Classical Noise III: Nonlinear Markoff Processes

MELVIN LAX
Bell Telephone Laboratories, Murray Hill, New Jersey

Our previous treatment of noise in the nonequilibrium steady state is extended to include nonstationary processes, and processes for which the quasilinear approximation is inadequate. By use of backward-equation methods, we show that

$$M_0(\mathbf{a}_0, t, t_0) = \left\langle \exp\left[-\int_{t_0}^{t} Q(\mathbf{a}(s), t-s)\, ds \right] \right\rangle$$

subject to $\mathbf{a}(t_0)=\mathbf{a}_0$ obeys the differential (integral) equation:

$$\frac{\partial M_0(\mathbf{a}_0, t, t_0)}{\partial t_0} = [Q(\mathbf{a}_0, t-t_0) - \sum_{n=1}^{\infty} \mathbf{D}_n(\mathbf{a}_0, t_0) : (\partial/\partial \mathbf{a}_0)^n] M_0,$$

where the $\mathbf{D}_n$ are the nth-order diffusion coefficients of the $\mathbf{a}(s)$ process, and $Q(\mathbf{a}(s), s)$ is an arbitrary function of $\mathbf{a}$ and s. The choice $D_n=0$, $n>2$, $D_2=D$, $D_1(a)=-\Lambda a$ makes $a(s)$ an Ornstein–Uhlenbeck (O.U.) process, i.e., white noise that has been filtered through an RC network with time constant $1/\Lambda$. The choice $Q(a(s), s)=k(t-s)[a(s)]^2$ squares the output and applies the time smoothing $k(t-s)$. For $k(s)=\exp(-2\beta s)$ [time smoothing through an RC network with time constant $(1/2\beta)$], an explicit solution is obtained for the characteristic function M_0. For arbitrary positive $k(s)$, we show that M_0 becomes independent of a_0 as $t\rightarrow\infty$ if $k(\infty)=0$, and M_0 becomes stationary if $\Lambda>0$ and

$$\int_0^{\infty} k(u)\, du < \infty.$$

CONTENTS

1. Introduction and Summary 359
2. Properties of Markoff Processes in the Large 361
3. A Class of Problems to be Solved 362
   A. Adiabatic Line Broadening 362
   B. Free Induction and Spin-Echo Experiments 363
   C. Nonlinear Transformations on Noise 363
   D. Domain Probabilities 364
   E. Random Walk with Absorbing Barriers 364
   F. Distribution of Spectral Components 364
4. Reduction of the Path Integral 364
5. Markoff Processes in the Small 365
6. Characteristic Functions and Linked Averages 366
7. Examples of Markoff Processes 368
   A. The Weiner Process 368
   B. The Fokker–Planck Process 368
   C. The Ornstein–Uhlenbeck (O.U.) Process 369
   D. The Poisson Process (Shot Noise) 371
   E. The Homogeneous Process 372
   F. Homogeneous Noise Plus Linear Decay 373
8. Smoothed Square-Law Response 374
   Forward-Equation Method 374
   Backward-Equation Method 375
      Case 1. Linearly Smoothed Weiner Process 377
      Case 2. Linearly Smoothed O.U. Process 377
      Case 3. Exponentially Smoothed O.U. Process 377
      Case 4. Exponentially Smoothed Weiner Process 378
      Case 5. Arbitrary Positive Smoothing Function 378

1. INTRODUCTION AND SUMMARY

Our previous treatment of fluctuations from the non-equilibrium steady state[1,2] was more general than previous work in that it did not assume linearity of the system, time reversibility of the system, Gaussian character of the random variables, Langevin forces, the Fokker–Planck approximation, or that fluctuations are from an equilibrium state.

We found that the assumptions[1] that the system is *Markoffian*, *stationary*, and *quasilinear* were sufficient to compute all autocorrelations $\langle \boldsymbol{\alpha}(t)\boldsymbol{\alpha}(u)\rangle$ of the deviations $\boldsymbol{\alpha}=\mathbf{a}-\mathbf{a}_0$ of a set of random variables $\mathbf{a}=[a_1, a_2, \cdots, a_n]$ from their steady-state values. The essential idea is that if one knows the solution $\langle \boldsymbol{\alpha}(t)\rangle_{\boldsymbol{\alpha}(u)}$ of the mean motion subject to the initial condition $\boldsymbol{\alpha}=\boldsymbol{\alpha}(u)$ at time u, one can compute the autocorrelation, and hence the fluctuations from

$$\langle \boldsymbol{\alpha}(t)\boldsymbol{\alpha}(u)\rangle = \langle \langle \boldsymbol{\alpha}(t)\rangle_{\boldsymbol{\alpha}(u)} \boldsymbol{\alpha}(u)\rangle_{\text{av}} \text{ over } \boldsymbol{\alpha}(u) \quad (1.1)$$

as proven in I(2.13). Thus the regression of a fluctuation obeys the "macroscopic" equations of motion for $\langle \boldsymbol{\alpha}(t)\rangle$, and the spectrum of the noise, is given by the Fourier transform of the time-dependent decay exhibited by $\langle \boldsymbol{\alpha}(t)\rangle_{\boldsymbol{\alpha}(u)}$. For a quasilinear system, this time dependence obeys a matrix equation I(3.6):

$$d\langle \boldsymbol{\alpha}(t)\rangle/dt = -\Lambda \langle \boldsymbol{\alpha}(t)\rangle, \quad (1.2)$$

leading to an exponential time dependence. The spectrum is thus known, but the magnitude of the noise requires a knowledge of $\langle \boldsymbol{\alpha}(u)\boldsymbol{\alpha}(u)\rangle$, i.e., of the fluctuations at one time. These can then be computed from the generalized Einstein relation I(5.18):

$$\boldsymbol{\Lambda}\langle \boldsymbol{\alpha}\boldsymbol{\alpha}\rangle + \langle \boldsymbol{\alpha}\boldsymbol{\alpha}\rangle \boldsymbol{\Lambda}^{\dagger} = 2\mathbf{D}, \quad (1.3)$$

where $\boldsymbol{\Lambda}^{\dagger}$ is the transpose of $\boldsymbol{\Lambda}$ and the diffusion matrix $\mathbf{D}$ can be computed from I(5.7),

$$2\mathbf{D} = \lim_{\Delta t\rightarrow 0} (\Delta t)^{-1} \langle [\boldsymbol{\alpha}(t+\Delta t)-\boldsymbol{\alpha}(t)][\boldsymbol{\alpha}(t+\Delta t)-\boldsymbol{\alpha}(t)]\rangle \quad (1.4a)$$

[1] M. Lax, Rev. Mod. Phys. **32**, 25 (1960). This paper contains an extensive bibliography that will not be duplicated here. This paper will be referred to as I in our series of papers on noise in classical systems. It treats Markoffian noise in the stationary state by quasilinear methods.

[2] M. Lax, Phys. Chem. Solids **14**, 248 (1960). This is Classical Noise II. It applies the methods of I on continuous parameters to trapping, diffusion, and carrier concentration noise.

or from the transition probabilities in the form[3] I(5.6),

$$2\mathbf{D}(\mathbf{a})=(\Delta t)^{-1}\int(\mathbf{a}'-\mathbf{a})(\mathbf{a}'-\mathbf{a})P(\mathbf{a}'t+\Delta t\mid \mathbf{a}t)\,d\mathbf{a}' \tag{1.4b}$$

and $\mathbf{D}=\mathbf{D}(\mathbf{a}_0)$. [The conditional probabilities P are defined in Sec. 2.] Similarly, the drift vector is defined by I(5.5):

$$\mathbf{A}(\mathbf{a})=\lim\,(\Delta t)^{-1}\int(\mathbf{a}'-\mathbf{a})P(\mathbf{a}',t+\Delta t\mid \mathbf{a}t)\,d\mathbf{a}' \tag{1.5}$$

and

$$(d/dt)\langle\mathbf{a}\rangle=\langle\mathbf{A}(\mathbf{a})\rangle\approx-\mathbf{\Lambda}\cdot(\mathbf{a}-\mathbf{a}_0), \tag{1.6}$$

where $\mathbf{a}_0$ is determined by $\mathbf{A}(\mathbf{a}_0)=0$.

In summary, if one is concerned with the frequency spectrum (or alternatively the autocorrelation $\langle\boldsymbol{\alpha}(t)\boldsymbol{\alpha}(u)\rangle$) of the noise of a classical, stationary nonlinear Markoffian process, then this problem is already solved in Eqs. (1.1)–(1.6) as long as *over the range of fluctuation* one can replace the system by a quasilinear one. In this case, we need not evaluate the complete conditional probability $P(\mathbf{a}'t'\mid \mathbf{a}t)$, but have simplified the problem of showing that the first and second moments $\langle\mathbf{a}(t)\rangle$; $\langle\mathbf{a}(t)\mathbf{a}(u)\rangle$ obey a closed system of equations that can be solved exactly.

The present paper, III, is written to cover some of the techniques available when the stationarity and quasilinearity approximations are no longer valid.[4–7] In particular we wish to be able to discuss cases when:

(1) the transition probabilities and/or the solution is nonstationary;

(2) the nonlinearity over the range of the fluctuations is large enough, that quasilinear approximations are invalid;

(3) the signal is passed through a nonlinear device, so that a knowledge of the complete distribution $P(\mathbf{a}'t'\mid \mathbf{a}t)$ is needed;

(4) the results depend on the random variables $\mathbf{a}(t)$ at more than two times. In particular, if the output of a nonlinear device is time-smoothed the distribution of outputs depends on the random variables over a continuum of times—all the past history!

Most of the problems we wish to solve can be reduced to an evaluation of the average

$$M=\left\langle\exp\left[-\int_{t_0}^{t}Q(\mathbf{a}(s),s)\,ds\right]\right\rangle, \tag{1.7}$$

which can be thought of as the characteristic function of the random variable in the exponent. The latter involves an arbitrary time smoothing, over an arbitrary nonlinear function Q of the $\mathbf{a}(s)$ over the history from t_0 to t. We shall use the label M_0, if we wish the average conditional on $\mathbf{a}(t_0)=\mathbf{a}_0$. By introducing a new unnormalized Markoff process $\hat{P}$ whose transition probability for small times is related to that for P by

$$\hat{P}(\mathbf{a}'t+\Delta t\mid \mathbf{a}t)=[1-Q(\mathbf{a},t)\Delta t]P(\mathbf{a}',t+\Delta t\mid \mathbf{a},t), \tag{1.8}$$

we show that

$$M_0=\int\hat{P}(\mathbf{a}t\mid \mathbf{a}_0,t_0)\,d\mathbf{a}=\hat{\phi}(0,t), \tag{1.9}$$

where

$$\hat{\phi}(\mathbf{y},t)=\int\exp(i\mathbf{y}\cdot\mathbf{a})\hat{P}(\mathbf{a}t\mid \mathbf{a}_0t_0)\,d\mathbf{a} \tag{1.10}$$

is the characteristic function associated with unnormalized probability $\hat{P}$. [It could be written more explicitly as $\hat{\phi}(\mathbf{y}t\mid \mathbf{a}_0t_0)$ to emphasize the initial condition.]

We then define the *characteristic function* associated with (unmodified) transitions in a small time interval by

$$\exp[-L(\mathbf{y},\mathbf{a},t)\Delta t]=\langle\exp(i\mathbf{y}\cdot\Delta\mathbf{a})\rangle$$
$$=\int d\Delta\mathbf{a}\,\exp(i\mathbf{y}\cdot\Delta\mathbf{a})P(\mathbf{a}+\Delta\mathbf{a},\;t+\Delta t\mid \mathbf{a},t). \tag{1.11}$$

In particular

$$-L(\mathbf{y},\mathbf{a},t)=\sum_{n=1}^{\infty}(i\mathbf{y})^n:\mathbf{D}_n(\mathbf{a},t), \tag{1.12}$$

where

$$n!\mathbf{D}_n(\mathbf{a},t)=(\Delta t)^{-1}\int(\mathbf{a}'-\mathbf{a})^nP(\mathbf{a}',t+\Delta t\mid a,t)\,da' \tag{1.13}$$

and

$$\mathbf{y}^3:\mathbf{D}_3\equiv\sum y_iy_jy_kD_{ijk3} \tag{1.14}$$

displays the use of the colon : to imply the contraction of the product of two tensors on all their indices. With L understood to have all $\mathbf{y}$'s to the left of all $\mathbf{a}$'s as in (1.12), we show that $\hat{P}$ and $\hat{\phi}$ obey[8] (5.13),

[3] To conform with the mathematical literature, the times in our probability functions increase as one moves from right to left. The opposite convention was used in I.

[4] A review of the literature on nonlinear random processes with extensive bibliography is given by Ralph Deutsch, *Nonlinear Transformations of Random Processes* (Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1962). See also Refs. 5–7, and the references contained therein.

[5] David Middleton, *Introduction To Statistical Communication Theory* (McGraw-Hill Book Company, Inc., New York, 1960).

[6] R. L. Stratonovich, *Topics in the Theory of Random Noise* (Gordon and Breach, Inc., New York, 1963), Vol. I.

[7] *Fluctuation Phenomena in Solids*, edited by R. E. Burgess (Academic Press Inc., New York, 1965).

[8] Such forward equations have been obtained by A. J. F. Siegert, IRE Trans. Inform. Theory **3**, 4 (1954). See also, Ref. 4, Chap. 7, and R. Kubo in *Fluctuation, Dissipation and Resonance in Magnetic Systems*, edited by D. ter Haar (Plenum Press, Inc., New York, 1962).

(6.18):

$$\partial\hat{P}/\partial t=-[Q(\mathbf{a},t)+L(\mathbf{y}_{op},\mathbf{a},t)]\hat{P}(\mathbf{a},t) \quad (1.15)$$

$$\partial\hat{\phi}/\partial t=-[Q(\mathbf{a}_{op},t)+L(\mathbf{y},\mathbf{a}_{op},t)]\hat{\phi}(\mathbf{y},t) \quad (1.16)$$

with

$$\mathbf{y}_{op}=i\partial/\partial\mathbf{a};\qquad \mathbf{a}_{op}=-i\partial/\partial\mathbf{y}. \quad (1.17)$$

We also show that $\hat{P}(\mathbf{a}t\mid\mathbf{a}_0t_0)$ obeys the "backward equation," (5.19),

$$\partial P/\partial t_0=[Q(\mathbf{a}_0,t_0)+L^{\dagger}(\mathbf{y}_{op},\mathbf{a}_0)]\hat{P}(\mathbf{a}t\mid\mathbf{a}_0t_0) \quad (1.18)$$

or

$$\frac{\partial\hat{P}}{\partial t_0}=\left[Q(\mathbf{a}_0,t_0)-\sum_{n=1}^{\infty}\mathbf{D}_n(\mathbf{a}_0,t_0):\left(\frac{\partial}{\partial\mathbf{a}_0}\right)^n\right]\hat{P}. \quad (1.19)$$

Since $\mathbf{a}$ is now only a parameter, an integration over $\mathbf{a}$ commutes with the operations in (1.18) and (1.19). Thus M_0 of (1.9) obeys (1.18) and (1.19) in its dependence on $\mathbf{a}_0$ and t_0.

If we are interested in $M_0(\mathbf{a}_0,t,t_0)$, (1.7), it is *much simpler to compute it directly* by the use of (1.19) with the initial condition $M_0(\mathbf{a}_0,t,t)=1$ than to calculate the more detailed $\hat{P}(at\mid a_0t_0)$ or $\hat{\phi}(\mathbf{y}t\mid\mathbf{a}_0t_0)$. When the $\mathbf{D}_n$ do not depend explicitly in t_0, and $Q=Q(a(s),t-s)$ then $M_0(\mathbf{a}_0,t,t_0)=M_0(\mathbf{a}_0,t-t_0)$ which obeys[9]

$$-\frac{\partial M_0(\mathbf{a}_0,u)}{\partial u}=\left[Q(\mathbf{a}_0,u)-\sum_{n=1}^{\infty}D_n(\mathbf{a}_0):\left(\frac{\partial}{\partial\mathbf{a}_0}\right)^n\right]M_0(\mathbf{a}_0,u). \quad (1.20)$$

Section 7 reviews our knowledge of linear processes: the Weiner process, the Ornstein–Uhlenbeck process, the Poisson process, the homogeneous process, and a process that includes all the preceding: homogeneous noise plus linear damping,[10] for a *set* of random variables $\mathbf{a}$. For this most general linear case, we determine the multitime property:

$$M_0=\left\langle \exp i\int_{t_0}^{t}\mathbf{q}(s)\cdot\mathbf{a}(s)\,ds\right\rangle, \quad (1.21)$$

where $\mathbf{q}(s)$ is an *arbitrary* (vector) function of the time.

Section 7B discusses the multidimensional Fokker–Planck process (i.e., $\mathbf{D}_n=0$ for $n>2$) as an example of a nonlinear process.

