# Optimal Paths and the Prehistory Problem for Large Fluctuations in Noise-Driven Systems

M. I. Dykman [a] and P. V. E. McClintock

*School of Physics and Materials, Lancaster University, Lancaster, LA1 4YB, United Kingdom*

V. N. Smelyanski

*Institute of Semiconductors, Ukrainian Academy of Sciences, Kiev, U.S.S.R.*

N. D. Stein and N. G. Stocks

*School of Physics and Materials, Lancaster University, Lancaster, LA1 4YB, United Kingdom*

(Received 10 September 1991)

A prehistory problem is formulated for large occasional fluctuations in noise-driven systems. It has been studied theoretically and experimentally, thereby illuminating the concept of optimal paths and making it possible to visualize and investigate them. The prehistory probability distribution measured for a white-noise-driven system, taken as an example, is shown to be in agreement with the theory.

PACS numbers: 05.40.+j, 02.50.+s, 05.20.−y

Fluctuations in physical systems can often be viewed [1] as arising because of external noise. Under stationary conditions, a weak-noise-driven system fluctuates mostly about its attractor (or attractors, if several of them coexist). However, there is also a small probability that the system will be found at a position in phase space far from an attractor. It is just these large deviations from the average that are responsible for a number of interesting physical phenomena, e.g., for switching in a variety of multistable systems (including multimode lasers, passive optically bistable systems, and Josephson junctions) and large-angle scattering (in particular, that of light) in nearly homogeneous media.

A convenient and powerful approach to the analysis of the tails of the probability density distribution $p(\mathbf{x})$ (where the components of the vector $\mathbf{x}$ enumerate the dynamical variables of a system) for systems driven by Gaussian noise is based [2–8] on the method of optimal fluctuation [9]. This approach exploits the idea that the tails of $p(\mathbf{x})$ must be formed by large occasional outbursts of noise $\mathbf{f}(t)$ that push the system far from the attractor. The probabilities of such large outbursts are small, and the value of $p(\mathbf{x}_f)$ for a given remote $\mathbf{x}_f$ will actually be determined by the probability of the *most probable* outburst among those bringing the system to $\mathbf{x}_f$. This particular realization is just the optimal fluctuation for the given $\mathbf{x}_f$. Because a realization (a path) of noise $\mathbf{f}(t)$ results in a corresponding realization of the dynamical variable $\mathbf{x}(t)$, there also exists an optimal path $\mathbf{x}_{\rm opt}(t;\mathbf{x}_f)$ along which the system arrives at $\mathbf{x}_f$, with an overwhelming probability. Although eminently reasonable and highly successful, such approaches have lacked a direct basis in experiment—the existence of optimal paths never having been demonstrated—and, to this extent, the use of the method of optimal fluctuation has amounted to an act of faith.

In this Letter, we propose a new approach to the investigation of rare events in noise-driven systems, addressing ourselves directly to the question of how one of these events (i.e., the arrival of the system at $\mathbf{x}_f$) comes to occur. In doing so, we evolve the system backwards in time from the chosen event and we define a new statistical quantity, the *prehistory probability density*, that describes the distribution of paths ending at $\mathbf{x}_f$. The ridge along the top of this distribution, representing the most probable path, is identified with the optimal path of the earlier studies; the width of the distribution provides a measure of the degree to which individual paths deviate, on average, from the optimal one.

If a point $x_f$ lies far from the attractor, so that $p(\mathbf{x}_f)$ is small, the time intervals between successive passages of $\mathbf{x}_f$ will be large; they will considerably exceed both the characteristic relaxation time of the system $\tau_r$ and the noise correlation time $\tau_c$. The arrivals of the system at $\mathbf{x}_f$ are therefore mutually uncorrelated. Since the moment of observation is the only instant of time singled out under stationary conditions, there arises the following *prehistory problem*: What is the probability density $p_h(\mathbf{x},t;\mathbf{x}_f,t_f)$ that the system was at a point $\mathbf{x}$ at time $t<t_f$, if at $t_f$ it is at $\mathbf{x}_f$? We stress that $p_h(\mathbf{x},t;\mathbf{x}_f,t_f)$ is not a standard two-time transition probability: It is given by a ratio of the *three-time* transition probability $w(\mathbf{x}_f,t_f;\mathbf{x},t;\mathbf{x}_i,t_i)$ (the probability density of the transition $\mathbf{x}_i\rightarrow\mathbf{x}\rightarrow\mathbf{x}_f$) to the *two-time* one, $w(\mathbf{x}_f,t_f;\mathbf{x}_i,t_i)$, with the initial instant $t_i$ having been set equal to $-\infty$ so that both $t_i$ and the initially occupied position $\mathbf{x}_i$ have dropped out from $p_h(\mathbf{x},t;\mathbf{x}_f,t_f)$, according to the above arguments. We note that a similar ratio of transition probabilities, although in effect for finite $t_f-t_i$, has recently been considered [10] in order to clarify the relationship between time irreversibility in thermodynamics and in cosmology. Obviously, $p_h(\mathbf{x},t;\mathbf{x}_f,t_f)\equiv p_h(\mathbf{x},t-t_f;\mathbf{x}_f,0)$.

