Fluctuations from the Nonequilibrium Steady State

MELVIN LAX

Bell Telephone Laboratories, Murray Hill, New Jersey

CONTENTS

HISTORICAL REVIEW

1. Introduction

WORK on fluctuations and noise in physical systems systems can be characterized by the kind of assumptions made. The previous literature[1–40] can be broken into two broad categories: microscopic and macroscopic. The microscopic treatments usually make two basic assumptions:

1. The system is describable by a Hamiltonian, i.e., its future is determined by a complete knowledge of all

[1] J. L. Lawson and G. E. Uhlenbeck editors, *Threshold Signals*, Massachusetts Institute of Technology, Radiation Laboratory Series (McGraw-Hill Book Company, Inc., New York, 1950), Vol. 24.

[2] E. B. Moullin, *Spontaneous Fluctuations of Voltage* (Oxford University Press, London, 1938) presents a useful summary of early experimental results in Nyquist noise, shot noise, and their interpretation.

[3] A. Van der Ziel, *Noise* (Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1954). References 1–3 emphasize the physical origins of noise and review the experimental literature.

[4] L. D. Landau and E. M. Lifshitz, *Statistical Physics* (Addison-Wesley Publishing Company, Inc., Reading, Massachusetts, 1958).

[5] C. Kittel, *Elementary Statistical Physics* (John Wiley & Sons, Inc., New York, 1958). References 4 and 5 provide modern accounts of fluctuation-dissipation theory and irreversible processes.

[6] R. C. Tolman, *The Principles of Statistical Mechanics* (Oxford University Press, London, 1938).

[7] R. H. Fowler, *Statistical Mechanics* (Cambridge University Press, London, 1936).

[8] *The Collected Works of J. Willard Gibbs* (Yale University Press, New Haven, 1948), Vol. 2. References 6–8 supply the fundamentals of thermodynamic and statistical mechanical fluctuation theory.

[9] S. R. deGroot, *Thermodynamics of Irreversible Processes* (Interscience Publishers, Inc., New York, 1951).

[10] I. Prigogine, *Etude Thermodynamique des phénomènes Irreversibles*, Dunod, Paris, 1947); *Introduction to Thermodynamics of Irreversible Process* (Charles C. Thomas, Springfield, Illinois, 1955). For related books see also R. T. Cox, *Statistical Mechanics of Irreversible Processes* (John Hopkins Press, Baltimore, 1955).

[11] J. L. Doob, *Stochastic Processes* (John Wiley & Sons, Inc., New York, 1953).

[12] M. S. Bartlett, *An Introduction to Stochastic Processes* (Cambridge University Press, London, 1955).

[13] N. Wiener, *The Extrapolation, Interpolation and Smoothing of Stationary Time Series* (John Wiley & Sons, Inc., New York, 1950).

[14] U. Grenander and M. Rosenblatt, *Statistical Analysis of Stationary Time Series* (John Wiley & Sons, Inc., New York, 1957). References 11–14 emphasize the mathematical aspects of random processes.

[15] J. J. Freeman, *Principles of Noise* (John Wiley & Sons, Inc., New York, 1958).

[16] J. S. Bendat, *Principles and Applications of Random Noise Theory* (John Wiley & Sons, Inc., New York, 1958).

[17] E. J. Schremp, *Vacuum Tube Amplifiers*, G. E. Valley and H. Wallman, editors, Massachusetts Institute of Technology, Radiation Laboratory Series (McGraw-Hill Book Company, Inc., New York, 1948), Vol. 18, Chap. 12.

[18] W. B. Davenport, Jr., and W. C. Root, *Introduction to the Theory of Random Signals and Noise* (McGraw-Hill Book Company, Inc., New York, 1958). References 15–18 emphasize communications aspects of random processes, filtering of signal from noise, prediction, etc.

[19] S. Chandrasekhar, Revs. Modern Phys. **15**, 1 (1943).

[20] G. E. Uhlenbeck and L. S. Ornstein, Phys. Rev. **36**, 823 (1930).

[21] M. C. Wang and G. E. Uhlenbeck, Revs. Modern Phys. **17**, 323 (1945).

[22] S. O. Rice, Bell System Tech. J. **23**, 282 (1944); **24**, 46 (1945). See W. R. Bennett, Proc. Inst. Radio Engrs. **44**, 609 (1956) on the need for an ensemble average in Eq. (2.3).

[23] M. Kac, Amer. Math. Monthly **54**, 369 (1947).

[24] J. L. Doob, Ann. Math. **43**, 351 (1942).

[25] N. Wax, *Noise and Stochastic Processes* (Dover Publications, New York, 1954) represents a collection of classic review papers, references 19–24.

[26] *Proceedings of the International Conference of Theoretical Physics, Kyoto and Tokyo, 1953* (Science Council of Japan, Ueno Park, Tokyo, 1954), especially sections on transport phenomena and irreversible processes.

[27] *Proceedings of the International Symposium on Transport Processes in Statistical Mechanics, Brussels* (Interscience Publishers, Inc., New York, 1958).

the microscopic variables at the present. Such a system is "Markoffian" in the trivial sense of being deterministic. (Quantum mechanical systems are also covered by these remarks. For these, determinism means that the probability of any future event is determined by a complete specification of the present state.)

2. The fluctuations are computed with reference to an *equilibrium* state.

By using such modest assumptions, Callen and coworkers,[37–40] Kubo,[41] Lax,[42] Ekstein and Rostoker,[43] and others[44] have been able to establish a generalized quantum mechanical Nyquist theorem relating noise in some variables to the admittance of the system for these same variables.[45] A corresponding classical proof for a single variable was given by Richardson.[46] Richardson also tried to calculate the noise in a "driven system," i.e., the fluctuations from a nonequilibrium steady state. He found that there was then no necessary connection between the noise and the impedance of the system. In other words, the conventional Nyquist theorem breaks down for nonequilibrium situations.

There has also been work on irreversible processes by Bergmann[47] and Lebowitz[48] in which the first assumption is modified. The system is treated in a Hamiltonian way, but the interaction with reservoirs is handled by a stochastic assumption.

Before leaving the microscopic domain, we should note that the fluctuations in one variable, or in any pair of variables at one time, can be computed by means of statistical mechanics (which also makes assumptions 1 and 2). See, for example, Tolman,[6] Fowler,[7] and Landau and Lifshitz.[4]

An excellent development of the theory of fluctuations of extensive thermodynamic parameters has been given by Greene and Callen[49] who provide formulas for second-, third-, and higher-order correlation moments in terms of derivatives of extensive parameters with respect to intensive parameters, i.e., the response of a system's extensive parameters to static applied forces (see Sec. 4). A knowledge of the complete response to a static force of arbitrary size provides a knowledge of all moments and hence a knowledge of the complete distribution function for a set of thermodynamic parameters. For the case of a single parameter, this distribution function has been written down explicitly by Magalinskii and Terletskii.[50]

Magalinskii and Terletskii also show that if one knows the average time-dependent response $\langle q(t)\rangle$ to the sudden turn on (or off) of a static force, one can write the complete joint distribution function[51] of $q(t)$ and $q(0)$. This work is interesting because it demonstrates that if the response is linear in the force, the distribution function is Gaussian. The proof, however, is based on the equilibrium assumption. For the nonequilibrium case, it may not be true. It is easy to demonstrate that Markoffian systems with linear responses are not necessarily Gaussian.

The work of Magalinskii and Terletskii[50] is not particularly helpful in computing noise, since it is precisely the response $\langle q(t)\rangle$ assumed known by them that we must obtain in order to compute the noise. This remark is characteristic of the results of microscopic theories: precise relationships are obtained, e.g., between noise and the response of the systems. Neither quantity can, however, be computed from the formal expressions so obtained.[45] Actual calculations of responses, e.g., electrical and thermal conductivity, are based on the development of a *transport* equation. While a transport description is detailed, we refer to it as *macroscopic*

---

[28] See R. B. Barnes and S. Silverman, Revs. Modern Phys. **6**, 162 (1934) for a discussion of Brownian motion as a natural limit to measuring processes.

[29] *Proceedings of the First, Second, and Third Berkeley Symposia on Mathematical Statistics and Probability*, Jerzy Neyman, editor (University of California Press, 1949, 1951, 1956).

[30] J. E. Moyal, J. Roy. Stat. Soc. Series 13, **2**, 150 (1949). An extensive review paper emphasizing mathematical and physical foundations of stochastic processes.

[31] E. M. Montroll and M. S. Green, Ann. Rev. Phys. Chem. **5**, 449 (1954). References 26–31 provide useful review papers.

[32] A. Van der Ziel, Proc. Inst. Radio Engrs. **46**, 589 (1958).

[33] A. Van der Ziel, Proc. Inst. Radio Engrs. **46**, 1019 (1958).

[34] K. M. van Vliet, Proc. Inst. Radio Engrs. **46**, 1004 (1958).

[35] R. L. Petritz, *Photoconductivity Conference* (John Wiley & Sons, Inc., New York, 1954), pp. 49–77.

[36] R. L. Petritz, Proc. Inst. Radio Engrs. **40**, 1440 (1952). References 32–36 include a review of the experimental literature on noise in solids.

[37] H. B. Callen and T. A. Welton, Phys. Rev. **83**, 34 (1951).

[38] Callen, Barasch, and Jackson, Phys. Rev. **88**, 1382 (1952).

[39] J. Jackson, Phys. Rev. **87**, 471 (1952).

[40] T. A. Kaplan, Phys. Rev. **102**, 1447 (1956).

[41] R. Kubo, Can. J. Phys. **34**, 1274 (1956); J. Phys. Soc. Japan **12**, 570 (1957).

[42] M. Lax, Phys. Rev. **109**, 1921 (1958).

[43] H. Ekstein and N. Rostoker, Phys. Rev. **100**, 1023 (1955).

[44] H. Mori, J. Phys. Soc. Japan **11**, 1029 (1956). J. Weber, Phys. Rev. **101**, 1619, 1620 (1956).

[45] References 36–44 consider the *mean* response of a system to *first* order in the driving forces. H. B. Callen, Phys. Rev. **111**, 367 (1958) has calculated the probability of any *specified* response to *first* order. W. Bernard and H. B. Callen, Revs. Modern Phys. **31**, 1017 (1959) have calculated the *mean* response to *second* order in the driving forces, and the fluctuations to first and second order in the driving forces. The first (second) order noise in a weakly driven system has been expressed in terms of third (fourth) order time-displaced moments at equilibrium. This result is a purely formal one, since it is no easier to compute equilibrium time-displaced higher moments than it is to compute the noise directly. However, the Bernard-Callen work represents the first microscopic calculation of noise in a nonequilibrium state.

[46] J. Richardson, Inst. Radio Engrs. Trans. on Inform. Theory **IT-1**, No, 1, 62 (1955).

[47] P. G. Bergmann and J. L. Lebowitz, Phys. Rev. **99**, 578 (1955).

[48] J. L. Lebowitz and P. G. Bergmann, Ann. Phys. (N. Y.) **1**, 1 (1957); J. L. Lebowitz and H. L. Frisch, Phys. Rev. **107**, 917 (1957).

[49] R. F. Greene and H. B. Callen, Phys. Rev. **83**, 1231 (1951).

[50] V. B. Magalinskii and I. P. Terletskii, Soviet Phys. JETP **7**, 501 (1958).

[51] H. Takahasi, J. Phys. Soc. Japan **7**, 439 (1952) gave an earlier proof that certain time-delayed second moments could be calculated from the response to a sudden force. Essentially the same results have also been obtained by R. H. Kraichnan, Phys. Rev. **113**, 1181 (1959).

since the time evolution of the system is computed without *complete* specification of all the microscopic variables. Since one is usually concerned with the calculation of the noise in a system whose transport properties have already been investigated, we assume that the basic transition probabilities giving rise to the transport equation are already known (usually from microscopic calculations of collision rates). Our aim is to show how this same information can be incorporated directly into a calculation of the noise. We therefore adopt a *macroscopic* viewpoint and propose to show how noise can be computed by using a minimum number of additional assumptions.

Previous work on the *macroscopic* treatment of noise generally makes several of the following assumptions:

1. The system is Markoffian. In other words there is some set of parameters $\boldsymbol{\alpha}=\alpha_1, \alpha_2, \alpha_3, \cdots$ (*less* than a complete set of microscopic variables) with the property that a knowledge of the present value of $\boldsymbol{\alpha}$ at $t=0$ determines the mean value of $\boldsymbol{\alpha}(t)$ for $t>0$ and even the distribution of values of $\boldsymbol{\alpha}(t)$. Furthermore, knowledge of $\boldsymbol{\alpha}(0)$ makes obsolete all information $\boldsymbol{\alpha}(t)$ for $t<0$.
2. The system is stationary, i.e., invariant with respect to a shift of the origin of time. Thus, all transition probabilities do not depend explicitly on time, and all probabilities of joint events depend only on their time separation.
3. The system is linear, i.e., its responses to external forces are linear in these forces.
4. The system, at least from a microscopic point of view, obeys time reversibility.
5. Fluctuations are from an equilibrium state.
6. The individual jumps from $\boldsymbol{\alpha}$ to $\boldsymbol{\alpha}'$ are sufficiently small to permit the Fokker-Planck approximation.
7. The variables $\boldsymbol{\alpha}(t)$ have a Gaussian distribution (are Gaussian). *Note.*—3+6 imply 7.
8. The fluctuations in $\boldsymbol{\alpha}(t)$ can be regarded as produced by random forces—the Langevin viewpoint.

The macroscopic treatment presented here uses the Markoffian and stationary assumptions that are common to all previous macroscopic work, plus an assumption of quasi-linearity, which is appreciably less restrictive than complete linearity. These assumptions are discussed in more detail in Sec. 2. Fluctuations in a variable of a nonlinear system can usually be handled by a quasi-linear treatment of fluctuations in the distribution function for that variable. Fluctuations in distribution functions are discussed in Sec. 12.

Since a number of extensive review papers[19–36] and books[1–18] exist, we indicate here only the relative importance of some previous work for that which follows later.

The macroscopic theory of fluctuations at one time was developed by Einstein[52] by inverting Boltzmann's relation (see Sec. 4). (A macroscopic justification for the Einstein point of view was given by Greene and Callen.[49])

Einstein Relation

The theory of Brownian motion was developed extensively from the Langevin and Fokker-Planck viewpoints.[53,54] Except for time reversibility, all eight of the given assumptions are used explicitly or implicitly. One of the key results of this early work was the Einstein relation between diffusion constant and mobility. Because of the importance of this relation to our work, we present two elementary derivations of this relation. Mobility μ can be defined as the terminal velocity of a charged particle per unit applied field,

$$v=\mu E, \tag{1.1}$$

and the diffusion constant can be defined in terms of the particle current induced by a concentration gradient

$$j_{\text{part}}=-D(\partial n/\partial x). \tag{1.2}$$

Einstein's[55] original idea is that if one sets up an electric field in an open circuit, a concentration gradient will build up big enough for the diffusion current to cancel the drift current,

$$j=ne\mu E-eD(\partial n/\partial x)=0,$$

or

$$n(x)\propto\exp(\mu Ex/D). \tag{1.3}$$

Under *thermal equilibrium* conditions, with the potential energy

$$V=-eEx, \tag{1.4}$$

Boltzmann's law indicates a distribution

$$n(x)\propto\exp(-V/kT). \tag{1.5}$$

Comparison of the last three equations yields the Einstein relation[55]

$$D=kT(\mu/e)=kTL', \tag{1.6}$$

where $L'=v/eE$ is the dc admittance to the applied force eE.

Conservation of particles yields the equation

$$\partial n/\partial t+\operatorname{div}(j/e)=\partial n/\partial t+v(\partial n/\partial x)-D(\partial^2 n/\partial x^2)=0, \tag{1.7}$$

where $v=\mu E$ is the drift velocity. The Green's function[56] of this equation describing the density at x and t of a

[52] A. Einstein, Ann. Physik **33**, 1275 (1910).

[53] P. Langevin, Compt. rend. **146**, 530 (1908). See also reference 20 and Sec. 8.

[54] A. D. Fokker, Ann. Physik **43**, 812 (1914); M. Planck, Sitzber preuss. Akad. Wiss. Physik. math. Kl. 324 (1917). See also references 19 and 21, and Sec. 5.

[55] A. Einstein, Ann. Physik **17**, 549 (1905); **19**, 371 (1906); *Investigations on the Theory of the Brownian Movement* (E. P. Dutton, and Company, Inc., New York, 1926); R. Fürth, *Schwankungserscheinungen in der Physik* (Vieweg and Sons, Brauschweig, Germany, 1920).

[56] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill Book Company, Inc., New York, 1953), Sec. 7.4.

pulse which originated at x' at time 0 is

$$n(x,x't)=(4\pi Dt)^{-\frac{1}{2}}\exp[-(x-x'-vt)^2/4Dt]. \quad (1.8)$$

An interesting direct verification of the validity of the Einstein relation for electrons and holes in a semiconductor[57] was made by the Haynes technique of observing the mean drift and spreading of a pulse of injected minority carriers.

Perrin[58] used the Brownian motion of colloidal particles to measure their diffusion constant. When combined with Stokes's law for the viscous drag on such particles,

$$R'=1/L'=6\pi a\eta, \quad (1.9)$$

where η is the viscosity of the surrounding medium and a the radius of the particle, the Einstein relation (1.6) yields a determination of Boltzmann's constant k, and through the gas constant R, of Avogadro's number $N=R/k$.

The second moment of Eq. (1.8) yields another definition of the diffusion constant,

$$D=(2T)^{-1}\langle[x(T)-x(0)-vT]^2\rangle, \quad (1.10)$$

in terms of the mean-square displacement of a particle in time T after the nonrandom drift part vT is subtracted. (Angular brackets represent an ensemble average.) The time T must be taken long compared to a microscopic collision time t in order for the diffusion equation (1.7) to be valid but is otherwise arbitrary. One customarily takes the limit as $T\to\infty$. [When complicated external forces are present, it is easier not to try to subtract the nonrandom part and take instead the limit $T\to 0$, since the nonrandom part of $x(T)-x(0)$ vanish linearly in T and does not contribute to the diffusion constant, cf. Eq. (5.7).]

For the case of no drift, $v=0$, we can, by a shift of origin, write

$$D=\lim_{T\to\infty}(2T)^{-1}\left\langle\left|\int_{-T/2}^{T/2}v(t)dt\right|^2\right\rangle. \quad (1.11)$$

The noise at frequency f associated with velocity fluctuations is by definition [see Eq. (2.3)]

$$G(v,f)=\lim_{T\to\infty}\frac{2}{T}\left\langle\left|\int_{-\frac{1}{2}T}^{\frac{1}{2}T}v(t)e^{-2\pi ift}dt\right|^2\right\rangle. \quad (1.12)$$

We now obtain the fundamental relation that the noise at frequency 0 in $v=dx/dt$ is equal to the diffusion constant for x:

$$G(v,0)=4D. \quad (1.13)$$

The Nyquist[59] theorem relating the noise at frequency $f=\omega/2\pi$ to the admittance $Y(\omega)$,

$$G(v,f)=4kT\,\mathrm{Re}Y(\omega), \quad (1.14)$$

reduces therefore at zero frequency to the Einstein relation when we remember that $Y(0)=L'$.

The previous remarks describe Brownian motion in the absence of restoring forces. In the presence of restoring forces $Y(\omega)\to 0$ as $\omega\to 0$ so that the noise at zero frequency vanishes (see Sec. 7). Equation (1.13) is obviously no longer true. The reason is that, in the presence of restoring forces, Eq. (1.11) should be used with $T\to 0$ as previously remarked. The derivation of Eq. (1.13) then fails, but D is still related to the dissipative part of the admittance via the Einstein relation as proven in Sec. 5.

Nyquist Theorem

The Wiener[60]-Khintchin[61] theorem [see Eq. (2.5)] permits Eq. (1.12) to be rewritten in the form

$$G(v,f)=4\,\mathrm{Re}\int_0^\infty e^{-i\omega t}\langle v(t)v(0)\rangle dt. \quad (1.15)$$

In the presence of a constant (velocity independent) relaxation time, the decay from a fixed initial value $v(0)$ has the form

$$\langle v(t)\rangle_{v0}=\exp(-t/\tau)v(0); \quad (1.16)$$

thus

$$G(v,f)=4\,\mathrm{Re}\tau(1+i\omega\tau)^{-1}\langle v(0)v(0)\rangle, \quad (1.17)$$

and

$$D=\tau\langle vv\rangle. \quad (1.18)$$

To establish the Nyquist and Einstein relations, for the case of a relaxation time, we must calculate the response to an applied field,[62–67]

$$dv/dt-v/\tau=(e/m)E\exp(i\omega t); \quad (1.19)$$

therefore

$$v(t)=\mu(\omega)E\exp(i\omega t), \quad (1.20)$$

with the mobility at frequency ω given by

$$\mu(\omega)=\frac{e}{m}\frac{\tau}{1+i\omega\tau}. \quad (1.21)$$

Comparison of Eqs. (1.17) and (1.18) with Eq. (1.21) yields

$$G(v,f)=4m\langle vv\rangle\,\mathrm{Re}Y(\omega);\quad Y(\omega)=\mu(\omega)/e, \quad (1.22)$$

$$D=m\langle vv\rangle\mu(0)/e. \quad (1.23)$$

---

[57] Transistor Teachers Summer School, Phys. Rev. **88**, 1368 (1952).

[58] J. Perrin, *Atoms* (Constable, London, 1916). See also reference 19.

[59] H. Nyquist, Phys. Rev. **32**, 110 (1928).

[60] N. Wiener, Acta. Math. **55**, 117 (1930); J. Math. Phys. **5**, 99 (1926).

[61] A. I. Khintchin, Math. Ann. **109**, 604 (1934).

[62] For early derivation of Nyquist noise from the Lorentz picture of an electron gas, see references 63–67.

[63] J. Bernamont, Ann. Phys. **7**, 71 (1937).

[64] D. A. Bell, J. Inst. Elec. Engrs. London **82**, 522 (1938).

[65] C. J. Bakker and G. Heller, Physica **6**, 262 (1939).

[66] E. Spenke, Wiss. Veröffentl. Siemens-Werken **18**, 54 (1939).

[67] The last paper has been criticized by A. Van der Ziel, J. Appl. Phys. **21**, 399 (1950) since it yields a result based on the classical rather than the quantum equipartition laws. This difficulty has since been eliminated by the quantum treatments in references 41–45. See also Eq. (7.12) and reference 68.

These are equivalent to the usual relations (1.14) and (1.6) with kT replaced by $m\langle vv\rangle$. (We write vv rather than v^2 because, in the three-dimensional case, the theorems remain valid with D_{rs}, a tensor computed from $\langle v_r v_s\rangle$.) In the equilibrium case, $\langle mv^2\rangle = kT$, but the above relations are valid for the nonequilibrium case as well for the case of a constant relaxation time. The ratio $C=\langle mv^2\rangle/kT$ represents a correction factor to the Einstein and Nyquist relations which appears in a generalized form later.

For the case in which the basic collision processes cannot be described by a relaxation process, the mobility must be calculated by solving a transport equation. This is usually difficult to do in practice. In principle, however, we have shown[68] that a small added ac electric field, $\mathbf{E}(t)=\mathbf{E}(0)\exp(i\omega t)$, perturbs a stationary distribution $f_0(\mathbf{v})$ into $f_0(\mathbf{v})+f_1(\mathbf{v},t)$, where

$$f_1(\mathbf{v},t)=-\frac{e}{m}\mathbf{E}(t)\int_0^\infty e^{-i\omega t}\int\frac{\partial f_0(\mathbf{v}')}{\partial \mathbf{v}'}d\mathbf{v}' \times W(\mathbf{v},\mathbf{v}',t)dt, \quad (1.24)$$

in which $W(\mathbf{v},\mathbf{v}',t)$ represents the probability of finding a particle with velocity $\mathbf{v}$ at time t if it started with velocity $\mathbf{v}'$ at time 0 in the absence of the added ac electric field. The mean added drift velocity due to the added field may be obtained by multiplying Eq. (1.24) by $\mathbf{v}$ and integrating. The result is a ("differential") mobility tensor

$$\mathfrak{y}(\omega)=-\frac{e}{m}\int_0^\infty e^{-i\omega t}dt\int\langle\mathbf{v}(t)\rangle_{\mathbf{v}'}\frac{\partial f_0(\mathbf{v}')}{\partial \mathbf{v}'}d\mathbf{v}', \quad (1.25)$$

where

$$\langle\mathbf{v}(t)\rangle_{\mathbf{v}'}=\int\mathbf{v}W(\mathbf{v},\mathbf{v}',t)d\mathbf{v} \quad (1.26)$$

is the mean velocity at time t for an electron which started with velocity $\mathbf{v}'$ at time 0.

If f_0 represents a thermal equilibrium distribution, then

$$\partial f_0(\mathbf{v}')/\partial\mathbf{v}'=-(m\mathbf{v}'/kT)f_0(\mathbf{v}'), \quad (1.27)$$

and we can write

$$\mathfrak{y}(\omega)=\frac{e}{kT}\int_0^\infty e^{-i\omega t}dt\langle\mathbf{v}(t)\mathbf{v}(0)\rangle, \quad (1.28)$$

where $v(0)=v'$, and the average is taken with respect to the equilibrium distribution f_0. Comparison with Eqs. (1.15) and (1.13) then yields the usual Nyquist and Einstein relations. Although we have used the language of a conduction problem, if $\mathbf{x}$ is any variable, and if $e\mathbf{E}=\mathbf{F}$ is the force which acts on it, then the noise in $\mathbf{v}=d\mathbf{x}/dt$ is related to the admittance $\mathbf{Y}=\mathbf{v}/\mathbf{F}$. For example, $\mathbf{x}$ could be a magnetization, $\mathbf{F}$ a magnetic field, and $\mathbf{Y}$ would be, aside from a frequency factor, a magnetic susceptibility.[69]

Callen and Greene[70,71] have given a particularly intuitive derivation of the Nyquist theorem for the thermal equilibrium case. Their derivation is based on the following syllogism:

1. The frequency dependence of the noise is determined by the Fourier transform of the function $\langle v(t)\rangle_{v0}$ which describes the "regression" from a spontaneous fluctuation of amount $v(0)$.
2. The regression of fluctuations obeys the macroscopic equations of motion.
3. The macroscopic equations of motion are describable in the frequency domain by the admittance of the system, ergo there must be a relation between the noise and the admittance.

The second point was carefully stated as an assumption by Onsager[72] in his derivation of the Onsager relations. deGroot[9] has argued that the only distinction between a fluctuation and a macroscopic deviation is one of size. Hence, if the system is linear, e.g., describable by a linear transport equation, one would expect macroscopic decay and microscopic regression to obey identical laws.

Callen and Greene do not make explicit use of the Onsager assumption. Instead they argue that the regression $\langle v(t)\rangle_{v0}$ can be observed in three logically distinct ways.

(a) Observe the system without disturbing it. Whenever the velocity $v(0)$ is observed, wait a time t and measure the velocity. The average of such measurements is the desired $\langle v(t)\rangle_{v0}$.

(b) Clamp a "microcanonical" constraint on the system for $t<0$ requiring v to take the value $v(0)$. Release the constraint at $t=0$ and observe at time t. Callen and Greene argue that (a) and (b) are equivalent because they both give equal weight to all microstates consistent with $v(0)$.

(c) For all $t<0$ add an extra external force chosen to make the mean v equal to $v(0)$. At $t=0$ release the extra force and observe $v(t)$. Callen and Greene argue that the step from (b) to (c) constitutes the usual replacement of a microcanonical ensemble by a canonical one, a change which "as we know, does not influence the macroscopic thermodynamics of the system."

[68] Equation (1.24) is based on Eq. (4.13) of reference 42 which is derived by assuming arbitrary transition probabilities $w_{\mathbf{v}'\mathbf{v}}$ from velocity $\mathbf{v}$ to velocity $\mathbf{v}'$, and in this sense generalizes the proofs mentioned in reference 62 which assume the existence of a relaxation time. Van der Ziel's remark (reference 6) that Spenke's derivation (reference 66) might have yielded the quantum-mechanical correction factor of Eq. (7.12) if he had taken into account inelastic collisions is shown here to be false since our present derivation includes inelastic collisions, but the result [Eq. (1.28)] yields the classical Nyquist theorem. Microscopic derivations yielding the quantum correction factor are given in references 41–45.

[69] R. Kubo and K. Tomita, J. Phys. Soc. Japan **9**, 888 (1954).

[70] H. B. Callen and R. F. Greene, Phys. Rev. **86**, 702 (1952).

[71] R. F. Greene and H. B. Callen, Phys. Rev. **88**, 1378 (1952) gives a generalization of the previous reference to many variables.

[72] L. Onsager, Phys. Rev. **37**, 405 (1931); **38**, 2265 (1931).

Approach (c) is actually used to compute the "regression." The added force ΔF for $t<0$ and zero for $t>0$ is represented by a Fourier integral, and the response to each frequency component computed by means of the admittance at that frequency. The final result is the Nyquist theorem Eq. (1.14).

Onsager and Machlup[73] have remarked, however, that a system whose future depends on the initial state but not on whether it arrived there by a spontaneous fluctuation, a constraint, or an applied force, is Markoffian. Indeed they prove that a set of Gaussian random variables obeying the linear regression equation are of necessity Markoffian, i.e., they claim that Callen and Greene have tacitly made the Markoffian assumption.

In my opinion, all of these remarks are correct and are related to the fact that macroscopic systems can be described by thermodynamic variables. The fact that in practice there is no distinction between microcanonical and canonical ensembles[74] is closely related to the Markoffian behavior of macroscopic systems for a suitably chosen[73] set of thermodynamic variables.

It is of interest that the microscopic proofs of the Nyquist theorem do not require the use of the Markoffian assumption (except in the trivial sense that the future is determined by a complete description of the present), but lean heavily on the assumption that one is dealing with fluctuations from an equilibrium state.

