T_N)/(2T_L - T_N)$, on both sides, this inequality can be transformed into

$$\frac{T_L - T_N}{\theta + g_L(\theta - T_N)/g} < \left(1 + \frac{g}{g_L}\right)\left(\frac{T_L - T_N}{2T_L - T_N}\right). \tag{76}$$

The left-hand side of this inequality is recognized from (70) to

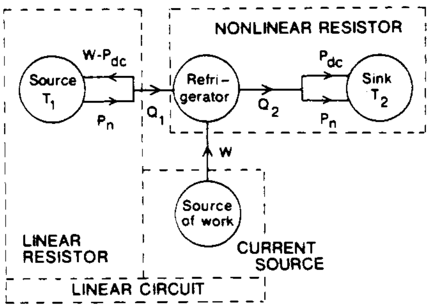

Fig. 8. One possible arrangement for using a nonlinear resistor as a refrigerator. $T_2 > T_1$.

be the first-order approximation to $-\beta V_T$, for small $(T_L - T_N)$. As $-\beta V_T$ is a positive quantity, (76) can be written as

$$\left(1 + \frac{g}{g_L}\right)\left(-\frac{1}{\beta V_T}\right) > 1 + \frac{T_L}{T_L - T_N}. \tag{77}$$

From (73) it follows that η is less than the Carnot efficiency

$$\eta < \frac{T_L - T_N}{T_L}. \tag{78}$$

An alternative arrangement for the heat engine would be one in which the nonlinear resistor is placed in contact with the heat source at the higher temperature and the linear load resistor is placed in contact with the sink at the lower temperature. In this case, both the dc and the noise power flow from the nonlinear resistor to the linear resistor. The thermal energy delivered to the sink is

$$Q_1 = -P_n \tag{79}$$

and since P_n is negative, it follows from (65b) that βV_T is positive. The efficiency of this heat engine may also be similarly calculated.

B. Nonlinear Resistors as Refrigerators

By definition, a refrigerator is a heat engine operated in reverse, wherein work is done to make thermal energy flow "uphill," from a body at lower temperature to one at a higher temperature. If the thermal energy extracted from a body at temperature T_1 is Q_1, and an amount Q_2 is delivered to another body at a higher temperature T_2, with the aid of work W performed on the refrigerator (see Fig. 8), then the first law of thermodynamics states that

$$Q_2 = Q_1 + W \tag{80}$$

and the second law of thermodynamics (in Carnot's theorem form) states that

$$\frac{Q_1}{W} \leqslant \frac{T_1}{T_2 - T_1}. \tag{81}$$

Once again, a nonlinear resistor can be used as a refrigerator, since the flow of thermal energy can take place electrically via noise.

Once again, several different refrigerator configurations are possible, only one of which is analyzed in detail here. In this arrangement, the nonlinear resistor is placed in contact with the heat source (the lower temperature reservoir from which heat is to be extracted) as indicated in Fig. 8. The work required is provided by an ideal dc current source, of magnitude I_{dc}, connected in parallel with two resistors. Obviously, the situation is still described by the circuit model of Fig. 6, but with $T_L < T_N$.

It is clear that thermal energy will be transported (as noise power) from the lower temperature resistor to the higher temperature one provided the noise temperature of the nolinear resistor falls below that of the linear resistor, i.e., from (24)

$$T_N(1 - \beta I_T/g) < T_L < T_N. \tag{82}$$

If T_L is sufficiently close to T_N, this condition can always be satisfied by a small bias current for which the model is applicable. However, this condition is not sufficient for ensuring that the linear resistor cools, because the source I_S causes a dissipation in the resistor. The source delivers the work W per unit time, given by

$$W = I_S V_T \tag{83}$$

of which a part is dissipated in the nonlinear resistor

$$P_{dc} = I_T V_T \tag{84}$$

and the remainder in the linear resistor

$$W - P_{dc} = g_L V_T^2. \tag{85}$$

As a result, the linear resistor is cooled if and only if the noise power transported away from it exceeds the dissipation in it

$$P_n > W - P_{dc}. \tag{86}$$

The purpose of the remainder of this section is to show that this condition can indeed be satisfied, and thus to find out the minimum dc bias needed to bring about the refrigeration.