Since prior to reaching $\mathbf{x}_f$ the system had been fluctuating for a long time, the probability density $p_h(\mathbf{x},t;\mathbf{x}_f,t_f)$ goes over into the stationary distribution for large $t_f-t$, that is, $p_h(\mathbf{x},t;\mathbf{x}_f,t_f)\rightarrow p(\mathbf{x})$ for $t_f-t\rightarrow\infty$. Because (by definition) the optimal path $\mathbf{x}_{\rm opt}(t-t_f;\mathbf{x}_f)$ is the most probable path for reaching $\mathbf{x}_f$, the function $p_h(\mathbf{x},t;\mathbf{x}_f,t_f)$ at a given $t-t_f$ should have a sharp maximum in $\mathbf{x}$ lying along that path $\mathbf{x}=\mathbf{x}_{\rm opt}(t-t_f;\mathbf{x}_f)$.

Therefore, by investigating the prehistory probability density $p_h(\mathbf{x},t;\mathbf{x}_f,0)$ experimentally one can find not only the optimal paths themselves, but also test immediately the general concepts of optimal path and optimal fluctuation and establish the range of parameters and the area of phase space within which these concepts are applicable to any given system. Before reporting, below, the results of an experiment on a particular system, we will discuss in a more general way the physics of the problem and the type of behavior to be anticipated.

By definition, the distribution $p_h(\mathbf{x},t;\mathbf{x}_f,t_f)$ is expressed in terms of the probability density functional $\Xi[\mathbf{x}(t)]$ of the paths $\mathbf{x}(t)$ of a noise-driven system as

$$p_h(\mathbf{x},t;\mathbf{x}_f,t_f)=\int_{x(-\infty)=\mathbf{x}_{\mathrm{eq}}}^{x(t_f)=\mathbf{x}_f}\mathcal{D}\mathbf{x}(\tilde{t})\,\delta(\mathbf{x}(t)-\mathbf{x})\Xi[\mathbf{x}(\tilde{t})]\left(\int_{x(-\infty)=\mathbf{x}_{\mathrm{eq}}}^{x(t_f)=\mathbf{x}_f}\mathcal{D}\mathbf{x}(\tilde{t})\,\Xi[\mathbf{x}(\tilde{t})]\right)^{-1}, \tag{1}$$

the value of $p_h(\mathbf{x},t;\mathbf{x}_f,t_f)$ being given by the relative numbers of those paths ending in $\mathbf{x}_f$ at $t_f$ that passed the point $x$ at the instant $t<t_f$. We note that the boundary condition for $t\to-\infty$ in (1) is arbitrary, strictly speaking: The system quickly gets randomized and will have forgotten its initial position prior to the fluctuation bringing it to $\mathbf{x}_f$ at $t_f$. In what follows, we assume that the attractor is a stable state, and that $\mathbf{x}_{\mathrm{eq}}$ in (1) is the position of this state.