We can shed some light on the success of the microscopic proofs by comparing Eq. (1.15) for the noise with Eq. (1.25) for the mobility. We see that both expressions are averages involving the *same* regression $\langle v(t)\rangle_{v_0}$. It is the weight factors $v(0)f_0$ and $\partial f_0/\partial v$ that are logically distinct. For the thermal equilibrium case they are proportional to one another via Eq. (1.27). In the absence of equilibrium, Eq. (1.27) is no longer obeyed, and there is no longer any necessary relation between the noise and the admittance—this is essentially Richardson's remark.[46]

We anticipate, however, that even for nonequilibrium problems, the Nyquist and Einstein relations still remain approximately correct. Our confidence is based on Wannier's[75-77] thorough investigation of the motion of gaseous ions in strong electric fields. He finds that for any one of the principal directions n, the diffusion and mobility tensors obey

$$eD_n \approx (\partial v_n/\partial E)m\langle \Delta v_n \Delta v_n\rangle, \tag{1.29}$$

where $\partial v_n/\partial E$ is the differential mobility for an added field, and $m\langle \Delta v_n \Delta v_n\rangle$ is twice the mean random energy along n (i.e., $\Delta v_n = v_n - \langle v_n\rangle$). Wannier finds this relation to be obeyed exactly for the case of a constant mean free time between collisions. For the case of a hard-sphere model, with mass ratio unity between the ions and the neutrals they collide with, Wannier finds, as a result of a long numerical calculation, that the longitudinal diffusion coefficient is 18% less than Eq. (1.29) would indicate. In other words, within 18%, one could have used for $f_0(\mathbf{v})$ a pseudo-Boltzmann distribution with a different mean velocity and temperature for each principle direction so that

$$\partial f_0/\partial v_n = -f_0 m(v_n - \langle v_n\rangle)/kT_n. \tag{1.30}$$

Since f_0 is properly normalized,

$$\langle v_n\rangle \int f_1(\mathbf{v})d\mathbf{v} = 0, \tag{1.31}$$

thus $\mathbf{v}(t)$ in Eq. (1.25) can be replaced by $\mathbf{v}(t) - \langle \mathbf{v}\rangle$. The use of Eq. (1.30) then reduces (1.25) to the Einstein relation (1.29) connecting the diffusion constant with the differential mobility when we identify

$$kT_n = m\langle (v_n - \langle v_n\rangle)^2\rangle \tag{1.32}$$

in accord with the pseudo-Boltzmann distribution.

The basic reason for the 18% discrepancy in the hard-sphere problem, for the nonequilibrium case, is that one has a constant mean free path, rather than a constant mean free time, i.e., $\tau = l/v$, so that Eq. (1.19) becomes nonlinear. The nonlinearity is appreciable since τ varies considerably over the range of thermal fluctuations of velocity.

Noise problems can therefore be divided into two categories: (1) those in which the parameters of the system do not change significantly when the variables fluctuate thermally, (2) those in which the changes induced in the parameters by fluctuations are large. Problems of the first category are handled by a quasi-linear approximation, i.e., we expand about some steady-state value, and treat the deviations by means of linear equations which represent a multivariable generalization of Eq. (1.19). Problems of the second category require that one discuss distribution functions of the variables in question. Instead of directly discussing fluctuations in the variables, we instead discuss fluctuations in the distribution functions—from which all the necessary physical results can be computed (Sec. 12). The fluctuation in the distribution functions is usually sufficiently small for them to be treated in a quasi-linear way, even if the original variables would have required a nonlinear treatment. Many distribution functions, indeed, obey strictly linear equations, as for example, in the problem considered by Wannier.

Our macroscopic treatment of linearized systems is closely related to the approach of Onsager[72] and Hashitsume,[78] whose papers supplied the inspiration and motivation for the work presented here. We differ from Hashitsume and from Onsager and Machlup[73]

[73] L. Onsager and S. Machlup, Phys. Rev. **91**, 1505, 1512 (1953).
[74] The relation between canonical and microcanonical ensembles is discussed by M. Lax, Phys. Rev. **97**, 1419 (1955). An example where the two disagree is described.
[75] G. H. Wannier, Phys. Rev. **83**, 281 (1951).
[76] G. H. Wannier, Phys. Rev. **87**, 795 (1952).
[77] G. H. Wannier, Bell System Tech. J. **32**, 170 (1953).
[78] N. Hashitsume, Progr. Theoret. Phys. (Kyoto) **15**, 369 (1956).

(also Siegel,[79] and Tisza and Manning[80–81]) in two ways. (1) We do not assume that we are dealing with Gaussian variables. We only attempt to calculate in detail the first and second moments of our variables (all that is needed for admittance and noise calculations). If one adds the assumption of Gaussian variables, one could immediately write down the complete distribution function from the computed first and second moments. We do not do this because the added information is, for our present purposes, superfluous, and may be wrong.[82] (2) We do not assume that we are dealing with fluctuations from an equilibrium state but allow nonequilibrium steady states.

2. Review of Our General Assumptions

Since noise represents a small fluctuation superimposed upon a signal, a system, although nonlinear "in the large," behaves in a "quasi-linear" way for small deviations from a particular bias position. (This remark may be verified in detail by examining the result Eq. (5.36) of an exact and elegant treatment by Kac and Siegert[83] of the square law detection of signal plus noise for the case in which the input signal is large compared to the noise: $s\gg1$.) Thus, the linear procedures we apply are applicable to nonlinear systems providing the signal (+ noise) remains in the neighborhood of some "bias" position. A discussion of the correction produced by nonlinearities is given in Sec. 14.

Once the restriction to quasi-linear processes is made, the requirement of a stationary random process is a rather weak one. We may be dealing with fluctuations from an equilibrium state, or a steady state, or even with fluctuations about some time dependent motion ("signal"). If $I(t)$ is a random variable representing such a *fluctuation*, i.e., such that

$$\langle I(t)\rangle=0, \tag{2.1}$$

the requirement of stationarity may be represented by the statement that the ensemble average,

$$\langle I(t+t')I(t')\rangle=\langle I(t)I(0)\rangle, \tag{2.2}$$

is independent of t', i.e., *as far as the noise is* concerned, there is no absolute origin of time. When we construct a Markoffian master equation, we interpret the requirement of stationarity to mean that the fundamental transition probabilities describing the stochastic process are independent of time.

The noise power $G(f)df$ in the frequency interval df, associated with the stationary random variable $I(t)$ obeying (2.1), may be defined by[22]

$$G(I,f)=\lim_{T\to\infty}\frac{2}{T}\left\langle\left|\int_{-\frac{1}{2}T}^{\frac{1}{2}T}I(t)e^{-2\pi ift}dt\right|^2\right\rangle, \tag{2.3}$$

where the factor 2 arises from the convention that the total dissipated "power" is given by

$$\langle I^2\rangle=\lim_{T\to\infty}\frac{1}{T}\int_0^T I^2(t)dt=\int_0^\infty G(f)df, \tag{2.4}$$

so that $G(f)$ includes contributions from both frequencies f and $-f$. (We abbreviate $G(I,f)$ by $G(f)$ or $G(I)$ if the other symbol is known from context.)

That $G(f)df$ as given by Eq. (2.3) actually represents the noise in the frequency interval df has been verified by Ekstein and Rostoker's illuminating analysis[43] of the process of noise measurement.

If one represents Eq. (2.3) as a product of two integrals, one in t, the other in t', then the integral contains $\langle I(t)I(t')\rangle$, which in view of Eq. (2.2) is a function only of $t-t'$. Thus, one may integrate over $t+t'$ for fixed $t-t'$. The result can, by using Eq. (2.2), be simplified to the form

$$G(f)=2\int_{-\infty}^{\infty}e^{-i\omega t}\langle I(t)I(0)\rangle dt, \tag{2.5}$$

where $\omega=2\pi f$. To obtain a result expressed only in terms of *future* values of t, we split the region of integration and make use of the relation

$$\langle I(t)I(0)\rangle=\langle I(0)I(-t)\rangle,$$

based on stationarity, to obtain

$$G(f)=G_+(f)+G_-(f), \tag{2.6}$$

$$G_+(f)=2\int_0^\infty e^{-i\omega t}\langle I(t)I(0)\rangle dt, \tag{2.7}$$

$$G_-(f)=2\int_0^\infty e^{i\omega t}\langle I(0)I(t)\rangle dt. \tag{2.8}$$

It is customary at this point to note that the averages in (2.7) and (2.8) are identical, so that $G_-(f)=G_+(-f)$, and $G(f)$ can be expressed as the $\cos\omega t$ transform of $4\langle I(t)I(0)\rangle$ (the Wiener-Khintchin theorem). In the case in which $I(t)$ represents a vector or a set of variables $[I_s(t)]$, it is clear that $\langle I_r(t)I_s(0)\rangle\neq\langle I_r(0)I_s(t)\rangle$. $G(f)$ is then a matrix with components $G_{rs}(f)$ and its parts obey

$$\mathbf{G}_-(f)=\mathbf{G}_+^{\dagger}(-f), \tag{2.9}$$

where † indicates the transposed matrix, i.e., $\mathbf{G}_-$ and $\mathbf{G}_+$ are Hermitian adjoints of one another, and their sum $\mathbf{G}$ is a Hermitian matrix.

Our problem of finding the noise spectrum is therefore reduced to the evaluation of a second moment $\langle I(t)I(0)\rangle$.

[79] A. Siegel, Phys. Rev. **102**, 953 (1956); **106**, 609 (1957).
[80] L. Tisza and I. Manning, Phys. Rev. **105**, 1695 (1957). See also a similar approach in reference 81.
[81] D. Falkoff, Ann. Phys. (N. Y.) **4**, 325 (1958).
[82] For example, the moments calculated in Eqs. (14.37) to (14.40) do not bear the relationship required for a Gaussian distribution.
[83] M. Kac and A. J. F. Siegert, J. Appl. Phys. **18**, 383 (1947).

It is not necessary, therefore, to have knowledge of the complete distribution function $W[I(0,)I(t)]$ but only of its second moments. It is certainly not necessary to restrict oneself at the outset to Gaussian random variables.[82] However, those statements about the second moments that are correct about all kinds of random variables remain correct for Gaussian variables. The assumption of Gaussian variables does not lead to incorrect results, provided no information beyond the second moments is used.

We prefer to take as our key assumption that we are dealing with a Markoffian random process. For a Markoffian system the present separates the past from the future: knowledge of the system at one instant, the present, permits a prediction of the "average" future and indeed of the probability of various possible futures. Information about the past is not needed, and indeed is not relevant when available. A Markoffian system in the random variable (or set of variables) α is therefore completely characterized by the conditional probability $P[\alpha(0)\,|\,\alpha,t]$ that the system will be in the state α at time t if it was in the state $\alpha(0)$ at time 0. This conditional probability is not arbitrary, however, it must obey the Smoluchowski consistency condition,[84]

$$P[\alpha(0)\,|\,\alpha t]=\int d\alpha' P[\alpha(0)\,|\,\alpha',t']P[\alpha'\,|\,\alpha,(t-t')], \quad (2.10)$$

for any t' in $0\leq t'\leq t$, namely, the probability of passing from $\alpha(0)$ to α in time t is composed of the probabilities of passing through some intermediate state α' in the time t' times the probability of a subsequent transition to the final state in the remaining time $t-t'$.

It may seem rather arbitrary to assume that we are dealing with a Markoffian system. After all, the future of a physical system can be predicted only if we know the initial positions and momenta of all the particles composing the system. Yet it is an empirical fact that for most physical systems of interest, the future of a small set of macroscopic variables can be predicted from their present values, e.g., for an electric circuit we must know the initial charges on the condensers and the initial currents, for heat flow in a rod we need know only the initial temperature distribution. Similarly, viscous fluid flow is adequately described by the Navier-Stokes equation, and the motion of added carriers in a semiconductor usually is described adequately by a diffusion equation with drift and recombination terms. A detailed discussion of the Markoffian assumption has been given by M. S. Green.[85]

We may understand the Markoffian behavior of most physical systems in the following way: each system has a small number of approximate integrals of the motion. We therefore may introduce a small set of new variables α which constitute approximate integrals of the motion and vary slowly, i.e., at macroscopically measurable rates, and the remainder constitute variables β that change at rates too fast to be observed on the time scale of the experiment. The equations for β then can be solved in the adiabatic approximation, i.e., treating the α's as constants. Because of the fast time constants of the β's, the latter rapidly come into equilibrium with the instantaneous values of the α. By neglecting the short time delay, the β's are functions of the instantaneous α's. Thus $d\alpha/dt=f(\alpha,\beta)\approx f[\alpha,\beta(\alpha)]=g(\alpha)$ is approximately a function only of α, and the α's predict their own future.

The Markoffian character of a physical system is then simply its ability to forget quickly the initial values of the nonrelevant variables—the β's. Strictly speaking, the equation for the α's should be written in the form $\Delta\alpha/\Delta t\simeq g(\alpha)$, where Δt may be small macroscopically but must be larger than the forgetting time for the β's.

Our problem then is not to explain how a system can behave in a Markoman way but to choose a sufficiently complete set of variables α so that a Markoffian description is possible. Clearly, the omission of any relevant slowly varying variable would prevent successful prediction of the future from the present. (For quantum systems, one often chooses α and β to be the diagonal and off-diagonal elements of the density matrix in a suitable representation.)

In practice, the choice of a set of relevant variables can be made from a knowledge of the relaxation rates in the physical system. For example, added carriers in a semiconductor may drift or diffuse in milliseconds and recombine in microseconds, but equilibration of velocities of these carriers takes place in mean free times of the order of 10^{-11} seconds. For an examination of noise in a frequency range well below 10^{11} cycles per second, the density of carries $n(\mathbf{r},t)$ obeying a diffusion equation should constitute an adequate description, whereas at higher frequencies, a phase-space description $f(\mathbf{r},\mathbf{p},t)$ would be necessary.

After a sufficiently long time, our system not only forgets the initial β's, it also forgets the initial α's, and approaches an equilibrium or a steady state,

$$\lim_{t\to\infty} P[\alpha(0)\,|\,\alpha,t]=W(\alpha). \quad (2.11)$$

The joint probability distribution for $\alpha(0)$ and $\alpha(t)$ can then be written as the probability of finding $\alpha(0)=\alpha_0$ times the probability that given $\alpha(0)=\alpha_0$, $\alpha(t)$ will take the value α a time t later:

$$W(\alpha_0,\alpha)=W(\alpha_0)\cdot P(\alpha_0\,|\,\alpha,t). \quad (2.12)$$

As we have previously remarked, we do not need

[84] This equation is discussed by Wang and Uhlenbeck, reference 21, and Chandrasekhar, reference 19. See also A. Kolmogoroff, Math. Ann. **104**, 415 (1931); **108**, 149 (1933).

[85] M. S. Green, J. Chem. Phys. **20**, 1281 (1952); **22**, 398 (1954). Another attempt to justify a Markoffian approach is given by N. G. Van Kampen, Physica **20**, 603 (1954). The most careful justification of a Markoffian approximation for quantum mechanical systems has been given by L. Van Hove, Physica **21**, 517 (1955).

this complete probability distribution but only its second moment,

$$\langle\boldsymbol{\alpha}(0)\boldsymbol{\alpha}(t)\rangle=\int\boldsymbol{\alpha}_0W(\boldsymbol{\alpha}_0)d\boldsymbol{\alpha}_0\langle\boldsymbol{\alpha}(t)\rangle_{\alpha_0}, \tag{2.13}$$

where

$$\langle\boldsymbol{\alpha}(t)\rangle_{\alpha_0}=\int P[\boldsymbol{\alpha}(0)\,|\,\boldsymbol{\alpha},t]\boldsymbol{\alpha}d\boldsymbol{\alpha} \tag{2.14}$$

is the conditional mean of $\boldsymbol{\alpha}(t)$, subject to the initial conditions $\boldsymbol{\alpha}(0)=\boldsymbol{\alpha}_0$.

METHOD OF SOLVING MARKOFFIAN NOISE PROBLEMS

3. Quasi-Linear Systems

Equations (2.13) and (2.14) show that our problem can be decomposed into two parts: the determination of the equilibrium or steady-state distribution $W(\boldsymbol{\alpha})$, and the average regression of a fluctuation $\langle\boldsymbol{\alpha}(t)\rangle_{\alpha_0}$. We now address ourselves to the latter problem with the following simplifications: (1) the variables $\boldsymbol{\alpha}(t)$ are already chosen to be deviations from the equilibrium or steady-state values, i.e., $\langle\boldsymbol{\alpha}(t)\rangle=0$; (2) the system behaves in a linear way with respect to these deviations $\boldsymbol{\alpha}$.

Equation (2.10) can be rewritten in the form

$$P(\boldsymbol{\alpha}_0\,|\,\boldsymbol{\alpha},t+\Delta t)=\int d\boldsymbol{\alpha}'P(\boldsymbol{\alpha}_0\,|\,\boldsymbol{\alpha}',t)P(\boldsymbol{\alpha}'\,|\,\boldsymbol{\alpha},\Delta t), \tag{3.1}$$

where the integral over $\boldsymbol{\alpha}'$ would be replaced by a sum if $\boldsymbol{\alpha}$ could take only discrete values. We have, in addition, the initial condition

$$\lim_{t\to0}P(\boldsymbol{\alpha}_0\,|\,\boldsymbol{\alpha},t)=\delta(\boldsymbol{\alpha}_0-\boldsymbol{\alpha}). \tag{3.2}$$

On multiplying Eq. (3.1) by $\boldsymbol{\alpha}$ and integrating over $\boldsymbol{\alpha}$, we obtain

$$\langle\boldsymbol{\alpha}(t+\Delta t)\rangle_{\alpha_0}=\int d\boldsymbol{\alpha}'P(\boldsymbol{\alpha}_0\,|\,\boldsymbol{\alpha}',t)\int P(\boldsymbol{\alpha}'\,|\,\boldsymbol{\alpha},\Delta t)\boldsymbol{\alpha}d\boldsymbol{\alpha}. \tag{3.3}$$

Our assumption that the system behaves in a linear way can now be expressed more precisely by the statement that the last integral in Eq. (3.3) is a linear function of the $\boldsymbol{\alpha}'$. In view of Eq. (3.2) we may take this linear function to have the form

$$\int P(\boldsymbol{\alpha}'\,|\,\boldsymbol{\alpha},\Delta t)\boldsymbol{\alpha}d\boldsymbol{\alpha}=\boldsymbol{\alpha}'-\Delta t\boldsymbol{\Lambda}\boldsymbol{\alpha}' \tag{3.4}$$

or

$$\int P(\boldsymbol{\alpha}'\,|\,\boldsymbol{\alpha},\Delta t)(\boldsymbol{\alpha}-\boldsymbol{\alpha}')d\boldsymbol{\alpha}=-\Delta t\boldsymbol{\Lambda}\boldsymbol{\alpha}'.$$

Equation (3.3) then reduces to

$$\langle\boldsymbol{\alpha}(t+\Delta t)\rangle_{\alpha_0}=\langle\boldsymbol{\alpha}(t)\rangle_{\alpha_0}-\boldsymbol{\Lambda}\Delta t\langle\boldsymbol{\alpha}(t)\rangle_{\alpha_0} \tag{3.5}$$

or

$$d\langle\boldsymbol{\alpha}(t)\rangle/dt=-\boldsymbol{\Lambda}\langle\boldsymbol{\alpha}(t)\rangle, \tag{3.6}$$

subject to $\langle\boldsymbol{\alpha}(0)\rangle=\boldsymbol{\alpha}_0$. Since $\boldsymbol{\alpha}$ constitutes a set of macroscopic variables, $\boldsymbol{\Lambda}$ is a matrix.

A formal solution of Eq. (3.6) can be written immediately:

$$\langle\boldsymbol{\alpha}(t)\rangle_{\alpha_0}=\exp(-\boldsymbol{\Lambda}t)\boldsymbol{\alpha}_0. \tag{3.7}$$

Comparison with Eq. (2.13) indicates that we can write the quantity we really wish to calculate in the form

$$\langle\boldsymbol{\alpha}(t)\boldsymbol{\alpha}(0)\rangle=\exp(-\boldsymbol{\Lambda}t)\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle, \tag{3.8}$$

where

$$\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle=\int\boldsymbol{\alpha}\boldsymbol{\alpha}W(\boldsymbol{\alpha})d\boldsymbol{\alpha} \tag{3.9}$$

has been written without the subscript zero, since $\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle$ is independent of time.

According to Eq. (2.7), with $I(t)$ replaced by $\boldsymbol{\alpha}(t)$, the positive frequency contribution to the noise can be written

$$\mathbf{G}_+(f)=2(i\omega+\boldsymbol{\Lambda})^{-1}\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle, \tag{3.10}$$

(where we write briefly ω, rather than ω times the unit matrix) and according to Eq. (2.9), the negative frequency contribution is

$$\mathbf{G}_-(f)=2\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle(-i\omega+\boldsymbol{\Lambda}^\dagger)^{-1}, \tag{3.11}$$

where $\boldsymbol{\Lambda}^\dagger$ is the transpose of $\boldsymbol{\Lambda}$. (If these results are generalized from real to complex variables, it is appropriate to interpret $\boldsymbol{\Lambda}^\dagger$ as the Hermitian adjoint matrix.)

If, for example, we were concerned with the noise in a variable $I(t)=\sum C_r\alpha_r$, it would be given by

$$G(I,f)=\textstyle\sum_{r,s}C_rG_{rs}(f)C_s, \tag{3.12}$$

with

$$G_{rs}(f)=2(i\omega+\boldsymbol{\Lambda})_{rk}^{-1}\langle\alpha_k\alpha_s\rangle+2\langle\alpha_r\alpha_k\rangle(-i\omega+\boldsymbol{\Lambda}^\dagger)_{ks}^{-1}, \tag{3.13}$$

with summation over k understood.

In order to obtain a better understanding of $\boldsymbol{\Lambda}$, we derive the value of $\boldsymbol{\Lambda}$ from an equation slightly less fundamental than (3.1). For many stochastic systems (Brownian motion is an exception), the probability $P(\boldsymbol{\alpha}'\,|\,\boldsymbol{\alpha},\Delta t)$ can be expanded for small Δt:

$$P(\boldsymbol{\alpha}'\,|\,\boldsymbol{\alpha},\Delta t)=\delta(\boldsymbol{\alpha}-\boldsymbol{\alpha}')(1-\Gamma_\alpha\Delta t)+\Delta tw_{\alpha\alpha'}, \tag{3.14}$$

where $w_{\alpha\alpha'}$ can be regarded as the transition probability per unit time from state α' to state α, and

$$\Gamma_\alpha=\int w_{\alpha'\alpha}d\boldsymbol{\alpha}' \tag{3.15}$$

is the total transition probability per unit time out of α. Thus Eq. (3.1) can be written in the familiar form

$$\partial P(\boldsymbol{\alpha},t)/\partial t=\int w_{\alpha\alpha'}d\boldsymbol{\alpha}'P(\boldsymbol{\alpha}',t)-\Gamma_\alpha P(\boldsymbol{\alpha},t) \tag{3.16}$$

$$=-\int T(\boldsymbol{\alpha},\boldsymbol{\alpha}')d\boldsymbol{\alpha}'P(\boldsymbol{\alpha}',t),$$

where

$$T(\boldsymbol{\alpha},\boldsymbol{\alpha}')=\Gamma_{\alpha'}\delta(\boldsymbol{\alpha}-\boldsymbol{\alpha}')-w_{\alpha\alpha'}. \tag{3.17}$$

The first term in (3.16) represents the rate of transitions from all $\boldsymbol{\alpha}'$ into $\boldsymbol{\alpha}$, whereas the second represents the total of transitions out of $\boldsymbol{\alpha}$. The solution of Eq. (3.16) appropriate to $P(\boldsymbol{\alpha}_0|\,\boldsymbol{\alpha}t)$ may be obtained by applying the initial condition

$$P(\boldsymbol{\alpha},0)=\delta(\boldsymbol{\alpha}_0-\boldsymbol{\alpha}). \tag{3.18}$$

For most noise problems, one starts by choosing a relevant set of variables $\boldsymbol{\alpha}$ and a set of transition probabilities $w_{\alpha\alpha'}$ that describe the stochastic process giving rise to the noise. It would therefore be convenient to express $\boldsymbol{\Lambda}$ in terms of $w_{\alpha\alpha'}$. By multiplying Eq. (3.16) by $\boldsymbol{\alpha}$ and integrating, we obtain Eq. (3.6) with $\boldsymbol{\Lambda}$ defined by

$$\begin{aligned}\boldsymbol{\Lambda}\boldsymbol{\alpha}'&=\int\boldsymbol{\alpha}T(\boldsymbol{\alpha},\boldsymbol{\alpha}')d\boldsymbol{\alpha}\\&=\Gamma_{\alpha'}\boldsymbol{\alpha}'-\int\boldsymbol{\alpha}w_{\alpha\alpha'}d\boldsymbol{\alpha}\end{aligned} \tag{3.19}$$

or

$$\boldsymbol{\Lambda}\boldsymbol{\alpha}'=-\int(\boldsymbol{\alpha}-\boldsymbol{\alpha}')w_{\alpha\alpha'}d\boldsymbol{\alpha} \tag{3.20}$$

so that $\boldsymbol{\Gamma}$ is a diagonal matrix, and both (3.19) and (3.20) are presumed to be linear in $\boldsymbol{\alpha}'$. (The vanishing of the constant term determines the steady state from which $\boldsymbol{\alpha}'$ is a fluctuation. See Sec. 5.) If $\boldsymbol{\alpha}$ is a discrete variable, integrals are to be replaced by sums.

To illustrate these remarks let us consider a simple example discussed in detail by Burgess[86]—the statistics of charge carrier fluctuations in semiconductors. When there are n carriers in the conduction band, let the rate of generation be $g(n)$, and the rate of recombination be $r(n)$, with n the only fluctuating variable. For example, for a strongly extrinsic n-type semiconductor with N donors and a negligible number of holes, one might take (with Burgess)

$$g=\gamma(N-n),\quad r=\rho n^2. \tag{3.21}$$

The macroscopic equation of motion is

$$dn/dt=g(n)-r(n). \tag{3.22}$$

The steady-state number of electrons n_0 obeys

$$g(n_0)=r(n_0) \tag{3.23}$$

so that in this strongly extrinsic case

$$n_0=-\tfrac{1}{2}K+(\tfrac{1}{4}K^2+KN)^{\frac{1}{2}}, \tag{3.24}$$

where $K=\gamma/\rho$ is the equilibrium constant for the ionization process. By expanding Eq. (3.22) about equilibrium, $n=n_0+\alpha$, we get

$$d\alpha/dt=-\Lambda\alpha;\quad \Lambda=r'(n_0)-g'(n_0)=2\rho n_0+\gamma \tag{3.25}$$

so that the decay rate is known from elementary considerations. If it were not, we could write the analog of Eq. (3.16), which gives a complete description of the stochastic process:

$$\begin{aligned}dP(n)/dt=&\,r(n+1)P(n+1)+g(n-1)P(n-1)\\&-[r(n)+g(n)]P(n),\end{aligned} \tag{3.26}$$

where $P(n)$ is the probability that there are precisely n electrons present. We see that $\Gamma=r(n)+g(n)$, and this equation can be written in the form

$$dP(n)/dt=-\textstyle\sum_{n'}T(n,n')P(n'), \tag{3.27}$$

where

$$T(n,n')=[r(n)+g(n)]\delta_{nn'}-r(n')\delta_{n',n+1}-g(n')\delta_{n',n-1}.$$

By using the notation

$$\langle f(n)\rangle=\textstyle\sum f(n')P(n'), \tag{3.28}$$

we can multiply Eq. (3.27) by n and sum to obtain

$$\begin{aligned}d\langle n\rangle/dt&=-\textstyle\sum_{n'}[\sum_n nT(n,n')]P(n')\\&=-\textstyle\sum_{n'}[r(n')-g(n')]P(n')\\&=-\langle r(n)-g(n)\rangle,\end{aligned} \tag{3.29}$$

thus recovering the macroscopic equation (3.22) and the results previously derived. Alternatively, one can work directly with the analog of Eq. (3.19) and linearize by using $n'=n_0+\alpha'$:

$$\begin{aligned}\textstyle\sum_n nT(n,n')&=r(n')-g(n')\\&=r(n_0)-g(n_0)+[r'(n_0)-g'(n_0)]\alpha'.\end{aligned} \tag{3.30}$$

In this way Λ appears as the coefficient of α' and the equilibrium *or steady-state* n_0 is determined by the vanishing of the constant term. If there were several random variables, the procedure would be exactly the same except that $\boldsymbol{\alpha}$ would be a set of variables and $\boldsymbol{\Lambda}$ a tensor as in Eq. (3.19).

We anticipate the methods of Sec. 5 by noting that

$$\begin{aligned}\textstyle\sum_n(n-n_0)^2T(n,n')=&-[g(n')+r(n')]\\&+2[r(n')-g(n')](n'-n_0).\end{aligned} \tag{3.31}$$

To the lowest nonvanishing under (for corrections see Sec. 14) the right-hand side, (3.31) becomes

$$-[g(n_0)+r(n_0)]+2[r'(n_0)-g'(n_0)](n'-n_0)^2$$

so that when Eq. (3.27) is multiplied by $(n-n_0)^2$ and summed over n, one obtains

$$d\langle\alpha^2\rangle/dt\simeq g(n_0)+r(n_0)-2\Lambda\langle\alpha^2\rangle, \tag{3.32}$$

where $\alpha=n-n_0$. The steady-state fluctuation obtained by setting Eq. (3.32) to zero is

$$\langle\alpha^2\rangle\simeq[g(n_0)+r(n_0)]/(2\Lambda)$$

or

$$\langle(n-n_0)^2\rangle\simeq g(n_0)/[r'(n_0)-g'(n_0)]. \tag{3.33}$$

This result agrees with the one obtained by Burgess[86] by the more difficult procedure of finding the steady-

[86] R. E. Burgess, Proc. Phys. Soc. (London) **B69**, 1020 (1956); **B68**, 661 (1955); Physica **20**, 1007 (1954).

state solution $P(n)$ of Eq. (3.26), locating its maximum at n_0, and approximating it by a Gaussian near $n=n_0$ so that

$$\langle(n-n_0)^2\rangle=-1/[d^2\ln P(n)/dn^2]_{n=n_0}. \quad (3.34)$$

For the special case defined by Eq. (3.21), Eq. (3.33) can be simplified to

$$\langle(\Delta n)^2\rangle=\frac{n_0^2}{2n_0+K}=\frac{n_0(N-n_0)}{2N-n_0}. \quad (3.35)$$

Equation (3.33) has been referred to as the generation-recombination theorem by Van Vliet and Blok[87] and has been generalized by them to the case of a set of occupancies n_i.

4. Fluctuations at a Given Instant from Equilibrium

We have given a formal expression [Eqs. (3.10–3.11)] for the noise associated with $\boldsymbol{\alpha}$ in terms of a matrix $\boldsymbol{\Lambda}$ that is known macroscopically [or via Eq. (3.20)] and the moments $\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle$ associated with a pair of variables at the same instant of time. When we are dealing with deviations from an equilibrium state, the fluctuation moments $\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle$ can be evaluated by thermodynamic methods. We therefore consider first the equilibrium case and return later to the more general problem of fluctuations from a nonequilibrium steady state.