When the $\mathbf{D}_n$'s or Q depend nonlinearly on $\mathbf{a}$, there are, of course, no general methods of solution. A *pièce de résistance* of the present paper is the reduction of

$$M_0=\left\langle \exp\left[i\int_{t_0}^{t}\mathbf{q}(s)\cdot\mathbf{a}(s)\,ds\right]\times\exp\left[-\lambda\int_{t_0}^{t}\mathbf{a}(s)\cdot\mathbf{k}(t-s)\cdot\mathbf{a}(s)\,ds\right]\right\rangle \quad (1.22)$$

to the solution of a set of coupled *ordinary* differential equations of Riccati form when $\mathbf{a}(s)$ is an Ornstein–Uhlenbeck process, and $\mathbf{q}(s)$ and $\mathbf{k}(t-s)$ are arbitrary functions of the time.

For the one-dimensional case, in particular, we show that

$$M_0=\left\langle \exp\left[-\lambda\int_{t_0}^{t}k(t-s)a(s)^2\,ds\right]\right\rangle \quad (1.23)$$

$$=[Y(t-t_0)]^{-\frac{1}{2}}\exp[-a_0^2R(t-t_0)], \quad (1.24)$$

$$Y(t-t_0)=\exp\left[4D\int_0^{t-t_0}R(u)\,du\right], \quad (1.25)$$

$$dR/du=\lambda k(u)-2\Lambda R-4DR^2, \quad (1.26)$$

where Λ is the decay constant and D the diffusion constant of the Ornstein–Uhlenbeck process. For the case of exponential smoothing, $k(u)=\exp(-2\beta u)$, we find an exact solution to the differential equation (1.26). Thus we have found the answer to a hitherto unsolved problem: *RC*-smoothed white noise (time constant $1/\Lambda$) is passed through a square-law device and the output is again *RC*-smoothed (time constant $1/2\beta$). Then M_0 is the characteristic function of the output.

The special cases $\beta=\Lambda=0$; $\beta=0$, $\Lambda>0$; $\beta>0$, $\Lambda>0$; $\beta>0$, $\Lambda=0$ all yield different answers as to (1) whether M_0 forgets $a_0=a(t_0)$ and (2) whether M_0 becomes stationary, i.e., independent of t. For *any* $k(u)>0$, we show that M_0 forgets a_0 if and only if $k(u)\to 0$ as $u\to\infty$. Moreover, M_0 becomes stationary, if and only if $\Lambda\neq 0$ and

$$\int_0^{\infty}k(u)\,du$$

converges.

We intend in paper IV of this series to discuss nonlinear random processes from the Langevin point of view. We shall show how Langevin processes can often be reduced to Markoff processes and thus made amenable to the techniques of the present paper.

2. PROPERTIES OF MARKOFF PROCESSES IN THE LARGE

Let $\mathbf{a}(t)\equiv[a_1(t),a_2(t),\cdots,a_N(t)]$ be an N-dimensional random process, and

$$P[\mathbf{a}(t_n)\mid\mathbf{a}(t_{n-1}),\cdots,\mathbf{a}(t_2),\mathbf{a}(t_1)]\equiv\frac{P[\mathbf{a}(t_n),\cdots,\mathbf{a}(t_1)]}{P[\mathbf{a}(t_{n-1}),\cdots,\mathbf{a}(t_1)]} \quad (2.1)$$

[9] This paper may be the first to make explicit use of backward equations as a means of dealing directly with the characteristic function $M_0(\mathbf{a}_0,t-t_0)$.

[10] This terminology is explained in the body of this paper. For definitions see also Refs. 1–8 or the collection of papers in N. Wax, *Noise and Stochastic Processes* (Dover Publications, Inc., New York, 1954).

be the conditional probability density[6] that $\mathbf{a}(t)$ take the value $\mathbf{a}(t_n)$ at time t_n given the values $\mathbf{a}(t_j)$ at the earlier times $t_n \geq t_{n-1} \geq t_{n-2} \geq \cdots \geq t_2 \geq t_1$. The random process $\mathbf{a}(t)$ is Markoffian if

$$P[\mathbf{a}(t_n) \mid \mathbf{a}(t_{n-1}), \cdots, \mathbf{a}(t_1)] = P[\mathbf{a}(t_n) \mid \mathbf{a}(t_{n-1})], \tag{2.2}$$

i.e., if the probability of any event depends *only* on the latest piece of information available. A Markoff process has no memory of earlier events.

If we now write briefly $\mathbf{a}(t_j) = \mathbf{a}_j$, the repeated use of the definition (2.1) of conditional probability density in terms of the full probability densities permits one to write

$$P(\mathbf{a}_n, \mathbf{a}_{n-1}, \cdots, \mathbf{a}_1) = P(\mathbf{a}_n \mid \mathbf{a}_{n-1}, \cdots, \mathbf{a}_1) P(\mathbf{a}_{n-1} \mid \mathbf{a}_{n-2}, \cdots, \mathbf{a}_1), \cdots, P(\mathbf{a}_2 \mid \mathbf{a}_1) P(\mathbf{a}_1), \tag{2.3}$$

which for Markoff processes reduces to

$$P(\mathbf{a}_n, \cdots, \mathbf{a}_1) = P(\mathbf{a}_n \mid \mathbf{a}_{n-1}) P(\mathbf{a}_{n-1} \mid \mathbf{a}_{n-2}) \cdots P(\mathbf{a}_2 \mid \mathbf{a}_1) P(\mathbf{a}_1). \tag{2.4}$$

The factorization (2.4) permits the probability of a compound event occurring at many times to be expressed entirely in terms of the transition probability $P(\mathbf{a}_n \mid \mathbf{a}_{n-1})$ and the initial distribution $P(\mathbf{a}_1)$. For example, if $V(t)$, a one-dimensional Markoff process (e.g., a noise voltage) is passed through a square-law rectifier:

$$I(t) = V(t)^2, \qquad V(t) > 0;$$
$$I(t) = 0, \qquad V(t) < 0. \tag{2.5}$$

The triple correlation is given by

$$\langle I(t_3) I(t_2) I(t_1) \rangle = \int_0^\infty \int_0^\infty \int_0^\infty dV_1\, dV_2\, dV_3 \times V_3^2 P(V_3 \mid V_2) V_2^2 P(V_2 \mid V_1) V_1^2 P(V_1). \tag{2.6}$$

Strictly speaking, P depends not only on $\mathbf{a}(t_n)$, but also on t_n. Processes will be called *stationary* if

$$P(\mathbf{a}, t_n, \mathbf{a}_{n-1} t_{n-1}, \cdots, \mathbf{a}_1, t_1) = P(\mathbf{a}_n, t_n + \tau, \mathbf{a}_{n-1}, t_{n-1} + \tau, \cdots, \mathbf{a}_1, t_1 + \tau), \tag{2.7}$$

i.e., if all probabilities are functions only of the time differences. In particular, stationarity implies that

$$P(\mathbf{a}_2 t_2 \mid \mathbf{a}_1 t_1) = P(\mathbf{a}_2, t_2 - t_1 \mid \mathbf{a}_1 0), \tag{2.8}$$

$$P(\mathbf{a}, t) = P(\mathbf{a}, 0) = P(\mathbf{a}), \tag{2.9}$$

$$\lim t \to \infty \qquad P(\mathbf{a}, t \mid \mathbf{a}_0 t_0) = P(\mathbf{a}). \tag{2.10}$$

Whereas a stationary process has transition probabilities that are functions only of the time difference, the converse follows from (2.4) only if $P(\mathbf{a}_1, t_1)$ is not time-dependent. Even if $P(\mathbf{a}_2 \mid \mathbf{a}_1)$ is a function only of $t_2 - t_1$, no time-independent solution may exist for $P(\mathbf{a}_1, t_1)$ if the process possesses some instability, or even neutral stability. A simple example of neutral stability is Brownian motion in an infinite domain.

If we set $n=3$ in (2.4) and divide by $P(\mathbf{a}_1)$ we obtain

$$P(\mathbf{a}_3 \mathbf{a}_2 \mid \mathbf{a}_1) = P(\mathbf{a}_3 \mid \mathbf{a}_2) P(\mathbf{a}_2 \mid \mathbf{a}_1). \tag{2.11}$$

Integrating over $\mathbf{a}_2$ we obtain the Chapman–Kolmogoroff condition

$$P(\mathbf{a}_3 \mid \mathbf{a}_1) = \int P(\mathbf{a}_3 \mid \mathbf{a}_2)\, d\mathbf{a}_2 P(\mathbf{a}_2 \mid \mathbf{a}_1) \tag{2.12}$$

on the transition probabilities.

Our equations are written in a form appropriate to continuous variables. They remain valid for the discrete case if integrals over $\mathbf{a}$ are replaced by sums.

It is customary to take all probabilities to be normalized. In particular this leads to the requirement that the transition probability obey

$$\int d\mathbf{a}_2 P(\mathbf{a}_2 \mid \mathbf{a}_1) = 1. \tag{2.13}$$

However, we shall carry through our analysis without imposing this condition, since processes *violating* (2.13) can be used to analyze nonlinear time-smoothing of processes that *do* obey (2.13) [see Sec. 4].

We shall call a process that obeys

$$P(\mathbf{a}_2 t_2 \mid \mathbf{a}_2, t_1) = P(\mathbf{a}_2 - \mathbf{a}_1, t_2 \mid 0 t_1) \tag{2.14}$$

a *homogeneous Markoff process* since it has no preferred origin in $\mathbf{a}$ space.

3. A CLASS OF PROBLEMS TO BE SOLVED

A number of important problems in the theory of random processes can be reduced to an expectation value of the form[11]

$$M = \left\langle \exp\left[-\int_{t_0}^{t} Q(\mathbf{a}(s), s)\, ds \right] \right\rangle \tag{3.1}$$

or to a Fourier transform of such an expression.

A. Adiabatic Line Broadening

The absorption of radiation at frequency ω by a system with electric (or magnetic) dipole moment $\mu(t)$ is given by

$$I(\omega) \propto T^{-1} \int_{-T/2}^{T/2} \mu(t'') \times \exp(-i\omega t'')\, dt'' \int_{-T/2}^{T/2} \mu(t')^* \exp(i\omega t')\, dt' \tag{3.2}$$

[11] See Ref. 4, Chap. 7.

or with $t''=t'+t$ and $\langle\ \rangle$ representing an ensemble average

$$I(\omega)=\int_{-\infty}^{\infty}\exp(-i\omega t)M(t)\,dt, \tag{3.3}$$

$$M(t)=\lim_{T\to\infty}T^{-1}\int_{-T/2}^{T/2}dt'\langle\mu(t+t')\mu(t')^*\rangle. \tag{3.4}$$

In the usual "adiabatic" theory of line broadening[12] (say of absorption by one gas atom in a background of foreign atoms) the influence of collisions on the dipole matrix element $\mu=\mu_{fi}$

$$\mu_{fi}(t)=(\psi_f(t),\mu\psi_i(t))$$
$$=\left(\exp\left[-i\int_{-\infty}^{t}H(s)\,ds\right]\psi_f,\ \mu\exp\left[-i\int_{-\infty}^{t}H(s)\,ds\right]\psi_i\right)$$
$$=\mu_{fi}\exp\left[i\int_{-\infty}^{t}\omega(s)\,ds\right], \tag{3.5}$$

$$\omega(s)=H_{ii}(s)-H_{ff}(s), \tag{3.6}$$

is taken to be purely that of a phase shift induced in the initial and final states (i and f) with no off-diagonal elements in the Hamiltonian H to correspond to "non-adiabatic" transitions. (The nonadiabatic modification of this theory has been given by Anderson[13] and by Byron and Foley.[14])

In the adiabatic theory, $\omega(s)$ is regarded as a random variable subject to jumps introduced by the reservoir of foreign atoms. For *spin diffusion*[15] the same model has been assumed with the atom in question being a "spin," and the foreign atoms other spins that interact with the first.

With (3.5), Eq. (3.4) can be rewritten in the form

$$M(t)=\text{average over } t' \text{ of}\left\langle\exp\left[i\int_{t'}^{t'+t}\omega(s)\,ds\right]\right\rangle \tag{3.7}$$

or

$$M(t)=\left\langle\exp\left[i\int_{0}^{t}\omega(s)\,ds\right]\right\rangle \tag{3.8}$$

where the last form is valid when $\omega(s)$ is a stationary random variable so that the previous average is independent of t'. Our result (3.8) has the form (3.1). See the evaluation of $M(t)$ in Secs. 7E and 7F.

---

[12] For a review of line broadening see R. G. Breene, Jr., *The Shift and Shape of Spectral Lines* (Pergamon Press, Inc., New York, 1961).

[13] P. W. Anderson, Phys. Rev. **76**, 647 (1949).

[14] F. W. Byron, Jr., and H. M. Foley, Phys. Rev. **134**, A625 (1964).

[15] J. R. Klauder and P. W. Anderson, Phys. Rev. **125**, 912 (1962).

B. Free-Induction and Spin-Echo Experiments

In a free-induction experiment[15] a short rf pulse at time $t\approx 0$ rotates the resonant spins from the z direction onto the x axis. Thus all spins start with the same phase at $t=0$. The resulting free-induction signal is proportional to $\mu(t)$ or

$$M(t)=\langle\mu(t)\rangle=\left\langle\exp\left[i\int_{0}^{t}\omega(s)\,ds\right]\right\rangle. \tag{3.9}$$

Thus the free-induction experiment measures directly the Fourier transform of the line shape $I(\omega)$.

A disadvantage of the free-induction experiment is that the signal $M(t)$ will decay rapidly if there is a spread of initial frequencies $\omega(0)$ because of a distribution of dc magnetic fields at the various sites. This inhomogeneous broadening which makes the phases $\omega(0)t$ differ from one another, is overcome in a spin-echo experiment in which a 180° pulse is applied at time τ to reverse the direction of each spin. This reverses the phase acquired up to time τ. For $t>\tau$ the phase is then

$$\int_{\tau}^{t}\omega(s)\,ds-\int_{0}^{\tau}\omega(s)\,ds. \tag{3.10}$$

If the frequencies have a spread but are not random in time, the phase distribution associated with this spread cancels at $t=2\tau$. Thus a peak is seen in the induced signal at $t=2\tau$, that is less than that at $t=0$ only if dynamic random fluctuations occur in ω. The observed signal at any time t is given by[15]

$$M(t)=\left\langle\exp\left[i\int_{0}^{t}m(s)\omega(s)\,ds\right]\right\rangle, \tag{3.11}$$

where

$$m(s)=-1,\qquad 0<s<\tau;$$
$$m(s)=+1,\qquad \tau<s<t.$$

More general spin-echo experiments, with more than two pulses can be described using more complicated functions $m(s)$. In general the echo occurs at the time t for which

$$\int_{0}^{t}m(s)\,ds=0 \tag{3.12}$$

and the effect of the initial frequency distribution cancels. For this general case, $a=\omega$, $Q=-im(s)a$ in Eq. (3.1)

C. Nonlinear Transformations on Noise

If a signal a is passed through a nonlinear device, and the result passed through a linear filter, the output signal takes the form[8]

$$S(t)=\int_{-\infty}^{t}k(t-s)V[a(s)]\,ds. \tag{3.13}$$

The characteristic function whose Fourier transform (on z) gives the probability distribution of S is

$$\langle \exp (izS) \rangle = \left\langle \exp \left\{ iz \int_{-\infty}^{t} k(t-s) V[a(s)] \, ds \right\} \right\rangle, \tag{3.14}$$

a result that also has the form (3.1). Special cases such as

$$S(t) = \int_0^t k(s) a^2(s) \, ds, \tag{3.15}$$

$$S(t) = \int_0^1 | a(s) | \, ds, \tag{3.16}$$

have been studied by Kac and Siegert,[16] by Kac,[17] by Siegert,[18] and others.[19]

D. Domain Probabilities

If we set[17]

$$Q(\mathbf{a}(s), s) = \begin{cases} 1 & \text{if } \mathbf{a}(s) \text{ in } D \\ 0 & \text{if } \mathbf{a}(s) \text{ not in } D \end{cases} \tag{3.17}$$

where D is some domain, e.g., $\sum [a_i(s)]^2 < R$ and

$$S(t, t_0) = \int_{t_0}^{t} Q(\mathbf{a}(s), s) \, ds, \tag{3.18}$$

then $S(t, t_0)/(t-t_0)$ is the fraction of time spent in the domain D, and $\langle \exp (izS) \rangle$ gives the characteristic function whose Fourier transform gives the probability distribution of S. As remarked by Deutsch,[11] the fluctuations in S measure the reliability with which a finite measurement time $t-t_0$ can be used to estimate the distribution of $\mathbf{a}(s)$.

E. Random Walk With Absorbing Barriers

The probability that $\mathbf{a}(s)$ never leaves domain D in the time interval (t_0, t) is given by[11,17]

$$\text{Prob } \{\mathbf{a}(s) \text{ in } D, t_0 < s < t\} = \lim_{z \to \infty} \left\langle \exp \left[-z \int_{t_0}^{t} V_D[\mathbf{a}(s)] \, ds \right] \right\rangle, \tag{3.19}$$

$$V_D(\mathbf{a}(s)) = \begin{cases} 0 & \mathbf{a}(s) \text{ in } D \\ 1 & \mathbf{a}(s) \text{ outside } D. \end{cases}$$

Equation (3.19) corresponds to a random walk with an absorbing barrier at the boundary of the domain D.