With substitutions from (56) and (85), the inequality (86) can be written as

$$(\beta V_T)^2 + \frac{T_N}{\theta}(\beta V_T) + \left(\frac{g}{g_L}\frac{(T_N - T_L)}{\theta} - \frac{\beta I_S}{g_L}\cdot\frac{T_N}{\theta}\right) < 0 \tag{87}$$

where (βV_T) itself is a function of the current source I_S and the temperature differential $(T_N - T_L)$, as given by (58). Elimination of (βV_T) between (58) and (87) shows that the condition for cooling can be met only for small temperature differentials, which meet the approximate condition

$$T_N - T_L < T_N \cdot \frac{\beta I_S}{g} \cdot \frac{g - \beta I_S}{g + g_L - \beta I_S}. \tag{88}$$

The efficiency of the refrigerator can also be calculated, and is found to satisfy the requirement (81) of the second law of thermodynamics.

VII. Conclusions

The highlights of a half century of engineering literature on thermal noise have been surveyed in Section II. The author's own work on thermal noise in nonlinear resistor has been summarized in Section III, and illustrated in Section IV by examples. The equivalent circuit model for a noisy nonlinear resistor, developed and illustrated in Sections V and VI, respectively, is the principal new work reported here. It must be emphasized that the model is approximate, and has a limited range of validity. It has allowed the solution of some interesting problems, and will hopefully find applications in the solution of still other problems.

One of the important questions to ask about the model is whether the results of Section III are necessary, i.e., whether the nonlinearity dependence of the noise source i_n has a significant bearing on the results deduced from the model. As the thermal noise theorem of Section III, embodied in (23), has been used as the nonlinear replacement for the linear Nyquist theorem given in (2), the question can be addressed by replacing (23) everywhere by (2). This results in the elimination of the T_N/θ term in several equations, such as (58), (59), (70), and (71), but does not materially influence the results of Sections V and VI-A. However, the refrigerator action and the inequality (87) now become impossible, leading to the conclusion that the nonlinearity dependence of noise is essential for the discussion of Section VI-B.

The proposed noise equivalent circuit model for nonlinear resistor is consistent with thermodynamic laws. In particular, the model leads to the conclusion that a nonlinear resistor cannot serve as Maxwell's demon, as demonstrated in (63). Earlier demonstrations [54] that a rectifier cannot become Maxwell's demon invoke a "dc component of fluctuations," and seem less appealing than the present one.

References

[1] M. S. Gupta, "Thermal fluctuations in driven nonlinear resistive systems," *Phys. Rev. A*, vol. 18, pp. 2725-2731, Dec. 1978.

[2] ——, "Thermodynamic limits to nonlinearity, II: Purely resistive dissipative systems," *J. Appl. Phys.* (submitted for publication).

[3] ——, *Electrical Noise: Fundamentals and Sources*. New York: IEEE Press, 1977.

[4] H. Nyquist, "Thermal agitation of electric charge in conductors," *Phys. Rev.*, vol. 32, pp. 110-113, July 1928.

[5] J. B. Johnson, "Thermal agitation of electricity in conductors," *Phys. Rev.*, vol. 32, pp. 97-107, July 1928.

[6] N. G. van Kampen, "Fluctuations in nonlinear systems," in *Fluctuation Phenomena in Solids*, R. E. Burgess, Ed. New York: Academic Press, 1965, pp. 139-177.

[7] W. Bernard and H. B. Callen, "Irreversible thermodynamics of nonlinear processes and noise in driven systems," *Rev. Mod. Phys.*, vol. 31, pp. 1017-1044, Oct. 1959.

[8] P. Guttler, "The Nyquist formula—Possibilities of its derivation," *Nachrichtentech.*, vol. 18, pp. 425-432, Nov. 1968.

[9] J.K.E. Tunaley, "Classical microscopic models of Johnson noise," *Amer. J. Phys.*, vol. 43, pp. 446-448, May 1975.

[10] F.N.H. Robinson, "Thermal noise, shot noise, and statistical mechanics," *Int. J. Electron.*, vol. 26, pp. 227-235, Mar. 1969.