Einstein[22] founded the macroscopic theory of thermodynamic fluctuations by suggesting that Boltzmann's principle expressing the entropy S as Boltzmann's constant times the logarithm of the probability of a given fluctuation,

$$S(\boldsymbol{\alpha})=k\ln W(\boldsymbol{\alpha}), \quad (4.1)$$

can be inverted to yield W if the entropy is known. In the neighborhood of equilibrium, the entropy may be expanded as

$$S=S_0-\tfrac{1}{2}\textstyle\sum_{ij}s_{ij}\alpha_i\alpha_j+\text{higher terms}. \quad (4.2)$$

The Einstein treatment of fluctuations then yields

$$W(\boldsymbol{\alpha})\simeq N\exp(-\tfrac{1}{2}\boldsymbol{\alpha}\cdot\mathbf{s}\cdot\boldsymbol{\alpha}/k), \quad (4.3)$$

where N is a factor chosen to provide normalization. By omitting the higher terms, the approximation of treating the α's as Gaussian variables is made. It is then a purely mathematical exercise[88] to calculate the moments of this distribution,

$$\langle\alpha_i\alpha_j\rangle=k(s^{-1})_{ij}$$

or

$$\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle=k\mathbf{s}^{-1}, \quad (4.4)$$

where $\mathbf{s}^{-1}$ is the matrix reciprocal to the one whose elements are s_{ij}.

[87] K. M. van Vliet and J. Blok, Physica **22**, 231 (1956).

[88] See, for example, S. O. Rice, reference 22, Eq. (2.9-1) and H. Cramer, *Mathematical Methods of Statistics* (Princeton University Press, Princeton, New Jersey, 1946).

In order to express these moments in more familiar language, we recall that thermodynamic forces are usually defined in terms of derivatives of one of the thermodynamic functions. By taking entropy as basic thermodynamic potential, the force conjugate to α_i and tending to restore the latter to equilibrium may, in Onsager's[73] notation, be defined by

$$X_i=\partial S/\partial\alpha_i=-\textstyle\sum_j s_{ij}\alpha_j, \quad (4.5)$$

therefore

$$s_{ij}=-\partial^2 S/\partial\alpha_i\partial\alpha_j=-(\partial X_i/\partial\alpha_j). \quad (4.6)$$

(If α_i is a fluctuation in a variable a_j so that

$$\alpha_j=a_j-\langle a_j\rangle, \quad (4.7)$$

then $\partial X_i/\partial\alpha_j$ means $\partial X_i/\partial a_j$ evaluated at the equilibrium value $a_j=\langle a_j\rangle$.)

The matrix equation (4.5) can be inverted to give

$$\alpha_i=-\textstyle\sum_i(s^{-1})_{ij}X_j \quad (4.8)$$

so that

$$(s^{-1})_{ij}=-\partial\alpha_i/\partial X_j=-\partial a_i/\partial X_j \quad (4.9)$$

expresses the change in the equilibrium a_i per unit change in applied force X_j from the equilibrium force.

The form of Eq. (4.2), containing no linear terms, implies in view of Eq. (4.5), that at equilibrium (when $\boldsymbol{\alpha}=0$) the forces X_i vanish. One might think, therefore, that Eq. (4.3) is limited in usefulness to describing the fluctuations of internal variables because external variables are either subject to an external force or to a constraint, in which case they could not fluctuate. For example, one could break up a liquid into cells and consider the fluctuation in the number of particles in each cell, but the total number of particles would be fixed.[89] Under such circumstances, one would also be restricted to considering fluctuations about an equilibrium state of uniform density.

Fortunately, the microscopic theory of fluctuations has been adequately developed from a canonical point of view, i.e., for systems interacting with reservoirs that apply external forces to the system.[6] An elegant and critical development of this theory has been given by Greene and Callen[49] which leads to a simple derivation of the macroscopic "Einstein" distribution function. The result takes the form

$$W(\boldsymbol{\alpha})=N\exp\frac{1}{k}[S(\boldsymbol{\alpha})-S_0-\sum_{i=1}^{r}X_i\alpha_i]. \quad (4.10)$$

Now $S(\boldsymbol{\alpha})-S_0$ possesses linear terms in $\boldsymbol{\alpha}$, but these terms are precisely canceled by $\sum X_i\alpha_i$. Thus, Einstein's distribution is precisely correct, with the linear terms omitted but quadratic and all higher-order terms kept.

One might expect that the second moments, Eq. (4.4) computed from Eq. (4.3), with neglect of the higher-order terms, would be good approximations to the exact second moments. The surprising result estab-

[89] M. Klein and L. Tisza, Phys. Rev. **76**, 1861 (1949).

lished by Greene and Callen is that these "approximate" second moments are *exact*. By the use of these moments, we have then in no way committed ourselves to the approximation that our variables are Gaussian (although for most macroscopic variables that might very well be a good approximation). In addition, we are permitted to use these moments when the parameters are subject to external forces $X_i \neq 0$.

Thermodynamic forces are more conventionally defined in terms of derivatives of the internal energy,

$$P_i = \partial U(a_1, a_2, \cdots, a_r)/\partial a_i. \tag{4.11}$$

The P_i represent external forces acting *on* the system. Work done *by* the system is then

$$\delta W = -\sum P_i da_i, \tag{4.12}$$

the heat flow into the system is

$$\delta Q = \delta U - \sum P_i da_i, \tag{4.13}$$

and the entropy change is given by

$$dS = \delta Q/T = (1/T)dU - \sum_{i=1}^{r} (P_i/T) da_i. \tag{4.14}$$

Thus, we may regard the entropy as a function of $U = a_0$, a_1, a_2, $\cdots$, a_r. The entropy definition of force then yields

$$X_i = \partial S/\partial a_i = -P_i/T, \tag{4.15}$$

and in accordance with (4.14), we may adopt the convention $P_0 = -1$, $X_0 = 1/T$. We should remember, however, that pressure is defined conventionally as a force of the system on the surroundings so that if $a_1 = V =$ volume, $P_1 = -P$.

By combining Eqs. (4.4) and (4.9), we have

$$\langle \alpha_i \alpha_j \rangle = -k(\partial a_i/\partial X_j)_X, \tag{4.16}$$

where the subscript X is a reminder that all X's but X_j are to be held fixed. Providing $j \neq 0$, i.e., $\alpha_j \neq \Delta E =$ the energy change, T is among the variables held fixed, and in view of (4.15), we can write

$$\langle \alpha_i \alpha_j \rangle = kT(\partial a_i/\partial P_j)_{T,P}, \tag{4.17}$$

where T and all but one P are now held fixed. [Under an adiabatic constraint, one holds S rather than T fixed.[90]]

In addition to the variables $a_1, \cdots, a_r$ that we allow to vary, there are additional variables $a_{r+1}, a_{r+2}, \cdots$ that are understood to be kept fixed; and these are kept fixed during the differentiations in (4.16) and (4.17). For example, energy fluctuations at constant volume are given from (4.16) by

$$\langle (\Delta E)^2 \rangle = -k[\partial U/\partial(1/T)]_V = kC_v T^2, \tag{4.18}$$

whereas if both energy and volume are allowed to vary, we use (4.16) to obtain a less familiar result,

$$\langle (\Delta E)^2 \rangle = -k[\partial U/\partial(1/T)]_{(P/T)} = kT\{TC_P - P[T(\partial V/\partial T)_P + P(\partial V/\partial P)_T]\}. \tag{4.19}$$

Equation (4.17) now may be used to obtain the volume fluctuations

$$\langle (\Delta V)^2 \rangle = -kT(\partial V/\partial P)_T \tag{4.20}$$

and the correlation

$$\langle \Delta E \Delta V \rangle = -kT(\partial U/\partial P)_T = kT[T(\partial V/\partial T)_P + P(\partial V/\partial P)_T], \tag{4.21}$$

whereas one must use (4.16) to obtain

$$\langle \Delta V \Delta E \rangle = -k[\partial V/\partial(1/T)]_{(P/T)}, \tag{4.22}$$

which can be reduced to (4.21) after some manipulation.

Let us now consider an example that is useful later: a nondegenerate n-type semiconductor with n electrons in the conduction band. If the conduction band is in contact with a reservoir of chemical potential μ, the fluctuation in the number of carriers, according to Eq. (4.17), is

$$\langle (\Delta n)^2 \rangle = kT(\partial n/\partial \mu)_T. \tag{4.23}$$

Shockley[91] has shown, however, that for a set of noninteracting electrons, in the nondegenerate case when Boltzmann statistics are applicable,[92] the mean number of electrons is given by

$$n = N_c \exp[(\mu - E_c)/kT], \tag{4.24}$$

where μ is also referred to as the Fermi level, E_c is the energy of the bottom of the conduction band, and N_c is the effective number of states in the conduction band. By combining (4.23) and (4.24), we obtain the expected result

$$\langle (\Delta n)^2 \rangle = n. \tag{4.25}$$

If there are N_t donor states or traps occupied, on the average by $\hat{n}$ electrons in contact with a reservoir of chemical potential (= Fermi level) μ_t, then, by using Fermi statistics,

$$\hat{n} = N_t/\{1 + \exp[(E_t - \mu_t)/kT]\} \tag{4.26}$$

and

$$\langle (\Delta \hat{n})^2 \rangle = kT(\partial \hat{n}/\partial \mu_t)_T = \hat{n}(N_t - \hat{n})/N_t. \tag{4.27}$$

If both n and $\hat{n}$ were allowed to vary simultaneously, we would expect a distribution function

$$W \propto \exp\left[-\frac{(\Delta n)^2}{2\langle (\Delta n)^2 \rangle} - \frac{(\Delta \hat{n})^2}{2\langle (\Delta \hat{n})^2 \rangle} \right], \tag{4.28}$$

the term in $\Delta n \Delta \hat{n}$ vanishing because n does not depend on μ_t nor $\hat{n}$ on μ. If, however, n does not have its private

[90] See Appendix A of reference 71 or Appendix B of reference 70.

[91] W. Shockley, *Electrons and Holes in Semi-Conductors* (D. Van Nostrand Company, Inc., Princeton, New Jersey, 1950).

[92] D. J. Oliver, Proc. Phys. Soc. (London) **B70**, 244 (1957), discusses the case in which the number of free carriers is so high as to require the use of Fermi statistics.

reservoir but derives its electrons from the donors, then we must impose the condition

$$\Delta n = -\Delta \hat{n}. \tag{4.29}$$

When this condition is imposed on (4.28), we find that the new fluctuations are given in terms of the old by

$$\langle(\Delta n)^2\rangle = \langle(\Delta \hat{n})^2\rangle = \frac{\langle(\Delta n)^2\rangle\langle(\Delta \hat{n})^2\rangle}{\langle(\Delta n)^2\rangle + \langle(\Delta \hat{n})^2\rangle} = \frac{n\hat{n}(N_t-\hat{n})}{nN_t+\hat{n}(N_t-\hat{n})}. \tag{4.30}$$

In the presence of N_{co} compensating centers, Eq. (4.30) is to be used in conjunction with

$$n+\hat{n} = N_t - N_{co}. \tag{4.31}$$

In the absence of compensating centers, Eq. (4.30) reduces to the result [Eq. (3.35)] previously obtained by nonthermodynamic methods. If, in the "kinetic" method, one chooses an equilibrium constant K (which measures the ratio of ionization to capture rates) that disagrees with the thermal equilibrium values, then Eq. (3.35) remains correct, and so does Eq. (4.30). Thus, the thermodynamic method has a more general validity than one might have suspected.

Since the cross section for electron capture is sensitive to the electron energy, one may drive K from its equilibrium value by changing the electron energy distribution via a strong electric field. The value of K then is found, by combining Eqs. (3.23) and (3.21), to be

$$K = \gamma/\rho = n_0^2/(N-n_0) = n_0(N-\hat{n}_0)/\hat{n}_0. \tag{4.32}$$

If n_0 and $\hat{n}_0$ are expressed in terms of μ and μ_t by Eqs. (4.24) and (4.26), we obtain

$$K = N_c \exp[(E_t-E_c)/kT] \exp[(\mu-\mu_t)/kT]. \tag{4.33}$$

The equilibrium value of K is obtained by setting $\mu=\mu_t$. A strong field causing K to differ from its equilibrium value then can be handled in the thermodynamic context by saying that it permits (and indeed insures) that $\mu \neq \mu_t$. It is as if we are dealing with a system in equilibrium with two reservoirs subject to the constraint $\Delta n = -\Delta \hat{n}$. A justification of this viewpoint is given in Sec. 12.

5. Fluctuations at One Time from a Steady State

Although, as shown in the last section, thermodynamic methods are easy to apply and are even applicable to some cases that would not ordinarily be considered in thermodynamic equilibrium, we must now return to the general case of fluctuations from a steady state where thermodynamic methods may be open to question.

Let us deal with a set of variables $\mathbf{a}$ whose steady-state values are $\mathbf{a}_0$ and fluctuations are $\boldsymbol{\alpha} = \mathbf{a} - \mathbf{a}_0$. Because of the Markoffian character of our system, the distribution function $P(\mathbf{a},t)$ of $\mathbf{a}$ at the time t obeys an equation of the form

$$P(\mathbf{a}, t+\Delta t) = \int d\mathbf{a}' P(\mathbf{a}',t) P(\mathbf{a}' \mid \mathbf{a},\Delta t), \tag{5.1}$$

where $P(\mathbf{a}' \mid \mathbf{a},\Delta t)$ is the probability of a transition from $\mathbf{a}'$ to $\mathbf{a}$ in time Δt. If we were to use the initial condition

$$P(\mathbf{a},0) = \delta(\mathbf{a}-\mathbf{a}^0), \tag{5.2}$$

the solution would be the conditional probability $P(\mathbf{a}^0 \mid \mathbf{a},t)$. However, we need not restrict ourselves to this initial condition. Our transition probability is assumed known and must obey the normalization and initial conditions

$$\int P(\mathbf{a}' \mid \mathbf{a},\Delta t) d\mathbf{a} = 1, \tag{5.3}$$

$$P(\mathbf{a}' \mid \mathbf{a},0) = \delta(\mathbf{a}-\mathbf{a}'). \tag{5.4}$$

As a consequence, any moment of $\mathbf{a}-\mathbf{a}'$ taken with respect to $P(\mathbf{a}' \mid \mathbf{a},\Delta t)$ vanishes when $\Delta t \to 0$. We assume, however, that the first two moments are expandable in powers of Δt; therefore, for small Δt,

$$\int P(\mathbf{a}' \mid \mathbf{a},\Delta t)(\mathbf{a}-\mathbf{a}') d\mathbf{a} = \mathbf{A}(\mathbf{a}')\Delta t, \tag{5.5}$$

$$\int P(\mathbf{a}' \mid \mathbf{a},\Delta t)(\mathbf{a}-\mathbf{a}')(\mathbf{a}-\mathbf{a}') d\mathbf{a} = 2\mathbf{D}(\mathbf{a}')\Delta t, \tag{5.6}$$

where $\mathbf{A}$ is the "drift vector" and $\mathbf{D}$ is the diffusion matrix, since it describes the mean square displacement in the time Δt,

$$\mathbf{D} = \langle \Delta\mathbf{a}\Delta\mathbf{a}\rangle/(2\Delta t). \tag{5.7}$$

If we multiply Eq. (5.1) by $\mathbf{a}$, and on the right-hand side use $\mathbf{a}-\mathbf{a}'+\mathbf{a}'$, we obtain

$$\langle \mathbf{a}(t+\Delta t)\rangle = \langle \mathbf{a}(t)\rangle + \langle \mathbf{A}[\mathbf{a}(t)]\rangle \Delta t \tag{5.8}$$

or

$$d\langle \mathbf{a}(t)\rangle/dt = \langle \mathbf{A}[\mathbf{a}(t)]\rangle. \tag{5.9}$$

If we multiply Eq. (5.1) on the left by $(\mathbf{a}-\mathbf{a}_0)(\mathbf{a}-\mathbf{a}_0)$ and on the right by its equivalent

$$(\mathbf{a}-\mathbf{a}_0)(\mathbf{a}-\mathbf{a}_0) = (\mathbf{a}'-\mathbf{a}_0)(\mathbf{a}'-\mathbf{a}_0) + (\mathbf{a}-\mathbf{a}')(\mathbf{a}-\mathbf{a}') + (\mathbf{a}-\mathbf{a}')(\mathbf{a}'-\mathbf{a}_0) + (\mathbf{a}'-\mathbf{a}_0)(\mathbf{a}-\mathbf{a}'), \tag{5.10}$$

we obtain

$$\langle \boldsymbol{\alpha}(t+\Delta t)\boldsymbol{\alpha}(t+\Delta t)\rangle = \langle \boldsymbol{\alpha}(t)\boldsymbol{\alpha}(t)\rangle + 2\langle \mathbf{D}(\mathbf{a})\rangle \Delta t + \langle A[\mathbf{a}(t)]\boldsymbol{\alpha}(t)\rangle \Delta t + \langle \boldsymbol{\alpha}(t) A[\mathbf{a}(t)]\rangle \Delta t \tag{5.11}$$

or

$$d\langle \boldsymbol{\alpha}(t)\boldsymbol{\alpha}(t)\rangle/dt = 2\langle \mathbf{D}(\mathbf{a})\rangle + \langle \mathbf{A}(\mathbf{a})\boldsymbol{\alpha}\rangle + \langle \boldsymbol{\alpha}\mathbf{A}(\mathbf{a})\rangle. \tag{5.12}$$

We now make use of the quasi-linear nature of most noise problems to insert $\mathbf{a}=\mathbf{a}_0+\boldsymbol{\alpha}$ and retain only the nonvanishing terms of lowest order. In particular, we may write

$$\mathbf{A}(\mathbf{a})\simeq\mathbf{A}(\mathbf{a}_0)-\boldsymbol{\Lambda}\boldsymbol{\alpha}, \tag{5.13}$$

$$\mathbf{D}(\mathbf{a})\simeq\mathbf{D}(\mathbf{a}_0)=\mathbf{D}, \tag{5.14}$$

where a comparison with Eq. (5.9) indicates that the steady-state values $\mathbf{a}_0$ should be chosen so that

$$\mathbf{A}(\mathbf{a}_0)=0, \tag{5.15}$$

and therefore

$$d\langle\boldsymbol{\alpha}\rangle/dt=-\boldsymbol{\Lambda}\langle\boldsymbol{\alpha}\rangle, \tag{5.16}$$

and

$$d\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle/dt=2\mathbf{D}-\boldsymbol{\Lambda}\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle-\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle\boldsymbol{\Lambda}^\dagger, \tag{5.17}$$

where $\boldsymbol{\Lambda}^\dagger$ is the transpose of $\boldsymbol{\Lambda}$. The steady-state moments $\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle$ then can be obtained by solving the equation

$$2\mathbf{D}=\boldsymbol{\Lambda}\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle+\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle\boldsymbol{\Lambda}^\dagger. \tag{5.18}$$

As we demonstrate later, this equation is a generalized Einstein relation between diffusion constants and mobilities.

The diffusion constant can be seen to be symmetric from its definition (5.6) as well as from (5.18).

Equation (5.18) can be solved readily in a representation in which $\boldsymbol{\Lambda}$ is diagonal and then transformed back to the original representation. It is easier to guess that the solution has the form

$$\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle=2\int_0^\infty \exp(-\boldsymbol{\Lambda}t)\mathbf{D}\exp(-\boldsymbol{\Lambda}^\dagger t)dt, \tag{5.19}$$

and verify that this result is correct by inserting into (5.18) and integrating one term by parts.

We show from time reversibility in the next section that $\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle\boldsymbol{\Lambda}^\dagger=\boldsymbol{\Lambda}\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle$ so that the moments for time reversible problems are more easily computed by means of

$$\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle=\boldsymbol{\Lambda}^{-1}\mathbf{D}=\mathbf{D}(\boldsymbol{\Lambda}^\dagger)^{-1} \tag{5.20}$$

from $\boldsymbol{\Lambda}$ and $\mathbf{D}$, and the latter are simply related to the first two moments of the transition probability [Eqs. (5.5) and (5.6)].

For those problems in which the transition probability is itself expandible in Δt

$$P(\mathbf{a}'\,|\,\mathbf{a},\Delta t)=\delta(\mathbf{a}-\mathbf{a}')(1-\Gamma\Delta t)+w_{\mathbf{a}\mathbf{a}'}\Delta t, \tag{5.21}$$

we have

$$\mathbf{A}(\mathbf{a}')=\int(\mathbf{a}-\mathbf{a}')w_{\mathbf{a}\mathbf{a}'}d\boldsymbol{\alpha}\simeq A(\mathbf{a}_0)-\boldsymbol{\Lambda}(\mathbf{a}'-\mathbf{a}_0) \tag{5.22}$$

and

$$\mathbf{D}(\mathbf{a}')=\tfrac{1}{2}\int(\mathbf{a}-\mathbf{a}')(\mathbf{a}-\mathbf{a}')w_{\mathbf{a}\mathbf{a}'}d\mathbf{a}, \tag{5.23}$$

therefore

$$\boldsymbol{\Lambda}\boldsymbol{\alpha}'=-\int(\boldsymbol{\alpha}-\boldsymbol{\alpha}')w_{\alpha,\alpha'}d\boldsymbol{\alpha}, \tag{5.24}$$

in agreement with our previous definition (3.20), and

$$\mathbf{D}=\tfrac{1}{2}\int(\boldsymbol{\alpha}-\boldsymbol{\alpha}')(\boldsymbol{\alpha}-\boldsymbol{\alpha}')w_{\alpha\alpha'}d\boldsymbol{\alpha} \tag{5.25}$$

is expressible directly in terms of the transition probabilities that describe our stochastic process.

When dealing with fluctuations from an equilibrium state, it is often more convenient to use the Einstein relation (5.18) to evaluate $\mathbf{D}$ from the fluctuations $\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle$ given by the thermodynamic formula (4.16).

A remark concerning the relationship between the procedure we have described and the Fokker-Planck method is in order here.[54] If the transition probability has the character that all moments higher than the second vanish faster than Δt,

$$\lim_{\Delta t\to 0}(\Delta t)^{-1}\int d\mathbf{a}P(\mathbf{a}'\,|\,\mathbf{a},\Delta t)\,|\,(\mathbf{a}-\mathbf{a}')^n\,|=0,\quad n>2, \tag{5.26}$$

then by expanding in powers of $\mathbf{a}'-\mathbf{a}$, the Smoluchowski integral equation (5.1) is converted into the Fokker-Planck differential equation,[93–95]

$$\frac{\partial P(\mathbf{a},t)}{\partial t}=-\sum_i\frac{\partial}{\partial a_i}[A_i(\mathbf{a})P]+\sum_{ij}\frac{\partial}{\partial a_i}\frac{\partial}{\partial a_j}[D_{ij}(\mathbf{a})P], \tag{5.27}$$

where terms of higher derivative than the second do not appear because of (5.26).

The condition (5.26) on the moments would be satisfied for an idealized Brownian motion for which the transition probability has the form

$$P(a'\,|\,a,\Delta t)\sim\exp[-(a-a')^2/(4D\Delta t)],$$

but it would not be satisfied for any process in which $P(a'\,|\,a,\Delta t)$ is itself expandable in Δt as in Eq. (5.21), for then *all* the moments vanish only linearly in Δt.

On introducing the linearization approximation, the Fokker-Planck equation becomes

$$\frac{\partial P(\boldsymbol{\alpha},t)}{\partial t}=\sum_{ij}\left[\frac{\partial}{\partial\alpha_i}(\Lambda_{ij}\alpha_j P)+D_{ij}\frac{\partial}{\partial\alpha_i}\frac{\partial}{\partial\alpha_j}P\right]. \tag{5.28}$$

We seek a stationary solution of the form

$$P(\boldsymbol{\alpha})\propto\exp[-\tfrac{1}{2}\boldsymbol{\alpha}\cdot\mathbf{S}\cdot\boldsymbol{\alpha}], \tag{5.29}$$

where $\mathbf{S}$ must be a symmetric matrix. Since

$$\partial P/\partial\boldsymbol{\alpha}=-\mathbf{S}\cdot\boldsymbol{\alpha}P, \tag{5.30}$$

the stationary Fokker-Planck equation becomes

$$0=\frac{\partial}{\partial\boldsymbol{\alpha}}(\boldsymbol{\Lambda}\cdot\boldsymbol{\alpha}P)+\frac{\partial}{\partial\boldsymbol{\alpha}}\cdot\mathbf{D}\cdot\frac{\partial}{\partial\boldsymbol{\alpha}}P \tag{5.31}$$

$$=\partial(\mathbf{L}\cdot\boldsymbol{\alpha}P)/\partial\boldsymbol{\alpha}, \tag{5.32}$$

[93] See references 19, 21, and 54. Attempts to provide a justification for the Fokker-Planck approximation have been given by Green, reference 85, and by references 94 and 95.

[94] N. G. Van Kampen, Physica **23**, 707 (1957).

[95] R. T. Cox, Revs. Modern Phys. **22**, 238 (1950); **24**, 312 (1952).

where

$$\mathbf{L}=\mathbf{\Lambda}-\mathbf{D}\cdot\mathbf{S}. \tag{5.33}$$

Thus

$$\mathbf{S}=\mathbf{D}^{-1}\cdot\mathbf{\Lambda} \tag{5.34}$$

makes $\mathbf{L}$ vanish, and yields a solution, provided $\mathbf{D}^{-1}\cdot\mathbf{\Lambda}$ is a symmetric matrix. (For systems having an underlying time reversibility, this is shown to be true in the next section.) In any case, the differentiations in Eq. (5.32) can be completed to yield

$$0=[\text{trace}\mathbf{L}-(\mathbf{L}\boldsymbol{\alpha}\cdot\mathbf{S}\boldsymbol{\alpha})]P. \tag{5.35}$$

But

$$\mathbf{L}\boldsymbol{\alpha}\cdot\mathbf{S}\boldsymbol{\alpha}=\boldsymbol{\alpha}\cdot\mathbf{L}^{\dagger}\mathbf{S}\cdot\boldsymbol{\alpha}=\tfrac{1}{2}[\boldsymbol{\alpha}\cdot\mathbf{L}^{\dagger}\mathbf{S}+\mathbf{S}^{\dagger}\mathbf{L}\cdot\boldsymbol{\alpha}], \tag{5.36}$$

where the symbol † stands for transpose, so that the Fokker-Planck equation is satisfied if

$$\text{trace}\mathbf{L}=\text{trace}(\mathbf{\Lambda}-\mathbf{DS})=0 \tag{5.37}$$

and

$$\mathbf{S}^{\dagger}\mathbf{L}+\mathbf{L}^{\dagger}\mathbf{S}=0 \tag{5.38}$$

or since $\mathbf{S}$ is symmetric, we have, by using Eq. (5.33),

$$\mathbf{S\Lambda}+\mathbf{\Lambda}^{\dagger}\mathbf{S}=2\mathbf{SDS}. \tag{5.39}$$

The trace of

$$\mathbf{\Lambda}+\mathbf{S}^{-1}\mathbf{\Lambda}^{\dagger}\mathbf{S}=2\mathbf{DS} \tag{5.40}$$

yields Eq. (5.37) so that we only need solve (5.39). The latter, however, can be rewritten in the form

$$\mathbf{\Lambda}\mathbf{S}^{-1}+\mathbf{S}^{-1}\mathbf{\Lambda}^{\dagger}=2\mathbf{D}, \tag{5.41}$$

displaying its identity with the result (5.18) obtained by our methods, when it is recalled that a distribution (5.29) leads to moments

$$\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle=\mathbf{S}^{-1}. \tag{5.42}$$

[In the equilibrium case $\mathbf{S}=\mathbf{s}/k$, cf. Eq. (4.4), where $\mathbf{s}$ is related to the entropy by (4.2).]

In summary, a linearized Fokker-Planck approximation is equivalent to treating the $\boldsymbol{\alpha}$ as Gaussian variables. Since $P(\boldsymbol{\alpha})$ is then completely specified by the moments, it could have been written down directly from the moments computed by our linearization procedure, which does not assume the $\boldsymbol{\alpha}$ are necessarily Gaussian.

We may further remark, that the time displaced moments $\langle\boldsymbol{\alpha}(t)\boldsymbol{\alpha}(0)\rangle$ have been calculated by a procedure which treats $\boldsymbol{\alpha}(t)$ as a linear function of the $\boldsymbol{\alpha}(0)$, as in Eq. (3.7), so that if the $\boldsymbol{\alpha}(0)$ are Gaussian variables so are the $\boldsymbol{\alpha}(t)$. The joint distribution of $\boldsymbol{\alpha}(0)$ and $\boldsymbol{\alpha}(t)$ is then a Gaussian, with known moments [Eq. (3.8)], and the Gaussian constructed with these moments is automatically a solution of the time-dependent Fokker-Planck equation (4.28).

Since the calculation of the noise spectrum requires only the knowledge of the moments and not the complete distribution, our discussion of the Fokker-Planck equation was merely a digression to show the connection with the work of others.

CONNECTIONS WITH ONSAGER, EINSTEIN, LANGEVIN, AND THEVENIN

6. Time Reversibility

We have succeeded in obtaining formal expressions for the noise associated with arbitrary quasi-linear, Markoffian processes. There are some simplifications that can be made for processes that have an underlying time-reversible basis—even when the macroscopic equations used possess an irreversible character.

Onsager[72] has remarked that in the absence of magnetic fields and Coriolis forces our statistical mechanical systems obey time reversibility, whose macroscopic consequence can be expressed in the form

$$\langle\boldsymbol{\alpha}(t)\boldsymbol{\alpha}\rangle=\langle\boldsymbol{\alpha}(-t)\boldsymbol{\alpha}\rangle. \tag{6.1}$$

(We have tacitly assumed that all the quantities $\boldsymbol{\alpha}$ are even under time reversal. See Sec. 10 for odd variables.) By applying stationarity, Eq. (2.2), we get

$$\langle\boldsymbol{\alpha}(t)\boldsymbol{\alpha}\rangle=\langle\boldsymbol{\alpha}\boldsymbol{\alpha}(t)\rangle. \tag{6.2}$$

Equation (6.2) says simply that $\langle\boldsymbol{\alpha}(t)\boldsymbol{\alpha}\rangle=\exp(-\mathbf{\Lambda}t)\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle$ is a symmetric matrix, i.e.,

$$\exp(-\mathbf{\Lambda}t)\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle=\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle\exp(-\mathbf{\Lambda}^{\dagger}t). \tag{6.3}$$

By applying this condition at small times, we find that

$$\mathbf{\Lambda}\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle=\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle\mathbf{\Lambda}^{\dagger} \tag{6.4}$$

is also a symmetric matrix (for steady-state as well as equilibrium processes).