F. Distribution of Spectral Components

The components of a one-dimensional random variable $a(s)$ in some orthogonal basis $\phi_n(s)$ (e.g., the terms of a Fourier series) are given by $\int \phi_n(s) a(s) \, ds$ and the joint characteristic function of a number of such components is given by[11]

$$M = \left\langle \exp \left[i \sum z_n \int_{t_0}^{t} \phi_n(s) a(s) \, ds \right] \right\rangle, \tag{3.20}$$

which also has the standard form (3.1).

4. REDUCTION OF THE PATH INTEGRAL

Our average of a functional (3.1) can be written in the form

$$M = \left\langle \exp \left[-\sum_{j=0}^{n-1} Q(\mathbf{a}_j, s_j) \, \Delta s_j \right] \right\rangle = \int \exp \left[-\sum_{j=0}^{n-1} Q(\mathbf{a}_j, s_j) \, \Delta s_j \right] P(\mathbf{a}_n, \mathbf{a}_{n-1}, \cdots, \mathbf{a}_1, \mathbf{a}_0) \, d\mathbf{a}_n, \cdots, d\mathbf{a}_0, \tag{4.1}$$

where $s_n = t$, $\Delta s_j = s_{j+1} - s_j$. For a Markoff process, the factorization (2.4) permits us to rewrite this result in the form

$$M = \int M_0 P(\mathbf{a}_0) \, d\mathbf{a}_0, \tag{4.2}$$

where $P(\mathbf{a}_0)$ is the probability density at $t = t_0$ and M_0 is the conditional average with $\mathbf{a}(t_0) = \mathbf{a}_0$:

$$M_0 = \prod_{j=0}^{n-1} \int \exp [-Q(\mathbf{a}_j, s_j) \, \Delta s_j] P(\mathbf{a}_{j+1} \mid \mathbf{a}_j) \, d\mathbf{a}_{j+1}. \tag{4.3}$$

Direct evaluation of M_0 by such functional ("path") integrations has only been performed[20] for Weiner processes, i.e., the special case in which

$$P(\mathbf{a}'t' \mid \mathbf{a}t) = [4\pi(t'-t)]^{-N/2} (\det \mathbf{D})^{-\frac{1}{2}} \times \exp [-(\mathbf{a}'-\mathbf{a}) \cdot \mathbf{D}^{-1} \cdot (\mathbf{a}'-\mathbf{a})/4(t'-t)] \tag{4.4}$$

and for homogeneous processes and slight generalizations thereof.[15] All such results, and many others can be obtained by the following procedure which reduces the problem to one for which standard analytical techniques are available: Define (for sufficiently small Δs_j)

$$\hat{P}(\mathbf{a}_{j+1}, s_{j+1} \mid \mathbf{a}_j, s_j) = [1 - Q(\mathbf{a}_j, s_j) \Delta s_j] P(\mathbf{a}_{j+1}, s_{j+1} \mid \mathbf{a}_j, s_j). \tag{4.5}$$

[16] M. Kac and A. J. F. Siegert, Phys. Rev. **70**, 449 (1946); J. Appl. Phys. **18**, 383 (1947); Ann. Math. Stat. **18**, 38 (1947).

[17] M. Kac, Trans. Am. Math. Soc. **59**, 401 (1946); *Berkeley Symposium on Mathematics, Statistics and Probability* (University of California Press, Berkeley, 1951), Vol. 2, p. 189.

[18] A. J. F. Siegert, Phys. Rev. **81**, 617 (1951); IRE Trans. Inform. Theory **3**, 38 (1957); **4**, 4 (1958).

[19] R. C. Emerson, J. Appl. Phys. **24**, 1168 (1953); A. Rosenbloom, J. Heilfron, and D. C. Trautman, IRE Natl. Conv. Record Part 4, 106 (1955); M. A. Meyer and D. Middleton, J. Appl. Phys. **25**, 1037 (1954); P. Erdos and M. Kac, Bull. Am. Math. Soc. **52**, 292 (1946); W. Feller, Ann. Math. Stat. **2**, 427 (1951); D. A. Darling and A. J. F. Siegert, IRE Trans. Inform. Theory **3**, 32 (1957); Ann Math. Stat. **24**, 624 (1953).

[20] R. H. Cameron and W. T. Martin, Trans. Am. Math. Soc. **58**, 184 (1945); Ann. Math. **45**, 386 (1944); J. Math. Phys. **23**, 195 (1944); Bull. Am. Math. Soc. **51**, 73 (1945); Am. J. Math. **66**, 281 (1944); E. W. Montroll, Commun. Pure Appl. Math. **5**, 415 (1952).

We may regard $P(\mathbf{a}_{j+1} \mid \mathbf{a}_j)$ as the transition probability of a new Markoffian process. It obeys the usual properties of transition probabilities except for a change in normalization:

$$\int \hat{P}(\mathbf{a}', t+\Delta t \mid \mathbf{a}, t)\, d\mathbf{a}' = 1 - Q(\mathbf{a}, t)\, \Delta t \quad (4.6)$$

to first order in Δt. Thus if we regard $\hat{P}(\mathbf{a}, t)$ as a density of systems, then $Q(\mathbf{a}, t)$ can be regarded as the rate at which they disappear.

The transition probability $\hat{P}(\mathbf{a}, t \mid \mathbf{a}_0 t_0)$ over a finite time interval can be decomposed into transition probabilities over "infinitesimal" intervals by integrating (2.4) over the intermediate values $\mathbf{a}_2, \cdots, \mathbf{a}_{n-1}$ or by repeated use of the Chapman–Kolmogoroff relation:

$$\hat{P}(\mathbf{a}_n \mid \mathbf{a}_0) = \int d\mathbf{a}_{n-1}, \cdots, d\mathbf{a}_1 \hat{P}(\mathbf{a}_n \mid \mathbf{a}_{n-1}), \cdots, \hat{P}(\mathbf{a}_2 \mid \mathbf{a}_1)\hat{P}(\mathbf{a}_1 \mid \mathbf{a}_0) \quad (4.7)$$

on the P probabilities. Comparison with (4.3) now yields the result

$$M_0 = \int \hat{P}(\mathbf{a}_n \mid \mathbf{a}_0)\, d\mathbf{a}_n = \int \hat{P}(\mathbf{a}, t \mid \mathbf{a}_0, t_0)\, d\mathbf{a}, \quad (4.8)$$

i.e., $\hat{P}(\mathbf{a} \mid \mathbf{a}_0)$ is the probability density that a system starting at system $\mathbf{a}_0$ at t_0 will arrive at $\mathbf{a}$ at time t taking into account the loss rate $Q(\mathbf{a}, t)$, and M_0 is the fraction of systems starting at $\mathbf{a}_0$ that survive *anywhere* when losses are included. Our problem has thus been reduced, in the Markoff case, to the determination of the transition probability $\hat{P}(\mathbf{a}t \mid \mathbf{a}_0 t_0)$ over finite time intervals.

5. MARKOFF PROCESSES IN THE SMALL

To obtain a differential equation for $\hat{P}(\mathbf{a}t \mid \mathbf{a}_0 t_0)$ we can write the Chapman–Kolmogoroff equation in the form

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

where $\hat{P}(\mathbf{a}, t)$ reduces to $\hat{P}(\mathbf{a}t \mid \mathbf{a}_0 t_0)$ if we adopt the initial condition

$$\hat{P}(\mathbf{a}, t_0) = \delta(\mathbf{a} - \mathbf{a}_0). \quad (5.2)$$

For small Δt, we shall assume that the moments of the transition probability are expandable in powers of Δt. To first order in Δt, we shall write

$$\int \hat{P}(\mathbf{a}, t+\Delta t \mid \mathbf{a}', t)\, d\mathbf{a} = 1 - Q(\mathbf{a}', t)\, \Delta t \quad (5.3)$$

and for $n \geq 1$ there is no distinction between the moments of the $\hat{P}$ and P transition probabilities:

$$\int \hat{P}(\mathbf{a}, t+\Delta t \mid \mathbf{a}', t)(\mathbf{a} - \mathbf{a}')^n\, d\mathbf{a} = n!\mathbf{D}_n(\mathbf{a}', t)\, \Delta t$$

and

$$\int P(\mathbf{a}, t+\Delta t \mid \mathbf{a}', t)(\mathbf{a} - \mathbf{a}')^n\, d\mathbf{a} = n!\mathbf{D}_n(\mathbf{a}', t)\, \Delta t, \quad (5.4)$$

where boldface $\mathbf{D}_n$ has n (suppressed) subscripts, e.g., the moment of $(a_1 - a_1')^2(a_2 - a_2')$ would be written D_{112} in expanded notation rather than $\mathbf{D}_3$.

The Taylor expansion of an arbitrary function $f(\mathbf{a})$ can also be written in a condensed notation

$$f(\mathbf{a}) = f(\mathbf{a}') + \sum_{n=1}^{\infty} (\mathbf{a} - \mathbf{a}')^n : \mathbf{f}^{(n)}(\mathbf{a}')/n! \quad (5.5)$$

where the : is a shorthand notation that tells us to multiply corresponding terms and add. The second-order term is, for example,

$$(1/2!) \sum_{i,j=1}^{N} (a_i - a_i')(a_j - a_j')\partial^{(2)} f/\partial a_i' \partial a_j'.$$

Let us now multiply Eq. (5.1) by $f(\mathbf{a})$ and integrate over $\mathbf{a}$. On the right-hand side, $f(\mathbf{a})$ can be placed under the $\int d\mathbf{a}'$ and then replaced by the right-hand side of Eq. (5.5). Making use of (5.3), (5.4) we obtain

$$\langle f(a) \rangle_{t+\Delta t} = \langle [1 - Q(a, t)\Delta t] f(a) \rangle_t + \Delta t \sum_{n=1}^{\infty} \int d\mathbf{a}' P(\mathbf{a}', t)\mathbf{f}^{(n)}(\mathbf{a}') : \mathbf{D}_n(\mathbf{a}', t)$$

or

$$(\partial/\partial t)\langle f(\mathbf{a}) \rangle = -\langle Q(\mathbf{a}, t) f(\mathbf{a}) \rangle + \sum_{n=1}^{\infty} \langle \mathbf{D}_n(\mathbf{a}, t) : (\partial/\partial \mathbf{a})^n f(\mathbf{a}) \rangle \quad (5.6)$$

$$= -\langle (Q + L^\dagger) f(\mathbf{a}) \rangle, \quad (5.7)$$

where

$$\langle f(\mathbf{a}) \rangle = \int f(\mathbf{a}) \hat{P}(\mathbf{a}, t)\, d\mathbf{a} \quad (5.8)$$

$$L^\dagger = -\sum_{n=1}^{\infty} \mathbf{D}_n(\mathbf{a}, t) : (\partial/\partial \mathbf{a})^n. \quad (5.9)$$

Using $\mathbf{A}$ as an abbreviation for $\mathbf{D}_1(\mathbf{a}, t)$ (the drift vector) and $\mathbf{D}$ as an abbreviation for $\mathbf{D}_2(\mathbf{a}, t)$ (the diffusion matrix) let us set $f(\mathbf{a}) = 1, \mathbf{a}, \mathbf{a}^2, \mathbf{a}^3, \cdots$ in Eq. (5.6) to obtain the useful moment relations,

$$\partial\langle 1 \rangle/\partial t = -\langle Q \rangle,$$

$$\partial\langle \mathbf{a} \rangle/\partial t = -\langle Q\mathbf{a} \rangle + \langle \mathbf{A} \rangle,$$

$$\partial\langle \mathbf{a}^2 \rangle/\partial t = -\langle Q\mathbf{a}^2 \rangle + 2\langle \mathbf{a}\mathbf{A} \rangle + 2(\mathbf{D}),$$

$$\partial\langle \mathbf{a}^3 \rangle/\partial t = -\langle Q\mathbf{a}^3 \rangle + 3\langle \mathbf{a}^2\mathbf{A} \rangle + 6\langle \mathbf{a}\mathbf{D} \rangle + 6\langle \mathbf{D}_3 \rangle. \quad (5.10)$$

When $\mathbf{a}$ is a set of variables, the right-hand side is understood to be completely symmetrized. Thus

$$\partial\langle a_ia_ja_k\rangle/\partial t=\langle a_ia_ja_kQ\rangle+\langle a_ia_jA_k\rangle+\langle A_ia_ja_k\rangle$$
$$+\langle a_iA_ja_k\rangle+2\langle a_iD_{jk}\rangle+2\langle a_jD_{ik}\rangle+2\langle a_kD_{ij}\rangle+6\langle D_{ijk}\rangle. \tag{5.11}$$

[The second-moment equations for a set of variables were derived directly in I(5.12), and higher moment equations for one variable written out explicitly in I(14.29).]

If in (5.8) we set $f(\mathbf{a})=\delta(\mathbf{a}-\mathbf{x})$ we find that

$$\hat{P}(\mathbf{x},t)=\langle\delta(\mathbf{a}-\mathbf{x})\rangle. \tag{5.12}$$

Inserting this choice for $f(\mathbf{a})$ into (5.6) and integrating by parts we obtain

$$\partial\hat{P}(\mathbf{x},t)/\partial t=-Q(\mathbf{x})\hat{P}$$
$$+\sum_{n=1}^{\infty}(-1)^n(\partial/\partial\mathbf{x})^n\colon[D_n(\mathbf{x},t)\hat{P}(\mathbf{x},t)] \tag{5.13}$$

or

$$\partial\hat{P}(\mathbf{a},t)/\partial t=-[Q(\mathbf{a})+L]\hat{P}(\mathbf{a},t),$$

where L is the Hermitian adjoint of $L^\dagger$:

$$-L=\sum_{n=1}^{\infty}(-1)^n(\partial/\partial\mathbf{a})^n\colon\mathbf{D}_n(\mathbf{a},t). \tag{5.14}$$

Associated with the operator L, we shall define a numerical function

$$-L(\mathbf{y},\mathbf{a},t)=\sum_{n=1}^{\infty}(i\mathbf{y})^n\colon\mathbf{D}_n(\mathbf{a},t). \tag{5.15}$$

With the understanding that the $\mathbf{y}$'s are always to the left of the $\mathbf{a}$'s, we can write the operator as

$$L=L(\mathbf{y}_{\text{op}},\mathbf{a},t);\qquad \mathbf{y}_{\text{op}}=i\partial/\partial\mathbf{a}. \tag{5.16}$$

Equation (5.13) is customarily referred to as a forward equation, since it refers to $\partial\hat{P}/\partial t$. We can rewrite the Chapman–Kolmogoroff equations in a form

$$\hat{P}(\mathbf{a}t\mid\mathbf{a}_0,t-\Delta t)=\int\hat{P}(\mathbf{a}t\mid\mathbf{a}'t_0)\,d\mathbf{a}'\hat{P}(\mathbf{a}'t_0\mid\mathbf{a}_0,t_0-\Delta t) \tag{5.17}$$

suitable to derive an equation for $\partial\hat{P}/\partial t_0$, i.e., a backward equation. In (5.17) we insert

$$\hat{P}(\mathbf{a},t\mid\mathbf{a}',t_0)=\hat{P}(\mathbf{a}t\mid\mathbf{a}_0t_0)$$
$$+\sum[(\mathbf{a}'-\mathbf{a}_0)^n/n!]\colon(\partial/\partial\mathbf{a}_0)^n\hat{P}(\mathbf{a}t\mid\mathbf{a}_0t_0) \tag{5.18}$$

and use (5.3), (5.4) to obtain the backward equation

$$[\partial\hat{P}(\mathbf{a}t\mid\mathbf{a}_0t_0)/\partial t_0]=Q(\mathbf{a}_0,t_0)\hat{P}$$
$$-\sum\mathbf{D}_n(a_0,t_0)\colon(\partial/\partial\mathbf{a}_0)^n\hat{P}. \tag{5.19}$$

If the random variable whose properties we wish to discuss has the form

$$\int_{t_0}^{t}Q(\mathbf{a}(s),t-s)\,ds, \tag{5.20}$$

rather than depending on $Q(\mathbf{a}(s),s)$, then Eq. (5.19) remains valid when t regarded as a parameter, and $Q(\mathbf{a}(s),t-s)$ enters the equation as $Q(\mathbf{a}(t_0),t-t_0)$, i.e.,

$$[\partial\hat{P}(\mathbf{a}t\mid\mathbf{a}_0t_0)/\partial t_0]=Q(\mathbf{a}_0,t-t_0)\hat{P}$$
$$-\sum_{n=1}^{\infty}\mathbf{D}_n(\mathbf{a}_0,t_0)\colon(\partial/\partial\mathbf{a}_0)^n\hat{P}. \tag{5.21}$$