The probability density functional $\Xi[\mathbf{x}(t)]$ for a system driven by Gaussian noise was analyzed in Ref. [11]. A simple approximate expression for $\Xi[\mathbf{x}(t)]$ that gives the main term in $p_h(\mathbf{x},t;\mathbf{x}_f,t_f)$ can be obtained by making use of Feynman and Hibbs' idea [12] of the direct interrelation between $\Xi[\mathbf{x}(t)]$ and the probability density functional of the noise $\mathcal{P}[f(t)]$ (this idea has already [3,9] been applied to the problem of optimal fluctuation). To illustrate the approach we shall consider in what follows the case of a system described by one dynamical variable, $x$, and assume the equation of motion to be of the form

$$\dot{x}(t)=-U'(x)+f(t), \tag{2}$$

where $f(t)$ is stationary, zero-mean, Gaussian noise with the frequency-dependent power spectrum (colored noise)

$$\Phi(\omega)=\int_{-\infty}^{\infty}dt\exp(i\omega t)\langle f(t)f(0)\rangle \tag{3}$$

[$\Phi^{-1}(\omega)$ is supposed to be of the form of a polynomial in $\omega^2$]. For such noise the probability density functional can be written as

$$\mathcal{P}[f(t)]=\exp\left[-\frac{1}{2D}\int_{-\infty}^{\infty}dt\,f(t)F\left(-i\frac{d}{dt}\right)f(t)\right], \tag{4}$$

$$F(\omega)=D/\Phi(\omega),\quad D=\Phi_{\max}(\omega).$$

We have singled out here explicitly the characteristic noise intensity $D$. It is assumed small (the criteria are given below).

The main term in the prehistory probability density $p_h(x,t;x_f,t_f)$ is determined by the probability of the most favorable realization of noise bringing the system to $x_f$ at $t_f$ from the vicinity of the equilibrium position $x_{\mathrm{eq}}$, via the point $x$, at the moment $t$. Following the arguments of Ref. [9] one can show that

$$p_h(x,t;x_f,t_f)\equiv p_h(x,t-t_f;x_f,0)$$
$$=C\exp[-\rho(x,t-t_f;x_f)/D], \tag{5}$$
$$\rho(x,t;x_f)=\tilde{\rho}(x,t;x_f)-\tilde{\rho}(x_f,0;x_f),\quad \tilde{\rho}(x,t;x_f)\gg D,$$

where $\tilde{\rho}(x,t;x_f)$ is given by the solution of the following variational problem:

$$\tilde{\rho}(x,t;x_f)=\min\mathcal{R}[f;x,t;x_f],$$
$$\mathcal{R}[f;x,t;x_f]=\frac{1}{2}\int_{-\infty}^{\infty}d\tilde{t}\,f(\tilde{t})F\left(-i\frac{d}{d\tilde{t}}\right)f(\tilde{t}) \tag{6}$$
$$+\int_{-\infty}^{0}d\tilde{t}\,\lambda(\tilde{t})[\dot{x}+U'(x)-f(\tilde{t})],$$

with boundary conditions

$$f(\pm\infty)=0,\quad \lambda(-\infty)=0,\quad x(-\infty)=x_{\mathrm{eq}},$$
$$x(t)=x,\quad x(0)=x_f. \tag{7}$$

The minimum in (6) is taken over $f(\tilde{t})$, $\lambda(\tilde{t})$, and $x(\tilde{t})$ independently. The second term in $\mathcal{R}$ does not contribute to $\tilde{\rho}(x,t;x_f)$; it has been added to allow for the interrelation (2) between the paths of the noise and of the system. We note that $\lambda(\tilde{t})$, the highest-order derivative in $F(-id/d\tilde{t})f(\tilde{t})$, and (for white-noise-driven systems) $\dot{x}$ are discontinuous at the singled-out moment $\tilde{t}=t$. The inequality in (5) provides the required criterion for the weakness of the noise.

The term $\tilde{\rho}(x_f,0;x_f)$ is subtracted from $\tilde{\rho}(x,t;x_f)$ in (5) to allow for normalization in (1). This term gives [9], to logarithmic accuracy, the stationary distribution of the system, $D\ln p(x_f)\simeq-\tilde{\rho}(x_f,0;x_f)$. Because the value of $\tilde{\rho}(x_f,0;x_f)$ is determined by the most probable path $x_{\mathrm{opt}}(t;x_f)$ that leads to $x_f$ without the additional limitation of passing a given point at a given time, $\tilde{\rho}(x,t;x_f)\geq\tilde{\rho}(x_f,0;x_f)$. It is obvious from (6) and (7) that only for $x$ coinciding with $x_{\mathrm{opt}}(t;x_f)$ (for a given $t$) does $\tilde{\rho}(x,t;x_f)$ equal $\tilde{\rho}(x_f,0;x_f)$. Thus the maximum of the prehistory probability density $p_h(x,t;x_f,0)$ is indeed achieved on the optimal path,

$$\rho(x_{\mathrm{opt}}(t;x_f),t;x_f)=0. \tag{8}$$

Away from the optimal path $p_h$ decreases exponentially, according to (5)-(7); for weak noise, its shape is