We now make connection with Onsager's notation to demonstrate that Eq. (6.4) is a statement of the Onsager relations. Onsager chooses to represent the response of the fluxes $\dot{\alpha}_i$ to the thermodynamic driving forces $X_i=\partial S/\partial\alpha_i=-\sum s_{ij}\alpha_j$ by

$$\dot{\alpha}_i=\sum L_{ij}X_j \tag{6.5}$$

or

$$\sum R_{ij}\dot{\alpha}_j=X_i, \tag{6.6}$$

where

$$\mathbf{R}\cdot\mathbf{L}=1. \tag{6.7}$$

R is a "resistance" and L is a "conductance" or mobility matrix. Combination of (6.5) with the definition (4.5) of X leads to

$$d\boldsymbol{\alpha}/dt=\mathbf{LX}=-\mathbf{Ls}\boldsymbol{\alpha}, \tag{6.8}$$

and comparison with Eq. (3.6) leads to the identification

$$\mathbf{\Lambda}=\mathbf{Ls}. \tag{6.9}$$

Thus

$$\mathbf{\Lambda}\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle=k\mathbf{L}\cdot\mathbf{C}, \tag{6.10}$$

where

$$\mathbf{C}=\mathbf{s}\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle/k \tag{6.11}$$

is a correction factor that reduces to unity in the classical thermal equilibrium case, cf. Eq. (4.4). Time reversibility expressed by Eq. (6.4) then leads to the

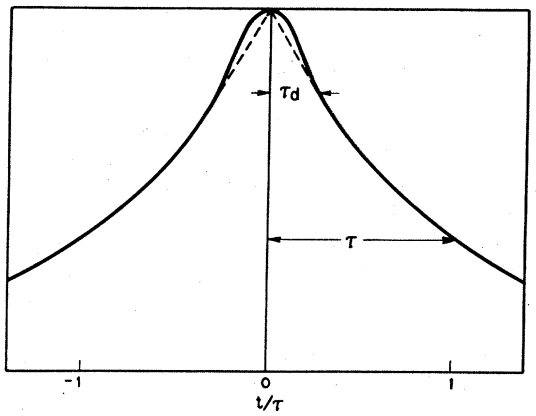

FIG. 1. The true regression of a fluctuation (solid curve) is compared to the Markoffian approximation (dashed curve). For $t>\tau_d$, the decay is $\exp(-t/\tau)$, where τ is a typical relaxation time and τ_d is the "duration of a collision" or the forgetting time of the system.

Onsager relations

$$\mathbf{LC}=\mathbf{C}^\dagger\mathbf{L}^\dagger \tag{6.12}$$

or

$$\mathbf{CR}^\dagger=\mathbf{RC}^\dagger, \tag{6.13}$$

i.e., $\mathbf{LC}$ and $\mathbf{RC}^\dagger$ are symmetric matrices. We refer to Eqs. (6.12) and (6.13) as Onsager relations since, in classical thermal equilibrium when $\mathbf{C}=1$, they reduce to the Onsager requirement that $\mathbf{L}$ and $\mathbf{R}=\mathbf{L}^{-1}$ be symmetric. Practically the whole of irreversible thermodynamics[9] consists at present in the application of the Onsager relations with $\mathbf{C}=1$.

In the time-reversible case, the Einstein relation Eq. (5.18) can be written, via (6.4) and (6.10), in the form

$$D=k\mathbf{LC}=kT\mathbf{L}'\cdot\mathbf{C}, \tag{6.14}$$

which is comparable to Eq. (1.6) with the correction factor $\mathbf{C}$. $\mathbf{L}$ is Onsager's conductance when $-\mathbf{X}$ is used as a driving force *on* the system. The more conventional choice of driving force is $\mathbf{P}=-T\mathbf{X}$, as in Eq. (4.15), so that the conventional admittance is

$$L'=L/T \tag{6.15}$$

(see also Sec. 7).

The factor kT in Eq. (6.14) arises because all fluctuations $\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle$ from an equilibrium state are proportional to kT. The fluctuations from a driven state may no longer be proportional to kT, and this finds expression in the correction factor $\mathbf{C}$.

In the presence of a magnetic field, time reversibility is obeyed if one simultaneously reverses the magnetic field,

$$\langle\boldsymbol{\alpha}(t,\mathbf{H})\boldsymbol{\alpha}(0,\mathbf{H})\rangle=\langle\boldsymbol{\alpha}(-t,-\mathbf{H})\boldsymbol{\alpha}(0,-\mathbf{H})\rangle, \tag{6.16}$$

which leads to the generalized Onsager relations

$$\boldsymbol{\Lambda}(\mathbf{H})\langle\boldsymbol{\alpha}(\mathbf{H})\boldsymbol{\alpha}(\mathbf{H})\rangle=\langle\boldsymbol{\alpha}(-\mathbf{H})\boldsymbol{\alpha}(-\mathbf{H})\rangle\boldsymbol{\Lambda}^\dagger(-H) \tag{6.17}$$

or

$$\mathbf{L}(\mathbf{H})\mathbf{C}(\mathbf{H})=\mathbf{C}^\dagger(-\mathbf{H})\mathbf{L}^\dagger(-\mathbf{H}). \tag{6.18}$$

We introduce the abbreviation

$$\boldsymbol{\beta}(t)=d\boldsymbol{\alpha}(t)/dt.$$

The time derivative of Eq. (6.1) leads to

$$\langle\boldsymbol{\beta}(t)\boldsymbol{\alpha}(0)\rangle=-\langle\boldsymbol{\beta}(-t)\boldsymbol{\alpha}(0)\rangle, \tag{6.19}$$

a time-reversal relationship typical of variables $\boldsymbol{\alpha}$ and $d\boldsymbol{\alpha}/dt$ which are, respectively, even and odd under time reversal. (The case when both types of variables are simultaneously present, e.g., when inertial effects are important, has been considered by Casimir,[96] and Machlup and Onsager,[73] and is discussed in Sec. 10.)

Equations (6.19) and (3.6)–(3.7) at $t=0$ lead to the contradictory results

$$\langle\boldsymbol{\beta}(0)\boldsymbol{\alpha}(0)\rangle=0, \tag{6.20}$$

$$\langle\boldsymbol{\beta}(0)\boldsymbol{\alpha}(0)\rangle=-\boldsymbol{\Lambda}\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle\neq 0. \tag{6.21}$$

The failure of Eq. (6.21) is related to the fact that reversible systems can be only approximately Markoffian in character (with less than a complete microscopic description),

$$\langle\Delta\boldsymbol{\alpha}\rangle/\Delta t=-\Lambda\langle\boldsymbol{\alpha}\rangle, \tag{6.22}$$

when Δt is greater than the forgetting time of the fast variables not included among the $\boldsymbol{\alpha}$.

In what sense, then, does the Markoffian solution approximate the correct one near $t=0$? With the help of stationary, Eq. (3.8) can be rewritten in the form

$$\langle\boldsymbol{\alpha}(t)\boldsymbol{\alpha}(u)\rangle=\exp[-\boldsymbol{\Lambda}|t-u|]\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle \quad \text{for } t>u. \tag{6.23}$$

For $t<u$, take the transpose and apply (6.23):

$$\langle\boldsymbol{\alpha}(t)\boldsymbol{\alpha}(u)\rangle=\langle\boldsymbol{\alpha}(u)\boldsymbol{\alpha}(t)\rangle_{\mathrm{tr}}=\{\exp[-\boldsymbol{\Lambda}|u-t|]\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle\}_{\mathrm{tr}}$$
$$=\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle\exp[-\boldsymbol{\Lambda}^\dagger|t-u|] \quad \text{for } t<u. \tag{6.24}$$

In view of Eq. (6.3), the expressions (6.23) and (6.24) are equivalent. But to make them so, we need the absolute value sign. As a consequence,

$$\langle\boldsymbol{\beta}(t)\boldsymbol{\alpha}(0)\rangle=-\boldsymbol{\Lambda}\exp(-\boldsymbol{\Lambda}t)\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle \quad \text{for } t>0, \tag{6.25}$$

$$\langle\boldsymbol{\beta}(t)\boldsymbol{\alpha}(0)\rangle=\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle\boldsymbol{\Lambda}^\dagger\exp(\boldsymbol{\Lambda}^\dagger t) \quad \text{for } t<0, \tag{6.26}$$

with the result that at $t=0$ there is a discontinuity in slope,

$$\langle\boldsymbol{\beta}(+0)\boldsymbol{\alpha}(0)\rangle=-\boldsymbol{\Lambda}\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle, \tag{6.27}$$

$$\langle\boldsymbol{\beta}(-0)\boldsymbol{\alpha}(0)\rangle=\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle\boldsymbol{\Lambda}^\dagger, \tag{6.28}$$

negative for $t>0$, positive for $t<0$, with a cusp at $t=0$. In the time-reversible case, Eq. (6.4) implies that these slopes are equal in magnitude, therefore the solution is indeed symmetric about $t=0$, and in the true state of affairs the cusp is slightly rounded to give zero slope at the origin (see Fig. 1).

The width of the region in which the true $\alpha(t)$ differs from the phenomenological cusp near $t=0$ is of the order τ_d=duration of a collision, or the forgetting time of the nonmacroscopic variables. On the other hand,

[96] H. B. G. Casimir, Revs. Modern Phys. **17**, 343 (1945).

for larger times, $\alpha(t)\simeq\exp(-t/\tau)$, where the relaxation time τ is the reciprocal of some eigenvalue of $\mathbf{\Lambda}$. Our phenomenological description then is useful provided $\tau_d\ll\tau$, i.e., the forgetting times must be short compared to typical relaxation times under study. In fact, one must always choose a set of variables $\boldsymbol{\alpha}$ for this condition to be obeyed.

7. Nyquist Theorem

The Nyquist theorem relates the noise associated with velocity fluctuations $\dot{\alpha}(t)$ at the frequency ω to the admittance of the system.

We may calculate the admittance by adding an external force[73] $-\mathbf{X}^e\equiv P/T$ to the right-hand side of (6.6), where T is the absolute temperature. [See Eq. (4.15).] On eliminating the internal force by means of (4.5), we obtain

$$\mathbf{R}(d\boldsymbol{\alpha}/dt)+\mathbf{s}\boldsymbol{\alpha}=-\mathbf{X}^e=\mathbf{P}/T. \tag{7.1}$$

To obtain the admittance at frequency ω, we assume that P has the time dependence $\exp(i\omega t)$ so that $(d\boldsymbol{\alpha}/dt)=i\omega\boldsymbol{\alpha}$, and we obtain

$$d\boldsymbol{\alpha}/dt=\mathbf{Y}(\omega)\mathbf{P}, \tag{7.2}$$

where

$$\mathbf{Y}(\omega)=[\mathbf{R}+(\mathbf{s}/i\omega)]^{-1}/T \tag{7.3}$$

is the desired admittance. It is convenient for later comparison to make use of (6.9) and (6.11) to rewrite the admittance in the form

$$\mathbf{Y}(\omega)=i\omega\mathbf{\Lambda}(i\omega+\mathbf{\Lambda})^{-1}[\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle/(kT)]\mathbf{C}^{-1}, \tag{7.4}$$

where $\mathbf{C}$ is defined by Eq. (6.11).

The noise associated with $\dot{\alpha}(t)$ is, by the original definition (2.3), equal to ω^2 times the noise associated with $\alpha(t)$. In view of Eq. (3.10), we can write

$$\begin{aligned}\omega^2\mathbf{G}_+(\boldsymbol{\alpha},f)&=2(-i\omega)(i\omega+\mathbf{\Lambda}-\mathbf{\Lambda})(i\omega+\mathbf{\Lambda})^{-1}\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle\\&=-2i\omega\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle+2i\omega\mathbf{\Lambda}(i\omega+\mathbf{\Lambda})^{-1}\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle.\end{aligned} \tag{7.5}$$

The negative frequency contribution is obtained, as usual, by reversing the sign of ω and taking the transpose. The first term cancels, and by comparison with (7.4), we can write

$$\omega^2\mathbf{G}(\boldsymbol{\alpha},f)=\mathbf{G}[(d\boldsymbol{\alpha}/dt),f]=2kT[\mathbf{Y}(\omega)\mathbf{C}+\mathbf{C}^\dagger\mathbf{Y}^\dagger(-\omega)]. \tag{7.6}$$

Since $\mathbf{Y}(\omega)$ is an admittance relating two real quantities $\boldsymbol{\alpha}$ and $\mathbf{P}$, we must have

$$\mathbf{Y}(-\omega)=\mathbf{Y}^*(\omega). \tag{7.7}$$

The Onsager relations (6.12)–(6.13), combined with the symmetry of $\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle$, imply that for the time-reversible case,

$$\mathbf{Y}(\omega)\mathbf{C}=\mathbf{C}^\dagger\mathbf{Y}^\dagger(\omega) \tag{7.8}$$

so that the Nyquist theorem can be simplified to

$$\mathbf{G}[(d\boldsymbol{\alpha}/dt),f]=4kT\,\mathrm{Re}\mathbf{Y}(\omega)\mathbf{C}. \tag{7.9}$$

Quantum-mechanical derivations of the Nyquist theorem, based on a microscopic point of view and thermal equilibrium, yield a Nyquist theorem in the form[41–45]

$$\mathbf{G}[(d\boldsymbol{\alpha}/dt),f]=4kT\,\mathrm{Re}[\mathbf{Y}(\omega)](\hbar\omega/2kT)\times\coth(\hbar\omega/2kT), \tag{7.10}$$

where the frequency-dependent correction factor is nearly unity for $\hbar\omega\ll kT$ but approaches $\hbar\omega/2kT$ for $\hbar\omega\gg kT$. It is interesting to note that these derivations do not make use of the Markoffian assumption, although they do apply only to fluctuations from an equilibrium state.

The reason that quantum mechanics simply introduces a universal correction factor can be understood as follows. The mean energy of a harmonic oscillator q_k of frequency ω_k is given in quantum mechanics by

$$\begin{aligned}E(\hbar\omega_k)&=\hbar\omega_k(\bar{n}_k+\tfrac{1}{2})\\&=\hbar\omega_k\{[\exp(\hbar\omega_k/kT)-1]^{-1}+\tfrac{1}{2}\}\\&=\tfrac{1}{2}\hbar\omega_k\coth(\hbar\omega_k/2kT),\end{aligned} \tag{7.11}$$

whereas its mean classical energy is kT. But the mean square displacement of q_k is proportional to its mean potential energy, and hence its mean total energy. Therefore, the ratio of quantum-mechanical to classical vibration amplitudes obeys

$$C=\langle q_k^2\rangle_{\text{Q.M.}}/\langle q_k^2\rangle_{\text{Cl}}=E(\hbar\omega_k)/kT. \tag{7.12}$$

Any parameter $\alpha(t)$ can be regarded as a linear combination of all the normal coordinates q_k. However, it is clear from the original definition [Eq. (2.3)] of the noise that only the Fourier component $\boldsymbol{\alpha}(\omega)$ is used, i.e., only the modes q_k with frequencies $\omega_k=\omega$ affect the noise at frequency ω. Hence, the correction factor is simply $E(\hbar\omega)/kT$ as shown in Eq. (7.10).

We see therefore that our correction factor $\mathbf{C}$, the ratio of the correct $\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle$ to its classically-computed thermal equilibrium value, yields (1) the quantum correction in the thermal equilibrium case, (2) the correction to be applied when the system is driven away from equilibrium.

The possibility that the usual Nyquist theorem would be violated for a driven system was stated explicitly by Richardson.[46] He argued that the regression of fluctuations obeyed one set of differential equations, describable by an impedance, and the response of a system to a perturbation was also describable by an impedance, not necessarily equal to the first one. While a Nyquist theorem would always exist relating the noise to the regression of fluctuations, Richardson questioned whether there was any connection between the noise and the impedance to perturbations exhibited by a driven system.

A Markoff system, however, cannot distinguish between the response for $t>0$, after the initial condition $\alpha(t)\,|_{t=0}\alpha=(0)$, and the response to a force $\delta P=(\partial P/\partial a)\cdot\alpha(0)$ for $t<0$ and 0 for $t>0$, since this force is calculated

to produce the time-independent displacement $\alpha(0)$ for $t<0$. If this force is analyzed into its frequency components, the response can be expressed in terms of the admittance of the system (see Callen and Green[70]). In this way the regression and hence the noise can be calculated in terms of the admittance—hence, a Nyquist theorem always exists for a Markoffian system. However, this Nyquist theorem for a driven system is generally of the form Eq. (7.6) containing the correction factor $\mathbf{C}=\mathbf{s}\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle/k$. This correction factor arises even in the Einstein relation (6.14).

The correction factor $\mathbf{C}$ appears in such a ubiquitous way because it represents a breakdown in the thermodynamic relationship between a fluctuation and a dc admittance:

$$\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle=kT\partial\mathbf{a}/\partial\mathbf{P}\cdot\mathbf{C}, \tag{7.13}$$

since $\mathbf{s}^{-1}/T$ represents the static ratio of displacement Δa to force $\Delta\mathbf{P}$ in conventional (energy language) notation.

There are many cases of driven systems, in which the fluctuations at one time $\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle$ are given by the thermodynamic formula, without a correction factor. These are usually cases in which the system is driven to a new steady position with a modified stiffness matrix, but the system behaves as if it were in equilibrium with respect to the degrees of freedom under consideration, although other degrees of freedom may be in a steady nonequilibrium state. We have already given an example of this for concentration fluctuations in the presence of a strong electric field, which leads to a nonequilibrium velocity distribution; yet the concentration fluctuations where calculable by thermodynamic methods [see Eq. (4.30)].

8. Langevin Method and Fluctuations in Intensive Parameters

We already have a formal solution of the noise problem in Eqs. (3.10) and (3.11), combined with Eq. (4.4) or (4.17) for the fluctuations from an equilibrium state, or Eqs. (5.19) and (5.20) for the fluctuations from a steady state. Our solution describes the fluctuations associated with extensive variables, the responses of the system. Following Langevin, it has been customary to regard these fluctuations in $\boldsymbol{\alpha}$ as produced by fluctuations in the corresponding intensive variables $\mathbf{P}$, e.g., current fluctuations may be regarded as produced by voltage fluctuations. The latter are representable in an equivalent circuit as noise voltage generators. The magnitude and spectrum of these noise generators are then defined by the requirement that they give rise to the noise already computed.

For example, we can replace Eq. (3.6) by

$$d\boldsymbol{\alpha}/dt+\boldsymbol{\Lambda}\boldsymbol{\alpha}=\mathbf{F}(t), \tag{8.1}$$

where

$$\langle\mathbf{F}(t)\rangle=0, \tag{8.2}$$

in order that Eq. (3.6) for the average response remain true. Our problem now is to give the random "force" $\mathbf{F}(t)$ an autocorrelation $\langle\mathbf{F}(t)\mathbf{F}(u)\rangle$ consistent with the requirements (6.23) and (6.24):

$$\langle\boldsymbol{\alpha}(t)\boldsymbol{\alpha}(u)\rangle=H(t-u)\exp(-\boldsymbol{\Lambda}|t-u|)\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle +H(u-t)\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle\exp(-\boldsymbol{\Lambda}^{\dagger}|t-u|), \tag{8.3}$$

where

$$H(t)=0,\quad t<0 \qquad =1,\quad t>0 \tag{8.4}$$

is the Heaviside function, whose derivative obeys

$$dH(t)/dt=\delta(t). \tag{8.5}$$

The autocorrelation in $\mathbf{F}(t)$ is given by

$$\langle\mathbf{F}(t)\mathbf{F}(u)\rangle=\langle(d/dt+\boldsymbol{\Lambda})\boldsymbol{\alpha}(t)(d/dt+\boldsymbol{\Lambda})\alpha(u)\rangle =(d/dt+\boldsymbol{\Lambda})\langle\boldsymbol{\alpha}(t)\boldsymbol{\alpha}(u)\rangle(\leftarrow d/du+\boldsymbol{\Lambda}^{\dagger}), \tag{8.6}$$

where $\leftarrow d/du$ is understood to act to the left. We now insert Eq. (8.3) and make use of the relationship

$$(d/dt+\boldsymbol{\Lambda})H(t)\exp(-\boldsymbol{\Lambda}t)=\delta(t) \tag{8.7}$$

to obtain

$$\langle\mathbf{F}(t)\mathbf{F}(u)\rangle=\delta(t-u)\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle(\leftarrow d/du+\boldsymbol{\Lambda}^{\dagger}) +(d/dt+\boldsymbol{\Lambda})\delta(t-u)\langle\alpha\alpha\rangle. \tag{8.8}$$

But

$$d/du\,\delta(t-u)=-d/dt\,\delta(t-u), \tag{8.9}$$

with the result that

$$\langle\mathbf{F}(t)\mathbf{F}(u)\rangle=\delta(t-u)(\boldsymbol{\Lambda}\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle+\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle\boldsymbol{\Lambda}^{\dagger}) \tag{8.10}$$

$$=2\mathbf{D}\delta(t-u) \tag{8.11}$$

in view of the Einstein relationship (5.18).

A random variable with a delta function autocorrelation is often described as completely random.[21] In view of the Weiner-Khintchin theorem, such a variable has a "white" (i.e., flat) noise spectrum.

The correctness of (8.11) may be verified by demonstrating that it does indeed give rise to the desired properties for $\boldsymbol{\alpha}$. Equation (8.1) can be rewritten in the form

$$\boldsymbol{\alpha}(t)=\exp(-\boldsymbol{\Lambda}t)\boldsymbol{\alpha}(0)+\int_0^t\exp(-\boldsymbol{\Lambda}s')\mathbf{F}(t-s')ds' \tag{8.12}$$

or

$$\boldsymbol{\alpha}(u)=\exp(-\boldsymbol{\Lambda}t)\boldsymbol{\alpha}(0)+\int_0^u\exp(-\boldsymbol{\Lambda}s)\mathbf{F}(t-s)ds. \tag{8.13}$$

For fixed $\boldsymbol{\alpha}(0)$, we then have that

$$\langle\boldsymbol{\alpha}(t)\boldsymbol{\alpha}(u)\rangle_{\alpha 0} =\exp(-\boldsymbol{\Lambda}t)\boldsymbol{\alpha}(0)\boldsymbol{\alpha}(0)\exp(-\boldsymbol{\Lambda}^{\dagger}u)+\int_0^t ds'\int_0^u ds \times\exp(-\boldsymbol{\Lambda}s')\langle\mathbf{F}(t-s')\mathbf{F}(u-s)\rangle\exp(-\boldsymbol{\Lambda}^{\dagger}s). \tag{8.14}$$

Now introduce Eq. (8.10) for the autocorrelation in $\mathbf{F}$. For the case $t>u$, first integrate over s' so that the

range of the delta function is covered:

$$\langle\boldsymbol{\alpha}(t)\boldsymbol{\alpha}(u)\rangle_{\alpha 0} = \exp(-\boldsymbol{\Lambda}t)\boldsymbol{\alpha}(0)\boldsymbol{\alpha}(0)\exp(-\boldsymbol{\Lambda}^\dagger u)+\int_0^u ds \times\exp[-\boldsymbol{\Lambda}(s+t-u)](\boldsymbol{\Lambda}\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle+\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle\boldsymbol{\Lambda}^\dagger)\times\exp(-\boldsymbol{\Lambda}^\dagger s). \quad (8.15)$$

The integrand is the exact differential of

$$-\exp[-\boldsymbol{\Lambda}(s+t-u)]\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle\exp(-\boldsymbol{\Lambda}^\dagger s),$$

therefore, the result can be rearranged in the form

$$\langle\boldsymbol{\alpha}(t)\boldsymbol{\alpha}(u)\rangle_{\alpha 0}=\exp(-\boldsymbol{\Lambda}|t-u|)\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle +\exp(-\boldsymbol{\Lambda}t)[\boldsymbol{\alpha}(0)\boldsymbol{\alpha}(0)-\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle]\exp(-\boldsymbol{\Lambda}^\dagger u). \quad (8.16)$$

We see that as t and u approach ∞ with $t-u$ fixed, the autocorrelation approaches the desired stationary value. Furthermore, an average over the initial values cancels the second term, and again yields the desired stationary answer (8.3).

Equation (8.12), also yields the macroscopically expected result,

$$\langle\boldsymbol{\alpha}(t)\rangle=\exp(-\boldsymbol{\Lambda}t)\boldsymbol{\alpha}(0), \quad (8.17)$$

so that Eq. (8.16) can be rewritten in terms of fluctuations about the average motion

$$\langle[\boldsymbol{\alpha}(t)-\langle\boldsymbol{\alpha}(t)\rangle][\boldsymbol{\alpha}(u)-\langle\boldsymbol{\alpha}(u)\rangle]\rangle_{\alpha 0} =\exp[-\boldsymbol{\Lambda}|t-u|]\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle-\exp(-\boldsymbol{\Lambda}t)\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle\times\exp(-\boldsymbol{\Lambda}^\dagger u). \quad (8.18)$$

Thus, the fluctuations about the average motion are not *stationary*, in spite of the fact that we are dealing with a linear system with time-independent transition probabilities. The reason for this is clear: Equation (8.18) must yield a vanishing result for the fluctuation when $t=u=0$ because of the imposed initial condition, whereas the stationary expression takes the maximum value $\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle$ at this time!

Our expression for the fluctuations in $\mathbf{F}(t)$ is already a simple and readily usable form. One may prefer, however, to discuss the fluctuations in $\mathbf{P}(t)$, the conventional force. Comparison with Eq. (7.1) indicates that

$$\mathbf{P}(t)=T\mathbf{R}\mathbf{F}(t) \quad (8.19)$$

so that

$$\langle\mathbf{P}(t)\mathbf{P}(u)\rangle=T^2\mathbf{R}\langle\mathbf{F}(t)\mathbf{F}(u)\rangle\mathbf{R}^\dagger =T^2\mathbf{R}(\mathbf{L}\mathbf{s}\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle+\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle\mathbf{s}^\dagger\mathbf{L}^\dagger)\mathbf{R}^\dagger\delta(t-u),$$

or, using $\mathbf{R}\mathbf{L}=\mathbf{L}^\dagger\mathbf{R}^\dagger=1$,

$$\langle\mathbf{P}(t)\mathbf{P}(u)\rangle=kT(\mathbf{C}\mathbf{R}'^\dagger+\mathbf{R}'\mathbf{C}^\dagger)\delta(t-u), \quad (8.20)$$

where $\mathbf{R}'=T\mathbf{R}$ is the conventional resistance and $\mathbf{C}=\mathbf{s}\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle/k$ is the usual correction factor. At equilibrium, $\mathbf{C}=1$ and $\mathbf{R}'^\dagger=\mathbf{R}'$, thus

$$\langle\mathbf{P}(t)\mathbf{P}(u)\rangle=2kT\mathbf{R}'\delta(t-u). \quad (8.21)$$

Application of the Wiener-Khintchin theorem, Eq. (2.5), leads to

$$\mathbf{G}(\mathbf{P},f)df=4kT\mathbf{R}'df, \quad (8.22)$$

the conventional statement of the Nyquist theorem for voltage fluctuations at equilibrium, whereas away from equilibrium we have

$$\mathbf{G}(\mathbf{P},f)=2kT(\mathbf{C}\mathbf{R}'^\dagger+\mathbf{R}'\mathbf{C}^\dagger). \quad (8.23)$$

Time reversibility in the form Eq. (6.13) states that

$$\mathbf{R}'\mathbf{C}^\dagger=\mathbf{C}\mathbf{R}'^\dagger \quad (8.24)$$

so that the two terms in Eq. (8.23) make an equal contribution.

The Nyquist theorem as used in practice is less general than that just derived, in that the latter describes the noise in a single variable rather than a set—but is more general in that the voltage noise source is not taken as flat but proportional to $R(\omega)$, the resistance at the frequency in question. The latter, more general statement is usually arrived at by a thermodynamic argument[59]: One connects a resistor R to a black box containing an impedance $Z=R(\omega)+iX(\omega)$. In order that the known transfer of power from R into the box be balanced by an equal transfer in the reverse direction, the box must be assigned an open-circuit noise source $4kTR(\omega)$.

This proof, while simple and elegant, leaves two questions unanswered: (1) How is it that one can assume the validity of the usual circuit diagram procedure—assuming that the voltage noise may be represented by a generator in series with the internal impedance—for the purpose of computing power transfer to a load? (2) In the case of a steady nonequilibrium state (e.g., current flow), what can one use to replace the thermodynamic argument?

These questions are answered in the next section.

9. Validity of Circuit Diagrams: Systems under Load

We have defined fluctuations in the intensive variables (e.g., voltages) by the requirement that they must give rise through the equations of motion to the directly computed noise in the extensive variables (currents). This artificial approach has led some authors[70] to refer to the fluctuations in the intensive variables as "fictitious" or "hypothetical."

The usefulness of considering "voltage" fluctuations is that they permit one to use Thevenin's theorem to compute the noise that arises when two complicated systems are hooked together (one regarded as the load of the other) in terms of parameters that describe each separate system. Thevenin's theorem[97] states that a complicated electrical network containing a variety of generators may be replaced by a single generator, whose

[97] E. A. Guillemin, *Communication Networks* (John Wiley & Sons, Inc., New York, 1935), Vol. II, p. 181.

voltage is the open-circuit output voltage of the system, in series with a single impedance—the internal impedance measured at the output terminals, with all the generators short circuited.

The purpose of the present section is to derive directly the noise in a loaded system and to show that the result has the form required for Thevenin's theorem to be valid for the thermal equilibrium case but not in general. (It is not entirely obvious that the internal impedance presented by the system to a noise generator is necessarily the same as that presented to a signal—although we anticipate that this should be true since we have explicitly made use of the fact that, for Markoffian systems, the regression of fluctuations obeys the macroscopic equations.)

We consider the noise in the parameter $\dot{\alpha}_1$. In other words, we regard the 11 terminals as the output terminals of our "network" and add a load across them, i.e., we modify $\mathbf{A}=\mathbf{Ls}$ to insert an added impedance z in circuit 1 so that our new impedance is

$$Z_{11}'=Z_{11}+z, \tag{9.1}$$

with all other elements Z_{ij} of the impedance matrix unmodified.