Moreover if the transition probabilities are time-independent, $\mathbf{D}_n$ is independent of t_0 and we may expect our solution to be a function only if $u=t-t_0$:

$$-\partial\hat{P}/\partial u=[Q(\mathbf{a}_0,u)-\sum_{1}^{\infty}\mathbf{D}_n(\mathbf{a}_0)\colon(\partial/\partial\mathbf{a}_0)^n]\hat{P}. \tag{5.22}$$

6. CHARACTERISTIC FUNCTIONS AND LINKED AVERAGES

The characteristic function ϕ of a *normalized* probability density function P is defined by

$$\phi(\mathbf{y},t)=\langle\exp(i\mathbf{y}\cdot\mathbf{a})\rangle=\int\exp(i\mathbf{y}\cdot\mathbf{a})P(\mathbf{a},t)d\mathbf{a}. \tag{6.1}$$

The moments $\langle\mathbf{a}^n\rangle$ are determined by the nth derivatives of ϕ at $\mathbf{y}=0$ since

$$\phi(\mathbf{y},t)=\sum_{n=0}^{\infty}(i\mathbf{y})^n\colon\langle\mathbf{a}^n\rangle/n!. \tag{6.2}$$

If $\mathbf{a}$ and $\mathbf{b}$ are two independent sets of random variables,

$$\langle\exp i\mathbf{y}\cdot(\mathbf{a}+\mathbf{b})\rangle=\langle\exp i\mathbf{y}\cdot\mathbf{a}\rangle\langle\exp i\mathbf{y}\cdot\mathbf{b}\rangle \tag{6.3}$$

or

$$\phi_{\mathrm{a+b}}(\mathbf{y})=\phi_{\mathrm{a}}(\mathbf{y})\phi_{\mathrm{b}}(\mathbf{y}),$$

i.e., the ln of $\phi_{\mathrm{a+b}}$ is an additive function. If we define

$$\langle\exp(i\mathbf{y}\cdot\mathbf{a})\rangle=\exp\sum_{n=1}^{\infty}(1/n!)(i\mathbf{y})^n\colon\boldsymbol{\mu}_n, \tag{6.4}$$

then for independent variables,

$$\boldsymbol{\mu}_n(\mathbf{a}+\mathbf{b})=\boldsymbol{\mu}_n(\mathbf{a})+\boldsymbol{\mu}_n(\mathbf{b}). \tag{6.5}$$

The $\boldsymbol{\mu}_n(\mathbf{a})$ are referred to as Thiele semi-invariants,[21] or cumulants,[22] or linked (L) moments and can be written symbolically as

$$\boldsymbol{\mu}_n(\mathbf{a})\equiv\langle\mathbf{a}^n\rangle^L. \tag{6.6}$$

---

[21] H. Cramer, *Mathematical Methods of Statistics* (Princeton University Press, Princeton, New Jersey, 1946), Sec. 15.8; M. G. Kendall and A. Stuart, *The Advanced Theory of Statistics* (Charles Griffin and Company, Ltd., London, 1958), Chap. 3.

[22] R. Kubo, J. Phys. Soc. Japan **17**, 1100 (1962).

The value of these moments must of course be obtained by expanding both sides of (6.4) and comparing coefficients of $\mathbf{y}^n$. For a single random variable a, we find, for example, that

$$\langle a\rangle^L=\langle a\rangle,$$

$$\langle a^2\rangle^L=\langle a^2\rangle-(\langle a\rangle)^2,$$

$$\langle a^3\rangle^L=\langle(a-\langle a\rangle)^3\rangle,$$

$$\langle a^4\rangle^L=\langle(a-\langle a\rangle)^4\rangle-3[\langle(a-\langle a\rangle)^2\rangle]^2, \quad (6.7)$$

so that $\langle a^n\rangle^L$ represents intrinsic correlations of nth order, i.e., those that do not arise from lower order correlations. If $\langle a^n\rangle^L=0$ for $n>m$ then all moments $\langle a^n\rangle$ for $n>m$ can be expressed in terms of the lower order correlations $\langle a^s\rangle$ for $s\leq m$.

With the notation (6.6), Eq. (6.4) can be rewritten in the elegant symbolic form

$$\langle\exp(i\mathbf{y}\cdot\mathbf{a})\rangle=\exp\{\langle[\exp(i\mathbf{y}\cdot\mathbf{a})-1]\rangle^L\}. \quad (6.8)$$

Although our diffusion coefficients are originally defined in terms of ordinary moments

$$n!\mathbf{D}_n(\mathbf{a},t)=\lim_{\Delta t\to 0}\langle[\mathbf{a}(t+\Delta t)-\mathbf{a}(t)]^n\rangle/\Delta t, \quad (6.9)$$

we can also use a linked-moment definition

$$n!\mathbf{D}_n(\mathbf{a},t)=\lim_{\Delta t\to 0}\langle[\mathbf{a}(t+\Delta t)-\mathbf{a}(t)]^n\rangle^L/\Delta t, \quad (6.10)$$

since the lower order moments to be subtracted off, yield higher powers of Δt. In both cases averages are taken subject to $\mathbf{a}(t)=\mathbf{a}$ as initial condition. With the notation $\Delta\mathbf{a}=\mathbf{a}(t+\Delta t)-\mathbf{a}(t)$ and (6.10) we can rewrite (5.15) for $L(\mathbf{y},\mathbf{a},t)$ in the form:

$$-L(\mathbf{y},\mathbf{a},t)=\langle[\exp(i\mathbf{y}\cdot\Delta\mathbf{a})-1]\rangle^L/\Delta t. \quad (6.11)$$

Thus the characteristic function of the transition probability

$$\langle\exp(i\mathbf{y}\cdot\Delta\mathbf{a})\rangle \equiv\int d\,\Delta\mathbf{a}\exp(i\mathbf{y}\cdot\Delta\mathbf{a})P(\mathbf{a}+\Delta\mathbf{a},t+\Delta t\mid\mathbf{a},t) \quad (6.12)$$

can be rewritten using (6.8) and (6.11) in the form

$$\langle\exp(i\mathbf{y}\cdot\Delta\mathbf{a})\rangle=\exp[-L(\mathbf{y},\mathbf{a},t)\Delta t]. \quad (6.13)$$

If the average (6.12) is taken against the unnormalized transition probability $\hat{P}(\mathbf{a}+\Delta\mathbf{a},t+\Delta t\mid\mathbf{a},t)$, the result is simply multiplied by the normalization:

$$\langle\exp(i\mathbf{y}\cdot\Delta\mathbf{a})\rangle =\exp[-Q(\mathbf{a},t)\Delta t]\exp[-L(\mathbf{y},\mathbf{a},t)\Delta t]. \quad (6.14)$$

These results are correct to order Δt, and permit the transition probability to be represented in the form

$$\hat{P}(\mathbf{a}+\Delta\mathbf{a},t+\Delta t\mid\mathbf{a},t) =(2\pi)^{-N}\int d\mathbf{y}\exp\{-i\mathbf{y}\cdot\Delta\mathbf{a}-[Q(\mathbf{a})+L(\mathbf{y},\mathbf{a})]\Delta t\}. \quad (6.15)$$

If this integral representation is inserted for $\hat{P}(\mathbf{a}_j\mid\mathbf{a}_{j-1})$ with an integration variable $\mathbf{y}_j$, the path integral (4.7) is converted into an integral over paths in the phase space $\mathbf{a}_j,\mathbf{y}_j$. If $\mathbf{a}$ is a position, then $\mathbf{y}$ can be thought of as a "momentum" and $L(\mathbf{y},\mathbf{a})$ as a kind of Hamiltonian operator. We found that $\mathbf{y}$ must be replaced by $i\partial/\partial\mathbf{a}$ when acting on P, just as the momentum operator in quantum mechanics is replaced when acting on the Schrödinger wave function.[23]

If we define the unnormalized characteristic function as

$$\hat{\phi}(\mathbf{y},t)=\langle\exp(i\mathbf{y}\cdot\mathbf{a})\rangle=\int\exp(i\mathbf{y}\cdot\mathbf{a})\hat{P}(\mathbf{a},t)\,d\mathbf{a} \quad (6.16)$$

then by (5.6) we obtain

$$\partial\hat{\phi}/\partial t=-\langle Q(\mathbf{a},t)\exp(i\mathbf{y}\cdot\mathbf{a})\rangle -\sum_{n=1}^{\infty}(i\mathbf{y})^n\colon\langle D_n(\mathbf{a},t)\exp(i\mathbf{y}\cdot\mathbf{a})\rangle. \quad (6.17)$$

Now $\mathbf{a}$ can be brought down from the exponent by acting with $-i\partial/\partial\mathbf{y}$ so that

$$\partial\hat{\phi}(\mathbf{y},t)/\partial t=-[Q(\mathbf{a}_{\text{op}},t)+L(\mathbf{y},\mathbf{a}_{\text{op}},t)]\hat{\phi}(y,t) \quad (6.18)$$

or

$$\partial\hat{\phi}/\partial t=[-Q(\mathbf{a}_{\text{op}})+\sum_{n=1}^{\infty}(i\mathbf{y})^n\colon D_n(\mathbf{a}_{\text{op}})]\hat{\phi}$$

where

$$\mathbf{a}_{\text{op}}=-i\partial/\partial\mathbf{y}.$$

If $\hat{P}$ is the analog of the Schrödinger wave function of quantum mechanics, then $\hat{\phi}$ is the analog of the corresponding momentum wave function,[23] and $L(\mathbf{y}_{\text{op}},\mathbf{a}_{\text{op}})$ is the analog of the Hamiltonian operator.

If $\hat{P}(\mathbf{a},t)$ is to be determined subject to the initial condition

$$\hat{P}(\mathbf{a},t_0)=\delta(\mathbf{a}-\mathbf{a}_0), \quad (6.19)$$

i.e., $\hat{P}(\mathbf{a},t)\equiv\hat{P}(\mathbf{a},t\mid\mathbf{a}_0,t_0)$ then the corresponding initial condition on $\hat{\phi}$ is

$$\hat{\phi}(\mathbf{y},t_0)=\exp(i\mathbf{y}\cdot\mathbf{a}_0), \quad (6.20)$$

and a comparison with (4.8) shows that the functional average (4.3) is simply given by

$$M_0=\hat{\phi}(0,t) \quad (6.21)$$

---

[23] A. Messiah, *Quantum Mechanics* (John Wiley & Sons, Inc., New York, 1961).

which depends on $\mathbf{a}_0$. The complete average M can be obtained by multiplying by $P(\mathbf{a}_0)$ and integrating over $\mathbf{a}_0$ as in (4.2), or using

$$M=\hat{\phi}(0, t) \quad \text{where} \quad \hat{\phi}(\mathbf{y}, t_0) \equiv \int \exp(i\mathbf{y}\cdot\mathbf{a}_0) P(\mathbf{a}_0)\, d\mathbf{a}_0 \tag{6.22}$$

is the initial condition.

7. EXAMPLES OF MARKOFF PROCESSES

A. The Weiner Process

This is a one-dimensional process in which $D_2=D$, a constant, and all other $D_n=0$. Thus the characteristic function obeys

$$\partial\phi/\partial t=-Dy^2\phi. \tag{7A.1}$$

The solution subject to the initial condition $\phi(y,0)=\exp(iya_0)$ of (6.20) is

$$\phi(y,t)=\exp(-Dy^2t)\exp(iya_0). \tag{7A.2}$$

Taking the inverse Fourier transform

$$P(at \mid a_0 0)=(4\pi Dt)^{-\frac{1}{2}}\exp[-(a-a_0)^2/4Dt]. \tag{7A.3}$$

Thus a is a Gaussian random variable with

$$\langle a\rangle=a_0,$$
$$\langle(a-a_0)^2\rangle=2Dt, \tag{7A.4}$$

so that D is the diffusion coefficient of this simple "Brownian-motion" process in which a is usually interpreted as a distance x.

The generalization of this result to a many dimensional process is immediate. The conditional probability already stated in Eq. (4.4) can be written down immediately from the knowledge that the process is Gaussian with

$$\langle\mathbf{a}\rangle=\mathbf{a}_0,$$
$$\langle(\mathbf{a}-\mathbf{a}_0)(\mathbf{a}-\mathbf{a}_0)\rangle=2\mathbf{D}t, \tag{7A.5}$$

where $\mathbf{D}$ is an $N\times N$ matrix.

B. The Fokker–Planck Process

This is a random process for which $\mathbf{D}_n=0$ for $n>2$ so that the probability density obeys the "Fokker–Planck" equation:

$$\frac{\partial P}{\partial t}=-\frac{\partial}{\partial a_i}[A_iP]+\frac{\partial}{\partial a_i}\frac{\partial}{\partial a_j}[D_{ij}P]\equiv-\frac{\partial}{\partial a_i}J_i, \tag{7B.1}$$

in which repeated indices are summed over, and A_i, D_{ij} can be functions of $\mathbf{a}$. If we regard

$$J_i\equiv A_iP-(\partial/\partial a_j)(D_{ij}P) \tag{7B.2}$$

as a probability current, then (7B.1) describes a conservation of probability. Steady-state solutions obey $\partial P/\partial t=0$ but need not have $J_i=0$. The condition $J_i=0$ is a detailed balance (D.B.) condition that will be obeyed in many problems. When it is, one can write

$$(\text{D.B.}):\quad \partial\ln P/\partial a_k=U_k\equiv-(D^{-1})_{ki}[\partial D_{ij}/\partial a_j-A_i]. \tag{7B.3}$$

In this case, one can express $\ln P$ in terms of the line integral

$$\ln P=\int_{\mathbf{a}_0}^{\mathbf{a}} U_k\, da_k, \tag{7B.4}$$

a result that is meaningful and independent of the path only if the integrability conditions curl $U_k=0$ are obeyed:

$$\partial U_k/\partial a_l=\partial U_l/\partial a_k. \tag{7B.5}$$

This is a condition on the coefficients D_{ij}, A_i that will permit such a detailed balance solution in the steady state.

For the special case of a linear Fokker–Planck process in which

$$D_{ij}=\text{const.},\qquad A_i=-\Lambda_{ij}a_j,$$

the integrability condition takes a form

$$(\mathbf{D}^{-1}\mathbf{\Lambda})_{kl}=(\mathbf{D}^{-1}\mathbf{\Lambda})_{lk}, \tag{7B.6}$$

equivalent to our previously described condition of time reversibility.[24] A close connection between detailed balance and time reversibility has been derived in another way by van Vliet.[25] The steady-state solution of the linear Fokker–Planck process is well known. Section 5 of Ref. 1, for example, presents the solution without assuming detailed balance or time reversibility. The general time-dependent solution will be obtained in our discussion of the Ornstein–Uhlenbeck process which follows in example C.

The one-dimensional Fokker–Planck process has a special simplicity in the steady state, because then

$$\partial J/\partial a=0 \quad\text{or}\quad J=\text{const.} \tag{7B.7}$$

The general solution of (7B.2) can then be written down in the well-known form[6]

$$P(a)=-\frac{J}{D(a)}\int_{a_1}^{a}da'\exp\int_{a'}^{a}\frac{A(b)}{D(b)}db+\frac{C}{D(a)}\exp\int_{a_1}^{a}\frac{A(b)}{D(b)}db, \tag{7B.8}$$

where (for any a_1) the integration constant C is fixed by normalization. The boundary conditions usually (but not always) require P (and J) to vanish at

[24] See II, Ref. 2, last paragraph of Appendix A.

[25] K. M. van Vliet, Phys. Rev. **133**, A1182 (1964); Erratum **138**, AB3 (1965).

infinity thus forcing J to be zero in this one-dimensional case.

The time-dependent problem is usually attacked[26] by seeking the eigenfunctions of the operator L

$$\partial P_n/\partial t = -LP_n = -\lambda_n P_n \tag{7B.9}$$

whether L has the Fokker–Planck, or a more general form.