[11] J. L. Lawson and G. E. Uhlenbeck, *Threshold Signals* (MIT Radiation Lab. Series, vol. 24). New York: McGraw-Hill, 1950, pp. 69-71.

[12] J. Mathews and M. A. Nicolet, "Current correlation function $\langle j(x, t)\, j(x', t')\rangle$ derived from a model based on Brownian motion," *Amer. J. Phys.*, vol. 44, pp. 448-454, May 1976.

[13] A. E. Siegman, "Thermal noise in microwave systems, Part I. Fundamentals," *Microwave J.*, vol. 4, pp. 81-90, Mar. 1961.

[14] B. M. Oliver, "Thermal and quantum noise," *Proc. IEEE*, vol. 53, pp. 436-454, May 1965.

[15] K. M. van Vliet, "Derivation of the fluctuation-dissipation theorem," *Phys. Rev.*, vol. 109, pp. 1021-1022, Feb. 15, 1958.

[16] R. Lenk, "A simple proof of the classical fluctuation dissipation theorem," *Phys. Lett.*, vol. 25A, pp. 198-199, Aug. 14, 1967.

[17] F. C. Williams, "Thermal fluctuations in complex networks," *J. Inst. Elec. Eng.* (London), vol. 81, pp. 751-760, Dec. 1937.

[18] R. Q. Twiss, "Nyquist's and Thevenin's theorems generalized for nonreciprocal linear networks," *J. Appl. Phys.*, vol. 26, pp. 599-602, May 1955.

[19] J. Weber, "Quantum theory of a damped electrical oscillator and noise," *Phys. Rev.*, vol. 90, pp. 977-982, June 1, 1953.

[20] A. E. Siegman, "Zero-point energy as the source of amplifier noise," *Proc. IRE*, vol. 49, pp. 633-634, Mar. 1961.

[21] N. F. Ramsey, "Thermodynamics and statistical mechanics at negative absolute temperatures," *Phys. Rev.*, vol. 103, no. 1, pp. 20-28, July 1, 1956.

[22] A. E. Siegman, "Thermal noise in microwave systems, Part II. Noise in high-frequency amplifiers, particularly masers," *Microwave J.*, vol. 4, pp. 66-73, Apr. 1961.
G. Convert, "On the extension of Nyquist theorem to systems at negative temperature," *Compt. Rend. Acad. Sci.* (Paris), vol. 261, pp. 58-61, July 5, 1965.

[23] H. A. Haus and R. B. Adler, "Optimum noise performance of linear amplifiers," *Proc. IRE*, vol. 46, pp. 1517-1533, Aug. 1958.

[24] R. V. Pound, "Spontaneous emission and the noise figure of maser amplifiers," *Ann. Phys.* (U.S.), vol. 1, pp. 24-32, Apr. 1957.

[25] D. C. Agouridis, "Thermal noise of transmission media," *IEEE Trans. Instrum. Meas.*, vol. IM-26, pp. 243-245, Sept. 1977.

[26] H. A. Haus, "Thermal noise in dissipative media," *J. Appl. Phys.*, vol. 32, pp. 493-500, Mar. 1961.

[27] M. Vanwormhoud and H. A. Haus, "Thermal noise in linear, lossy, electromagnetic media," *J. Appl. Phys.*, vol. 33, pp. 2572-2577, Aug. 1962.

[28] H. B. Callen and T. A. Welton, "Irreversibility and generalized noise," *Phys. Rev.*, vol. 83, pp. 34-40, July 1, 1951.

[29] P. T. Landsberg and E.A.B. Cole, "The steady state fluctuation dissipation theorem," *Physica*, vol. 37, pp. 309-319, 1967.

[30] M. Lax, "Fluctuations from the nonequilibrium steady state," *Rev. Mod. Phys.*, vol. 32, pp. 25-64, Jan. 1960.

[31] N. G. van Kampen, "Fluctuations in closed and open nonlinear systems," in *Statistical Physics: Proceedings of the International Conference Budapest, Hungary, August 1975*, L. Pal and P. Szepfalusy, Eds. Amsterdam, The Netherlands: North Holland, 1976, pp. 29-48.

[32] J. M. Richardson, "Noise in driven systems," *IRE Trans. Informat. Theory*, vol. IT-1, pp. 62-65, Mar. 1955.