The positive frequency part of the noise in $\dot{\alpha}_1$ is, according to Eq. (7.14), given by

$$G_+(\dot{\alpha})_{11}=2kT\sum_j Y_{1j}'(\omega)C_{j1}', \tag{9.2}$$

where the elements of $\mathbf{Y}'$, the matrix reciprocal to $\mathbf{Z}'$, can be computed easily as

$$\mathbf{Z}(d\boldsymbol{\alpha}/dt)=\mathbf{P}-\mathbf{z}(d\boldsymbol{\alpha}/dt), \tag{9.3}$$

$$d\boldsymbol{\alpha}/dt=\mathbf{YP}-\mathbf{Yz}(d\boldsymbol{\alpha}/dt), \tag{9.4}$$

$$\dot{\alpha}_1=\sum_j Y_{1j}P_j-Y_{11}z\dot{\alpha}_1, \tag{9.5}$$

when we make use of the fact that $\mathbf{z}$ has only an 11 element, with the result that

$$\dot{\alpha}_1=\sum_j Y_{1j}'P_j, \tag{9.6}$$

with

$$Y_{1j}'=Y_{1j}/(1+Y_{11}z), \tag{9.7}$$

or

$$\dot{\alpha}_1=V_{oc}/(Z+z), \tag{9.8}$$

where the open-circuit voltage V_{oc} and the internal impedance Z are customarily defined by

$$V_{oc}=\sum_j(Y_{1j}/Y_{11})P_j, \tag{9.9}$$

$$Z=1/Y_{11}. \tag{9.10}$$

The positive frequency noise is therefore given by

$$G_+(\dot{\alpha})_{11}=2kT\sum_j(Y_{1j}C_{j1}'/Y_{11})/(Z+z). \tag{9.11}$$

Since the negative frequency noise is simply the complex conjugate, we rationalize and take twice the real part to obtain

$$G(\dot{\alpha})_{11}=4kT\frac{\mathrm{Re}[(Z^*+z^*)\sum_j Y_{1j}C_{j1}'/Y_{11}]}{|Z+z|^2}. \tag{9.12}$$

Thevenin's theorem, Eq. (9.8), suggests that we should expect an answer in the form

$$G(\dot{\alpha})_{11}=G(V_{oc})/|Z+z|^2, \tag{9.13}$$

where the noise source associated with the open-circuit voltage depends only on the parameters of the system and not the load. Equation (9.12) does not have this required form. For the *equilibrium* case, $C_{j1}=\delta_{j1}$ so that only the term $j=1$ survives, and Eqs. (9.12) and (9.13) lead to a voltage source,

$$G(V)=4kT[\mathrm{Re}(Z)+\mathrm{Re}(z)], \tag{9.14}$$

which may be interpreted readily by assigning $\mathrm{Re}[Z(\omega)]$ as a noise source associated with the system and $\mathrm{Re}(z)$ as a noise source associated with the load. (The latter could be made to vanish by choosing z pure imaginary!) The use of circuit analysis and Thevenin's theorem therefore seems to be valid for the case of thermal equilibrium.

In the nonthermal equilibrium case, Eq. (9.12) appears, at least in some cases, to contradict Thevenin's theorem, for the term in z^* leads to a noise source that depends both on the system and on the load. For the special case in which $z=\Delta s_{11}/i\omega$ is pure reactive, $\mathbf{L}$ is not modified, nor is $\mathbf{D}$, so that the Einstein relation in the form (7.16) indicates that $\mathbf{C}=(k\mathbf{L})^{-1}\mathbf{D}$ is unmodified i.e., $\mathbf{C}'=\mathbf{C}$. The first term in Eq. (9.12) then can be shown to be equal to the open-circuit noise source that would be computed by the usual circuit analysis via Eq. (9.9):

$$\langle|V_{oc}(\omega)|^2\rangle=\langle(\mathbf{YP})_1^*(\mathbf{YP})_1\rangle/|Y_{11}|^2 \tag{9.15}$$

$$=(\mathbf{Y}^*\langle\mathbf{P}(\omega)^*\mathbf{P}(\omega)\rangle\mathbf{Y})_{11}/|Y_{11}|^2. \tag{9.16}$$

By making use of Eq. (8.23), our result can be written

$$\langle|V_{oc}(\omega)|^2\rangle=2kT(\mathbf{Y}^*(\mathbf{CR}'^\dagger+\mathbf{R}'\mathbf{C}^\dagger)\mathbf{Y})_{11}/|Y_{11}|^2 \tag{9.17}$$

or, using $\mathbf{R}=\frac{1}{2}(\mathbf{Z}+\mathbf{Z}^*)$ and the Onsager relations Eq. (7.8),

$$\langle|V_{oc}(\omega)|^2\rangle=4kT\,\mathrm{Re}(\mathbf{YC})_{11}/|Y_{11}|^2, \tag{9.18}$$

which is simply $|Z|^2$ times the current noise shown in Eq. (7.6). The first term in the numerator of Eq. (9.12) reduces to this form when one writes $Z^*=1/Y_{11}^*$. Thus, we have arrived at the identical *open-circuit* voltage noise in three ways: (1) by considering the current noise in the presence of a load, (2) by combining the individual voltage generators via the circuit Eq. (9.9), and (3) by multiplying the current noise by the absolute square of the internal impedance.

For the case in which z has a real part, $\mathbf{C}'\neq\mathbf{C}$. For the general validity of a circuit viewpoint, one should demonstrate that the total voltage fluctuation, minus the part we have already computed, depends only on z and not on the system. This I have not yet succeeded in doing. In the presence of a load the system generator for the nonequilibrium case appears to be modified by the load. In the thermal equilibrium case this difficulty does not occur.

The general problem of the fluctuations of intensive variables has always been a thorny one. There seem to be methods available, at least in principle, for a direct calculation of the fluctuations of intensive thermodynamic variables.[68] It is not clear therefore that one should *define* pressure fluctuations by the requirement that they give rise to the observed volume fluctuations through an equilibrium relationship,[4] e.g.,

$$\Delta P=(\partial P/\partial V)_T\Delta V, \tag{9.19}$$

since this procedure assumes that the equilibrium equation of state $P=P(V,T)$ holds during the fluctuations as well.

I avoid the difficulty of a direct calculation of the fluctuations in intensive variables by arguing that one always measured fluctuations in extensive variables. The usefulness of discussing fluctuations in intensive variables is that they make it easy to compute the new fluctuations in extensive variables when a load is placed on the system. The proper procedure then is to calculate the fluctuations in the extensive parameters without and with a load. The way in which the fluctuations depend on the load then indicate that in the *thermal equilibrium* case there is a reasonable way to define fluctuations in the intensive parameters to yield the correct load dependence, and this definition is consistent with the assumption that the intensive and extensive parameters are connected by the phenomenological equations.

At high frequencies an equation such as (9.19) must cease to be valid, but in that case ΔV does not provide an adequate number of variables for a Markoffian description to be valid. Our present view, in short, is that the fluctuations in the intensive variables can be computed from those in the extensive variables by the phenomenological relations, providing a complete set of variables has been used which yields a Markoffian system. This remark is valid only for the equilibrium case, and an equation such as (9.12) must be used for the general nonequilibrium case.

EXAMPLES

10. Systems Possessing Inertia

Because the systems previously discussed involved first-order time derivatives and were described in terms of a resistance and stiffness matrix, it might appear that new methods are required to handle systems possessing inertia. The latter are commonly described by differential equations second order in the time. It is well known, however, that (barring the anomalous case in which one cannot solve for the second derivatives in terms of those of lower order) second-order systems can be reduced to first-order systems by doubling the number of variables—the new ones are the set of first time derivatives. The extension to systems possessing inertia is then a purely formal one except for considerations of time reversibility.

The general prototype of a system possessing inertia is a system of N coupled oscillators. We may adopt

$$\mathbf{m}'(d^2\mathbf{q}/dt^2)+\mathbf{R}'(d\mathbf{q}/dt)+\mathbf{s}'\mathbf{q}=\mathbf{P} \tag{10.1}$$

as out phenomenological equations, where $\mathbf{P}$, $\mathbf{m}'$, $\mathbf{R}'$, $\mathbf{s}'$ are the conventional (energy language) external driving force, mass, resistance, and stiffness matrices, respectively. These quantities differ by a factor T from the corresponding entropy language quantities used before:

$$\mathbf{m}'=T\mathbf{m},\quad \mathbf{R}'=T\mathbf{R},\quad \mathbf{s}'=T\mathbf{s}, \tag{10.2}$$

with $\mathbf{P}=-T\mathbf{X}_e$ as in Eq. (7.1). To simplify the notation in this section, we drop the primes in Eq. (10.1).

We now convert to a set of first-order equations by doubling the number of variables,

$$\begin{aligned}\alpha_i&=q_i,\quad i=1,2\cdots N,\\ \alpha_{N+i}&=p_i=[\mathbf{m}(d\mathbf{q}/dt)]_i,\end{aligned} \tag{10.3}$$

so that $\boldsymbol{\alpha}=(\mathbf{q},\mathbf{p})$ obey the set of first-order equations

$$d\mathbf{q}/dt=\mathbf{m}^{-1}\mathbf{p}, \tag{10.4}$$

$$d\mathbf{p}/dt=-\boldsymbol{\gamma}\mathbf{p}-\mathbf{s}\mathbf{q}+\mathbf{P}, \tag{10.5}$$

where

$$\boldsymbol{\gamma}=\mathbf{R}\mathbf{m}^{-1}, \tag{10.6}$$

so that $d\boldsymbol{\alpha}/dt=-\boldsymbol{\Lambda}\boldsymbol{\alpha}$ with

$$\boldsymbol{\Lambda}=\begin{pmatrix}0, & -\mathbf{m}^{-1}\\ \mathbf{s}, & \boldsymbol{\gamma}\end{pmatrix}, \tag{10.7}$$

where we treat $\boldsymbol{\alpha}$ as a two-component quantity with components $\mathbf{q}$ and $\mathbf{p}$.

Time reversibility can be expressed in the form

$$\langle\mathbf{q}(t)\mathbf{q}(0)\rangle=\langle\mathbf{q}(0)\mathbf{q}(t)\rangle, \tag{10.8}$$

$$\langle\mathbf{p}(t)\mathbf{p}(0)\rangle=\langle\mathbf{p}(0)\mathbf{p}(t)\rangle, \tag{10.9}$$

$$\langle\mathbf{p}(t)\mathbf{q}(0)\rangle=-\langle\mathbf{p}(-t)\mathbf{q}(0)\rangle=-\langle\mathbf{p}(0)\mathbf{q}(t)\rangle \tag{10.10}$$

or in the form

$$[\boldsymbol{\Lambda}\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle]_{ij}=E_iE_j[\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle\boldsymbol{\Lambda}^\dagger]_{ij}, \tag{10.11}$$

where $E_i=+1$ for $1\leq i\leq N$, $E_i=-1$ for $N+1\leq i\leq 2N$.

It follows from (10.10) that at $t=0$

$$\langle\mathbf{p}\mathbf{q}\rangle=0. \tag{10.12}$$

If the fluctuations at one time differ from their thermal equilibrium value, we can [cf. Eq. (6.11)] write them as the thermal equilibrium value times a correction factor,

$$\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle=kT\begin{pmatrix}\mathbf{s}^{-1}, & 0\\ 0, & \mathbf{m}\end{pmatrix}\begin{pmatrix}\mathbf{C} & 0\\ 0 & \mathbf{B}\end{pmatrix}, \tag{10.13}$$

where the off-diagonal elements have been set equal to zero because of Eq. (10.12). Then

$$\boldsymbol{\Lambda}\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle=\begin{pmatrix}0, & -\mathbf{B}\\ \mathbf{C}, & \mathbf{R}\mathbf{B}\end{pmatrix}. \tag{10.14}$$

Time reversibility in the form (10.11) is equivalent to the statement that the diagonal part of $\mathbf{\Lambda}\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle$ is symmetric, but its off-diagonal parts are antisymmetric with respect to one another:

$$\mathbf{RB}=(\mathbf{B}^\dagger\mathbf{R}^\dagger), \tag{10.15}$$

and

$$\mathbf{C}^\dagger=-(-\mathbf{B})=\mathbf{B}. \tag{10.16}$$

According to our development of the Langevin point of view in Eq. (8.11), the noise may be calculated by introducing a source whose autocorrelation is

$$\langle\mathbf{P}(t)\mathbf{P}(u)\rangle=2\mathbf{D}\delta(t-u), \tag{10.17}$$

where

$$2\mathbf{D}=\mathbf{\Lambda}\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle+\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle\mathbf{\Lambda}^\dagger$$

$$=\begin{pmatrix}0, & 0\\ 0, & \mathbf{RC}^\dagger+\mathbf{CR}^\dagger\end{pmatrix}, \tag{10.18}$$

in agreement with Eq. (8.20). It is pleasing that we can verify directly, without such heuristic arguments as presented by Hashitsume,[78] for a simple harmonic oscillator that no noise source enters the relation [(Eq. (10.4)] between velocity and momentum.

It may be of interest to verify the Nyquist theorem directly without benefit of the Langevin viewpoint. The admittance relating $d\mathbf{q}/dt$ to $\mathbf{P}$ can be obtained directly from Eq. (10.1):

$$\mathbf{Y}(\omega)=[\mathbf{R}+i\omega\mathbf{m}-i(\mathbf{s}/\omega)]^{-1}. \tag{10.19}$$

The positive frequency component of the noise may be computed from the positive Fourier component Eq. (2.7) of the autocorrelation,

$$\langle\boldsymbol{\alpha}(t)\boldsymbol{\alpha}(0)\rangle=\exp(-\mathbf{\Lambda}t)\begin{pmatrix}\langle\mathbf{qq}\rangle, & \langle\mathbf{qp}\rangle\\ \langle\mathbf{pq}\rangle, & \langle\mathbf{pp}\rangle\end{pmatrix}, \tag{10.20}$$

$$G_+(\boldsymbol{\alpha})=\frac{2}{(i\omega+\mathbf{\Lambda})}\begin{pmatrix}\langle\mathbf{qq}\rangle, & 0\\ 0, & \langle\mathbf{pp}\rangle\end{pmatrix}, \tag{10.21}$$

where we have discarded the vanishing cross terms,

$$(i\omega+\mathbf{\Lambda})=\begin{pmatrix}i\omega, & -\mathbf{m}^{-1}\\ \mathbf{s}, & i\omega+\boldsymbol{\gamma}\end{pmatrix} \tag{10.22}$$

and

$$(i\omega+\mathbf{\Lambda})^{-1}=\begin{pmatrix}\mathbf{Y}(\mathbf{R}+i\omega\mathbf{m})/i\omega, & \mathbf{Y}/i\omega\\ -\mathbf{mYs}/i\omega, & \mathbf{mY}\end{pmatrix}, \tag{10.23}$$

where $\mathbf{Y}$ is the admittance, Eq. (10.19).

The noise associated with $\mathbf{q}$ may be computed from Eqs. (10.21) and (10.23) by taking the 11 component. After discarding an irrelevant term $2/i\omega$, which cancels out of the total noise, the positive frequency part can be written

$$\mathbf{G}_+(\mathbf{q})=(2/\omega^2)\mathbf{Y}(\omega)\mathbf{s}\langle\mathbf{qq}\rangle \tag{10.24}$$

or

$$\mathbf{G}_+(\mathbf{q})=(2/\omega^2)\mathbf{Y}(\omega)kT\mathbf{C}, \tag{10.25}$$

where the last form has used Eq. (10.13). Since the noise in $d\mathbf{q}/dt$ is ω^2 times that in $\mathbf{q}$, Eq. (10.25) agrees with the usual Nyquist theorem

$$\mathbf{G}_+(d\mathbf{q}/dt)=2kT\mathbf{Y}(\omega)\mathbf{C}. \tag{10.26}$$

As a test of consistency, we note that the 22 component of Eqs. (10.21) and (10.23) yields, for the noise in $\mathbf{p}$,

$$G_+(\mathbf{p})=2\mathbf{m}\mathbf{Y}(\omega)\langle\mathbf{pp}\rangle. \tag{10.27}$$

Since $d\mathbf{q}/dt=\mathbf{m}^{-1}\mathbf{p}$, we can write

$$\mathbf{G}_+(d\mathbf{q}/dt)=\mathbf{m}^{-1}\mathbf{G}_+(\mathbf{p})(\mathbf{m}^\dagger)^{-1} \tag{10.28}$$

$$=2\mathbf{Y}(\omega)\langle\mathbf{pp}\rangle/\mathbf{m}^\dagger. \tag{10.29}$$

In view of Eq. (10.13), the noise in $d\mathbf{q}/dt$ also can be written as

$$\mathbf{G}_+(d\mathbf{q}/dt)=2kT\mathbf{Y}(\omega)\mathbf{B}^\dagger, \tag{10.30}$$

which agrees completely with the usual theorem Eq. (10.26) if one remembers that $\mathbf{B}^\dagger=\mathbf{C}$ in accord with Eq. (10.16).

11. Continuous Parameter Example: Ambipolar Drift of Carrier Concentration Fluctuations

Although the methods described relate to a set of parameters α_i depending on a discrete index i, we have not made special use of the discreteness of this index; therefore, i could be replaced by a continuous index x. The only change in the formalism is that sums over i are replaced by integrals over x.

In order that our statements do not appear too abstract, we illustrate them by the example of the influence of ambipolar drift of carrier concentration fluctuations discussed by Hill and van Vliet.[98] This influence becomes important at electric fields high enough for the drift time between electrodes to be comparable to or shorter than the recombination time.

In the presence of a constant current generator, concentration fluctuations generate voltage fluctuations at the output terminals proportional to the induced resistance fluctuations:

$$\Delta V(t)=\Delta R(t)I \tag{11.1}$$

or

$$\frac{\Delta V}{V}=\frac{\Delta R}{R}=-\frac{\Delta Y}{Y}=-\frac{\Delta P\mu_p+\Delta N\mu_n}{P\mu_p+N\mu_n}, \tag{11.2}$$

where P and N are the total number of holes and electrons in the region between the electrodes, and ΔP, ΔN are the corresponding fluctuations. Under ambipolar conditions, charge neutrality prevails (except for times of the order of a dielectric relaxation time $\sim10^{-11}$ sec) so that

$$\Delta N=\Delta P, \tag{11.3}$$

[98] J. E. Hill and K. M. van Vliet, Physica **24**, 709 (1958); (a) M. Lax and P. Mengert, "Influence of trapping, diffusion, and recombination on carrier concentration fluctuations," Proceedings Second Conference on Semiconductor Surfaces, December 2–4, 1959, J. Phys. Chem. Solids (to be published).

and the voltage autocorrelation can be written

$$\langle \Delta V(t)\Delta V(0)\rangle = V^2[(1+b)/(P+Nb)]^2 \times \langle \Delta P(t)\Delta P(0)\rangle, \quad (11.4)$$

where $b=\mu_n/\mu_p$ is the mobility ratio, and $\Delta P(t)$ is the total hole fluctuation between the electrodes. For a one-dimensional geometry, Hill and van Vliet write

$$\Delta P(t)=\int_0^L \Delta p(x,t)dx, \quad (11.5)$$

where $\Delta p(x,t)$, the fluctuation in local concentration, constitutes the basic set of parameters $\alpha(x)$ of the problem, and L is the distance between the electrodes. The equation analogous to

$$\alpha_i(t)=\sum_j[\exp(-\mathbf{\Lambda}t)]_{ij}\alpha_j(0) \quad (11.6)$$

for the continuous case is

$$\alpha_x(t)=\int[\exp(-\mathbf{\Lambda}t)]_{xx'}dx'\alpha_{x'}(0). \quad (11.7)$$

By assuming, for the moment, that the "Green's function,"

$$[\exp(-\mathbf{\Lambda}t)]_{xx'}=K(x,x',t), \quad (11.8)$$

is known for this problem, we can write

$$\Delta p(x,t)=\int K(x,x'',t)dx''\Delta p(x'',0) \quad (11.9)$$

so that the correlation at different times,

$$\langle \Delta p(x,t)\Delta p(x',0)\rangle = \int K(x,x'',t)dx''\langle \Delta p(x'',0)\Delta p(x',0)\rangle, \quad (11.10)$$

is reduced as usual to a knowledge of the correlations at one time. Since fluctuations in added carrier concentration (electron-hole pairs) occur without a change in Coulomb energy, it will be a good approximation to regard the added pairs as noninteracting. Thus there will be no correlation between fluctuations in added carrier concentrations at different points. The Debye-Hückel theory of Coulomb correlations[98a] is shown in Appendix C of reference 98(a) to reduce *approximately* to the same delta function autocorrelation that would be obtained by the methods of Sec. 12 in the absence of two-body forces. Thus we take

$$\langle \Delta p(x'',0)\Delta p(x',0)\rangle = \langle (\Delta P)^2\rangle L^{-1}\delta(x''-x'), \quad (11.11)$$

where the coefficient of the delta function has been chosen, considering Eq. (11.5) so that the fluctuation in the total number of carriers is $\langle(\Delta P)^2\rangle$. Equation (11.10) now yields

$$\langle \Delta p(x,t)\Delta p(x',0)\rangle = \langle (\Delta P)^2\rangle L^{-1}K(x,x',t). \quad (11.12)$$

A correlation at $t\neq 0$ arises for $x'\neq x$ because a single carrier can propagate from x' to x in time t.

In view of Eq. (11.5), the autocorrelation for the total number of carriers is

$$\langle \Delta P(t)\Delta P(0)\rangle = \langle (\Delta P)^2\rangle \Phi(t), \quad (11.13)$$

$$\Phi(t)=\frac{1}{L}\int_0^L dx\int_0^L dx' K(x,x',t), \quad (11.14)$$

and the voltage noise, from (11.4), (11.14), and (2.5), is

$$G(V)=V^2\left(\frac{1+b}{P+Nb}\right)^2\langle (\Delta P)^2\rangle\int_0^\infty 4\cos\omega t\Phi(t)dt. \quad (11.15)$$

The corresponding current noise may be obtained by multiplying by $(I/V)^2$. If holes and electrons fluctuated independently, one would have, by Eq. (4.25),

$$\langle(\Delta N)^2\rangle = N; \quad \langle(\Delta P)^2\rangle = P \quad (11.16)$$

(assuming there are no fluctuations in bound electrons). Under conditions of charge neutrality, by the argument leading to Eq. (4.30), we obtain

$$\langle(\Delta P)^2\rangle = \langle(\Delta N)^2\rangle = NP/(N+P) \quad (11.17)$$

(even when the quasi-Fermi levels of the two bands are unequal).

To determine $\Phi(t)$ we must obtain K. We now show that $K(x,x',t)$ represents the response at x at time t due to a unit pulse which originates at x' at time 0. In view of Eq. (11.8), as $t\to 0$,

$$K(x,x',t)\to(1)_{xx'}=\delta(x-x'). \quad (11.18)$$

Since we only use K for $t\geq 0$, we may define K to vanish for $t<0$, i.e., we may set

$$K(x,x',t)=[\exp(-\mathbf{\Lambda}t)]_{xx'}H(t), \quad (11.19)$$

where $H(t)$ is the Heaviside jump function [cf. Eqs. (8.3) and (8.4)] so that

$$(\partial/\partial t+\mathbf{\Lambda})K(x,x',t)=\delta(x-x')\delta(t), \quad (11.20)$$

in accord with Eq. (8.5). In other words, K is the response at x and t to a pulse originating at x' at time 0.

The phenomenological equations for ambipolar diffusion have been given by van Roosbroeck.[99] For the one-dimensional case, these equations take the form

$$(\partial/\partial t+1/\tau+v_a\partial/\partial x-D_a\partial^2/\partial x^2)\langle \Delta p(x,t)\rangle = 0. \quad (11.21)$$

Thus we may interpret $\mathbf{\Lambda}$ as the operator

$$\mathbf{\Lambda}=1/\tau+v_a\partial/\partial x-D_a\partial^2/\partial x^2. \quad (11.22)$$

[The reader may have anticipated, in view of our matrix notation that $\mathbf{\Lambda}\cdot\boldsymbol{\alpha}$ would become an integral instead of the sum which appeared in the discrete case,

$$\mathbf{\Lambda}\boldsymbol{\alpha}(x)=\int\Lambda_{xx'}dx'\alpha(x'). \quad (11.23)$$

[99] W. van Roosbroeck, Phys. Rev. **91**, 282 (1953).

The integral notation is indeed more general. Differential operators are certainly special cases of integral operators. For example, if we take

$$\Lambda_{xx'}=(1/\tau)\delta(x-x')-v_a\delta'(x-x')-D_a\delta''(x-x'), \quad (11.24)$$

where δ' and δ'' are first and second derivatives of a delta function, then insertion into Eq. (11.23) and integration by parts lead to the same result as the direct use of Eq. (11.22).]

The parameters v_a and D_a in (11.21) are the ambipolar drift velocity and diffusion constant defined by[99]

$$v_a=e\mu_a E, \quad \mu_a=(N-P)\mu_n\mu_p/(N\mu_n+P\mu_p) \quad (11.25)$$

and

$$D_a=(N+P)D_pD_n/(ND_n+PD_p), \quad (11.26)$$

where the individual diffusion constants and mobilities are related by the Einstein relation

$$D_n=kT\mu_n/e; \quad D_p=kT\mu_p/e. \quad (11.27)$$

However, the ambipolar constants are not so related.

We must now solve Eq. (11.20) with Eq. (11.22) for $\mathbf{\Lambda}$. For the case in which diffusion is unimportant, the solution (with neglect of end effects) can be written immediately as

$$K(x,x',t)=\exp(-t/\tau)\delta(x-x'-v_at) \quad (11.28)$$

so that Eq. (11.14) yields

$$\Phi(t)=\exp(-t/\tau)[1-(t/T_a)] \quad \text{for } t<T_a$$
$$=0 \quad \text{for } t>T_a, \quad (11.29)$$

where

$$T_a=L/v_a \quad (11.30)$$

is the time required to drift between electrodes.

If we convert Eq. (11.15) to current noise by multiplying by I^2/V^2, and eliminate one I by means of

$$IT_a=[e(P\mu_p+N\mu_n)V/L^2](L^2/\mu_aV)$$
$$=e(P+nN)^2/(|N-P|b), \quad (11.31)$$

where $b=\mu_n/\mu_p$, we obtain the current noise in the form given by Hill and van Vliet,[98]

$$G(I)=2eI_{eq}, \quad (11.32)$$

where

$$I_{eq}=\frac{2I(b+2+b^{-1})}{|(N/P)-(P/N)|}W \quad (11.33)$$

and

$$W=W(\omega\tau,\omega T_a)=\int_0^{T_a}\cos\omega t e^{-t/\tau}[1-(t/T_a)]d(t/T_a)$$

$$=\frac{(\tau/T_a)^2}{1+(\omega\tau)^2}\left\{\frac{T_a}{\tau}-\frac{1-(\omega\tau)^2}{1+(\omega\tau)^2}+\frac{\exp(-T_a/\tau)}{1+(\omega\tau)^2}\right.$$

$$\left.\times[(1-(\omega\tau)^2)\cos(\omega T_a)-2\omega\tau\sin(\omega T_a)]\right\}. \quad (11.34)$$

The noise spectrum exhibits oscillations with a period $\Delta\omega T_a=2\pi$.

For the case in which diffusion is included, the Green's function becomes

$$K(x,x',t)=(4\pi D_at)^{-\frac{1}{2}}\exp[-(t/\tau)-(x-x'-v_at)^2/(4D_at)]. \quad (11.35)$$

For the purposes of evaluating $\Phi(t)$ by Eq. (11.14), it is more convenient to re-express Eq. (11.35) in terms of the integral representation

$$K(x,x',t)=(2\pi)^{-1}\exp(-t/\tau)$$

$$\times\int dk\exp[ik(x-x'-v_at)-D_ak^2t] \quad (11.36)$$

with the result that

$$\Phi(t)=\frac{L}{2\pi}\int_{-\infty}^{\infty}dk\left[\frac{\sin(kL/2)}{(kL/2)}\right]^2$$

$$\times\exp[-(ikv_a+D_ak^2+1/\tau)t] \quad (11.37)$$

and

$$\int_0^{\infty}\cos\omega t\Phi(t)dt$$

$$=\frac{1}{\pi}\int_{-\infty}^{\infty}dx\left(\frac{\sin x}{x}\right)^2$$

$$\times\frac{(1/\tau)+(4D_a/L^2)x^2}{[(1/\tau)+(4D_a/L^2)x^2]^2+[\omega+(2x/T_a)]^2}. \quad (11.38)$$

Equations (11.37) and (11.38) exhibit a distribution of relaxation times $(1/\tau)+D_ak^2$ and a distribution of center frequencies kv_a. For large enough L, the right-hand side of Eq. (11.38) reduces to the familiar $\tau/(1+\omega^2\tau^2)$. Equation (11.38) is evaluated exactly in reference 98(a).

As a closing remark, we note that the Langevin viewpoint leads to an equation of the form

$$(\partial/\partial t+\mathbf{\Lambda})\alpha(x,t)=F(x,t), \quad (11.39)$$

where the random noise source $F(x,t)$ obeys Eq. (8.11),

$$\langle F(x,t)F(x',t')\rangle=2D(x,x')\delta(t-t'), \quad (11.40)$$

and Eq. (5.18) requires that

$$2D(x,x')=\mathbf{\Lambda}\langle\alpha(x,0)\alpha(x',0)\rangle+\text{transpose} \quad (11.41)$$

or, for this example,

$$2D(x,x')=[(1/\tau)+v_ad/dx-D_ad^2/dx^2]$$
$$\times\langle\Delta p(x,0)\Delta p(x',0)\rangle+\text{transpose} \quad (11.42)$$
$$=2\langle(\Delta P)^2\rangle/L[(1/\tau)\delta(x-x')-D_a\delta''(x-x')],$$

where the term in $v_a\delta'(x-x')$ is odd in $x-x'$ and cancels when the transpose on x and x' is added. This result

might have been difficult to deduce by more intuitive methods.[100]

In this section we have demonstrated how easy it is to apply our methods to variables α with a continuous index x. Our argument was incomplete in that the fluctuations at different points at the same time were assumed to be uncorrelated on the basis of heuristic arguments. Even if we accept these arguments, the conclusion, Eq. (11.11), that the correlation is proportional to $\delta(x''-x')$ is not guaranteed since derivatives of delta functions could occur. Hashitsume[78] has indeed shown that the usual expression for the entropy, valid near an equilibrium state, leads to the delta function form. Our heuristic arguments suggest that, for a problem in which only one-body forces and one-body collisions are important, this form remains valid even in the nonequilibrium case. In the next section we make a direct calculation of the fluctuations at one time by means of Eq. (5.18), and verify the delta function form for fluctuations from a steady state. [In the case when the total number of carriers is fixed, we find a correction term of nondelta function form. In the Hill-van Vliet experiment this correction term is unnecessary because there is no constraint on the total number of added electron-hole pairs.]