For the one-dimensional Fokker–Planck case, these eigenfunctions obey an orthogonality condition

$$\int P_m(a)P_n(a)\,da/W(a) = \delta_{mn}, \tag{7B.10}$$

where the weight factor

$$W(a) = \frac{C}{D(a)}\exp\int^a \frac{A(b)}{D(b)}\,db \tag{7B.11}$$

is the steady solution when $J=0$ is imposed. To prove (7B.10) we make the transformation

$$P = \exp\left(\int f\,da\right)Q, \tag{7B.12}$$

$$\exp\left(-\int f\,da\right)\partial/\partial a \exp\int f\,da = \partial/\partial a + f, \tag{7B.13}$$

$$L' \equiv \exp\left(-\int f\,da\right)L\exp\left(\int f\,da\right)$$

$$= -\frac{\partial}{\partial a}\left(D\frac{\partial}{\partial a}\right) + \left[A - \left(2fD + \frac{\partial D}{\partial a}\right)\right]\frac{\partial}{\partial a} + m(a) \tag{7B.14}$$

$$m(a) = \left(\frac{\partial}{\partial a}+f\right)A - \left(\frac{\partial}{\partial a}+f\right)\left(\frac{\partial}{\partial a}+f\right)D. \tag{7B.15}$$

Then L' will be Hermitian (of Sturm–Liouville type) if the coefficient of $\partial/\partial a$ vanishes, i.e., if

$$2f = [A - \partial D/\partial a]/D \tag{7B.16}$$

or

$$\exp 2\int f\,da = D^{-1}\exp\int (A/D)\,da \equiv W(a). \tag{7B.17}$$

Thus $P = W^{\frac{1}{2}}Q$ and the Hermiticity of L' guarantees that its eigenfunctions Q_n obey the unweighted orthogonality condition

$$\int Q_m(a)Q_n(a)\,da = \delta_{mn} \tag{7B.18}$$

which demonstrates (7B.10). We shall assume completeness[27]:

$$\sum_n Q_n(a)Q_n(a') = \delta(a-a'). \tag{7B.19}$$

The general time-dependent solution using (7B.9) and (7B.19) can be written in the form

$$P(at \mid a_0t_0) = \sum_n \exp[-\lambda_n(t-t_0)]P_n(a)P_n(a_0)/W(a_0). \tag{7B.20}$$

The lowest eigenvalue is $\lambda_0 = 0$ with $P_0(a)$ the steady-state solution, which is equal to $W(a)$ when $J=0$, so that in this case

$$P(at, a_0t_0) \equiv P(at \mid a_0t_0)P_0(a_0)$$

$$= \sum_n \exp[-\lambda_n(t-t_0)]P_n(a)P_n(a_0). \tag{7B.21}$$

For $V = V(a)$ = any function of a,

$$\langle V(t)V(0)\rangle = \sum_n \exp(-\lambda_n t)\left[\int V(a)P_n(a)\,da\right]^2. \tag{7B.22}$$

Continuity Theorem (7B.1): Almost all sample functions of a one-dimensional Fokker–Planck process with bounded $D(a) \le D_m$ are continuous. We shall make this theorem plausible by using (6.15) to obtain

$$P(a+\Delta a, t+\Delta t \mid a, t)$$

$$= (4\pi\Delta tD)^{-\frac{1}{2}}\exp[-(\Delta a - A\Delta t)^2/4D\Delta t] \tag{7B.23}$$

for the transition probability. The drift term $A(a)\Delta t$ causes a differentiable change in a so that we shall ignore $A(a)$ and concentrate on $D(a)$. Over the time interval $0 \le t \le T$, choose $\Delta t = T/N$. Using the inequality[28]

$$\int_\lambda^\infty \exp(-\tfrac{1}{2}\xi^2)\,d\xi \le \int_\lambda^\infty \frac{\xi}{\lambda}\exp(-\tfrac{1}{2}\xi^2)\,d\xi = \frac{\exp(-\frac{1}{2}\lambda^2)}{\lambda},$$

$$P(\Delta a > (4D_0T)^{\frac{1}{2}}N^{-\frac{1}{4}})$$

$$\le \pi^{-\frac{1}{2}}(D_m/D_0)N^{-\frac{1}{4}}\exp[-D_0N^{\frac{1}{2}}/D_m]. \tag{7B.24}$$

Thus for all N intervals Δt, the probability for the largest Δa obeys

$$P((\Delta a)_{\text{largest}} > (4D_0T)^{\frac{1}{2}}N^{-\frac{1}{4}})$$

$$\le N\pi^{-\frac{1}{2}}(D_m/D_0)N^{-\frac{1}{4}}\exp(-D_0N^{\frac{1}{2}}/D_m) \tag{7B.25}$$

so that as $N\to\infty$, even the largest Δa goes to zero with probability one.

C. The Ornstein–Uhlenbeck (O.U.) Process

The O.U. process[29] is originally defined as a velocity $a = u$ subject to a damping linear in the velocity plus a random force (see the Langevin noise source de-

[26] N. G. van Kampen, Phys. Rev. **110,** 319 (1958); also Ref. 6.

[27] See, e.g., P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill Book Company, Inc., New York, 1953), Sec. 6.3, Eq. (6.3.11), and p. 729.

[28] Our proof is a slight generalization of that given by J. L. Doob, *Stochastic Processes* (John Wiley & Sons, Inc., New York, 1953), Sec. VIII.2.

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

scription in I, Sec. 8). In our present language, the linear damping [see Eq. (5.10)] means

$$\mathbf{D}_1(\mathbf{a}) \equiv \mathbf{A}(\mathbf{a}) = -\mathbf{\Lambda}\cdot\mathbf{a}, \qquad (7C.1)$$

where we have generalized to the multivariable case at the minor expense of using a matrix decay constant $\mathbf{\Lambda}$. The "Brownian motion" random force results in simple diffusion with

$$\mathbf{D}_2(\mathbf{a}) = \mathbf{D} = \text{a constant matrix}; \qquad \mathbf{D}_n = 0 \text{ for } n > 2. \qquad (7C.2)$$

(The relation between the Langevin and Markoff descriptions of random processes will be developed in paper IV of this series.)

Thus our "Hamiltonian" (5.16) has the form

$$L(\mathbf{y}, \mathbf{a}, t) = i\mathbf{y}\cdot\mathbf{\Lambda}\cdot\mathbf{a} + \mathbf{y}\cdot\mathbf{D}\cdot\mathbf{y}; \qquad \mathbf{a} = -i\partial/\partial\mathbf{y} \qquad (7C.3)$$

so that by (6.18), the characteristic function $\phi(\mathbf{y}, t)$ obeys

$$\partial\phi/\partial t = -\mathbf{y}\cdot\mathbf{\Lambda}\cdot\partial\phi/\partial\mathbf{y} - \mathbf{y}\cdot\mathbf{D}\cdot\mathbf{y}\phi \qquad (7C.4)$$

subject to

$$\phi(\mathbf{y}, 0) = \exp(i\mathbf{y}\cdot\mathbf{a}_0), \qquad (7C.5)$$

the initial condition (6.20).

This partial differential equation can be solved by first eliminating the "drift terms" (those proportional to $\mathbf{\Lambda}$). The transformation that does this leads to an exact solution of the case in which $\mathbf{D}=0$. For this special case, it is well known that such a linear partial differential equation of the first order is equivalent to the set of ordinary differential equations:

$$dt = (\mathbf{y}\cdot\mathbf{\Lambda})^{-1}\cdot d\mathbf{y} \qquad (7C.6)$$

or

$$d\mathbf{y}/dt = \mathbf{y}\cdot\mathbf{\Lambda}.$$

[A physicist's way of arriving at this drift equation would be to write a "material" derivative in the form

$$\frac{d\phi}{dt} \equiv \frac{\partial\phi}{\partial t} + \frac{d\mathbf{y}}{dt}\frac{\partial\phi}{\partial\mathbf{y}}.$$

Comparison with (7C.4) then yields $d\mathbf{y}/dt$.] The general solution of the ordinary differential equations is

$$\mathbf{y}(t) = \mathbf{y}(0)\exp(\mathbf{\Lambda}t)$$

or

$$\mathbf{y}(t)\exp(-\mathbf{\Lambda}t) = \mathbf{y}(0) = \text{const.} \qquad (7C.7)$$

Thus the general solution for ϕ is

$$\phi = \text{arbitrary function of } \mathbf{y}\exp(-\mathbf{\Lambda}t). \qquad (7C.8)$$

The solution obeying the initial condition (7C.5) is then

$$\phi(\mathbf{y}, t) = \exp[i\mathbf{y}\cdot\exp(-\mathbf{\Lambda}t)\cdot\mathbf{a}_0] \qquad (7C.9)$$

and the corresponding density is

$$P(\mathbf{a}, t) = \delta(\mathbf{a} - [\exp(-\mathbf{\Lambda}t)]\cdot\mathbf{a}_0). \qquad (7C.10)$$

Thus

$$\langle\mathbf{a}(t)\rangle = \exp(-\mathbf{\Lambda}t)\mathbf{a}_0 \qquad (7C.11)$$

in agreement with the first moment equation of (5.10) when $Q=0$, and, when $\mathbf{D}=0$, no fluctuations are present.

The results suggest, that when $\mathbf{D}\neq 0$ we make the transformation

$$\mathbf{y} = \mathbf{z}\cdot\exp(\mathbf{\Lambda}t) \qquad (7C.12)$$

$$\phi(\mathbf{y}, t) = \phi(\mathbf{z}\exp(\mathbf{\Lambda}t), t) \equiv \psi(\mathbf{z}, t) \qquad (7C.13)$$

with

$$\psi(\mathbf{z}, 0) = \phi(\mathbf{z}, 0) = \exp(i\mathbf{z}\cdot\mathbf{a}_0) \qquad (7C.14)$$

as the initial condition. Since

$$\left.\frac{\partial\psi}{\partial t}\right|_z = \frac{\partial\phi}{\partial t} + \left.\frac{d\mathbf{y}}{dt}\right|_z\frac{\partial\phi}{\partial\mathbf{y}} = \frac{\partial\phi}{\partial t} + \mathbf{z}\exp(\mathbf{\Lambda}t)\mathbf{\Lambda}\frac{\partial\phi}{\partial\mathbf{y}}$$

$$= (\partial\phi/\partial t) + \mathbf{y}\cdot\mathbf{\Lambda}(\partial\phi/\partial\mathbf{y}) = -\mathbf{y}\cdot\mathbf{D}\cdot\mathbf{y}\phi, \qquad (7C.15)$$

the drift terms disappear and

$$\partial\psi/\partial t = -\mathbf{z}\exp(\mathbf{\Lambda}t)\mathbf{D}\exp(\mathbf{\Lambda}^\dagger t)\mathbf{z}\psi(\mathbf{z}, t), \qquad (7C.16)$$

where $\mathbf{\Lambda}^\dagger$ is the transpose of $\mathbf{\Lambda}$. The solution of this equation with the initial condition (7C.14) is then

$$\psi(\mathbf{z}, t) = \exp\left[-\mathbf{z}\cdot\int_0^t \exp(\mathbf{\Lambda}s)\mathbf{D}\exp(\mathbf{\Lambda}^\dagger s)\,ds\cdot\mathbf{z}\right] \times\exp(i\mathbf{z}\cdot\mathbf{a}_0), \qquad (7C.17)$$

or in terms of the original variable $\mathbf{y}$,

$$\phi(\mathbf{y}, t) = \exp[i\mathbf{y}\cdot\langle\mathbf{a}\rangle^L - (1/2!)\mathbf{y}\cdot\langle\mathbf{aa}\rangle^L\cdot\mathbf{y}], \qquad (7C.18).$$

using (6.8) where the linked averages are given by

$$\langle\mathbf{a}\rangle^L = \langle\mathbf{a}(t)\rangle = \exp(-\mathbf{\Lambda}t)\mathbf{a}_0, \qquad (7C.19)$$

$$\langle\mathbf{aa}\rangle^L = \langle[\mathbf{a}(t) - \langle\mathbf{a}(t)\rangle][\mathbf{a}(t) - \langle\mathbf{a}(t)\rangle]\rangle$$

$$= 2\int_0^t \exp(-\mathbf{\Lambda}u)\mathbf{D}\exp(-\mathbf{\Lambda}^\dagger u)\,du, \qquad (7C.20)$$

where $u = t - s$. The stationary result can be obtained by taking the limit as $t\to\infty$. Thus $\langle\mathbf{a}\rangle\to 0$, and writing $\langle\boldsymbol{\alpha\alpha}\rangle$ for the second moment fluctuation in the stationary case:

$$\langle\boldsymbol{\alpha\alpha}\rangle = 2\int_0^\infty \exp(-\mathbf{\Lambda}t)\mathbf{D}\exp(-\mathbf{\Lambda}^\dagger t)\,dt. \qquad (7C.21)$$

We have previously obtained this result (I, Sec. 5) by solving the Einstein relation

$$2\mathbf{D} = \mathbf{\Lambda}\langle\boldsymbol{\alpha\alpha}\rangle + \langle\boldsymbol{\alpha\alpha}\rangle\mathbf{\Lambda}^\dagger, \qquad (7C.22)$$

which is a consequence of stationarity of the second moments. In the presence of underlying time reversibility, I, Sec. 6, or if $\mathbf{\Lambda}^{-1}\mathbf{D}$ is symmetric,[24]

$$\langle\boldsymbol{\alpha\alpha}\rangle = \mathbf{\Lambda}^{-1}\mathbf{D}. \tag{7C.23}$$

If (7C.22) is inserted into (7C.20), the integrand is found to be a perfect differential so that

$$\langle\mathbf{aa}\rangle^{L} = \langle\boldsymbol{\alpha\alpha}\rangle - \exp(-\mathbf{\Lambda}t)\langle\boldsymbol{\alpha\alpha}\rangle\exp(-\mathbf{\Lambda}^{\dagger}t), \tag{7C.24}$$

which explicitly displays the approach to the stationary case. Equation (7C.24) is in turn a special case of our earlier result

$$\begin{aligned}\langle[\mathbf{a}(t)-\langle\mathbf{a}(t)\rangle][\mathbf{a}(u)-\langle\mathbf{a}(u)\rangle]\rangle_{\mathbf{a}_0} &= \exp(-\mathbf{\Lambda}\,|\,t-u\,|)\langle\boldsymbol{\alpha\alpha}\rangle \\ &\quad - \exp(-\mathbf{\Lambda}t)\langle\boldsymbol{\alpha\alpha}\rangle\exp(-\mathbf{\Lambda}^{\dagger}u),\end{aligned} \tag{7C.25}$$

I(8.18), obtained by Langevin techniques.

The O.U. process and the Weiner process are *quasi-stationary* in the sense that their transition probabilities or "Hamiltonians" $L(\mathbf{y}, \mathbf{a}, t) = L(\mathbf{y}, \mathbf{a})$ do not depend explicitly on the time. The Weiner process possesses no stationary limit in the sense of (2.10) whereas the linear decay causes the O.U. process to possess such a stationary limit.

The O.U. process is linear and Gaussian as well as stationary. Thus it illustrates

Doob's Theorem: A random process that is stationary, Gaussian and Markoffian possesses an autocorrelation of the form

$$\langle\mathbf{a}(t)\mathbf{a}(0)\rangle = \exp(-\mathbf{\Lambda}t)\langle\mathbf{a}(0)\mathbf{a}(0)\rangle. \tag{7C.26}$$

Doob[30] states this theorem for a one-dimensional random process. It was extended to the many-dimensional case by Kac.[31] The proofs are based on the Chapman–Kolmogoroff relation (2.12) and are manipulative rather than informative. Let us therefore state the

Generalized Doob Theorem: A random process that is Gaussian and Markoffian must be a linear, Fokker–Planck process, i.e., $\mathbf{D}_n = 0$ *for* $n>2$, $\mathbf{D}_2 = \mathbf{D} =$ *independent of* $\mathbf{a}$, $\mathbf{D}_1 = -\mathbf{\Lambda}\cdot\mathbf{a}$, *where* $\mathbf{D}$ *and* $\mathbf{\Lambda}$ *can be time-dependent.*
Proof: If the $\mathbf{a}$ process is Gaussian, then the transition probability has the form

$$P(\mathbf{a}+\Delta\mathbf{a}, t+\Delta t\,|\,\mathbf{a}, t) \propto \exp[\mathbf{a}\cdot\boldsymbol{\mu}\cdot\mathbf{a} + \mathbf{a}\cdot\boldsymbol{\nu}\cdot\Delta\mathbf{a} + \Delta\mathbf{a}\cdot\boldsymbol{\lambda}\cdot\Delta\mathbf{a}], \tag{7C.27}$$

where $\boldsymbol{\mu}$, $\boldsymbol{\nu}$, and $\boldsymbol{\lambda}$ may depend on time, but not on $\mathbf{a}$ or $\Delta\mathbf{a}$. The characteristic function (6.12) of the transition probability then necessarily has the form

$$\exp[-L(\mathbf{y}, \mathbf{a}, t)\Delta t] \sim \exp[(i\mathbf{y}\cdot\mathbf{\Lambda}\cdot\mathbf{a} + \mathbf{y}\cdot\mathbf{D}\cdot\mathbf{y})\Delta t] \tag{7C.28}$$

since the Fourier transform of a Gaussian in $\Delta\mathbf{a}$ is a Gaussian in $\mathbf{y}$. Moreover, the form of (7C.27) requires that $\mathbf{D}$ be independent of $\mathbf{a}$, and the coefficient of $\mathbf{y}$ be linear in $\mathbf{a}$. With a suitable (possibly time-dependent) choice of origin, this coefficient can be written $\mathbf{\Lambda}\cdot\mathbf{a}$. This completes our proof. The original Doob theorem then follows if we note that stationarity forces $\mathbf{\Lambda}$ and $\mathbf{D}$ to be independent of time, reducing the process to an O.U. process. Then

$$\begin{aligned}\langle\mathbf{a}(t)\mathbf{a}(0)\rangle &= \int \mathbf{a}\, d\mathbf{a} P(\mathbf{a}t\,|\,\mathbf{a}_0 0)\, da_0 \mathbf{a}_0 P(\mathbf{a}_0) \\ &= \int \langle\mathbf{a}(t)\rangle_{\mathbf{a}_0}\mathbf{a}_0\, d\mathbf{a}_0 P(\mathbf{a}_0),\end{aligned}$$

which reduces to the desired form (7C.26) when the conditional mean $\langle\mathbf{a}(t)\rangle_{\mathbf{a}(0)}$ is given the appropriate value $\exp(-\mathbf{\Lambda}t)\mathbf{a}(0)$ for an O.U. process (7C.29).