Gaussian near the maximum,

$$\rho(x,t;x_f) \simeq [x - x_{\rm opt}(t;x_f)]^2/2\sigma(t;x_f) \,, \qquad |x - x_{\rm opt}(t;x_f)| \ll |x_{\rm eq} - x_f| \,. \tag{9}$$

The dispersion of $p_h(x,t;x_f,0)$ is determined by the value of $D\sigma(t;x_f)$. Equation (9), in addition to the exponential in $p_h(x,t;x_f,0)$, also gives the prefactor so that near the maximum,

$$p_h(x,t;x_f,0) = [D\sigma(t;x_f)/2\pi]^{1/2} \times \exp\{-[x - x_{\rm opt}(t;x_f)]^2/2D\sigma(t;x_f)\} \,. \tag{9a}$$

This is because the function $C$ in (5) is smooth on the scale $\Delta x \sim D^{1/2}$. To analyze $p_h(x,t;x_f,0)$ near the maximum one can therefore replace $C(x,t;x_f)$ by $C(x_{\rm opt}(t;x_f),t;x_f)$; the latter is given by the condition $\int dx\, p_h(x,t;x_f,0) = 1$.

The explicit form of the dispersion parameter $\sigma(t;x_f)$ in (9) can be obtained in some limiting cases. In what follows we give the results for the case of white noise [$F(\omega) = 1$ in (4) and (6)]. The variational problem [(6) and (7)] then reduces to

$$\ddot{x}(\tilde{t}) - U''(x)U'(x) = 0 \quad [F(\omega) = 1] \tag{10}$$

and the optimal path is known to be given by

$$\dot{x}_{\rm opt} = U'(x_{\rm opt}) \,. \tag{10a}$$

The latter equation is also the solution of (10) for $\tilde{t} \le t$ in (6). In the range $0 \ge \tilde{t} > t$ for small $|x - x_{\rm opt}(t;x_f)|$, Eq. (10) can be linearized in $x(\tilde{t}) - x_{\rm opt}(\tilde{t};x_f)$, and after some manipulations one arrives at an expression for $\sigma(t;x_f)$ of the form

$$\sigma(t;x_f) \equiv \tilde{\sigma}(x_{\rm opt}(t;x_f);x_f) \,, \qquad \tilde{\sigma}(x;x_f) = [U'(x)]^2 \int_x^{x_f} dy\, [U'(y)]^{-3} \,. \tag{11}$$

The distribution $p_h(x,t;x_f,0)$ can be seen from (9) and (11) to be extremely sharp for small $|t|$:

$$\rho(x,t;x_f) = \tfrac{1}{2}[x - x_f - U'(x_f)]t^2/|t| \,, \qquad |U''(x_f)(x_f - x)| \ll |U'(x_f)|, \quad |t| \ll [U''(x_{\rm eq})]^{-1} \,. \tag{12}$$

It coincides in shape with the distribution of a particle diffusing in the absence of an external force: The dispersion $D|t|$ increases linearly with the time interval $|t|$. In the opposite case of large $|t|$ the value of $x_{\rm opt}(t;x_f)$ lies close to the equilibrium position $x_{\rm eq}$ where $U'(x)$ vanishes. The dispersion parameter $\sigma(t;x_f)$ as given by (11) is then equal to $[2U''(x_{\rm eq})]^{-1}$, i.e., to the dispersion (divided by $D$) of the stationary distribution $p(x)$.