GENERALIZATIONS OF OUR METHOD TO CONTINUOUS PARAMETER AND NONLINEAR SYSTEMS

12. Fluctuations in Distribution Functions

We have been concerned largely with the fluctuations in a set of parameters $\mathbf{a}$ or in their deviations $\boldsymbol{\alpha}=\mathbf{a}-\mathbf{a}_0$ from the steady-state value $\mathbf{a}_0$. We could concern ourselves equally well with fluctuations[101] in the distribution functions $n(\mathbf{a},t)$ such that $n(\mathbf{a},t)d\mathbf{a}$ represents the number of systems between $\mathbf{a}$ and $\mathbf{a}+d\mathbf{a}$ at time t. If we call $n_0(\mathbf{a})$ the steady-state value of the distribution function, then we may write

$$\mathbf{a}_0=\int \mathbf{a}n_0(\mathbf{a})d\mathbf{a}, \tag{12.1}$$

$$\boldsymbol{\alpha}(t)=\int \mathbf{a}\Delta n(\mathbf{a},t)d\mathbf{a}, \tag{12.2}$$

where

$$\Delta n(\mathbf{a},t)=n(\mathbf{a},t)-n_0(\mathbf{a}), \tag{12.3}$$

with the result that

$$\langle\boldsymbol{\alpha}(t)\boldsymbol{\alpha}(0)\rangle=\int\int \mathbf{a}\mathbf{a}'d\mathbf{a}d\mathbf{a}'\langle\Delta n(\mathbf{a},t)\Delta n(\mathbf{a}'0)\rangle. \tag{12.4}$$

Equation (12.4) demonstrates that the noise in $\mathbf{a}$ can be calculated readily if the fluctuations in $n(\mathbf{a},t)$ are known. In particular, one must know the time-displaced correlation function whose importance in physical applications has been emphasized by Van Hove,[102] Vineyard,[103] Glauber,[104] and others.[105]

The introduction of distribution functions $n(\mathbf{a})$ rather than the original variables $\mathbf{a}$, like the introduction of second quantization into quantum mechanics, sets up a more powerful formalism than necessary for most problems. There are three occasions, however, when such a formalism is desirable, if not absolutely necessary:

(1) When we are concerned with fluctuations in quantities that are more naturally expressible in terms of $n(\mathbf{a})$ than $\mathbf{a}$. For example, in the previous section we were more concerned with the total number of carriers in a region than with the position of each, and it was more natural to work with $\Delta p(x)$ than with x.

(2) When the fluctuations in $\mathbf{a}$ are so large that the quasi-linear approximations made in Eqs. (5.13) and (5.14) are no longer valid. It is still generally true that the fluctuations in $n(\mathbf{a})$ are sufficiently small for linearization to be valid when the $n(\mathbf{a})$ are the random variables. For example, the velocities of conduction electrons in a solid vary widely compared to their average (drift) value. Thus, it is not legitimate to adopt a phenomenological equation of the form

$$d\mathbf{v}/dt=-\mathbf{v}/\tau \tag{12.5}$$

unless a relaxation time exists and is independent of velocity.[63] (Even if a relaxation time $\tau=\tau(\mathbf{v})$ existed, one could not linearize Eq. (12.5) by setting $\mathbf{v}=\mathbf{v}_0+\Delta\mathbf{v}$ because $\Delta\mathbf{v}\gg\mathbf{v}_0$.) However, the distribution function $n(\mathbf{v})$ can be treated by a linearization procedure.

(3) When pair forces and/or two-body collisions are involved. For example, if the Coulomb forces between electrons are important in influencing their motion or in producing collisions between pairs of electrons, one cannot use a one-body transition probability $w_{v'v}$ but needs instead a two-body transition probability $w_{v'''v'';v'v}$, and it is then simpler to follow the changes in $n(\mathbf{v},t)$ than in the velocities of a single or a pair of particles.

The procedure now described requires no originality. It is a straightforward application of the methods introduced in this paper to the variables $n(a,t)$ rather than a. (We understand that a can also represent a set of variables, but do not indicate this by boldface type

[100] M. Solow and R. L. Petritz, "Theory of noise in a multidimensional semi-conductor with a p-n junction," Navord Rept. 5762, thesis, Catholic University (1958). See also Van der Ziel, reference 33.

[101] For applications of the Nyquist theorem to fluctuations in continuous systems. See S. M. Rytov, Soviet Phys. JETP **6**, 130 (1958); Soviet Phys. Doklady **1**, 555 (1956).

[102] L. Van Hove, Phys. Rev. **95**, 249 (1954).

[103] G. H. Vineyard, Phys. Rev. **110**, 999 (1958).

[104] R. J. Glauber, Phys. Rev. **98**, 1692 (1955).

[105] The time-displaced correlation function was apparently introduced by G. I. Taylor, Proc. London Math. Soc. Ser. 2, **20**, 196 (1920), and has been extensively applied to the theory of turbulence. See, for example, G. K. Batchelor, *Theory of Homogeneous Turbulence* (Cambridge University Press, New York, 1953).

since no confusion can occur.) We indicate the procedure for the case of one-body collision. Since the changes in $n(a,t)$ are governed by the changes in a, we expect to express them in terms of the same transition probability $w_{a'a}$, that was used in our previous description which followed the motion of a single "particle." In that description, $w_{a'a}$ represented the transition probability for a "collision" which carries the particle from state a (which is singly occupied before the collision) to state a' (which is empty before the collision—since the particle was then in a).

In our present description there are $n(a)$ and $n(a')$ particles in states a and a' before the collision, and the transition rate $w_{a'a}$ is replaced by

$$x(a',a)=[1+\epsilon n(a')]w_{a'a}n(a), \tag{12.6}$$

where the first factor takes account of quantum statistics, with $\epsilon=0$, -1, or $+1$ for Boltzmann, Fermi-Dirac, and Einstein-Bose statistics, respectively.[106] We treat the "states" a as having discrete values, and use sums. These sums can be converted later to integrals for the continuous case.

The "master" transition probability from $n(a)$ to $n'(a)$ now can be written

$$W(n',n)=\sum_{a,a'\neq a} x(a',a)\delta[n'(a'),n(a')+1]$$
$$\times\delta[n'(a),n(a)-1]\prod_{a''\neq a,a'}\delta[n'(a''),n(a'')], \tag{12.7}$$

where $\delta(m,n)=1$ for $m=n$ and 0 otherwise is a Kronecker delta function.

We now perform the calculation defined by Eq. (5.5) to obtain the bth component of the first moment of the transition probability:

$$A_b(n)=\sum_{n'}[n'(b)-n(b)]W_{n'n}. \tag{12.8}$$

If one inserts Eq. (12.7) and sums first over n', the only terms in the sum which contribute are those for which $a=b$ and $a'=b$,

$$A_b(n)=-\sum_{a'}x(a',b)+\sum_a x(b,a). \tag{12.9}$$

Therefore, from Eqs. (5.9) and (12.6),

$$d\langle n(b,t)\rangle/dt=-\sum_{a'}\langle[1+\epsilon n(a')]n(b)\rangle w_{a'b}$$
$$+\sum_a\langle[1+\epsilon n(b)]n(a)\rangle w_{ba}. \tag{12.10}$$

This phenomenological equation is of the usual transport type. The mean rate of increase of particles in b is equal to the difference between the mean rates of flow into and out of b. The approach to a steady state of the solutions of an equation like (12.10), at least for the Boltzmann case, has been discussed by Siegert[107] and Watanabe.[108] The relation between the steady state and the state of minimum entropy production has been pointed out by Klein and Meijer.[109–114]

The steady-state values $n_0(b)$ are, as usual, determined by setting the right-hand side of Eq. (12.10) to zero. One is tempted to make the terms on the right-hand side cancel in pairs by detailed balance:

$$\frac{w_{ba}n_0(a)}{1+\epsilon n_0(a)}=\frac{w_{ab}n_0(b)}{1+\epsilon n_0(b)}. \tag{12.11}$$

However, if one requires that three states a, b, c be consistent with one another under this requirement, one finds that

$$f(b,a)=w_{ba}/w_{ab} \tag{12.12}$$

must satisfy the consistency requirement

$$f(c,a)=f(c,b)f(b,a), \tag{12.13}$$

a functional equation whose only solution is of the form

$$f(b,a)=g(b)/g(a). \tag{12.14}$$

Thus if the ratio of forward to reverse transition probabilities has the form (12.14), then there is a steady-state solution of the form

$$\frac{n_0(a)}{1+\epsilon n_0(a)}=\lambda g(a)$$

or

$$n_0(a)=\{[\lambda g(a)]^{-1}-\epsilon\}^{-1}, \tag{12.15}$$

where the proportionality constant λ is determined by normalization (i.e., by the requirement that the total number of particles be given correctly). Even in the nonequilibrium case, we can choose to define a quasi-Fermi[63] level μ by means of

$$\lambda=\exp(\mu/kT). \tag{12.16}$$

Under thermal equilibrium conditions,

$$g(a)=\exp[-E(a)/kT], \tag{12.17}$$

and Eq. (12.15) yields the conventional equilibrium result

$$n_0(a)=\{-\epsilon+\exp[(E(a)-\mu)/kT]\}^{-1}. \tag{12.18}$$

A nonequilibrium steady state may obey Eqs. (12.14)

[106] S. Vonsovsky, J. Phys. U.S.S.R. **10**, 367 (1946). See also Tolman, reference 6, Eq. (100.31).
[107] A. J. F. Siegert, Phys. Rev. **76**, 1708 (1949).
[108] S. Watanabe, Revs. Modern Phys. **27**, 26, 40, 179 (1955).
[109] M. J. Klein and P. H. E. Meijer, Phys. Rev. **96**, 250 (1954). The principle of minimum entropy production was formulated by Prigogine, reference 10. Discussions and applications of this principle are given in references 110–114.
[110] M. J. Klein, Phys. Rev. **98**, 1736 (1955).
[111] P. H. E. Meijer, Phys. Rev. **103**, 839 (1956).
[112] J. M. Ziman, Can. J. Phys. **34**, 12A, 1256 (1956), demonstrates that the minimum entropy production principle is equivalent to the variational principle formulated by M. Kohler, Z. Physik **124**, 772 (1948); **125**, 679 (1949); Ann. Physik (6) **6**, 18 (1949) and applied by him to the solution of transport problems. Later uses of this variational principle include references 113 and 114.
[113] E. H. Sondheimer, Proc. Roy. Soc. (London) **A203**, 75 (1950); **A234**, 391 (1956).
[114] D. J. Howarth and E. H. Sondheimer, Proc. Roy. Soc. (London) **A219**, 53 (1953).

to (12.16) but not (12.17) or (12.18). In that case, a steady state obeying detailed balance,[115] Eq. (12.11), but not thermal equilibrium, Eq. (12.18), is obtained. In general one obtains a steady state in which neither equilibrium nor detailed balance occur.

In any case, for small deviations from a steady state, we set

$$n(a)=n_0(a)+\Delta n(a), \tag{12.19}$$

and rewrite Eq. (12.10) in the form

$$d\langle\Delta n(b,t)\rangle/dt=-\sum_a \Lambda_{ba}\langle\Delta n(a,t)\rangle, \tag{12.20}$$

where

$$\Lambda_{ba}=-[1+\epsilon n_0(b)]w_{ba}+\epsilon n_0(b)w_{ab}, \quad b\neq a$$
$$\Lambda_{bb}=\sum_{a\neq b}\{[1+\epsilon n_0(a)]w_{ab}-w_{ba}\epsilon n_0(a)\}. \tag{12.21}$$

The equation for the second moments, after linearization, according to Eq. (5.17), is

$$d\langle\Delta n(b)\Delta n(c)\rangle/dt=2D_{bc}-\Lambda_{ba}\langle\Delta n(a)\Delta n(c)\rangle - \langle\Delta n(b)\Delta n(a)\rangle\Lambda_{ca}, \tag{12.22}$$

where summation over the repeated index a is understood. The diffusion matrix D_{bc} is defined as in Eq. (5.6) in terms of the second moment of the transition probability,

$$2D_{bc}=\sum_{n'}[n'(b)-n(b)][n'(c)-n(c)]W_{n'n}. \tag{12.23}$$

By using the definition (12.7) of $W_{n'n}$, one obtains

$$2D_{bc}=-[1+\epsilon n(b)]w_{bc}n(c)-[1+\epsilon n(c)]w_{cb}n(b) \tag{12.24}$$

for $c\neq b$. The two terms in D_{bc} are equal under detailed balance but not otherwise. For $c=b$, we obtain

$$2D_{bb}=\sum_a[1+\epsilon n(b)]w_{ba}n(a) + \sum_a[1+\epsilon n(a)]w_{ab}n(b). \tag{12.25}$$

The two sums in Eq. (12.25), become equal when we replace $n(a)$ and $n(b)$ by their steady-state values, since it is just the difference of these terms in Eq. (12.10) that represents a deviation from a steady state.

The steady-state second moments $\langle\Delta n(a)\Delta n(c)\rangle$ now are chosen so that the right-hand side of Eq. (12.22) vanishes, i.e., so that the Einstein relation is obeyed. In a formal way, this is accomplished by Eq. (5.19) or (5.20). In practice, however, for an *arbitrarily* chosen w_{ba}, we cannot expect to invert easily the matrix Λ. We expected, by intuitive arguments in Sec. 11, that there would be no correlation between fluctuations in different states. We therefore try a solution of the form

$$\langle\Delta n(a)\Delta n(c)\rangle=F(c)\delta(a,c) \tag{12.26}$$

to see under what condition it works. [We use a Kronecker delta function since a and b could refer to discrete states, e.g., quantum states. In the continuous case when sums are replaced by integrals, $\delta(a,c)$ is replaced by $\delta(a-c)$, the Dirac delta function.] For the Boltzmann case $\epsilon=0$, we find that this supposition is always successful in the form

$$\langle\Delta n(a)\Delta n(c)\rangle=n_0(c)\delta(a,c). \tag{12.27}$$

For the Fermi-Dirac and Einstein-Bose case, we find that the ansatz

$$\langle\Delta n(a)\Delta n(c)\rangle=[1+\epsilon n_0(c)]n_0(c)\delta(a,c) \tag{12.28}$$

is successful, providing the steady state is such as to obey detailed balance, Eq. (12.11), i.e., when the ratio of transition probabilities obeys Eq. (12.14). (If the ratio of transition probabilities is *arbitrary*, then an ansatz of the form (12.26) can be shown to fail, even though it is successful for the Boltzmann case.)

The fluctuations (12.27) and (12.28) are precisely what one would obtain if one assumed the validity of the customary entropy formula,

$$S=-k\sum_a\{n(a)\ \ln n(a)-\epsilon[1+\epsilon n(a)] \times\ln[1+\epsilon n(a)]\}, \tag{12.29}$$

and evaluated the fluctuations from the second derivatives of the entropy via Eqs. (4.2) and (4.4), i.e., by assuming the validity of the Einstein-Boltzmann point of view away from equilibrium. This suggests that the usual entropy formula is valid away from equilibrium for the Boltzmann case. For the Fermi and Einstein cases, the same remark applies provided one is in the neighborhood of a state of detailed balance.[115]

Conversely, for the Fermi and Einstein cases, if one is not near a state of detailed balance, Eq. (12.29) does not yield the correct fluctuations, and its use as an expression for the entropy is certainly open to question.

The formula

$$\lambda\partial n_0/\partial\lambda=kT\partial n_0/\partial u=(1+\epsilon n_0)n_0 \tag{12.30}$$

permits us to combine Eqs. (12.27) and (12.28) into the single equation

$$\langle\Delta n(a)\Delta n(c)\rangle=kT(\partial n_0(c)/\partial\mu)\delta(a,c) \tag{12.31}$$

in a form identical to that used in the thermal equilibrium case, Eq. (4.23).

If we denote the total number of systems by

$$N=\sum n(a),$$

a sum on a and c in Eq. (12.31) yields

$$\langle(\Delta N)^2\rangle=kT(\partial N_0/\partial u), \tag{12.32}$$

a result which we anticipate to be correct if particles are interchanged with a reservoir ("grand canonical ensemble"). However, we have postulated only those transitions $w_{a'a}$ which take a particle from a to a' without changing the total number of particles in the system; therefore, N should not fluctuate. The resolution of this paradox consists in noticing that we have found so far only a *particular* solution of the Einstein

[115] M. J. Klein, Phys. Rev. **97**, 1446 (1955), claims to show by an example, however, that the principle of detailed balance cannot hold in nonequilibrium steady states.

relation. This solution is not unique if one can find a solution of the homogeneous equation

$$\sum \Lambda_{ba}\langle \Delta n(a)\Delta n(c)\rangle = 0. \tag{12.33}$$

[The last term in Eq. (12.22) then also vanishes since it is the transpose of (12.33).]

By regarding $\langle \Delta n(a)\Delta n(c)\rangle$ as a function of the first variable a, we see that such a *homogeneous* solution would automatically obey Eq. (12.20) with $d\langle \Delta n(b,t)\rangle/dt = 0$. Thus, we could generate a new steady-state solution,

$$n_0(a)_{\text{new}} = n_0(a) + A\langle \Delta n(a)\Delta n(c)\rangle_{\text{hom}}, \tag{12.34}$$

with arbitrary A. We have tacitly assumed in the past that our system possesses a unique steady-state solution. And indeed, we continue to assume that our steady solutions are unique *except for normalization*.[116]

Since the Boltzmann equation is linear, a new steady-state solution can be obtained by taking $\Delta n(a) \propto n_0(a)$ or, by symmetry,

$$\langle \Delta n(a)\Delta n(c)\rangle_{\text{hom}} = n_0(a)n_0(c). \tag{12.35}$$

Equation (12.27) now can be replaced by

$$\langle \Delta n(a)\Delta n(c)\rangle = n_0(c)\delta(a,c) - (n_0(a)n_0(c)/N_0), \tag{12.36}$$

where the factor $-1/N_0$ has been chosen so that N does not fluctuate from its steady value N_0. For the Fermi and Einstein cases, we can obtain new steady solutions by changing the total number of particles by means of a change of λ in Eq. (12.15) or the quasi-Fermi level μ. For this case, $\Delta n(a) \propto \partial n_0(a)/\partial\mu$, and we obtain the final result

$$\begin{aligned}\langle \Delta n(a)\Delta n(c)\rangle &= kT(\partial n_0(c)/\partial\mu)\delta(a,c) - (kT\partial n_0(a)/\partial\mu) \\ &\quad \times (kT\,\partial n_0(c)/\partial\mu)/(kT\partial N_0/\partial\mu)\end{aligned} \tag{12.37}$$

or

$$\begin{aligned}\langle \Delta n(a)\Delta n(c)\rangle &= n_0(c)[1+\epsilon n_0(c)]\delta(a,c) \\ &\quad - \frac{n_0(a)[1+\epsilon n_0(a)]n_0(c)[1+\epsilon n_0(c)]}{\sum_b n_0(b)[1+\epsilon n_0(b)]}.\end{aligned} \tag{12.38}$$

The added term is of order $1/N$ and therefore is unimportant in calculating the fluctuations in any small portion of a large system. However, this term does affect fluctuations in appreciable parts of a system.

Consider, for example, a set of free electrons in different velocity states plus a set of bound electrons in traps. Let us calculate the fluctuations in the total number of free electrons, assuming that Boltzmann statistics is adequate for the free states, but that Fermi statistics is necessary for the bound states. Then we must sum Eq. (12.38) on a and c over the free states. By using

$$\sum_{\text{free}} n_0(c)[1-n_0(c)] \approx \sum n_0(c) = n, \tag{12.39}$$

$$\sum_{\text{bound}} n_0(b)[1-n_0(b)] = N_t(\hat{n}/N_t)[1-(\hat{n}/N_t)], \tag{12.40}$$

where N_t is the number of traps, n the steady number of free carriers, and $\hat{n}$ the steady number of bound carriers, Eq. (12.38) yields

$$\langle (\Delta n)^2\rangle = n - \frac{n^2}{n+\hat{n}[1-(\hat{n}/N_t)]}, \tag{12.41}$$

in agreement with Eqs. (4.30) and (3.35), the latter equations having been derived by two different methods for a two-state system.

As an exercise for the reader, we suggest the derivation of Eq. (12.38) from the distribution fuction

$$\delta[\sum \Delta n(a)] \exp\{-\tfrac{1}{2}\sum[\Delta n(a)]^2/\langle[\Delta n(a)]^2\rangle\}, \tag{12.42}$$

where $\langle[\Delta n(a)]^2\rangle$ is taken from the grand canonical result Eq. (12.31), and the constraint is imposed by the multiplying delta function.

The method which first suggested Eq. (12.38) to the author is discussed in Sec. 13. These various approaches were taken in order to make absolutely certain of the strange looking correction term in Eq. (12.38). After completing the manuscript, I found that in the thermal equilibrium case this correction was known to Fowler.[117]

The results of this section are, in a certain sense, obvious. For "particles" which move independently of one another, the occupancy of a given state is compounded of independent events, each "particle" having the same probability of occupying a given state. If the supply of "particles" is inexhaustible, we find that each state has a Poisson probability distribution of occupancy, and there is no correlation between the states. The second moment (12.27) is characteristic of such a Poisson distribution.

If one deals with a fixed total number of "particles" N, then we have the usual probability game of distributing N balls among a set of boxes, box a having probability $n_0(a)/N$, which leads to the standard multinomial distribution whose second moments are in accord with Eq. (12.36). The constraint $\delta(\sum_a n(a),N)$ thus leads to correlations between the occupancies. In fact, the multinomial distribution is simply the Poisson distribution, with the foregoing constraint as a multiplying factor, renormalized to unit total probability. A formal proof of the validity of the multinomial distribution for this Boltzmann problem has been given by Mathews, Shapiro, and Folkoff.[117a]

The probability distribution for any small number of states, obtained by summing the multinomial distribution over the occupancy of all other states can be

[116] Uniqueness has been established under fairly general conditions by Lebowitz and Bergmann, reference 48.

[117] See Fowler, reference 7, Eqs. (2115) and (2123); (a) Mathews, Shapiro, and Fulkoff, Bull. Am. Phys. Soc. Ser. II, 4, 15 (1959).

written down immediately by regarding all *other* states combined as a single state. For example, the distribution in occupancy of a single state a is a binomial distribution with probability $p=n_0(a)/N$ that state a is occupied and $q=1-p$ that the particle is in any other state. In particular, the second moment of this binomial distribution, Npq, agrees with Eq. (12.36) for $c=a$, whereas Eq. (12.36) with $c\neq a$ can be derived from a "trinomial" distribution which considers states a, c, and the remainder as the three possibilities.

The preceding results for the Boltzmann case lead us to suggest simple solutions to the problems for Fermi and Bose statistics when detailed balance[115] is obeyed. For the case of an inexhaustible reservoir, there will be no correlation between the states. For the Fermi case, state a is full or empty with probabilities $n_0(a)$ and $1-n_0(a)$, respectively. For the Bose case, the probability of occupancy $n(a)$ is

$$P[n(a)]=\{n_0(a)/[1+n_0(a)]\}^{n(a)}/[1+n_0(a)],$$

which is simply the result for thermal equilibrium with the equilibrium occupancy $\bar{n}(a)$ replaced by the steady, nonequilibrium occupancy $n_0(a)$.

The probability of a complete set of occupancies $\{n(a)\}$ is obtained by multiplying together the probabilities for each state a. If the total number of particles is fixed, one must multiply by the constraint $\delta(\sum_a n(a),N)$ and restore the normalization. Equations (12.28) and (12.38) yield second moments in agreement with these conclusions for the case of infinite and finite reservoirs, respectively.

13. Transformations on Random Variables

In solving problems involving a set of random variables, it is sometimes convenient to make a linear transformation to a new set of random variables in such a way that no correlations exist between the new variables.

Let us consider first a single variable $I(t)$ with autocorrelation

$$\langle I(t)I^*(s)\rangle=r(t,s), \tag{13.1}$$

where we have not assumed stationarity $r(t,s)=r(t-s)$. Suppose that the variables t and s are limited to some domain B. Let $f_n(t)$ be the set of orthonormal eigenfunctions of the integral equation[118,119]

$$\int_B r(t,s)f_n(s)ds=\nu_n f_n(t) \tag{13.2}$$

with eigenvalues ν_n. Then the random variables

$$c_n=\int f_n^*(t)I(s)ds \tag{13.3}$$

constitute an uncorrelated set for

$$\langle c_m c_n^*\rangle=\int f_m^*(t)dt\int ds\langle I(t)I^*(s)\rangle f_n(s)$$
$$=\nu_n\int f_m^*(t)f_n(t)dt=\nu_n\delta_{m,n}. \tag{13.4}$$

In general, $I(t)$ is real, and the $f_n(t)$ can be chosen real; however, it is sometimes convenient to make a complex choice. For example, if

$$r(t,s)=r(t-s), \tag{13.5}$$

and the domain of t is $-\infty<t<\infty$, as is appropriate for a stationary random process, then the integral, Eq. (13.2), has the set of eigenfunctions $\exp(-i\omega t)$ which form a continuum on ω, and one may verify by direct computation that if

$$i(\omega)=\int_{-\infty}^{\infty}\exp(-i\omega t)I(t)dt, \tag{13.6}$$

then

$$\langle i(\omega)i^*(\omega')\rangle=\pi\delta(\omega'-\omega)G(\omega), \tag{13.7}$$

where

$$G(\omega)=2\int_{-\infty}^{\infty}\exp(-i\omega t)\langle I(t)I(0)\rangle dt \tag{13.8}$$

is the noise spectrum. Equation (13.7) states that the different Fourier components of a stationary random variable are uncorrelated. This is the reason for the efficacy of the Fourier approach to noise introduced by Schuster[120] and Wiener,[60] and successfully exploited by Rice.[22]

It should be remarked that the construction we have given of uncorrelated variables does not lead to a unique result. After normalizing our variables to $c_n/(\nu_n)^{\frac{1}{2}}$ or $i(\omega)/[\pi G(\omega)]^{\frac{1}{2}}$, we can make any orthogonal or unitary transformation,

$$d_m=\sum O_{mn}c_n/(\nu_n)^{\frac{1}{2}}, \tag{13.9}$$

$$g_n=\int U_n(\omega)i(\omega)/[\pi G(\omega)]^{\frac{1}{2}}d\omega, \tag{13.10}$$

and the resulting variables d_m or g_n will be uncorrelated and have mean square unity (i.e., orthonormalized).

This lack of uniqueness is precisely what enables us to start with a set of variables $\alpha_i(t)$ and construct a new set $\beta_m(t)$ such that

$$\langle\beta_m(t)\beta_n(u)\rangle=\delta_{mn}F(t-u), \tag{13.11}$$

i.e., variables that are uncorrelated at $t=u=0$ and remain uncorrelated for arbitrary time differences.

A procedure for accomplishing (13.11) may be outlined briefly as follows:

[118] K. Karhunen, Ann. Acad. Sci. Fennicae Ser. A. I. **34** (1946); *ibid.* **37** (1947).

[119] M. Kac and A. J. F. Siegert, Ann. Math. Statistics **18**, 438 (1947). See also reference 14.

[120] A. Schuster, Proc. Roy. Soc. (London) **A77**, 136 (1906).

(1) Find an orthogonal transformation which diagonalizes the real symmetric matrix $\langle\alpha_i\alpha_j\rangle$. (Such a solution always exists.[121]) The new variables are uncorrelated but not normalized.

(2) Make a scale change in the new variables to provide normalization,

$$\langle\alpha_m'\alpha_n'\rangle=\delta_{mn}. \tag{13.12}$$

(3) The combined action of (1) and (2) constitutes a similarity transformation S under which

$$\langle\boldsymbol{\alpha}'\boldsymbol{\alpha}'\rangle=\mathbf{S}\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle\mathbf{S}^\dagger, \tag{13.13}$$

$$\mathbf{D}'=\mathbf{S}\mathbf{D}\mathbf{S}^\dagger, \tag{13.14}$$

$$\boldsymbol{\Lambda}'=\mathbf{S}\boldsymbol{\Lambda}\mathbf{S}^{-1}. \tag{13.15}$$

Hence, the symmetry of $\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle$ and $\mathbf{D}$ is preserved, but that of $\boldsymbol{\Lambda}$ is altered.

(4) The Einstein relation, Eq. (5.18), and time reversibility, Eq. (6.4), take the forms

$$2\mathbf{D}'=\boldsymbol{\Lambda}'+\boldsymbol{\Lambda}'^\dagger \tag{13.16}$$

and

$$\boldsymbol{\Lambda}'=\boldsymbol{\Lambda}'^\dagger \tag{13.17}$$

such that $\boldsymbol{\Lambda}'=\mathbf{D}'$ is a real symmetric matrix.

(5) Hence, an *orthogonal* transformation *exists*,[121]

$$\beta_r=\textstyle\sum_m O_{rm}\alpha_m', \tag{13.18}$$

which diagonalizes $\boldsymbol{\Lambda}'$:

$$\boldsymbol{\Lambda}'\beta_r=\lambda_r\beta_r. \tag{13.19}$$

(6) The orthogonality of the transformation preserves the orthonormality [Eq. (13.12)] of the fluctuations,

$$\langle\beta_r\beta_s\rangle=\delta_{rs}, \tag{13.20}$$

with the result that

$$\begin{aligned}\langle\beta_r(t)\beta_s\rangle&=\exp(-\boldsymbol{\Lambda}'t)\langle\beta_r\beta_s\rangle\\&=\exp(-\lambda_r t)\delta_{rs}\end{aligned} \tag{13.21}$$

has the desired lack of correlation shown in Eq. (13.11).

(7) In the absence of time reversibility, $\boldsymbol{\Lambda}'$ is not symmetric. It can be diagonalized by β_r, but the nonorthogonality of the transformation destroys Eq. (13.20). One can then introduce the eigenvectors γ_r of the transposed matrix

$$\boldsymbol{\Lambda}^\dagger\gamma_r=\lambda_r\gamma_r \tag{13.22}$$

which form a set orthogonal to the β's,[122]

$$\langle\beta_r\gamma_s\rangle=\delta_{rs}, \tag{13.23}$$

and

$$\langle\beta_r(t)\gamma_s\rangle=\exp(-\lambda_r t)\delta_{rs}, \tag{13.24}$$

$$\langle\beta_r\gamma_s(t)\rangle=\exp(-\lambda_s t)\delta_{rs}. \tag{13.25}$$

Occasionally, $\boldsymbol{\Lambda}$ and $\boldsymbol{\Lambda}^\dagger$ can be diagonalized by the same set of eigenvectors and the complication of a biorthogonal set of random variables need not be introduced.

To illustrate these remarks we consider the Boltzmann transport equation,

$$\partial f/\partial t=\int w(\mathbf{v},\mathbf{v}')f(\mathbf{v}')d\mathbf{v}'-[1/\tau(\mathbf{v})]f(\mathbf{v}), \tag{13.26}$$

where

$$1/\tau(\mathbf{v})=\int w(\mathbf{v}',\mathbf{v})d\mathbf{v}' \tag{13.27}$$

is the mean collision rate. We take as our steady-state solution a Boltzmann distribution,

$$f_B(v)=N(m/2\pi kT)^{\frac{3}{2}}\exp(-\tfrac{1}{2}mv^2/kT), \tag{13.28}$$

normalized to a total of N particles, and assume that detailed balance holds:

$$w(\mathbf{v},\mathbf{v}')f_B(v')=w(\mathbf{v}',\mathbf{v})f_B(v). \tag{13.29}$$

Our theory of fluctuations in distribution functions, Eq. (12.27), yields for the fluctuations at one time

$$\langle\Delta f(\mathbf{v})\Delta f(\mathbf{v}')\rangle=f_B(v)\delta(\mathbf{v}-\mathbf{v}'), \tag{13.30}$$

if we neglect the correction in Eq. (12.36) for the case of a fixed total number of particles.