D. The Poisson Process (Shot Noise)

This is a discrete process with $a(0)=0$ and $a(t)=$ an integer $=$ the number of events that have occurred in the time interval $(0, t)$. Then events are assumed to occur at random at a rate ν per unit time. Thus

$$\partial P(a, t)/\partial t = \nu[P(a-1, t) - P(a, t)], \tag{7D.1}$$

where the first term represents transitions into state a from $a-1$, and the second term represents transitions out of a (into $a+1$). For the case $a=0$,

$$\partial P(0, t)/\partial t = -\nu P(0, t). \tag{7D.2}$$

Thus

$$P(0, t) = \exp(-\nu t). \tag{7D.3}$$

Setting

$$P(a, t) = \exp(-\nu t) Q(a, t), \tag{7D.4}$$

Eq. (7D.1) reduces to

$$\partial Q(a, t)/\partial t = \nu Q(a-1, t). \tag{7D.5}$$

Using $Q(0, t)=1$ and iterating we get

$$Q(a, t) = (\nu t)^a/a!, \qquad P(a, t) = (\nu t)^a \exp(-\nu t)/a!. \tag{7D.6}$$

All of these results are standard, and the moments $\langle a^n\rangle$ can be computed from $P(a, t)$. It is more instructive, however, to rewrite (7D.1) as

$$\begin{aligned}\partial P(a, t)/\partial t &= \nu[\exp(-\partial/\partial a) - 1]P(a, t) \\ &= \nu\sum_1^{\infty}[(-1)^n/n!](\partial/\partial a)^n P(a, t) \\ &= \sum(-1)^n(\partial/\partial a)^n D_n P(a, t)\end{aligned} \tag{7D.7}$$

so that

$$D_n = \nu/n!, \tag{7D.8}$$

[30] J. L. Doob, Ann. Math. **43**, 351 (1942).
[31] See M. C. Wang and G. E. Uhlenbeck, Rev. Mod. Phys. **17**, 323 (1945), Appendix II.

and

$$-L(y, a, t)=\sum_1^{\infty}(iy)^nD_n=\nu[\exp(iy)-1]. \quad (7D.9)$$

Thus

$$\partial\phi/\partial t=\nu(e^{iy}-1)\phi. \quad (7D.10)$$

With the initial condition $P(a, 0)=\delta(a)$ or $\phi(y, 0)=1$ we have

$$\langle\exp iya(t)\rangle=\phi(y, t)=\exp[\nu t(e^{iy}-1)] \quad (7D.11)$$

or

$$\exp\langle e^{iya}-1\rangle^L=\exp[\nu t(e^{iy}-1)], \quad (7D.12)$$

and

$$\langle a^n\rangle^L=\nu t; \qquad n\geq 1 \quad (7D.13)$$

so that all linked moments, $n\geq 1$, in a Poisson process take the same value!

E. The Homogeneous Process

A homogeneous process is a random process that is independent of the choice of origin of $\mathbf{a}$. Thus all higher order probability densities $P(\mathbf{a}_n, \mathbf{a}_{n-1}, \cdots, \mathbf{a}_1)$ are functions only of the differences $\mathbf{a}_i-\mathbf{a}_j$. A homogeneous Markoff process can be defined most succintly by the requirement that

$$L(\mathbf{y}, \mathbf{a}, t)=L(\mathbf{y}, t) \quad (7E.1)$$

independent of $\mathbf{a}$. (If the process is also stationary, L will be independent of t.)

As a simple example of a homogeneous process, we note the one-dimensional Poisson process with jump η:

$$\partial P/\partial t=\nu[P(a-\eta, t)-P(a, t)], \quad (7E.2)$$

which may be immediately generalized to a normalized distribution $g(\eta)$ of possible jumps

$$\int g(\eta)d\eta=1,$$

$$\partial P/\partial t=\nu\int g(\eta)d\eta[\exp(-\eta\partial/\partial a)-1]P(a, t)$$

$$=\sum_{n=1}^{\infty}(-1)^n(\partial/\partial a)^nD_nP(a, t), \quad (7E.3)$$

where

$$D_n=(\nu/n!)\int g(\eta)\, d\eta\eta^n.$$

Thus

$$-L(y, t)=\sum_1^{\infty}(iy)^nD_n=\nu\int g(\eta)d\eta[\exp(iy\eta)-1]. \quad (7E.4)$$

This process will be stationary or not, according as $g(\eta)$ does or does not depend explicitly on t. In any case

$$\partial\phi/\partial t=-L(y, t)\phi(y, t) \quad (7E.5)$$

has the explicit solution

$$\phi(y, t)=\exp\left[-\int_0^t L(y, s)\, ds\right]\exp(iya_0) \quad (7E.6)$$

if we take $\phi(y, 0)=\exp(iya_0)$ appropriate to $a(0)=a_0$. Thus

$$P(a, t\mid a_0, 0)=\frac{1}{2\pi}\int\exp[-iy(a-a_0)]\, dy \times\exp\left[-\int_0^t L(y, s)\, ds\right] \quad (7.E7)$$

In the adiabatic theory of line broadening discussed in Sec. 3A, we are concerned with

$$M(t)=\left\langle\exp\left(i\int_0^t\omega(s)\, ds\right)\right\rangle. \quad (7E.8)$$

If we interpret

$$a(t)-a(0)=\int_0^t\omega(s)\, ds \quad (7E.9)$$

and η as the phase shift $\int\omega\, ds$ induced in each collision, and write

$$L(y)=\nu\int g(\eta)\, d\eta[1-\exp(iy\eta)], \quad (7E.10)$$

then with $g(\eta)$ independent of the time,

$$M(t)=\langle\exp[i(a-a_0)]\rangle$$

$$=\phi(1, t)\exp(-ia_0)=\exp[-L(1)t]$$

$$=\exp\left[-\nu t\int g(\eta)\, d\eta(1-e^{i\eta})\right]. \quad (7E.11)$$

If now we write

$$\nu=nv\sigma, \quad (7E.12)$$

where n is the density of foreign atoms producing the collision broadening, v a typical relative velocity, and σ the total cross section for collision, then $\sigma g(\eta)d\eta$ is the cross section for collision with phase shift in η, $\eta+d\eta$. If we define

$$\sigma_r+i\sigma_i=\sigma\int g(\eta)\, d\eta(1-e^{i\eta}), \quad (7E.13)$$

then

$$M(t)=\exp[-nvt(\sigma_r+i\sigma_i)], \qquad t>0,$$

$$=\exp[-nv\sigma_r\mid t\mid-inv\sigma_i t], \qquad \text{all } t, \quad (7E.14)$$

where the last form has used $M(-t)=M(t)^*$ in order that the spectral distribution

$$I(\omega)=\int \exp\left[-i(\omega-\omega_{if})t\right] dM(t) \tag{7E.15}$$

$$=\frac{2nv\sigma_r}{(nv\sigma_r)^2+(\omega-\omega_{if}+nv\sigma_i)^2} \tag{7E.16}$$

be real. The usual Lorentz line shape,[13] with width $nv\sigma_r$ and frequency shift $nv\sigma_i$ is thus seen to be a simple consequence of the (generalized) Poisson matrix of the collisions, i.e., the assumption that the duration of each collision is so short that we have a succession of independent phase shifts occurring at a certain rate, and with a certain distribution.

F. Homogeneous Noise Plus Linear Damping

We can always decompose L into a drift part and a noise or diffusion part:

$$L(\mathbf{y},\mathbf{a},t)=-i\mathbf{y}\cdot\mathbf{A}+K(\mathbf{y},\mathbf{a},t), \tag{7F.1}$$

where

$$-K(\mathbf{y},\mathbf{a},t)=\sum_{n=2}^{\infty}(i\mathbf{y})^n:\mathbf{D}_n(\mathbf{a},t). \tag{7F.2}$$

The process we wish to consider in this subsection involves linear damping and homogeneous noise, i.e.,

$$L(\mathbf{y},\mathbf{a},t)=i\mathbf{y}\cdot\mathbf{A}(t)\cdot\mathbf{a}+K(\mathbf{y},t). \tag{7F.3}$$

Except for the nonlinear Fokker–Planck case, this process includes all previous cases A to E. Moreover, we shall ask a more general question,

$$M_0=\left\langle \exp i\int_0^t \mathbf{q}(s)\cdot\mathbf{a}(s)\,ds\right\rangle, \tag{7F.4}$$

than solved in the previous cases. [The most important applications will be stationary, i.e., $\mathbf{A}(t)=\mathbf{A}$; $K(\mathbf{y},t)=K(\mathbf{y})$.] It follows from (6.21) and (6.18) with $Q=-i\mathbf{q}(t)\cdot\mathbf{a}$ that

$$M_0=\hat{\phi}(0,t), \tag{7F.5}$$

where

$$[\partial\hat{\phi}(\mathbf{y},t)]/\partial t=(\mathbf{q}-\mathbf{y}\cdot\mathbf{A})\cdot(\partial\hat{\phi}/\partial\mathbf{y})-K(\mathbf{y},t)\hat{\phi}. \tag{7F.6}$$

We now follow our procedure in treating the O.U. process. The solution of the "dynamical" equation

$$d\mathbf{y}/dt=\mathbf{y}\cdot\mathbf{A}(t)-\mathbf{q}(t) \tag{7F.7}$$

can be written

$$\mathbf{y}(t)=\mathbf{y}(0)\cdot\mathbf{m}(t)+\mathbf{y}_0(t), \tag{7F.8}$$

where $\mathbf{m}(t)$ obeys

$$d\mathbf{m}/dt=\mathbf{m}\cdot\mathbf{A}(t). \tag{7F.9}$$

If $\mathbf{A}$ does not depend explicitly on the time

$$\mathbf{m}(t)=\exp(\mathbf{A}t). \tag{7F.10}$$

Otherwise,

$$\mathbf{m}(t)=1+\int_0^t \mathbf{A}(s_1)\,ds_1+\int_0^t ds_2\int_0^{s_2} ds_1\mathbf{A}(s_1)\mathbf{A}(s_2)+\cdots$$

$$+\int_0^t ds_n\int_0^{s_n} ds_{n-1}\cdots\int_0^{s_2} ds_1\mathbf{A}(s_1)\mathbf{A}(s_2)\cdots\mathbf{A}(s_n)+\cdots$$

$$=\sum_{n=0}^{\infty}\frac{1}{n!}\int_0^t\int_0^t\cdots\int_0^t ds_n\cdots ds_1 T_{\rightarrow}[\mathbf{A}(s_1)\mathbf{A}(s_2)\cdots\mathbf{A}(s_n)]$$

$$=T_{\rightarrow}\left[\exp\int_0^t ds\mathbf{A}(s)\right], \tag{7F.11}$$

where $T_{\rightarrow}$ is the usual time-ordering symbol[32] that places the matrices from left to right in order of increasing time. The other term in (7F.8) obeys

$$d\mathbf{y}_0/dt=-\mathbf{q}(t)+\mathbf{y}_0(t)\cdot\mathbf{A}(t), \tag{7F.12}$$

and is given explicitly by

$$\mathbf{y}_0(t)=-\int_0^t du\mathbf{q}(u)\mathbf{m}(u,t), \tag{7F.13}$$

where

$$\mathbf{m}(u,t)\equiv\mathbf{m}(u)^{-1}\mathbf{m}(t)=T_{\rightarrow}\left[\exp\int_u^t ds\mathbf{A}(s)\right]. \tag{7F.14}$$

Equation (7F.8) suggests the transformation

$$\mathbf{y}=\mathbf{z}\cdot\mathbf{m}(t)+\mathbf{y}_0(t) \tag{7F.15}$$

$$\psi(\mathbf{z},t)=\hat{\phi}(\mathbf{y},t)=\hat{\phi}(\mathbf{z}\cdot\mathbf{m}(t)+\mathbf{y}_0(t),t) \tag{7F.16}$$

with

$$\psi(\mathbf{z},0)=\hat{\phi}(\mathbf{z},0)=\exp(i\mathbf{z}\cdot\mathbf{a}_0). \tag{7F.17}$$

As in the O.U. case, the first-derivative terms have been made to disappear:

$$\partial\psi(\mathbf{z},t)/\partial t=-K((\mathbf{z}\cdot\mathbf{m}(t)+\mathbf{y}_0(t),t)\psi(\mathbf{z},t) \tag{7F.18}$$

or

$$\psi(\mathbf{z},t)=\exp\left[-\int_0^t K(\mathbf{z}\cdot\mathbf{m}(s)+\mathbf{y}_0(s),s)\,ds\right]\times\exp(i\mathbf{z}\cdot\mathbf{a}_0). \tag{7F.19}$$

Inserting $\mathbf{z}=[\mathbf{y}-\mathbf{y}_0(t)]\cdot\mathbf{m}(t)^{-1}$ we obtain

$$\hat{\phi}(\mathbf{y},t)=\exp\left[-\int_0^t K\left(\mathbf{y}\cdot\mathbf{m}(t,s)+\int_s^t du\mathbf{q}(u)\cdot\mathbf{m}(u,s),s\right)ds\right]$$

$$\times\exp\left[i\mathbf{y}\cdot\mathbf{m}(t)^{-1}\cdot\mathbf{a}_0+i\int_0^t du\mathbf{q}(u)\mathbf{m}^{-1}(u)\cdot\mathbf{a}_0\right]. \tag{7F.20}$$

[32] See, e.g., F. J. Dyson, Phys. Rev. **75**, 486, 1736 (1949); R. P. Feynman, *ibid.* **84**, 108 (1951); M. L. Goldberger and E. N. Adams, J. Chem. Phys. **20**, 240 (1952). Also Ref. 22, S. S. Schweber, H. Bethe, and F. de Hoffman, *Mesons and Fields* (Row, Peterson and Company, New York, 1956), Vol. I, Sec. 16.

Thus

$$M_0=\hat{\phi}(0,t)=\exp\left(i\int_0^t ds\mathbf{q}(s)\cdot\langle\mathbf{a}(s)\rangle\right)$$
$$\times\exp\left[-\int_0^t K\left(\int_s^t du\mathbf{q}(u)\cdot\mathbf{m}(u,s),s\right)ds\right],\quad(7\text{F}.21)$$

where

$$\langle\mathbf{a}(s)\rangle=\mathbf{m}(s)^{-1}\cdot\mathbf{a}_0\quad(7\text{F}.22)$$

is the mean "position" in the original random process (with $\mathbf{q}=0$) in the sense that

$$d\langle\mathbf{a}(s)\rangle/ds=\mathbf{\Lambda}(s)\langle\mathbf{a}(s)\rangle.\quad(7\text{F}.23)$$

If we set $\mathbf{q}(s)=\mathbf{y}\delta(t-s)$, we obtain

$$\left\langle\exp\left[i\int_0^t\mathbf{q}(s)\cdot\mathbf{a}(s)\,ds\right]\right\rangle=\langle\exp[i\mathbf{y}\cdot\mathbf{a}(t)]\rangle$$
$$=\exp[i\mathbf{y}\cdot\langle\mathbf{a}(t)\rangle]$$
$$\times\exp\left[-\int_0^t K(\mathbf{y}\cdot\mathbf{m}(t,s),s)\,ds\right].\quad(7\text{F}.24)$$

But this result is just $\phi(\mathbf{y},t)$ which can also be obtained by setting $\mathbf{q}=0$ in $\hat{\phi}(\mathbf{y},t)$. Finally,

$$P(\mathbf{a}t\mid\mathbf{a}_0,0)=\frac{1}{(2\pi)^N}\int d\mathbf{y}\exp[-i\mathbf{y}\cdot(\mathbf{a}-\langle\mathbf{a}(t)\rangle_{\mathbf{a}_0})]$$
$$\times\exp\left[-\int_0^t K(\mathbf{y}\cdot\mathbf{m}(t,s),s)\,ds\right].\quad(7\text{F}.25)$$

If we set

$$K(\mathbf{y},s)=\mathbf{y}\cdot\mathbf{D}\cdot\mathbf{y},$$
$$\mathbf{\Lambda}(s)=\mathbf{\Lambda},$$
$$\mathbf{m}(t,s)=\exp[(s-t)\mathbf{\Lambda}],\quad(7\text{F}.26)$$

and set $u=t-s$, our results (7F.24) specialized immediately to the O.U. process results (7C.18)–(7C.20).