We note that the dispersion $D\tilde{\sigma}(x;x_f)$ can be a *nonmonotonic* function of $x = x_{\rm opt}(t;x_f)$. The nonmonotonicity is most pronounced when the equilibrium position $x_{\rm eq}$ is that of a metastable state and the point $x_f$ lies close to the local maximum $x_s$ of the potential $U(x)$ [$U'(x_s) = 0$, $U''(x_s) < 0$]. In this case the parameter $\tilde{\sigma}(x;x_f)$ can be written as

$$\tilde{\sigma}(x;x_f) \simeq \sigma_1(x) + \sigma_2(x)(x_f - x_s)^{-2} \,, \qquad |x_f - x_s| \ll |x_{\rm eq} - x_s| \,, \tag{13}$$

$$\sigma_2(x) = -\tfrac{1}{2}[U'(x)]^2[U''(x_s)]^{-3}, \quad \sigma_2 > 0 \,.$$

Note that $\sigma_{1,2}(x)$ depend smoothly on $x_f$ for $|x - x_s| \gg |x_f - x_s|$. For $x$ far from both $x_f$ and $x_{\rm eq}$ the function $\tilde{\sigma}(x;x_f)$ increases sharply as $x_f$ approaches $x_s$. The evolution of $\tilde{\sigma}(x;x_f)$ with varying $x_f$ predicted by (13) for the simple potential

$$U(x) = -\tfrac{1}{2}x^2 + \tfrac{1}{4}x^4 \tag{14}$$

is shown by the curves in Fig. 1; the optimal path predicted by (10a) is plotted in the inset.

These ideas have been tested experimentally with an analog electronic circuit model [13] of (2) and (14), driven by weak quasiwhite noise. Starting with the system in close vicinity to one of the attractors, successive blocks of $x(t)$ time series were digitized with a Nicolet LAB80 data processor, and examined. The moment at which $x(t)$ eventually reached a preset value $x_f$ was noted, and the path by which it had reached that point was recorded. The process was then repeated, so as to build up an ensemble average of the paths leading to $x_f$. A typical example of the resultant prehistory probability density $p_h(x,t;x_f,0)$ is shown in Fig. 2.

It is immediately evident that the distribution is sharply peaked along a certain path $x(t)$, and it seems reason-

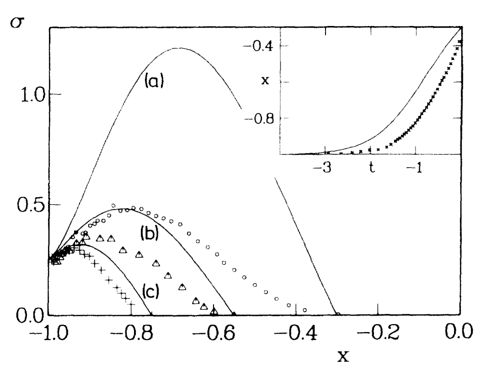


FIG. 1. Dispersion of the prehistory probability density $p_h(x,t;x_f,0)$. The dispersion parameter $\tilde{\sigma}(x;x_f)$ as a function of $x$ measured in the analog experiment (data points) is compared with the theoretical prediction (curves) based on (11) for (a) $x_f = -0.30$, $D = 0.0701$ (circles); (b) $x_f = -0.55$, $D = 0.0265$ (triangles); (c) $x_f = 0.75$, $D = 0.0085$ (plusses). Inset: The optimal path (curve) predicted by (10a) is compared with a path along the ridge (data points) of the experimental distribution for $x_f = -0.30$, $D = 0.0701$.

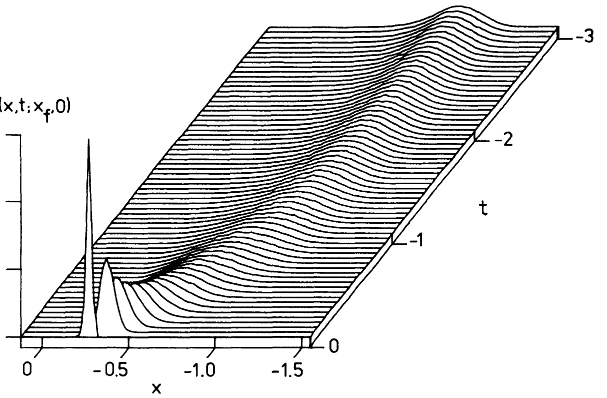


FIG. 2. The prehistory probability density $p_h(x,t;x_f,0)$ for (12) and (14) measured in the analog electronic experiment for a final position $x_f=-0.30$ with $D=0.0701$.