Since our moments are already in diagonal form, we may start our procedure with step (2) by introducing

$$\Delta f(\mathbf{v})=[f_B(v)]^{\frac{1}{2}}g(\mathbf{v}) \tag{13.31}$$

to obtain the normalized moments,

$$\langle g(\mathbf{v})g(\mathbf{v}')\rangle=\delta(\mathbf{v}-\mathbf{v}'). \tag{13.32}$$

Our phenomenological equation now takes the form

$$\partial g/\partial t=\int L(\mathbf{v},\mathbf{v}')g(\mathbf{v}',t)d\mathbf{v}'-g(\mathbf{v},t)/\tau(\mathbf{v}), \tag{13.33}$$

where

$$L(\mathbf{v},\mathbf{v}')=[f_B(v)]^{-\frac{1}{2}}w(\mathbf{v},\mathbf{v}')[f_B(v')]^{\frac{1}{2}} \tag{13.34}$$

is a real *symmetric* matrix because of the detailed balance condition (13.29), as promised in step (4). But real symmetric matrices always possess a set of orthonormal eigenvectors,[121] which we may assume to be complete, i.e., there exists a complete set of functions $\psi_{nl}(\mathbf{v})$ obeying

$$\begin{aligned}\int L(\mathbf{v},\mathbf{v}')d\mathbf{v}'\psi_{nl}(\mathbf{v}')-\psi_{nl}(\mathbf{v})/\tau(\mathbf{v})\\=-\lambda_{nl}\psi_{nl}(\mathbf{v})\end{aligned} \tag{13.35}$$

and

$$\int\psi_{nl}(\mathbf{v})\psi_{n'l'}(\mathbf{v})d\mathbf{v}=\delta_{nn'}\delta_{ll'}, \tag{13.36}$$

where the eigenvalues λ_{nl}, of the original equation [Eq.

[121] H. Weyl, *The Theory of Groups and Quantum Mechanics*, translated by H. P. Robertson (E. P. Dutton and Company, Inc., New York, 1931), Chap. 1, Sec. 5.

[122] Morse and Feshbach, reference 56, p. 884.

(13.26)] have been shown to be semipositive

$$\lambda_{nl} \geq 0 \tag{13.37}$$

by Siegert.[107] We use a double index notation, n for radial dependence, l for angular dependence, for reasons which shortly become evident. In any case, we now can write

$$\Delta f(\mathbf{v},t)[f_B(v)]^{-\frac{1}{2}} = g(\mathbf{v},t) = \sum c_{nl}(t)\psi_{nl}(\mathbf{v}), \tag{13.38}$$

with

$$\langle c_{nl}c_{n'l'}\rangle = \int \psi_{nl}(\mathbf{v})\psi_{n'l'}(\mathbf{v}')d\mathbf{v}d\mathbf{v}'\langle g(\mathbf{v})g(\mathbf{v}')\rangle$$

$$= \delta_{nn'}\delta_{ll'}. \tag{13.39}$$

A comparison between Eqs. (13.33), (13.35), and (13.38) also yields

$$\partial\langle c_{nl}(t)\rangle/\partial t = -\lambda_{nl}\langle c_{nl}(t)\rangle. \tag{13.40}$$

Thus, the autocorrelation is

$$\langle c_{nl}(t)c_{n'l'}(0)\rangle = \exp(-\lambda_{nl}t)\delta_{nn'}\delta_{ll'}, \tag{13.41}$$

and the fluctuations of $\Delta f(\mathbf{v},t)$ are now completely determined via Eq. (13.38).

Actually, one member of our set must be

$$\psi_{00}(v) = [f_B(v)]^{\frac{1}{2}}/N^{\frac{1}{2}}, \tag{13.42}$$

corresponding to the original Boltzmann distribution, with eigenvalue $\lambda_{00}=0$, and we assume that this is the only stationary solution. The condition that there be no fluctuation in the total number of particles can, by using Eqs. (13.38), (13.42), and (13.36), be written in the form

$$0 = \int \Delta f d\mathbf{v} = \sum c_{nl}(t)\int \psi_{nl}(\mathbf{v})\psi_{00}(v)d\mathbf{v} = c_{00}(t) \tag{13.42}$$

so that the variable c_{00} does not fluctuate. If we assume the validity of Eq. (13.39) for all other variables, we now have

$$\langle g(\mathbf{v})g(\mathbf{v}')\rangle = \sum \psi_{nl}(\mathbf{v})\psi_{n'l'}(\mathbf{v}')\langle c_{nl}c_{n'l'}\rangle$$
$$= \sum_{nl\neq 00} \psi_{nl}(\mathbf{v})\psi_{nl}(\mathbf{v}')$$
$$= \delta(\mathbf{v}-\mathbf{v}') - [f_B(v)f_B(v')]^{\frac{1}{2}}/N, \tag{13.43}$$

where we have assumed completeness in replacing the unrestricted sum by $\delta(\mathbf{v}-\mathbf{v}')$. Finally, we obtain

$$\langle \Delta f(\mathbf{v})\Delta f(\mathbf{v}')\rangle$$
$$= [f_B(v)f_B(v')]^{\frac{1}{2}}\delta(\mathbf{v}-\mathbf{v}') - f_B(v)f_B(v')/N, \tag{13.44}$$

in agreement with Eq. (12.36) for the correction due to the constraint of a fixed total number of particles.

For the common case in which $w(\mathbf{v},\mathbf{v}')$ is invariant against a simultaneous rotation of $\mathbf{v}$ and $\mathbf{v}'$, i.e., is a function only of the angle γ between the vectors

$$w(\mathbf{v},\mathbf{v}') = w(v,v';\cos\gamma), \tag{13.45}$$

$$L(\mathbf{v},\mathbf{v}') = L(v,v';\cos\gamma), \tag{13.46}$$

we may take the ψ_{nl} to be of the form

$$\psi_{nl}(\mathbf{v}) = [F_{nl}(v)/v]Y_l(\cos\theta), \tag{13.47}$$

where we adopt the convention

$$\int Y_l(\cos\theta)Y_{l'}(\cos\theta)d\Omega = \delta_{ll'} \tag{13.48}$$

or

$$Y_l(\cos\theta) = (2l+1/4\pi)^{\frac{1}{2}}P_l(\cos\theta). \tag{13.49}$$

Then the F_{nl} obey an integral equation of the form

$$\int K_l(v,v')F_{nl}(v')dv' - F_{nl}(v)/\tau(v) = -\lambda_{nl}F_{nl}, \tag{13.50}$$

and

$$\int F_{nl}(v)F_{n'l}(v)dv = \delta_{nn'}, \tag{13.51}$$

with the kernel given by

$$K_l(v,v') = vv'\int L(v,v';\cos\gamma)P_l(\cos\gamma)d\Omega. \tag{13.52}$$

These results justify our double indexing procedure.

With the help of Eqs. (13.38) and (13.41), the autocorrelation of any variable of the form

$$I(t) = \int h(\mathbf{v})f(\mathbf{v},t)d\mathbf{v} \tag{13.53}$$

can be written in the form

$$\langle \Delta I(t)\Delta I(0)\rangle = \sum_{n,l}{}' \left| \int [f_B(\mathbf{v})]^{\frac{1}{2}}h(\mathbf{v})\psi_{nl}(\mathbf{v})d\mathbf{v}\right|^2$$
$$\times \exp(-\lambda_{nl}t) \tag{13.54}$$

or, when Eqs. (13.45) to (13.47) hold,

$$\langle \Delta I(t)\Delta I(0)\rangle = \sum_{l,n}{}' \left| \int v[f_B(v)]^{\frac{1}{2}}h_l(v)F_{nl}(v)dv\right|^2$$
$$\times \exp(-\lambda_{nl}t), \tag{13.55}$$

where

$$h_l(v) = \int h(\mathbf{v})Y_l(\cos\theta)d\Omega. \tag{13.56}$$

The foregoing sums are primed to indicate the omission of the $n=0$, $l=0$ term which does not contribute to the fluctuations. The fluctuations in $I(t)$ contain a sum of decay times weighted by a factor which depends on the

extent to which $h(\mathbf{v})$ couples the lowest mode ψ_{00} to the others.

Unfortunately, the calculation of Eq. (13.55) requires a detailed knowledge of all the eigenstates and eigenvalues associated with the scattering probabilities $w(\mathbf{v},\mathbf{v}')$. When the scattering is nearly elastic, as is often the case, it is possible for the terms in $l\neq 0$ to make the relaxation approximation

$$\sum_n F_{nl}(v)F_{nl}(v')\exp(-\lambda_{nl}t) \approx \delta(v-v')\exp[-\lambda_l(v)t], \quad (13.57)$$

by means of which Eq. (13.35) can be simplified to

$$\langle \Delta I(t)\Delta I(0)\rangle \approx \sum_l \int f_B(v)h_l(v)\exp[-\lambda_l(v)t]v^2dv. \quad (13.58)$$

Equation (13.57) is equivalent to the statement that the random variable

$$F_l(v,t)=\sum F_{nl}(v)c_{nl}(t) \quad (13.59)$$

decays approximately as

$$\langle F_l(v,t)\rangle \approx \exp[-\lambda_l(v)t]F_l(v,0). \quad (13.60)$$

Comparison with Eq. (13.50) indicates that $F_l(v,t)$ obeys

$$\partial F_l/\partial t=\int K_l(v,v')F_l(v')dv'-F_l(v)/\tau(v). \quad (13.61)$$

The assumption of nearly elastic collisions is equivalent to the statement that $K_l(v,v')$ has a sharp peak near $v'=v$. Thus, it is legitimate to write

$$\int K_l(v,v')F_l(v')dv' \approx F_l(v)\int K_l(v,v')dv' \quad (13.62)$$

which immediately leads to Eq. (13.60) with

$$\lambda_l(v)=1/\tau(v)-\int K_l(v,v')dv' \quad (13.63)$$

$$=\int w(\mathbf{v}',\mathbf{v})d\mathbf{v}' \times[1-vf_B(v)P_l(\cos\gamma)/v'f_B(v')] \quad (13.64)$$

$$\approx \langle 1-P_l(\cos\gamma)\rangle/\tau(v), \quad (13.65)$$

if we set $v'\approx v$ inside the brackets.

The relaxation approximation, Eq. (13.60), breaks down for the case $l=0$, because $F_0(v,t)$ is normalized and does not decay, but approaches equilibrium by drift and diffusion of the speed or energy. [Equation (13.65) would yield $\lambda_0(v)=0$.] For this case, $F_l(v')$ in Eq. (13.61) is expanded about $v'=v$ up to terms of second order, leading to an approximation of Fokker-Planck type,

$$\partial F_0/\partial t=\left[\int K_0(v,v')dv'-1/\tau(v)\right]F_0(v) +\left[\int K_0(v,v')(v'-v)dv'\right]F_0'(v) +\tfrac{1}{2}\left[\int K_0(v,v')(v'-v)^2dv'\right]F_0''(v). \quad (13.66)$$

The eigenvalue problem obtained by setting $\partial F_0/\partial t=-\lambda_0 F_0$ involves a second-order differential equation that can be reduced easily to Sturm-Liouville form. Exact and approximate methods of finding eigenvalues and eigenfunctions of such second-order equations are well known,[56] and are not discussed here.

If the fluctuations in v were small compared to mean v, then we would have a slightly nonlinear system, and the much simpler procedures of the next section could be used.

14. Slightly Nonlinear Systems

We turn to the consideration of slightly nonlinear systems for several reasons. (1) We can extend the scope of the quasi-linear methods we have introduced and estimate the errors in a strictly linear approach. (2) We can explain the contradictory results already obtained by MacDonald,[123] van Kampen,[124] Davies,[125] and Alkemade.[126] (3) We can reconsider the physical assumptions made by the preceding authors.

We consider in detail the problem proposed by MacDonald: a condenser C in series with a nonlinear resistance $I=I(V)$ such that, with $V=q/C$,

$$dq/dt=-I(q/C)\equiv A(q)\simeq-(\Lambda q+Bq^2+\Gamma q^3). \quad (14.1)$$

With van Kampen, we emphasize that this problem is more difficult than the purely mathematical one of noise of known statistical properties passing through a nonlinear device.[127] The problem of Brownian motion in a nonlinear field of force considered by Kramers[128] is in the same category, because the random force is is unchanged. We are concerned specifically with the case in which the dissipative element is nonlinear. In in this case, the random noise source is itself modified, in an as yet unknown way, by the nonlinearity.

Alkemade considered an idealized vacuum diode by kinetic methods. We discuss his results in due course. The remaining three authors all make the Brownian

[123] D. K. C. MacDonald, Phys. Rev. **108**, 541 (1957).
[124] N. G. Van Kampen, Phys. Rev. **110**, 319 (1958).
[125] R. O. Davies, Physica **24**, 1055 (1958).
[126] C. T. J. Alkemade, Physica **24**, 1029 (1958).
[127] D. M. Middleton, J. Appl. Phys. **22**, 1143, 1153 (1951), provides references to the literature on noise passing through nonlinear devices.
[128] H. A. Kramers, Physica **7**, 284 (1940).

motion *approximation*. If we write

$$D_n(q)=\lim_{\Delta t\to 0}\frac{1}{n!}\int (q'-q)^n P(q|q',\Delta t)dq'/(\Delta t), \quad (14.2)$$

where

$$D_1(q)=A(q); \quad D_2(q)=D(q),$$

then the Brownian motion approximation consists in the assumption that all higher moments of the transition probability $P(q|q',\Delta t)$ vanish sufficiently rapidly for $D_n(q)$ to equal zero for $n>2$. In short, MacDonald, van Kampen, and Davies accept the validity of the Fokker-Planck equation,

$$\partial P/\partial t=-\partial/\partial q[A(q)P(q)]+\partial^2/\partial q^2[D(q)P(q)]. \quad (14.3)$$

Davies and van Kampen, moreover, *assume* that the *equilibrium* condition implies that, even in the presence of the nonlinearity, Eq. (14.3) possesses the *usual* equilibrium solution $P(q,t)\equiv W(q)$ with

$$W(q)=(2\pi kTC)^{-\frac{1}{2}}\exp[-(q-\bar{q})^2/(2kTC)], \quad (14.4)$$

with

$$\bar{q}=0, \quad (14.5)$$

whereas MacDonald makes the milder assumption

$$\langle q^2\rangle=kTC, \quad (14.6)$$

which he justified by an argument based on the second law of thermodynamics.

The neglect of $\bar{q}$ in $W(q)$ [Eq. (14.4)] or in the second moment [Eq. (14.6)] is certainly not consistent when *odd* nonlinearities are present, i.e., if $B\neq 0$ in Eq. (14.1).

The condition that $W(q)$ be a stationary solution of the Fokker-Planck equation [Eq. (14.4)] is taken by van Kampen and Davies to be

$$A(q)=[W(q)]^{-1}\partial[W(q)D(q)]/\partial q, \quad (14.7)$$

$$A(q)=-D(q)(q-\bar{q})/\sigma^2+D'(q), \quad (14.8)$$

where σ^2 is an abbreviation for

$$\langle(q-\bar{q})^2\rangle=kTC=\sigma^2. \quad (14.9)$$

Van Kampen eliminates $A(q)$ from Eq. (14.3) by means of (14.7), and then solves by a perturbation from the linear case for the cases $D(q)=1+\gamma q^2$ and $D(q)=1+\beta q$. We show that van Kampen's perturbation treatment is flawless. The discrepancy between van Kampen and the other authors arises because he determines $D(q)$ from $A(q)$ by means of Eq. (14.8) with $D'(q)$ and $\bar{q}$ omitted. These terms are small, leading only to second-order corrections, but that is precisely the size of the effect to be computed.

In fact, if we write

$$D(q)=D+Eq+Fq^2, \quad (14.10)$$

then Eq. (14.8) leads to the conditions

$$\Gamma\sigma^2=F, \quad B\sigma^2=E-F\bar{q},$$
$$\Lambda\sigma^2=D-E\bar{q}-2F\sigma^2, \quad \bar{q}=-(E\sigma^2)/D, \quad (14.11)$$

expressing Λ, B, Γ, and $\bar{q}$ in terms of D, E, and F. These conditions, which have not been stated before, can be inverted approximately to express all quantities in terms of Λ, B, Γ. We write the results correct to fourth order,

$$\bar{q}/\sigma\approx -b(1-3g+b^2),$$
$$D\approx\Lambda\sigma^2[1+2g-b^2(1-4g+b^2)], \quad (14.12)$$
$$E\approx\Lambda\sigma^2 b(1-g), \quad F\approx\Lambda g,$$

where

$$b=B\sigma/\Lambda, \quad g=\Gamma\sigma^2/\Lambda \quad (14.13)$$

are dimensionless small quantities of first and second order, respectively. [In an application to a real problem, it would not be proper to use these relations beyond the second order of accuracy without including contributions of higher-order nonlinearities.]

For purposes of comparison with experiment, the admittance $G(V)$ may be defined by

$$I(V)=G(V)V=[G_0+G_1V+G_2V^2]V. \quad (14.14)$$

Comparison with Eq. (14.1) then yields

$$\Lambda=G_0/C; \quad B=G_2/C^2; \quad \Gamma=G_2/C^3, \quad (14.15)$$

$$b=(G_1/G_0)(kT/C)^{\frac{1}{2}}; \quad g=(G_2/G_0)(kT/C). \quad (14.16)$$

The nonlinear effects then approach zero as $C\to\infty$, and indeed are small for practical C's and moderate ratios of G_1/G_0 and G_2/G_0.

In our computations, we do not start by assuming the Fokker-Planck equation or the equilibrium distribution, Eq. (14.4), because we do not agree with the assumptions involved. For most problems, the higher moments $D_n(q)$ exist for $n>2$. For linearized systems, the "Brownian motion" approximation which neglects them is adequate, because, as we have seen, the linear calculation of noise simply does not involve these higher moments. However, as we show soon a calculation of the noise to the second order in the nonlinearity (and this is the lowest nonvanishing increment in the Brownian approximation) involves both D_3 and D_4.

We do not use the equilibrium Gaussian distribution (14.4) for two reasons. First, we wish our calculations to be valid for the nonequilibrium steady state; second, even for the equilibrium case, Eq. (14.4) is likely not to be precisely true. We prove the latter point by means of two examples. The first, due to John Hopfield of these Laboratories, makes use of an ideal rectifier with infinite impedance for voltages in one direction, and with essential zero impedance for voltages above some threshold in the other direction. The voltage fluctuations (and hence the charge fluctuations) on the condenser cannot be a Gaussian because voltage fluctuations above the threshold in the "easy" direction cannot occur.

This example suggests that, for slightly nonlinear systems, there can be a small but significant deviation from the equilibrium distribution. We demonstrate the

latter remark by considering a low-density electron gas colliding with impurity scatterers which interact with the electric field produced by the charge on the condenser. The Hamiltonian for such a system can be written

$$H=(p^2/2m)+\sum V(x-x_j)-exE+(q^2/2C), \quad (14.17)$$

where

$$E=V/d=q/(dC), \quad (14.18)$$

and d is the distance over which the voltage V is applied. The interaction Hamiltonian over kT can be rewritten as

$$H_{\text{int}}/kT\equiv -eqx/(dkTC)=-\lambda(x/d)[q/(kTC)^{\frac{1}{2}}], \quad (14.19)$$

where

$$\lambda=e/(kTC)^{\frac{1}{2}}. \quad (14.20)$$

In terms of the dimensionless variables $x'=x/d$ and $q'=q/(kTC)^{\frac{1}{2}}$, we see that λ is a proper measure of the smallness of the coupling. In terms of these dimensionless variables, the Boltzmann distribution takes the form,

$$P(q')\propto\int\exp[-\tfrac{1}{2}(q'-\lambda x')^2-\tfrac{1}{2}f(x')]dx', \quad (14.21)$$

of a convolution of two distributions $\exp(-\frac{1}{2}q^2)$ and $\exp[-\frac{1}{2}f(x)]$ (we hereafter drop the primes). It is well known that the characteristic function (i.e., the Fourier transform) of $P(q)$ is the product of the characteristic functions of these two distributions, and the Thiele semi-invariants are additive,[129] so that, dropping the primes,

$$\langle q\rangle=0+\lambda\langle x\rangle, \quad (14.22)$$

$$\langle q^2\rangle=1+\lambda^2\langle x^2\rangle, \quad (14.23)$$

$$\langle(q-\langle q\rangle)^3\rangle=0+\lambda^3\langle(x-\langle x\rangle)^3\rangle, \quad (14.24)$$

$$\langle(q-\langle q\rangle)^4\rangle-3[\langle(q-\langle q\rangle)^2\rangle]^2 =0+\lambda^4[\langle(x-\langle x\rangle)^4\rangle-3\langle(x-\langle x\rangle)^2\rangle^2], \quad (14.25)$$

results which can also be obtained by integrating $q^n=(q-\lambda x+\lambda x)^n$ against Eq. (14.21) for $n=1, 2, 3, 4$. Thus, the third, fourth, and fifth semi-invariants vanish vanish as λ^3, λ^4, λ^5, etc. To accuracy λ^2, the distribution function is therefore a Gaussian. But the first and second moments deviate from the expected values by terms of order λ and λ^2, respectively. We may expect that the moments in x are of order of magnitude unity. When the distribution in x is uniform and centered,

$$\langle x^2\rangle=\int_{-\frac{1}{2}}^{\frac{1}{2}}x^2dx=\tfrac{1}{12},$$

and the correction in the second moment of the charge is $(\lambda^2/12)$. Now the largest feasible value of

$$\lambda^2=e^2/(kTC),$$

obtained by setting $T=1°\text{K}$ and $C=1\,\mu\mu\text{f}(10^{-x}f)$ is 10^{-3}. For practical purposes, one could neglect the shift in $\langle q^2\rangle$. Unfortunately, as we see, the fractional change in the noise due to nonlinearities is of order λ^2 in the case of a rectifier. The nonlinear problem is therefore a very delicate one. The assumptions usually made, which are quite reasonable, are no longer good enough, because one now is attempting to compute the fluctuations to a higher order of accuracy.

The starting point of our Markoffian calculation is, as usual, the transition probability integral equation [Eq. (5.1)]. For simplicity of notation, we restrict ourselves to one random variable. The equations for the higher moments then can be derived by methods precisely the same as those used to get Eq. (5.12) for the second moments. With $\alpha=a-a_0$, we obtain, for the first five moments,

$$d\langle\alpha\rangle/dt=\langle A(a)\rangle, \quad (14.26)$$

$$d\langle\alpha^2\rangle/dt=2\langle D(a)\rangle+2\langle A(a)\alpha\rangle, \quad (14.27)$$

$$d\langle\alpha^3\rangle/dt=6\langle D_3(a)\rangle+6\langle D(a)\alpha\rangle+3\langle A(a)\alpha^2\rangle, \quad (14.28)$$

$$d\langle\alpha^4\rangle/dt=24\langle D_4(a)\rangle+24\langle D_3(a)\alpha\rangle +12\langle D(a)\alpha^2\rangle+4\langle A(a)\alpha^3\rangle, \quad (14.29)$$

$$d\langle\alpha^5\rangle/dt=120\langle D_5(a)\rangle+120\langle D_4(a)\alpha\rangle +60\langle D_3(a)\alpha^2\rangle+20\langle D(a)\alpha^3\rangle+5\langle A(a)\alpha^4\rangle, \quad (14.30)$$

where $D_n(a)$ is defined by Eq. (14.2) with q replaced by a. We use the letter a rather than q because the present analysis is a general one, and is applicable to other examples as well as to MacDonald's[123] example of a condenser in series with a nonlinear resistance. [Indeed the foregoing equations are applicable to the case where α represents a set of variables. It is only necessary to symmetrize the right-hand side. The numerical coefficients actually indicate the number of distinct terms that would be obtained if all indices were unequal. For example, there are $4!=24$ permutations of 4 indices, hence the coefficient 24 in the first two terms of Eq. (14.29). In the third term, such permutations as $D_{12}\alpha_3\alpha_4$ and $D_{12}\alpha_4\alpha_3$ are not distinct, and there are only 12 terms in the symmetrized expression. In other words, if we were to write out all 12 terms, each such numerical coefficient would be unity.]

If one expands all moments to the second order,

$$D_n(a)=D_n+E_n\alpha+F_n\alpha^2. \quad (14.31)$$

Then the preceding equations, with the unit of time chosen so that $\Lambda=1$, take the form

$$d\langle\alpha\rangle/dt=-\langle\alpha\rangle-B\langle\alpha^2\rangle-\Gamma\langle\alpha^3\rangle, \quad (14.32)$$

$$d\langle\alpha^2\rangle/dt=2D+2E\langle\alpha\rangle-2(1-F)\langle\alpha^2\rangle -2B\langle\alpha^3\rangle-2\Gamma\langle\alpha^4\rangle, \quad (14.33)$$

$$d\langle\alpha^3\rangle/dt=6D_3+6(D+E_3)\langle\alpha\rangle+6(E+F_3)\langle\alpha^2\rangle -3(1-2F)\langle\alpha^3\rangle-3B\langle\alpha^4\rangle-3\Gamma\langle\alpha^5\rangle, \quad (14.34)$$

[129] See H. Cramer, reference 88.

$$d\langle\alpha^4\rangle/dt=24D_4+24(D_3+E_4)\langle\alpha\rangle +12(D+2E_3+2F_4)\langle\alpha^2\rangle +12(E+2F_3)\langle\alpha^3\rangle-4(1-3F)\langle\alpha^4\rangle -4B\langle\alpha^5\rangle-4\Gamma\langle\alpha^6\rangle, \quad (14.35)$$

$$d\langle\alpha^5\rangle/dt=120D_5+120(D_4+E_5)\langle\alpha\rangle +60(D_3+2E_4+2F_5)\langle\alpha^2\rangle +20(D+3E_3+6F_4)\langle\alpha^3\rangle+20(E+3F_3)\langle\alpha^4\rangle -5(1-4F)\langle\alpha^5\rangle-5B\langle\alpha^6\rangle-5\Gamma\langle\alpha^7\rangle. \quad (14.36)$$

The steady-state moments may be obtained by equating the time derivatives to zero and solving for $\langle\alpha^n\rangle$. One may argue that one has more unknowns $\langle\alpha\rangle$ to $\langle\alpha^7\rangle$ than equations. However, one may get a zeroth approximation by setting all the nonlinear terms equal to zero:

$$\langle\alpha\rangle^{(0)}=0;\quad \langle\alpha^2\rangle^{(0)}=D;\quad \langle\alpha^3\rangle^{(0)}=2D_3;$$
$$\langle\alpha^4\rangle^{(0)}=6D_4+3D^2; \quad (14.37)$$
$$\langle\alpha^5\rangle^{(0)}=24D_5+20D_3D.$$

One may then iterate the resulting equations and calculate $\langle\alpha^4\rangle$ to first order, $\langle\alpha^3\rangle$ to second order, $\langle\alpha^2\rangle$ to third order, and $\langle\alpha\rangle$ to fourth order. Actually, we are content with one order of accuracy less than this, with the result that

$$\langle\alpha\rangle^{(3)}=-BD+2(B^2-\Gamma)D_3+(3EB^2-BF-2\Gamma E +8B\Gamma D-5B^3D)D+(12B\Gamma-6B^3)D_4, \quad (14.38)$$

$$\langle\alpha^2\rangle^{(2)}=D-2BD_3+(F-3EB+5B^2D-3\Gamma D)D +6(B^2-\Gamma)D_4, \quad (14.39)$$

$$\langle\alpha^3\rangle^{(1)}=2D_3-5BD^2+2ED-6BD_4. \quad (14.40)$$

For purposes of later comparison with the thermal equilibrium case, we note that if one regards

$$\langle(\Delta\alpha)^2\rangle=\langle\alpha^2\rangle-[\langle\alpha\rangle]^2 \quad (14.41)$$

as known, one can compute this quantity to second order and express D to second order as

$$D=\langle(\Delta\alpha)^2\rangle+2BD_3-6(B^2-\Gamma)D_4 -[F-3EB+(4B^2-3\Gamma)\langle(\Delta\alpha)^2\rangle]\langle(\Delta\alpha)^2\rangle, \quad (14.42)$$

and $\langle\alpha\rangle$ is given to third order by

$$\langle\alpha\rangle^{(3)}=-B[\langle(\Delta\alpha)^2\rangle(1+B^2\langle(\Delta\alpha)^2\rangle)] -\Gamma[-5B\langle(\Delta\alpha)^2\rangle^2+2E\langle(\Delta\alpha)^2\rangle -6BD_4+2D_3], \quad (14.43)$$

where the first and second brackets represent $\langle\alpha^2\rangle$ and $\langle\alpha^3\rangle$, respectively. MacDonald's Eq. (27) for $\langle q^3\rangle$ takes account of only our first term $5B\Gamma\langle(\Delta\alpha)^2\rangle^2$. The last two terms are discarded by the Brownian motion approximation, and the term in E is discarded by his procedure which replaces $D(q)$ by an average value—his $\bar{F}$. We see however, from the second of Eqs. (14.11) that, under the thermal equilibrium approximation (14.11), we have to second order

$$E\approx B\langle(\Delta\alpha)^2\rangle \quad (14.44)$$

so that the term in E is comparable in importance to the one term kept by MacDonald. The term in B^3 in Eq. (14.43) does not appear in MacDonald's Eq. (27) because he sets $\langle q^2\rangle=kTC$, whereas we set $\langle(q-\bar{q})^2\rangle=kTC$. It is still an open question as to which if either, of these assumptions is correct. [See remark (5) at the end of this section.]