8. SMOOTHED SQUARE-LAW RESPONSE

If a nonlinear device produces an output signal $\mathbf{b}$ that is nonlinearly related to the corresponding input signal $\mathbf{a}$ at the same instant of time then the probability distribution of the output is immediately determined by that of the input through the nonlinear transformation, e.g.,

$$P_{\text{out}}(\mathbf{b}'t'\mid\mathbf{b}t)d\mathbf{b}'=P_{\text{in}}(\mathbf{a}'t'\mid\mathbf{a}t)d\mathbf{a}'$$

or

$$P_{\text{out}}(\mathbf{b}',t'\mid\mathbf{b}t)=P_{\text{in}}(\mathbf{a}'(\mathbf{b}'),t'\mid\mathbf{a}(\mathbf{b}),t)J(\mathbf{a}'/\mathbf{b}'),\quad(8.1)$$

where $\mathbf{a}=\mathbf{a}(\mathbf{b})$ is the inverse of the nonlinear transformation $\mathbf{b}=\mathbf{b}(\mathbf{a})$ performed by the device and $J(\mathbf{a}'/\mathbf{b}')$ is the Jacobian of this (inverse) transformation.

A much more complicated problem arises if a time-smoothing operation is then performed on the output signal, since the new random variable is then a function of the original variables at many (a continuum of) time points. It is simplest to deal with the nonlinearity and time smoothing together. For example, we might be interested in the statistical behavior of

$$S=\int^t k(t-s)V(\mathbf{a}(s))\,ds\quad(8.2)$$

which we can ascertain by investigating $\langle\exp(-\lambda S)\rangle$. The procedure for solving such a problem is already given by (6.18) with $Q(\mathbf{a},s)=\lambda k(t-s)V(\mathbf{a})$ with t regarded as a parameter. To have a specific problem in mind, let us attempt to evaluate

$$M_0=\left\langle\exp\left[i\int_{t_0}^t\mathbf{q}(s)\mathbf{a}(s)\,ds\right]\right.$$
$$\left.\times\exp\left[-\lambda\int_{t_0}^t\mathbf{a}(s)\cdot\mathbf{k}(t-s)\cdot\mathbf{a}(s)\,ds\right]\right\rangle.\quad(8.3)$$

With $\mathbf{q}=0$, our nonlinear device is "square law," but with $\mathbf{q}\neq0$ it is a general quadratic device. Moreover, $\mathbf{k}$ is a matrix that combines different components of $\mathbf{a}$.

Forward-Equation Method

If the original random process is described by the function $L(\mathbf{y},\mathbf{a},t)$ of (6.13) then we must solve the problem

$$\frac{\partial\hat{\phi}}{\partial s}=\mathbf{q}\cdot\frac{\partial\hat{\phi}}{\partial\mathbf{y}}+\lambda\frac{\partial}{\partial\mathbf{y}}\cdot\mathbf{k}(t-s)\cdot\frac{\partial\hat{\phi}}{\partial\mathbf{y}}-L\left(\mathbf{y},-i\frac{\partial}{\partial\mathbf{y}},s\right)\hat{\phi},\quad(8.4)$$

where

$$\hat{\phi}=\hat{\phi}(\mathbf{y},s)=\hat{\phi}(\mathbf{y},s,t)\quad(8.5)$$

depends on t as a parameter, and our desired result is

$$M_0=\phi(0,t,t),\quad(8.6)$$
$$\hat{\phi}(\mathbf{y},t_0,t)=\exp(i\mathbf{y}\cdot\mathbf{a}_0).\quad(8.7)$$

A subsequent average over the equilibrium distribution of $\mathbf{a}_0$ can be performed if desired.

For homogeneous noise with linear decay (Sec. 7F), our equation simplifies to

$$\frac{\partial\hat{\phi}}{\partial s}=(\mathbf{q}-\mathbf{y}\cdot\mathbf{\Lambda})\frac{\partial\hat{\phi}}{\partial\mathbf{y}}+\lambda\frac{\partial}{\partial\mathbf{y}}\cdot\mathbf{k}\cdot\frac{\partial\hat{\phi}}{\partial\mathbf{y}}-K(\mathbf{y},s)\hat{\phi}.\quad(8.8)$$

This is equivalent to solving the Schrödinger equation for a (multidimensional) oscillator with time-varying mass moving in an arbitrary potential $K(\mathbf{y})$. Exact solutions are, of course, only possible with selected forms of the potential. The most important case, corresponding to *Gaussian noise* is

$$K(\mathbf{y},s)=\mathbf{y}\cdot\mathbf{D}\cdot\mathbf{y},\quad(8.9)$$

i.e., the Ornstein–Uhlenbeck process.

We now make the *ansatz*

$$\hat{\phi}=[X(s)]^{-\frac{1}{2}}\exp\{-[\mathbf{y}\cdot\mathbf{C}(s)\cdot\mathbf{y}+\mathbf{B}(s)\cdot\mathbf{y}]\},\quad (8.10)$$

subject to the initial conditions

$$X(0)=1,\qquad \mathbf{C}(0)=0,\qquad \mathbf{B}(0)=-i\mathbf{a}_0.\quad (8.11)$$

Remembering that $X(s)=X(s,t)$ depends on t as a parameter, our answer is determined once we know X:

$$M_0=[X(t,t)]^{-\frac{1}{2}},\quad (8.12)$$

and X, $\mathbf{B}$, and $\mathbf{C}$ are determined by

$$X^{-1}dX/ds=4\lambda\,\mathrm{Tr}\,(\mathbf{k}\cdot\mathbf{C})+2\mathbf{q}\cdot\mathbf{B}-2\lambda\mathbf{B}\cdot\mathbf{k}\cdot\mathbf{B},$$

$$d\mathbf{B}/ds=2\mathbf{q}\cdot\mathbf{C}-(\mathbf{\Lambda}+4\lambda\mathbf{k}\cdot\mathbf{C})\mathbf{B},$$

$$d\mathbf{C}/ds=-[\mathbf{\Lambda}\cdot\mathbf{C}+\mathbf{C}\cdot\mathbf{\Lambda}^{\dagger}]-4\lambda\mathbf{C}\cdot\mathbf{k}\cdot\mathbf{C}+\mathbf{D},\quad (8.13)$$

where $\mathbf{\Lambda}^{\dagger}$ is the transpose of $\mathbf{\Lambda}$. These equations are valid even for the nonstationary case in which $\mathbf{\Lambda}$ and $\mathbf{D}$ depend explicitly on s. For the remainder of this section, we restrict ourselves to the stationary case, however.

The equations for $\mathbf{C}$ (a set of coupled Riccati equations) are not coupled to those for $\mathbf{B}$ and X. Thus these equations must be solved first. The equations for $\mathbf{B}$ and X can then be integrated immediately. To illustrate the procedure, we shall develop in detail the solution for the *one-dimensional case*—a not entirely trivial problem.

The transformation

$$4\lambda kC=Y'/Y\quad (8.14)$$

reduces

$$dC/ds=D-2\Lambda C-4\lambda k\cdot C^2\quad (8.15)$$

to the simple form

$$Y''+[2\Lambda-(k'/k)]Y'=4\lambda DkY,\quad (8.16)$$

where primes are derivatives with respect to s and $k=k(t-s)$ is now a scalar function of s. Since the normalization of Y clearly has no effect on C, we shall take our initial conditions in the form

$$Y(t_0)=1,\qquad Y'(t_0)=0.\quad (8.17)$$

The exact solution for $B(s)$ is then given by

$$Y(s)B(s)=-ia_0\exp[-\Lambda(s-t_0)]$$

$$+2\int_{t_0}^{s}ds'\exp[-\Lambda(s-s')]q(s')C(s')Y(s')\,ds',\quad (8.18)$$

and the solution for X yields our answer in the form

$$M_0=[X(t,t)]^{-\frac{1}{2}}=[Y(t,t)]^{-\frac{1}{2}}\exp\left[-\int_{t_0}^{t}q(s)B(s)\,ds\right]$$

$$\times\exp\left[\lambda\int_{t_0}^{t}k(t-s)B(s)^2\,ds\right].\quad (8.19)$$

Note that $Y^{-\frac{1}{2}}$ is already the solution to the square-law problem ($q=0$) with the initial condition $\mathbf{a}_0=0$. If $q=0$, but $a_0\neq 0$, we have

$$M_0=[Y(t,t)]^{-\frac{1}{2}}\exp[-R(t)a_0^2],\quad (8.20)$$

$$R(t)=\lambda\int_{t_0}^{t}k(t-s)\exp[-2\Lambda(s-t_0)]\,ds/[Y(s,t)]^2.\quad (8.21)$$

[As a check on our arithmetic, we set $\lambda=0$ and $Y=1$ in (8.15), (8.18), and (8.19) to obtain

$$\left\langle\exp i\int_{t_0}^{t}q(s)a(s)\,ds\right\rangle$$

$$=\exp\left[i\int_{t_0}^{t}q(s)\,ds a_0\exp[-\Lambda(s-t_0)]\right]$$

$$\times\exp\left\{-\frac{D}{\Lambda}\int_{t_0}^{t}\int_{t_0}^{t}q(s)q(s')\{\exp[-\Lambda(s-s')]\right.$$

$$\left.-\exp[\Lambda(2t_0-s-s')]\}\,ds\,ds'\right\},\quad (8.22)$$

a result which may be compared with (7F.21) on setting $K=Dy^2$ in the latter.]

For the special case of exponential smoothing $k(t-s)=\exp[-2\beta(t-s)]$ Eq. (8.16) for $Y(s,t)$ can be solved exactly. Indeed, our first calculations were performed in this way. But we were not always able to perform the integration in (8.21) to obtain $R(t)$. Moreover, the procedure was cumbersome because one solves for $Y(s,t)$ when one is only interested in $Y(t,t)\equiv Y(t)$. Both of these difficulties are immediately removed by the use of the backward Eq. (5.22) since that procedure leads directly to an equation for the desired object $Y(t)\equiv Y(t,t)$ and moreover, we shall show that

$$R(t)=(4D)^{-1}Y'(t)/Y(t).\quad (8.23)$$

Backward-Equation Method

According to (4.8), $M_0(t,t_0)$ can be obtained by integrating $\hat{P}(\mathbf{a}t\mid\mathbf{a}_0t_0)$ over $\mathbf{a}$. Since $\mathbf{a}$ is only a parameter in the backward Eq. (5.21), we may integrate this equation over $\mathbf{a}$ without changing its form:

$$\frac{\partial M_0(\mathbf{a}_0,t,t_0)}{\partial t_0}=[Q(\mathbf{a}_0,t-t_0)$$

$$-\sum_{n=1}^{\infty}\mathbf{D}_n(\mathbf{a}_0,t_0):(\partial/\partial\mathbf{a}_0)^n]M_0.\quad (8.24)$$

For the case of time-independent transition probabilities $\mathbf{D}_n(\mathbf{a}_0,t_0)=\mathbf{D}_n)\mathbf{a}_0)$, we expect that $M_0=M_0(\mathbf{a}_0,u)$

where $u=t-t_0$ and

$$-\frac{\partial M_0(\mathbf{a}_0, u)}{\partial u}=\left[Q(\mathbf{a}_0, u)-\sum_{n=1}^{\infty}\mathbf{D}_n(\mathbf{a}_0):\left(\frac{\partial}{\partial \mathbf{a}_0}\right)^n\right]M_0. \tag{8.25}$$

For a (multidimensional) Ornstein–Uhlenbeck process

$$-\frac{\partial M_0}{\partial u}=\left[Q(\mathbf{a}_0,u)+(\mathbf{\Lambda a}_0)\cdot\frac{\partial}{\partial \mathbf{a}_0}-\frac{\partial}{\partial \mathbf{a}_0}\cdot\mathbf{D}\cdot\frac{\partial}{\partial \mathbf{a}_0}\right]M_0. \tag{8.26}$$

If, moreover, we are concerned with the average

$$M_0=\left\langle \exp i\int_{t_0}^{t}\mathbf{q}(t-s)\cdot\mathbf{a}(s)\,ds \times\exp-\lambda\int_{t_0}^{t}\mathbf{a}(s)\cdot\mathbf{k}(t-s)\cdot\mathbf{a}(s)\,ds\right\rangle, \tag{8.27}$$

then comparing with (5.20) and (5.21),

$$Q(\mathbf{a}_0, u)=-i\mathbf{q}(u)\cdot\mathbf{a}_0+\lambda\mathbf{a}_0\cdot\mathbf{k}(u)\cdot\mathbf{a}_0. \tag{8.28}$$

We can now assume a solution of the form

$$M_0(\mathbf{a}_0, u)=[Y(u)]^{-\frac{1}{2}}\exp[-\mathbf{a}_0\cdot\mathbf{R}(u)\cdot\mathbf{a}_0-\mathbf{S}(u)\cdot\mathbf{a}_0], \tag{8.29}$$

explicitly displaying the $\mathbf{a}_0$ dependence. The coefficients then obey

$$Y'(u)/Y(u)=4\mathbf{R}(u):\mathbf{D}+\mathbf{S}(u)\cdot\mathbf{D}\cdot\mathbf{S}(u), \tag{8.30}$$

$$\mathbf{S}'(u)=\mathbf{S}\cdot[\mathbf{\Lambda}-4\mathbf{D}\cdot\mathbf{R}]-\mathbf{q}(u), \tag{8.31}$$

$$\mathbf{R}'(u)=\lambda\mathbf{k}(u)-4\mathbf{R}\cdot\mathbf{D}\cdot\mathbf{R}-(\mathbf{R}\cdot\mathbf{\Lambda}+\mathbf{\Lambda}^{\dagger}\cdot\mathbf{R}), \tag{8.32}$$

where $\mathbf{\Lambda}^{\dagger}$ is the transpose of $\mathbf{\Lambda}$. Thus $\mathbf{S}$ is expressible in terms of $\mathbf{R}$ using (7F.7), (7F.13):

$$\mathbf{S}(u)=-\int_0^u\mathbf{q}(u')\,du' \times T_{\rightarrow}\exp\int_{u'}^{u}ds\,[\mathbf{\Lambda}-4\mathbf{D}\cdot\mathbf{R}(s)] \tag{8.33}$$

and Y can be obtained if both $\mathbf{S}$ and $\mathbf{R}$ are known:

$$Y(u)=\exp 4\int_0^u\mathbf{R}(u'):\mathbf{D}\,du'\exp\int_0^u\mathbf{S}(u')\cdot\mathbf{D}\cdot\mathbf{S}(u')\,du'. \tag{8.34}$$

The problem is thus reduced to obtaining $\mathbf{R}$ by solving Eq. (8.32) which is an equation for $\mathbf{R}$ alone.

It is easier to solve for $\mathbf{R}$ if we transform to new coordinates

$$\mathbf{a}_0=\mathbf{U}\cdot\mathbf{b}_0; \tag{8.35}$$

$$\mathbf{a}_0\cdot\mathbf{R}\cdot\mathbf{a}_0=\mathbf{b}_0\cdot\mathbf{V}\cdot\mathbf{b}_0, \qquad \mathbf{R}=\mathbf{U}^{\dagger}\mathbf{V}\mathbf{U}; \tag{8.36}$$

$$\mathbf{R}\cdot\mathbf{D}\cdot\mathbf{R}=\mathbf{U}^{\dagger}\mathbf{V}\mathbf{U}\mathbf{D}\mathbf{U}^{\dagger}\mathbf{V}\mathbf{U}; \tag{8.37}$$

in such a way that the new $\mathbf{D}$ is proportional to the unit matrix

$$\mathbf{D}'=\mathbf{U}\cdot\mathbf{D}\cdot\mathbf{U}^{\dagger}=\tfrac{1}{4}\mathbf{1}, \tag{8.38}$$

which is accomplished by the choice

$$\mathbf{U}=\tfrac{1}{2}(\mathbf{D})^{-\frac{1}{2}}=\mathbf{U}^{\dagger}, \tag{8.39}$$

where the symmetric (inverse) square root is understood ($\mathbf{U}^{\dagger}$ is the transpose of $\mathbf{U}$). Then

$$d\mathbf{V}/du=\mathbf{K}+\mathbf{\Gamma}^{\dagger}\mathbf{\Gamma}-(\mathbf{V}+\mathbf{\Gamma}^{\dagger})\cdot(\mathbf{V}+\mathbf{\Gamma}); \tag{8.40}$$

$$\mathbf{K}=(\mathbf{U}^{\dagger})^{-1}\cdot\lambda\mathbf{k}\cdot\mathbf{U}^{-1}; \tag{8.41}$$

$$\mathbf{\Gamma}=\mathbf{U}\cdot\mathbf{\Lambda}\cdot\mathbf{U}^{-1}, \qquad \mathbf{\Gamma}^{\dagger}=(\mathbf{U}^{\dagger})^{-1}\cdot\mathbf{\Lambda}^{\dagger}\cdot\mathbf{U}^{\dagger}. \tag{8.42}$$

The further transformation

$$\mathbf{V}=\mathbf{Z}^{-1}\cdot d\mathbf{Z}/du-\Gamma^{\dagger} \tag{8.43}$$

transforms our coupled Riccati equations into a set of linear (coupled) second-order differential equations:

$$d^2\mathbf{Z}/du^2+(d\mathbf{Z}/du)(\mathbf{\Gamma}-\mathbf{\Gamma}^{\dagger})=\mathbf{Z}\cdot(\mathbf{K}+\mathbf{\Gamma}^{\dagger}\mathbf{\Gamma}). \tag{8.44}$$