[33] W. Bernard and H. B. Callen, "Irreversible thermodynamics of a nonlinear R-C system," *Phys. Rev.*, vol. 118, pp. 1466-1470, June 15, 1960.

[34] B.D.O. Anderson, "Nonlinear networks and Onsager-Casimir reversibility," *IEEE Trans. Circuits Syst.*, vol. CAS-27, pp. 1051-1058, Nov. 1980.

[35] A. van der Ziel, "Nyquist's theorem for nonlinear resistors," *Solid-State Electron.*, vol. 16, pp. 751-752, June 1973.

[36] A. G. van Nie, "Representation of linear passive noise 1-ports by two correlated noisy sources," *Proc. IEEE*, vol. 60, pp. 751-753, June 1972.

[37] J. B. Gunn, "Thermodynamics of nonlinearity and noise in diodes," *J. Appl. Phys.*, vol. 39, pp. 5357-5361, Nov. 1968.

[38] H. C. Torrey and C. A. Whitmer, *Crystal Rectifiers* (MIT Radiation Lab. Series, vol. 15). New York: McGraw-Hill, 1946.

[39] A. M. Tremblay, B. Patton, P. C. Martin, and P. F. Maldague, "Microscopic calculation of the nonlinear current fluctuations of a metallic resistor: The problem of heating in perturbation theory," *Phys. Rev. A*, vol. 19, pp. 1721-1740, Apr. 1979.

[40] M. A. Mikulinski, "A generalization of the fluctuation-dissipation theorem for the case of excess noise," *Phys. Lett.*, vol. 66A, pp. 440-442, June 26, 1978.

[41] N. Enders and F. Landstorfer, "Uncooled resistive two-pole networks with noise temperatures below 290 K." *Nachrichtentech Z.*, vol. 28, pp. 107-110, Mar. 1975.

[42] P. D. Strum, "Noise in a nonlinear conductance," *Proc. IRE*, vol. 42, p. 1320, Aug. 1954.

[43] L. O. Chua, *Introduction to Nonlinear Network Theory*. New York: McGraw-Hill, 1969, p. 201.

[44] W. Shockley, *Electrons and Holes in Semiconductors*. New York: Van Nostrand, 1950.

[45] A. van der Ziel, "Noise in solid-state devices and lasers," *Proc. IEEE*, vol. 58, pp. 1178-1206, Aug. 1970.

[46] A. Ferendeci and W. H. Ko, "A two-term analytical approximation of tunnel-diode static characteristics," *Proc. IRE*, vol. 50, no. 8, pp. 1852-1853, Aug. 1962.

[47] R. A. Pucel, "The equivalent noise current of Esaki diodes," *Proc. IRE*, vol. 49, no. 6, pp. 1080-1081, June 1961.

[48] B. G. King and G. E. Sharpe, "Measurement of the spot noise of germanium, gallium antimonide, gallium arsenide, and silicon Esaki diodes," *IEEE Trans. Electron Devices*, vol. ED-11, no. 6, pp. 273-285, June 1964.

[49] *IEEE Standard Dictionary of Electrical and Electronics Terms*. New York: IEEE/Wiley-Interscience, 1972, p. 593.

[50] H. K. Gummel and J. L. Blue, "A small-signal theory of avalanche noise in IMPATT diodes," *IEEE Trans. Electron Devices*, vol. ED-14, pp. 569-580, Sept. 1967.

[51] A. van der Ziel, "H. F. thermal noise in space-charge-limited solid-state diodes," *Solid-State Electron.*, vol. 9, pp. 1139-1140, Nov./Dec. 1966.

[52] A. van der Ziel and K. M. van Vliet, "H. F. thermal noise in space-charge-limited solid-state diodes, II," *Solid-State Electron.*, vol. 11, pp. 508-509, Apr. 1968.

[53] J. C. Yater, "Power conversion of energy fluctuations," *Phys. Rev. A*, vol. 10, pp. 1361-1369, Oct. 1974.

[54] L. Brillouin, "Can the rectifier become a thermodynamical demon?" *Phys. Rev.*, Ser. 2, vol. 78, pp. 627-628, June 1, 1950.