If no nonlinear terms are present, Eqs. (14.32) to (14.36) have a set of uncoupled normal modes with

$$\langle\alpha^n\rangle\propto\exp(-nt). \quad (14.45)$$

In the presence of nonlinearities, there are a set of modes with characteristic decays $\exp(-\lambda_n t)$ with $\lambda_n=n+$small corrections due to nonlinearities. These modes may be obtained by solving Eqs. (14.32) to (14.36). The resulting secular determinant for the eigenvalues λ is

$$\begin{vmatrix} 1-\lambda, & B, & \Gamma, & 0, & 0 \\ -2E, & (2-2F)-\lambda, & 2B, & 2\Gamma, & 0 \\ -6(D+E_3), & -6(E+F_3), & 3-6F-\lambda, & 3B, & 3\Gamma \\ -24(D_3+E_4), & 12(D+2E_3+2F_4), & 12(E+2F_3), & 4-12F-\lambda, & 4B \\ -120(D_4+E_5), & -60(D_3+2E_4+2F_5), & -20(D+3E_3+6F_4), & -20(E+3F_3), & 5-20F-\lambda \end{vmatrix}=0,$$

after omitting terms in $\langle\alpha^6\rangle$ and $\langle\alpha^7\rangle$. This determinant is sufficiently large to yield λ_1 to fourth order, λ_2 to third order, etc. For practical purposes, we need λ_1 only to second order. Therefore, we may set the upper left-hand 3×3 determinant equal to zero. Since one wishes the result only to second-order accuracy, it is more convenient to apply directly the Brillouin-Wigner perturbation formalism,[130]

$$\lambda_n=H_{nn}+\sum_{m\neq n}{}'\frac{H_{nm}H_{mn}}{\lambda-H_{mm}} +\sum_{\substack{m\neq n\\ p\neq n}}{}'\frac{H_{nm}H_{mp}H_{pn}}{(\lambda-H_{mm})(\lambda-H_{pp})}+\cdots, \quad (14.47)$$

where the primed sum means that the numerator contains no diagonal elements. This formalism can be shown to be valid for our problem even when H is a nonsymmetric or non-Hermitian matrix. In our case, some of the off-diagonal elements are large. To get

[130] See E. Feenberg, Phys. Rev. **74**, 206 (1948), Eq. (3), or Morse and Feshbach, reference 56, Eq. (9.1.90). The original references are L. Brillouin, J. phys. radium **3**, 373 (1932); E. P. Wigner, Math. u. naturw. Anz. ungar. Akad. Wiss. **53**, 475 (1935).

results accurate to second order, it is necessary to use the first three terms of this expansion,

$$\lambda_1=1+\frac{H_{12}H_{21}}{\lambda_1-H_{22}}+\frac{H_{13}H_{31}}{\lambda_1-H_{33}}+\frac{H_{12}H_{23}H_{31}}{(\lambda_1-H_{22})(\lambda_1-H_{33})}. \quad (14.48)$$

Since the numerators are already of second order, we need only a zeroth-order approximation in the denominators,

$$\lambda_1\approx 1+\frac{B(-2E)}{(-1)}+\frac{\Gamma(-6)(D+E_3)}{(-2)}+\frac{B(2B)(-6)(D+E_3)}{(-1)(-2)}, \quad (14.49)$$

$$\lambda_1\approx 1+2EB+3\Gamma D-6B^2D.$$

A similar treatment of λ_2 yields

$$\lambda_2\approx 2-2F+10BE-12\Gamma D+36B^2D, \quad (14.50)$$

although we need λ_2 only to zeroth order.

We now seek a solution of the complete time-dependent equations subject to the initial condition,

$$\langle\alpha^n(t)\rangle|_{t=0}=[\alpha(0)]^n. \quad (14.51)$$

The solution of $\langle\alpha(t)\rangle$ clearly can be written in the form,

$$\langle\alpha(t)\rangle=\langle\alpha\rangle+c_1\exp(-\lambda_1 t)+c_2\exp(-\lambda_2 t)+\cdots, \quad (14.52)$$

where $\langle\alpha\rangle$ is the steady-state value given by (14.38), and enough terms are kept to yield second-order accuracy in the final result. The two coefficients c_1 and c_2 can be determined from the two initial conditions,

$$\langle\alpha(0)\rangle=\alpha(0), \quad (14.53)$$

$$d\langle\alpha(t)\rangle/dt|_{t=0}=-\alpha(0)-B\alpha^2(0)-\Gamma\alpha^3(0), \quad (14.54)$$

with the result that

$$c_2=B(\alpha^2-D)+\Gamma(\alpha^3-3\alpha D)+\alpha(6B^2D-2EB), \quad (14.55)$$

$$c_1=\alpha-\langle\alpha\rangle-c_2, \quad (14.56)$$

where α is written briefly for $\alpha(0)$.

The autocorrelation in the fluctuations is then given by

$$\langle\alpha(t)\alpha(0)\rangle-\langle\alpha\rangle^2=\langle c_1\alpha\rangle\exp(-\lambda_1 t)+\langle c_2\alpha\rangle\exp(-\lambda_2 t), \quad (14.57)$$

and by the Wiener-Khintchin theorem, the noise in α is

$$G(\alpha,f)=4\langle c_1\alpha\rangle\frac{\lambda_1}{\lambda_1^2+\omega^2}+4\langle c_2\alpha\rangle\frac{\lambda_2}{\lambda_2^2+\omega^2}. \quad (14.58)$$

With the help of the steady-state moments, Eqs. (14.38) to (14.40), we find that to second order,

$$\langle c_1\alpha\rangle=D+(F-3EB+2B^2D-3\Gamma D)D-4BD_3+12(B^2-\Gamma)D_4, \quad (14.59)$$

$$\langle c_2\alpha\rangle=2B^2D^2+2BD_3+6(\Gamma-B^2)D_4. \quad (14.60)$$

The terms in D_4 occur even to this order, because one needs knowledge of the steady-state moment $\langle\alpha^4\rangle$. One may specialize to the Fokker-Planck approximation by setting $D_3=D_4=0$ in the foregoing expressions. If one further makes use of the equilibrium assumption (whose validity has already been questioned) in the form (14.12), then the preceding formulas reduce to

$$\langle c_1\alpha\rangle=\sigma^2(1-2b^2), \quad (14.61)$$

$$\langle c_2\alpha\rangle=2\sigma^2b^2, \quad (14.62)$$

$$\lambda_1=\Lambda(1+3g-4b^2), \quad (14.63)$$

where b and g are dimensionless B and Γ defined by Eq. (14.13), and in the noise, Eq. (14.58), it is sufficient to use $\lambda_2=2\Lambda$.

We are now in a position to compare our results with those of other authors who have made equilibrium assumptions. For the case $b=0$, Davies and MacDonald find an autocorrelation which can be written in our notation with $\Lambda=1$ as

$$\langle\alpha(t)\alpha(0)\rangle=\sigma^2(1-3gt)\exp(-t). \quad (14.64)$$

Our result for comparison is

$$\langle\alpha(t)\alpha(0)\rangle=\sigma^2\exp[-(1+3g)t], \quad (14.65)$$

which reduces to Eq. (14.64) on expanding in g.

Van Kampen works with units $\langle q^2\rangle=1$ and writes

$$D=1+\gamma q^2, \quad (14.66)$$

and quotes an eigenvalue to second order as

$$\lambda_1=1+\gamma-3\gamma^2. \quad (14.67)$$

Since his $\gamma=$ our $F=$ our g, in view of Eq. (14.12) his result $\lambda_1\approx 1+g$ disagrees with MacDonald, Davies, and myself. We get $\lambda_1\approx 1+3g$. The error is made in assuming the validity of van Kampen's Eq. (6) [see our discussion following Eq. (14.8)], which leads one to assume that $\Lambda=D=1$, when in fact, Eq. (14.11) for this case ($\sigma^2=1$, $E=0$) leads to

$$\Lambda=D-2F=1-2g, \quad (14.68)$$

and Eq. (14.63) now yields van Kampen's result,

$$\lambda_1\approx(1-2g)(1+3g)\approx 1+g, \quad (14.69)$$

to second order. In order to verify van Kampen's term in γ^2, we have made a calculation to fourth order of accuracy for the case in which $b=0$. This is fairly easy to do since one only needs the odd equations among (14.32) to (14.36). The result is

$$\lambda_1=\Lambda(1+3g-3g^2), \quad (14.70)$$

$$\lambda_3=\Lambda(3+21g+3g^2), \quad (14.71)$$

$$\langle\alpha(t)\alpha(0)\rangle=\sigma^2[(1-\tfrac{3}{2}g^2)\exp(-\lambda_1 t)+\tfrac{3}{2}g^2\exp(-\lambda_3 t)]. \quad (14.72)$$

The coefficients $\frac{3}{2}g^2$ and $1-\frac{3}{2}g^2$ agree with van Kampen. To the desired accuracy, $\lambda_3=3$ is in agreement. The term in λ_1 of order g^2 disagrees, however, with van Kampen's result (14.67) if one sets $g=\gamma$, $\Lambda=1-2\gamma$. At the moment, however, we are comparing two ways of obtaining a Fokker-Planck equilibrium approximation, and the results should agree except for arithmetic or typographical errors. Since the term of order g^2 is not likely to be of importance, we pursue this point no further.

Van Kampen also considers the case

$$D(q)=1+\beta q, \tag{14.73}$$

and quotes

$$\lambda_1=1-2\beta^2, \tag{14.74}$$

where his β equals our b. This appears to disagree with our Eq. (14.63). However, to convert to his units, according to (14.12) we must set

$$\Lambda\approx(1+b^2)D/\sigma^2\approx(1+b^2)(1)(1+b^2), \tag{14.75}$$

using

$$\sigma^2=\langle q^2\rangle-\langle q\rangle^2\approx 1-b^2, \tag{14.76}$$

since he worked with units in which $\langle q^2\rangle=1$ rather than $\langle(\Delta q)^2\rangle=\sigma^2=1$. As a result of Eq. (14.63), we then have

$$\lambda_1\approx(1+2b^2)(1-4b^2)\approx 1-2b^2, \tag{14.77}$$

in agreement with van Kampen.

For the case when both even and odd nonlinearities are present simultaneously, the complete spectrum has not previously been obtained. Our result, Eq. (14.58), may be used in general form with Eqs. (14.49), (14.59), and (14.60), or with the thermal equilibrium approximation using Eqs. (14.61) to (14.63). MacDonald and Alkemade however, have obtained some limiting results with which we can compare. Because of the Wiener-Khintchin theorem and the usual asymptotic relations between a function and its Fourier transform,[131]

$$\lim_{\omega\to\infty}\omega^2G(\alpha,\omega)/4=\lim_{t\to 0}\langle[\alpha(t)-\alpha(0)]^2\rangle/(2t)$$
$$=-\langle\dot\alpha(0)\alpha(0)\rangle. \tag{14.78}$$

Our general result for this quantity, obtained from either the first or third expressions, is

$$\lim_{\omega\to\infty}\omega^2G(\alpha,\omega)/4=D(1+F-EB) \tag{14.79}$$
$$=\Lambda\sigma^2(1+3g-2b^2). \tag{14.80}$$

Terms in D_3 and D_4 have canceled in Eq. (14.79).

Equation (14.80) is obtained from (14.79) by using the thermal equilibrium assumption (14.12). By comparison, MacDonald obtains the correction factor $(1+3g-5b^2)$ due to a disagreement in $\langle\alpha\rangle$ and $\langle\alpha^3\rangle$ which we have already discussed following Eq. (14.43).

Alkemade[126] has made a kinetic theory treatment of the ideal rectifier:

$$I=I_0[\exp(eV/kT)-1] \tag{14.81}$$
$$\approx I_0[(e/kT)V+\tfrac{1}{2}(e/kT)^2V^2+\tfrac{1}{6}(e/kT)^3V^3]. \tag{14.82}$$

The use of Eqs. (14.14) to (14.16) yields

$$b=\tfrac{1}{2}e/(kTC)^{\frac{1}{2}};\quad g=e^2/(6kTC). \tag{14.83}$$

Thus, to second order, $3g-2b^2$ cancels and

$$\lim_{\omega\to\infty}\omega^2G(\alpha,\omega)/4=\Lambda\sigma^2=eI_0, \tag{14.84}$$

in agreement with Alkemade's conclusion that there is no correction to order e^2/kTC.

Since we have obtained explicit results consistently to the second order of nonlinearity, it is quite easy to apply them to a variety of slightly nonlinear problems, including nonequilibrium examples. Some words of caution, however:

(1) A slightly nonlinear $I(V)$ relationship does not guarantee that from a microscopic viewpoint the nonlinearity is slight, e.g., in an electron gas the mean drift velocity may be small compared to the fluctuation $(kT/m)^{\frac{1}{2}}$ even though the gross $I(V)$ relation is nearly linear.

(2) In the nonequilibrium case, the noise is not determined by $I(V)$ alone, but depends independently on D, E, F, etc.

(3) Even with the equilibrium assumption (14.11), the noise to second order depends on D_3 and D_4, i.e., on some details of fluctuation mechanism, and not simply $I(V)$.

(4) When it is legitimate to make the Brownian motion approximation, Eq. (14.27) ff lead to a set of exact steady-state relations

$$\langle A(a)\alpha^{n+1}\rangle+(n+1)\langle D(a)\alpha^n\rangle=0. \tag{14.85}$$

The case $n=0$ is a generalization of the Einstein relation to nonlinear problems, and verifies MacDonald's result

$$\langle F(q)\rangle=\langle G(q)q^2\rangle/(kTC). \tag{14.86}$$

Equations (14.85) can be used to compute the steady moments when $A(a)$ and $D(a)$ are known.

(5) Suppose, however, that $D(a)$ is not known, then these equations can be used to compute $D(a)$ providing the moments are known from equilibrium considerations. We have indicated the results obtained by assuming the Gaussian distribution Eq. (14.4) but have not established its validity. Indeed, Eq. (14.21) ff contradict (14.4). It may not be possible to settle this discrepancy by a purely thermodynamic argument: The usual derivation of a Boltzmann distribution presupposes that our system is coupled *arbitrarily weakly* to a reservoir. The corrections to the usual moments obtained from Eq. (14.21) are in fact associated with the coupling energy. Unfortunately, these corrections can be comparable to those associated with nonlinear effects. MacDonald's second-law argument that $\langle q^2\rangle$

[131] Morse and Feshbach, reference 56, p. 462.

cannot exceed kTC or one could, by the use of switches, transfer energy from the nonlinear resistor to a linear one is now difficult to maintain precisely since the coupling energy that is disturbed while opening and closing the switches can be comparable to the energy one supposes one could transfer.

15. Summary

We may summarize by describing the procedure for solving a typical noise problem. First, one must find a set of parameters $\mathbf{a}=a_1, a_2, a_3\cdots$ sufficiently complete such that a Markoffian description of the system is possible with these parameters. (This is usually possible although it may require that a_j be a continuous function of its index. If, for example, the density $n(x)$ or the velocity distribution function $f(v)$ constitute the set, then the index x or v is continuous.) Let $\mathbf{a}_0$ represent the steady-state values of $\mathbf{a}$, and $\boldsymbol{\alpha}=\mathbf{a}-\mathbf{a}_0$ the fluctuations.

Then our Markoffian assumption combined with a quasi-linear assumption tells us that the variables $\boldsymbol{\alpha}$ must obey an equation of the form

$$d\langle\boldsymbol{\alpha}\rangle/dt+\boldsymbol{\Lambda}\langle\boldsymbol{\alpha}\rangle=0. \qquad (3.6), (5.16)$$

This represents the phenomenological description of our system, and the elements of the matrix $\boldsymbol{\Lambda}$ are the "circuit parameters." For a continuous system, $\boldsymbol{\Lambda}$ usually becomes a differential operator that is already known. (In any case, a method of obtaining $\boldsymbol{\Lambda}$ is defined shortly.)

By means of the phenomenological equation, the autocorrelation is given by

$$\langle\boldsymbol{\alpha}(t)\boldsymbol{\alpha}(0)\rangle=\exp(-\boldsymbol{\Lambda}t)\langle\boldsymbol{\alpha}(0)\boldsymbol{\alpha}(0)\rangle. \qquad (3.8)$$

The positive frequency component of the noise, defined by

$$\mathbf{G}_+(f)=2\int_0^\infty e^{-i\omega t}\langle\boldsymbol{\alpha}(t)\boldsymbol{\alpha}(0)\rangle dt, \qquad (2.7)$$

is given by

$$\mathbf{G}_+(f)=2(i\omega+\boldsymbol{\Lambda})^{-1}\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle. \qquad (3.10)$$

The negative frequency component of the noise may be obtained by reversing the sign of the frequency and taking the transpose

$$\mathbf{G}_-(f)=\mathbf{G}_+{}^\dagger(-f), \qquad (2.9)$$

with the result that the total noise in $\boldsymbol{\alpha}$ is

$$\mathbf{G}(\boldsymbol{\alpha},f)=2(i\omega+\boldsymbol{\Lambda})^{-1}\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle+2\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle(-i\omega+\boldsymbol{\Lambda}^\dagger)^{-1}. \qquad (15.1), (3.13)$$

The corresponding noise in $\dot{\alpha}$ can be obtained by multiplying by ω^2 and dropping some canceling terms,

$$\mathbf{G}[(d\boldsymbol{\alpha}/dt),f]=2i\omega\boldsymbol{\Lambda}(i\omega+\boldsymbol{\Lambda})^{-1}\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle - \langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle(-i\omega+\boldsymbol{\Lambda}^\dagger)^{-1}2i\omega\boldsymbol{\Lambda}^\dagger. \qquad (15.2), (7.5)$$

Thus, our problem is reduced to a knowledge of the correlations between variables α_i and α_j at the same time. For the thermal equilibrium case, these fluctuations are given by the thermodynamic formula

$$\langle\alpha_i\alpha_j\rangle=-k\partial a_i/\partial X_j=kT(\partial a_i/\partial P_j), \qquad (4.16), (4.17)$$

where

$$X_j=\partial S/\partial a_j, \quad P_j=\partial U/\partial a_j \qquad (4.11), (4.15)$$

are the thermodynamic forces in the entropy (S) and energy (U) languages, respectively, and are related by

$$\mathbf{X}=-\mathbf{P}/T. \qquad (4.15)$$

The correlations also can be summarized in the equivalent matrix equation

$$\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle=k\mathbf{s}^{-1}, \qquad (4.4)$$

where the $\mathbf{s}$ are the negative second derivatives of the entropy

$$S=S_0-\tfrac{1}{2}\boldsymbol{\alpha}\cdot\mathbf{s}\cdot\boldsymbol{\alpha}. \qquad (4.2)$$

For the nonequilibrium case, the second moments of $\boldsymbol{\alpha}$ are to be determined from the Einstein relation

$$2\mathbf{D}=\boldsymbol{\Lambda}\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle+\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle\boldsymbol{\Lambda}^\dagger, \qquad (5.18)$$

i.e.,

$$\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle=2\int_0^\infty \exp(-\boldsymbol{\Lambda}t)\mathbf{D}\exp(-\boldsymbol{\Lambda}^\dagger t)dt, \qquad (5.19)$$

where the diffusion matrix $\mathbf{D}$ (and $\boldsymbol{\Lambda}$) can be computed from the basic transition probability of the Markoff process,

$$w_{\mathbf{a}'\mathbf{a}}=P(\mathbf{a}\,|\,\mathbf{a}',\Delta t)/\Delta t, \qquad (5.21)$$

by means of the relationships

$$\mathbf{A}(\mathbf{a})=\int(\mathbf{a}'-\mathbf{a})w_{\mathbf{a}'\mathbf{a}}d\mathbf{a}'\approx A(\mathbf{a}_0)-\boldsymbol{\Lambda}\boldsymbol{\alpha}, \qquad (15.3), (5.22)$$

$$\mathbf{D}(\mathbf{a})=\tfrac{1}{2}\int(\mathbf{a}'-\mathbf{a})(\mathbf{a}'-\mathbf{a})w_{\mathbf{a}'\mathbf{a}}d\mathbf{a}'\approx\mathbf{D}(\mathbf{a}_0), \qquad (15.4), (5.23)$$

and the steady-state $\mathbf{a}_0$ obeys

$$\mathbf{A}(\mathbf{a}_0)=0. \qquad (5.15)$$

The reason for these definitions is that the first and second moments of the Smoluchowski equation (5.1) lead to the rigorous equations

$$d\langle\mathbf{a}(t)\rangle/dt=\langle\mathbf{A}(\mathbf{a})\rangle, \qquad (5.9)$$

$$d\langle\boldsymbol{\alpha}(t)\boldsymbol{\alpha}(t)\rangle/dt=2\langle D(\mathbf{a})\rangle+\langle\mathbf{A}(\mathbf{a})\boldsymbol{\alpha}\rangle+\langle\boldsymbol{\alpha}\mathbf{A}(\mathbf{a})\rangle. \qquad (5.12)$$

The linearization of the first of these equations, with our choice of $\mathbf{a}_0$, leads to our original assumed phenomenological equations, and the linearization of the second yields

$$d\langle\boldsymbol{\alpha}(t)\boldsymbol{\alpha}(t)\rangle/dt=2\mathbf{D}-\boldsymbol{\Lambda}\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle-\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle\boldsymbol{\Lambda}^\dagger, \qquad (5.17)$$

whose steady-state solution yields the Einstein relation.

Can the above results for the autocorrelation be

derived from an equation of Langevin type,

$$d\boldsymbol{\alpha}/dt+\boldsymbol{\Lambda}\boldsymbol{\alpha}=\mathbf{F}(t), \tag{8.1}$$

if the random force $\mathbf{F}(t)$ is given suitable properties? Yes! By using this equation as the definition of $\mathbf{F}$, one derives the condition

$$\langle\mathbf{F}(t)\mathbf{F}(u)\rangle=2\mathbf{D}\delta(t-u), \tag{8.11}$$

usually assumed in the Langevin approach.

The condition of time reversibility,

$$\langle\boldsymbol{\alpha}(t)\boldsymbol{\alpha}(0)\rangle=\langle\boldsymbol{\alpha}(0)\boldsymbol{\alpha}(t)\rangle, \tag{6.2}$$

implies

$$\exp(-\boldsymbol{\Lambda}t)\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle=\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle\exp(-\boldsymbol{\Lambda}^{\dagger}t), \tag{6.3}$$

which leads to the generalized Onsager relations or

$$\boldsymbol{\Lambda}\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle=\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle\boldsymbol{\Lambda}^{\dagger}, \tag{6.4}$$

and the Einstein equation is simplified for the time reversible case to yield the moments directly:

$$\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle=\boldsymbol{\Lambda}^{-1}\mathbf{D}. \tag{5.20}$$

Comparison with the Onsager phenomenological equations in the presence of an external force,

$$\mathbf{R}\langle d\boldsymbol{\alpha}/dt\rangle+\mathbf{s}\langle\boldsymbol{\alpha}\rangle=-\mathbf{X}^{e}=\mathbf{P}^{e}/T, \tag{7.1}$$

yields the admittance at frequency ω

$$\mathbf{Y}(\omega)=(d\boldsymbol{\alpha}/dt)/\mathbf{P}=[\mathbf{R}+(\mathbf{s}/i\omega)]^{-1}/T, \tag{7.3}$$

and

$$\boldsymbol{\Lambda}=\mathbf{L}\mathbf{s}, \quad \text{where } \mathbf{L}=\mathbf{R}^{-1}. \tag{6.9}$$

By introducing a correction factor

$$\mathbf{C}=\mathbf{s}\langle\boldsymbol{\alpha}\boldsymbol{\alpha}\rangle/k, \tag{6.11}$$

which reduces to unity when the moments take their equilibrium value, the Onsager, Einstein, and Nyquist theorems take the corrected forms

$$\mathbf{L}\mathbf{C}=(\mathbf{L}\mathbf{C})^{\dagger}=\mathbf{C}^{\dagger}\mathbf{L}^{\dagger}, \tag{6.12}$$

$$2\mathbf{D}=k[\mathbf{L}\mathbf{C}+(\mathbf{L}\mathbf{C})^{\dagger}]=k(\mathbf{L}\mathbf{C}+\mathbf{C}^{\dagger}\mathbf{L}^{\dagger}), \tag{15.3), (6.14}$$

$$\mathbf{G}[(d\boldsymbol{\alpha}/dt),f]=2kT[\mathbf{Y}(\omega)\mathbf{C}+\mathbf{C}^{\dagger}\mathbf{Y}^{\dagger}(-\omega)] \tag{7.6}$$

$$=4kT\,\mathrm{Re}\mathbf{Y}(\omega)\mathbf{C}, \tag{7.9}$$

where the last form makes use of time reversibility.

The external forces,

$$\mathbf{P}(t)=T\mathbf{R}\mathbf{F}(t)=\mathbf{R}'\mathbf{F}(t), \tag{8.19}$$

where $\mathbf{R}'$ is the conventional resistance, obey

$$\langle\mathbf{P}(t)\mathbf{P}(u)\rangle=kT(\mathbf{R}'\mathbf{C}^{\dagger}+\mathbf{C}^{\dagger}\mathbf{R}'^{\dagger})\delta(t-u), \tag{8.20}$$

with the result that the Nyquist law in the intensive variable takes the corrected form

$$\mathbf{G}(\mathbf{P},f)df=2kT(\mathbf{R}'\mathbf{C}^{\dagger}+\mathbf{C}\mathbf{R}'^{\dagger})df \tag{8.23}$$

$$=4kT\mathbf{R}'\mathbf{C}^{\dagger}df, \tag{15.4), (8.24}$$

where the last step assumes time reversibility.

The validity of discussing fluctuations in the forces depends on whether the usual circuit diagram analysis for computing the resulting fluctuations in the variables is correct. This is tested in Sec. 9 by computing the modified fluctuations in the extensive variables when a load is added to the system. For the case of thermal equilibrium, the resulting fluctuations agree with the fluctuations in the P's and the usual circuit analysis. For the steady nonequilibrium case, in the presence of the load, a "voltage" generator is required whose properties, in general, depend on *both* the system *and* the load, in contradiction to the usual application of Thevenin's theorem.

As a discrete example involving variables odd under time reversal, the motion of a set of coupled oscillators is discussed in Sec. 10. As an example of fluctuations in continuous systems, we discuss in Sec. 11 the Hill-van Vleit work on ambipolar drift of carrier concentration fluctuations.

In Sec. 12 on fluctuations in distribution functions, we use the Einstein relation to *prove* that

$$\langle\Delta n(a)\Delta n(a')\rangle=n_0(a)[1+\epsilon n_0(a)]\delta(a-a'), \tag{12.28}$$

where $n_0(a)$ is the steady-state distribution function and $\epsilon=0,\ -1,\ +1$ for Boltzmann, Fermi, or Einstein-Bose statistics. This result justified the assumption by Hill and van Vliet (Sec. 11) that concentration fluctuations at different positions but *the same time* are uncorrelated. It is in *contradiction* with the assumption that concentration fluctuations at different positions but the same *frequency* are uncorrelated. Van der Ziel[33] and Solow[100] make the milder assumption that the *Langevin sources* giving rise to density fluctuations at a given frequency and different positions are uncorrelated. This may be true in their case, but we have shown in Eq. (11.42) that this source correlation contains a term in $\delta''(x-x')$ as well as $\delta(x-x')$ with the result that such a Langevin treatment must be given extra care. For the Fermi and Einstein cases, Eq. (15.39) can be justified only for fluctuations from a steady state obeying detailed balance. For the Boltzmann case, there is no such restriction. *In all cases, we have assumed that the basic stochastic process is a one-body transition probability $w_{a'a}$ from state a to a'.* When the total number $\Sigma n(a)$ is constrained not to fluctuate, the foregoing equation is replaced by

$$\langle\Delta n(a)\Delta n(a')\rangle$$

$$=\{n_0(a)[1+\epsilon n_0(a)]\delta(a-a')\}$$

$$-\frac{n_0(a)[1+\epsilon n_0(a)]n_0(a')[1+\epsilon n_0(a')]}{\sum n_0(b)[1+\epsilon n_0(b)]}. \tag{12.38}$$

Some physical arguments supported by Eqs. (12.28) and (12.38) have led us to conclude that the complete distribution of occupancies has the same form in the steady state as we know it to possess in the equilibrium

state with the $n_0(a)$ interpreted as the steady, rather than equilibrium, occupancies (providing detailed balance is obeyed in the Fermi and Bose cases).

In Sec. 13, we show that uncorrelated random variables can be obtained by (1) orthonormalizing the fluctuations at one time, and (2) making an orthogonal transformation to diagonalize $\mathbf{\Lambda}$, which as a result of (1) is symmetric in the time reversible case. The Boltzmann equation is considered as an example.

The slightly nonlinear case, in which the first two moments of the transition probability can be written approximately as

$$A(a)\approx-(\Lambda\alpha+B\alpha^2+\Gamma\alpha^3), \tag{14.1}$$

$$D(a)\approx D+E\alpha+F\alpha^2, \tag{14.10}$$

is treated in Sec. 14 by a set of moment equations. The resulting noise spectrum is found to be

$$G(\alpha,f)=4\langle c_1\alpha\rangle\frac{\lambda_1}{\lambda_1^2+\omega^2}+4\langle c_2\alpha\rangle\frac{\lambda_2}{\lambda_2^2+\omega^2}, \tag{14.58}$$

where

$$\langle c_1\alpha\rangle=D+(F-3EB+2B^2D-3\Gamma D)D -4BD_3+12(B^2-\Gamma)D_4, \tag{14.59}$$

$$\langle c_2\alpha\rangle=2B^2D^2+2BD_3+6(\Gamma-B^2)D_4, \tag{14.60}$$

$$\lambda_1\approx 1+2EB+3\Gamma D-6B^2D, \tag{14.49}$$

$$\lambda_2\approx 2-2F+10BE-12\Gamma D+36B^2D, \tag{14.50}$$

to the second order of accuracy in units in which $\Lambda=1$.

The moments D_n are defined by $D_n=D_n(a_0)$ with

$$D_n(a)=\int w_{a'a}(a'-a)^n da'/n!. \tag{14.2}$$

In the Brownian motion approximation, one assumes $D_n(a)\equiv 0$ for $n>2$. If, in addition, one imposes an equilibrium condition, the preceding results simplify to

$$\langle c_1\alpha\rangle=\sigma^2(1-2b^2), \tag{14.61}$$

$$\langle c_2\alpha\rangle=2\sigma^2b^2, \tag{14.62}$$

$$\lambda_1=\Lambda(1+3g-4b^2), \tag{14.63}$$

where

$$\sigma^2=\langle(\alpha-\langle\alpha\rangle)^2\rangle, \tag{14.9}$$

$$b=B\sigma/\Lambda;\quad g=\Gamma\sigma^2/\Lambda, \tag{14.13}$$

and it is sufficient to take $\lambda_2=2\Lambda$ in the noise. These results can be used to explain the discrepancies between MacDonald, van Kampen, Davies, and Alkemade. However, they are based on an equilibrium assumption, Eq. (14.8), whose validity is questionable. In fact, we have work in progress which indicates that even at equilibrium there is no unique relationship between $D(q)$ and $A(q)$ so that a knowledge of the nonlinear response $I(V)$ is not adequate to determine the nonlinear corrections to the noise.

ACKNOWLEDGMENTS

The author wishes to thank H. C. Montgomery for discussions of specific noise problems which provided the stimulation for this work and J. J. Hopfield for discussion of some of the ideas presented here.