Equation (8.34) can now be simplified to

$$Y(u)=\exp\int_0^u[\mathrm{Tr}\,\mathbf{V}(u')]\,du' \times\exp\int_0^u\mathbf{S}(u')\cdot\mathbf{D}\cdot\mathbf{S}(u')\,du'. \tag{8.45}$$

The initial condition $M_0(\mathbf{a}_0, 0)=1$ implies

$$\mathbf{R}(0)=\mathbf{S}(0)=Y'(0)=0, \qquad Y(0)=1. \tag{8.46}$$

For the *one-dimensional case* our equations reduce to

$$dR/du=\lambda k(u)-2\Lambda R-4R^2D, \tag{8.47}$$

$$dV/du=4\lambda Dk(u)-2\Lambda V-V^2, \tag{8.48}$$

$$d^2Z/du^2=[4\lambda Dk(-u)+\Lambda^2]Z(u), \tag{8.49}$$

$$4DR(u)=V(u)=Z^{-1}dZ/du-\Lambda, \tag{8.50}$$

$$Y(u)=\exp(-\Lambda u)Z(u)\exp D\int_0^u[S(u')]^2\,du'. \tag{8.51}$$

The combination $Z\exp(-\Lambda u)$ obeys the simple equation

$$[Z\exp(-\Lambda u)]''+2\Lambda[Z\exp(-\Lambda u)]' =4\lambda Dk(u)[Z\exp(-\Lambda u)].$$

For the case in which $q=S=0$, we can write this as

$$Y''(u)+2\Lambda Y'(u)=4\lambda Dk(u)Y(u), \tag{8.52}$$

$$4DR(u)=Y'(u)/Y(u). \tag{8.53}$$

Exponential smoothing then leads to the differential equation

$$Y''(u)+2\Lambda Y'(u)=4\lambda D\exp(-2\beta u)\,Y(u), \quad (8.54)$$

an equation similar to but not identical to Eq. (8.10) obeyed by $Y(s,t)$. Our initial conditions are:

$$Y(0)=1, \qquad Y'(0)=0 \quad (8.55)$$

Case 1. $\Lambda=\beta=0$, *Linearly Smoothed, Squared Weiner Process*

$$Y''(u)=4\lambda DY(u), \quad (8.56)$$

$$Y(u)=\cosh[(4\lambda D)^{\frac{1}{2}}u], \quad (8.57)$$

$$R(u)=(4D)^{-1}(4\lambda D)^{\frac{1}{2}}\tanh[(4\lambda D)^{\frac{1}{2}}u], \quad (8.58)$$

$$M_0=\left\langle \exp\left[-\lambda\int_{t_0}^{t}a(s)^2\,ds\right]\right\rangle=M_0(t-t_0),$$

$$M_0(u)=\cosh[(4\lambda D)^{\frac{1}{2}}u]$$
$$\times\exp\{-\lambda a_0^2(4\lambda D)^{-\frac{1}{2}}\tanh[(4\lambda D)^{\frac{1}{2}}u]\}, \quad (8.59)$$

with $u=t-t_0$. Thus we find that as $u=t-t_0\rightarrow\infty$, $Y(u)$ continues to depend on u, i.e., our output random variable never becomes stationary. Moreover, $R(u)$ does not vanish as $u\rightarrow\infty$, so that *memory of a_0 is retained forever.* Cameron and Martin,[20] have previously evaluated $M_0(a)$ for this special case by nontrivial path integral techniques.

Case 2. $\beta=0$, $\Lambda\neq0$, *Linearly Smoothed Squared O.U. Process*

$$Y''+2\Lambda Y=4\lambda DY, \quad (8.60)$$

$$Y(u)=(2\Gamma)^{-1}\exp(-\Lambda u)[(\Gamma+\Lambda)\exp(\Gamma u)$$
$$+(\Gamma-\Lambda)\exp(-\Gamma u)],$$
$$=\Gamma^{-1}(4\lambda D)^{\frac{1}{2}}e^{-\Lambda u}\cosh(\Gamma u+\theta], \quad (8.61)$$

where

$$\Gamma=(\Lambda^2+4\lambda D)^{\frac{1}{2}},$$

$$e^{\theta}=(\Gamma+\Lambda)/(\Gamma^2-\Lambda^2)^{\frac{1}{2}},$$

$$\tanh\theta=\Lambda/\Gamma,$$

$$R(u)=(4D)^{-1}Y'/Y=\Gamma[\tanh(\Gamma u+\theta)-(\Lambda/\Gamma)], \quad (8.62)$$

$$M_0(u)=[Y(u)]^{-\frac{1}{2}}$$
$$\times\exp\{-\lambda a_0^2(\Gamma/4\lambda D)[\tanh(\Gamma u+\theta)-\tanh\theta]\}. \quad (8.63)$$

Thus the linearly smoothed squared O.U. process remains nonstationary as $u=t-t_0\rightarrow\infty$, and moreover it always remembers the starting value $a(t_0)=a_0$ even though the O.U. process with $\Lambda>0$ has a built in mechanism for forgetting this value.

Case 3. $\beta\neq0$, $\Lambda\neq0$, *Exponentially Smoothed O.U. Process*

$$Y''+2\Lambda Y'=4\lambda D\exp(-2\beta u)\,Y. \quad (8.64)$$

Let

$$x=[4\lambda D/\beta^2]^{\frac{1}{2}}e^{-\beta u}, \qquad \frac{d}{du}=\frac{dx}{du}\frac{d}{dx}=-\beta x\frac{d}{dx}, \quad (8.65)$$

$$\frac{d^2Y}{dx^2}+\frac{1-2p}{x}\frac{dY}{dx}-Y=0, \qquad p\equiv\Lambda/\beta, \quad (8.66)$$

from which we can conclude[33]

$$Y=x^p[AI_p(x)+BK_p(x)],$$

where I_p and K_p are the modified Bessel functions. With $x_0=[4\lambda D/\beta^2]$, we can make $dY/dx=0$ at $u=0$, i.e., $x=x_0$ by choosing

$$Y(x)=C\{-(d/dx_0)[x_0^pK_p(x_0)]x^pI_p(x)$$
$$+(d/dx_0)[x_0^pI_p(x_0)]x^pK_p(x)\}$$
$$=Cx_0^{p-1}x^p[K_{p-1}(x_0)I_p(x)+I_{p-1}(x_0)K_p(x)],$$

where we have used the relations[33]

$$(x^{-1}\,d/dx)^m x^pI_p(x)=x^{p-m}I_{p-m}(x),$$

$$(x^{-1}\,d/dx)^m x^pK_p(x)=(-1)^m x^{p-m}K_{p-m}(x), \quad (8.67)$$

for $m=1$. If we make use of the Wronskian relation[33]

$$K_{p-1}(x)I_p(x)+I_{p-1}(x)K_p(x)=1/x, \quad (8.68)$$

we can obey the initial condition $Y=1$ when $u=0$ or $x=x_0$ by setting $C=x_0^{2(1-p)}$ to get

$$Y(x)=x_0(x/x_0)^p[K_{p-1}(x_0)I_p(x)+I_{p-1}(x_0)K_p(x)]. \quad (8.69)$$

Using (8.53), (8.65), and (8.67), we obtain

$$4DR(u)$$
$$=\beta x^2\frac{I_{p-1}(x_0)x^{p-1}K_{p-1}(x)-K_{p-1}(x_0)x^{p-1}I_{p-1}(x)}{I_{p-1}(x_0)x^pK_p(x)+K_{p-1}(x_0)x^pI_p(x)}. \quad (8.70)$$

---

[33] I. M. Ryshik and I. S. Gradstein, *Tables of Series, Products and Integrals* (Plenum Press, New York and VEB Deutscher Verlag der Wissenschaften, Berlin, 1963), p. 329, formula 6; E. Jahnke and F. Emde, *Tables of Functions* (Dover Publications, Inc., New York, 1943), Sec. VIII.7. G. M. Watson, *Bessel Functions* (Cambridge University Press, Cambridge, England, 1948), 2nd ed.

The exact relations[33]

$$K_p(x)=(\pi/2\sin p\pi)[I_{-p}(x)-I_p(x)], \quad (8.71)$$

$$\Gamma(p)\Gamma(1-p)=(\pi/\sin p\pi), \quad (8.72)$$

and the behavior[34] as $u\rightarrow\infty$ or $x\rightarrow 0$:

$$I_p(x)\underset{x\rightarrow 0}{\rightarrow}[\Gamma(1+p)]^{-1}(x/2)^p[1+(p+1)^{-1}(x/2)^2+\cdots] \quad (8.73)$$

yields the asymptotic behavior

$$Y(x)\underset{x\rightarrow 0}{\rightarrow}\Gamma(p)(2/x_0)^{p-1}I_{p-1}(x_0)\times\{1+(1-p)^{-1}(x/2)^2 - (1-p)^{-1}\Gamma(2-p)(x/2)^{2p}[\Gamma(1+p)]^{-1} + 2(x/2)^{2p}K_{p-1}(x_0)/[I_{p-1}(x_0)\Gamma(1+p)\Gamma(p)]\}. \quad (8.74)$$

Thus as $u\rightarrow\infty$, Y achieves the *stationary value*

$$Y(u=\infty)=\Gamma(p)(2/x_0)^{p-1}I_{p-1}(x_0) \quad (8.75)$$

and $R(u)$ approaches zero rapidly:

$$4DR(u)=Y'(u)/Y(u)=2\lambda D(\Lambda-\beta)^{-1}\exp(-2\beta u) - 2(\Lambda/\beta)(\Lambda-\beta)^{-1}\Gamma(2-p)(\lambda D/\beta^2)^{\Lambda/\beta}\exp(-2\Lambda u)/\Gamma(1+p) - 4\Lambda K_{p-1}(x_0)(\lambda D/\beta^2)^{\Lambda/\beta}\exp(-2\Lambda u)/[I_{p-1}(x_0)\Gamma(1+p)\Gamma(p)], \quad (8.76)$$

showing that the smoothed O.U. process rapidly forgets a_0, i.e., M_0 becomes truly stationary! This stationary result is what will be observed experimentally, since the duration of the measurement is usually large compared to $1/\beta$ or $1/\Lambda$.

Case 4. $\beta\neq 0$, $\Lambda=0$, Exponentially Smoothed Weiner Process

Since we have allowed $\Lambda u\rightarrow\infty$ in the asymptotic formula (8.76), this result can not be applied to the $\Lambda=0$ case. We can, however, set $p=\Lambda/\beta=0$ in (8.69) or (8.70), to obtain

$$Y(x)=x_0[K_1(x_0)I_0(x)+I_1(x_0)K_0(x)], \quad (8.77)$$

$$4DR=\beta x[I_1(x_0)K_1(x)-K_1(x_0)I_1(x)]/[I_1(x_0)K_0(x)+K_1(x_0)I_1(x)]. \quad (8.78)$$

The behavior as $x\rightarrow 0$

$$xK_1(x)\rightarrow 1, \qquad K_0(x)=-\ln(x/2), \quad (8.79)$$

leads to the asymptotic ($u\rightarrow\infty$) behavior

$$Y(u)\underset{u\rightarrow\infty}{\rightarrow}x_0I_1(x_0)(u+\tau), \quad (8.80)$$

$$4DR(u)\rightarrow(u+\tau)^{-1}, \quad (8.81)$$

where

$$\tau=\beta^{-1}\{[K_1(x_0)/I_1(x_0)]-\tfrac{1}{2}\ln(\lambda D/\beta^2)\}. \quad (8.82)$$

Thus Y does not become stationary, and R vanishes very slowly as $u\rightarrow\infty$.

Case 5. Arbitrary Positive Smoothing Function $k(s-t)$

Let us summarize our results in the form

$$M_0(t-t_0)=\left\langle\exp\left[-\lambda\int_{t_0}^{t}k(t-s)a(s)^2\,ds\right]\right\rangle \quad (8.83)$$

$$=[Y(t-t_0)]^{-\frac{1}{2}}\exp[-a_0^2R(t-t_0)], \quad (8.84)$$

$$Y(t-t_0)=\exp\left[4D\int_0^{t-t_0}R(u)\,du\right]. \quad (8.85)$$

Stationarity properties of M_0 can then be stated as follows:

M_0 becomes independent of a_0 if $\lim_{u\rightarrow\infty}R(u)=0$ (8.86)

$$Y(t-t_0)\underset{t_0\rightarrow-\infty}{\longrightarrow}\text{constant} \quad \text{if} \quad \int_0^\infty R(u)\,du<\infty. \quad (8.87)$$

We see immediately that the second condition can be obeyed ($Y\rightarrow$constant) only if the first is obeyed ($M_0\rightarrow$ independent of a_0); but the converse is not necessarily true, since $R(u)$ may fall off too slowly for the integral to be convergent [as in (8.81)]. A necessary condition for $R(\infty)$ to vanish is $k(\infty)=0$, but this condition is not sufficient, since the equation

$$dR/du=-2\Lambda R-4DR^2 \quad (8.88)$$

admits two special solutions $R=0$ and $R=-(\Lambda/2D)$. The first of these solutions is stable, but the second can readily be shown to be unstable. If one even arrived at $R<-(\Lambda/2D)$, one would find a runaway solution $R\rightarrow-\infty$. We shall now show that $k(u)>0$ is sufficient to prevent such runaway solutions. Since most smoothing functions obey $k>0$, and we have not found the *necessary* condition on k to prevent

[34] This special case is worked out by Deutsch, Ref. 4 Chap. 7, using forward equation techniques.

runaway solutions, we shall suppose $k>0$ in all that follows.

Since $k(0)>0$, dR/du is positive at $u=0$, so that R increases from zero and achieves a positive value. Thenceforth, R can never become negative, for if R were to become zero at $u=u_1$, we would have $dR/du_1=\lambda k(u_1)>0$. Moreover, if $k\rightarrow 0$ as $u\rightarrow\infty$ then R must likewise approach zero, for if R remains larger than (say) $R_1>0$, then for sufficiently large u, dR/du would become negative and lead to a contradiction. The same argument essentially tells us that the condition $k\rightarrow 0$ is necessary as well as sufficient for $R(u)$ to vanish as $u\rightarrow\infty$, i.e., for M_0 in (8.83) to "forget" a_0^2. A similar argument tells us that if $k\rightarrow$constant as $a\rightarrow\infty$ then R must also approach a constant, which explains the results in Cases 1 and 2 above.

We must next investigate (8.87), i.e., whether $R(u)$ is integrable. We shall first show that if $\Lambda>0$, R is integrable if and only if k is integrable. Integrate Eq. (8.47) for R from $u=0$ to $u=\infty$ to obtain

$$2\Lambda\int_0^\infty R(u)\,du+4D\int_0^\infty R^2(u)\,du=\lambda\int_0^\infty k(u)\,du-R(\infty). \tag{8.89}$$

Suppose now, that k is integrable, then

$$R(\infty)<\lambda\int_0^\infty k(u)\,du<\infty.$$

Thus the right-hand side of (8.89) is finite, and each of the positive integrals on the left-hand side must converge.

Suppose now that the integral of k diverges. We shall show that the integral of R diverges by proving that

$$Y=\exp\left[4D\int_0^t R\,du\right]$$

cannot approach a constant. Integration of Eq. (8.52) yields

$$dY/dt+2\Lambda Y=2\Lambda+4\lambda D\int_0^t k(u)\,Y(u)\,du. \tag{8.90}$$

If $Y(u)\rightarrow$constant as $u\rightarrow\infty$, the right-hand side diverges, which forces dY/dt to diverge and contradicts the possibility of Y remaining finite when $\int k\,du$ diverges.

We shall now show that for $\Lambda=0$, R is never integrable. The transformation

$$R=[4Du+T]^{-1} \tag{8.91}$$

yields

$$dT/du=-\lambda k(4Du+T)^2. \tag{8.92}$$

Thus T is a monotonic decreasing function of u. For $u>u_1$, $T(u)<T(u_1)$ and

$$R(u)\geq[4Du+T(u_1)]^{-1}. \tag{8.93}$$

Thus $R(u)$ is not integrable. Our results may be summarized in the

Asymptotic Theorem: *Within the class of positive smoothing functions $k(t-s)$, the average*

$$M_0=\left\langle\exp\left[-\lambda\int_{t_0}^t k(t-s)\,a(s)^2\,ds\right]\right\rangle$$

[*where $a(s)$ is an Ornstein–Uhlenbeck process subject to $a(t_0)=a_0$*] *becomes independent of a_0 as $t-t_0\rightarrow\infty$ if and only if $k(u)\rightarrow 0$ as $u\rightarrow\infty$, and M_0 becomes independent of the time in this limit, if and only if $\Lambda\neq 0$* (the *O.U. process does not reduce to a Weiner process*) *and the integral*

$$\int_0^\infty k(u)\,du<\infty$$

converges.