able to associate this path with the expected optimal path for the noise-driven system. Except in the limit of very small $t$, the *shape* of this ridge is well described by (10a), as shown in the inset to Fig. 1. The *slope* of the theoretical curve for $t \rightarrow 0$ differs noticeably from the experiment, however, because (10a) is only valid for $|x-x_s| \gg D^{1/2}$; it is this that causes the displacement between the calculated and experimental curves. A more surprising feature of the distribution in Fig. 2, which had not been predicted in advance of the measurements, is the marked broadening and flattening that occurs in the intermediate range of $x$ between $x_{\mathrm{eq}}$ and $x_f$; but the latter behavior is, of course, very much in accord with the arguments given above. A direct comparison of the measured and calculated dispersion is given for three values of $x_f$ in Fig. 1. Although the theory clearly provides a good description of the general behavior of $\tilde{\sigma}(x)$, it is also evident that the calculated increase in $\tilde{\sigma}$ is significantly larger than in the measurements. One of the reasons is that the theory holds provided $(\tilde{\sigma}D)^{1/2} \ll |x_s - x_{\mathrm{eq}}|$, i.e., $|x_f - x_s| \gg D^{1/2}$, according to (13); whereas, for small $|x_f - x_s|$, the Gaussian approximation (9) is inapplicable.

In conclusion, we have presented the first calculations and measurements of a new statistical quantity, the prehistory probability density, that illuminates and extends the concept of optimal paths and which has enabled us to demonstrate their physical reality. The approach has been shown to work well, and to yield interesting results, in relation to the simple example (2) and (14) driven by white noise; but we emphasize that it is in no sense restricted to systems of this kind. Rather, it should be widely applicable to the investigation of rare events in fluctuating systems in general.

The research has been supported by the Ukrainian Academy of Sciences, by the Science and Engineering Research Council (United Kingdom), and by the European Community.

[a]Present address: Department of Physics and Astronomy, San Francisco State University, 1600 Holloway Avenue, San Francisco, CA 94132.


[1] *Noise in Nonlinear Dynamical Systems*, edited by F. Moss and P. V. E. McClintock (Cambridge Univ. Press, Cambridge, 1989), Vols. 1–3.

[2] M. I. Freidlin and A. D. Ventzel, *Random Perturbations in Dynamical Systems* (Springer-Verlag, Berlin, 1984); E. Ben Jacob, D. J. Bergman, B. J. Matkowsky, and Z. Schuss, Phys. Rev. A **26**, 2805 (1982); R. L. Kautz, Phys. Rev. A **38**, 2066 (1988).

[3] M. I. Dykman and M. A. Krivoglaz, Zh. Eksp. Teor. Fiz. **77**, 60 (1979) [Sov. Phys. JETP **50**, 30 (1979)]; in *Soviet Physics Reviews*, edited by I. M. Khalatnikov (Harwood Academic, New York, 1984), Vol. 5, p. 265.

[4] U. Weiss, Phys. Rev. A **25**, 2444 (1982).

[5] V. F. Baibuz, V. Yu Zitserman, and A. N. Drozdov, Physica (Amsterdam) **127A**, 173 (1984).

[6] J. F. Luciani and A. D. Verga, Europhys. Lett. **4**, 255 (1987); J. Stat. Phys. **50**, 567 (1988).

[7] A. J. Bray and A. J. McKane, Phys. Rev. Lett. **62**, 493 (1989); A. J. McKane, Phys. Rev. A **40**, 4050 (1989); A. J. McKane, H. C. Luckock, and A. J. Bray, Phys. Rev. A **41**, 644 (1990).

[8] H. S. Wio, P. Colet, M. San Miguel, L. Pesquera, and M. A. Rodriguez, Phys. Rev. A **40**, 7312 (1989).

[9] M. I. Dykman, Phys. Rev. A **42**, 2020 (1990); M. I. Dykman, P. V. E. McClintock, N. D. Stein, and N. G. Stocks, Phys. Rev. Lett. **67**, 933 (1991).

[10] L. S. Schulman, Physica (Amsterdam) **177A**, 373 (1991).

[11] R. Phythian, J. Phys. A **10**, 777 (1977); L. Pesquera, M. A. Rodriguez, and E. Santos, Phys. Lett. **94A**, 287 (1983).

[12] R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York, 1965).

[13] L. Fronzoni, in *Noise in Nonlinear Dynamical Systems* (Ref. [1]), Vol. 3, pp. 222–242; P. V. E. McClintock and F. Moss, *ibid.*, pp. 243–274.