Journal of Statistical Physics, Vol. 42, Nos. 1/2, 1986

The First, The Biggest, and Other Such Considerations

Katja Lindenberg[1] **and Bruce J. West**[2]

We investigate the relation between the underlying dynamics of randomly evolving systems and the extrema statistics for such systems. Independent processes, Fokker–Planck processes and Lévy processes are considered.

KEY WORDS: Extremes; first passage times; success; failure; independent processes; Fokker–Planck processes; Lévy processes.

1. INTRODUCTION

1.1. Engineering Applications of Extrema Statistics

The study of the statistics of extremes is of importance in a variety of problems in engineering and applied physics. One of the quantities that arises frequently in this context is the time when a process first reaches a particular large threshold value, i.e., the first passage time to that threshold. An example of a first passage time problem is the time required for a mechanical structure to first reach a critical breaking amplitude due to random external excitations such as wind, ocean waves, earthquakes, etc.[1-12] Another example, encountered in communication theory, is the time at which the fluctuations cause the current or the voltage in a noisy electrical circuit to reach a predetermined critical value, i.e., the so-called "false alarm" problem.[13-15] A related problem is that of the noise-induced extreme values attained by the system variables in a given time. For example, the distribution of the maximum vibrations of a mechanical structure is critical in determining the accumulated fatigue damage within that structure.[3] Also, in maintenance policy decisions the correlation of the occurrence of maxima with failure rates has proven very useful.[16,17]

[1] Department of Chemistry, University of California at San Diego, La Jolla, California 92093.
[2] Center for Studies of Nonlinear Dynamics (affiliated with the University of California, San Diego), La Jolla Institute, 3252 Holiday Court, Suite 208, La Jolla, California 92037.

0022-4715/86/0100-0201$05.00/0

The most extensive analysis of the statistics of extreme values has been done for independent processes. The first successful application of extreme value theory to failure predictions seems to have been made by Weibull[2] with his statistical theory of brittle fracture. Gumbel[1] and others analyzed the extreme values of floods, earthquakes, and other natural phenomena as well as failure rates and other operations research problems. A complete discussion of the asymptotic distribution of extreme events for independent processes has been given by Gumbel.[1]

The analysis of events that are sufficiently rare so as to be treated as independent has heretofore usually been approached as a problem in statistics and the processing of time series. The relation between the statistics of extremal events and the underlying dynamics of the process has not been addressed in most engineering applications. The underlying dynamics necessarily introduces correlations between events and determines the actual statistics of the rare events. In the latter sense the dynamics play an important role even for events that are sufficiently rare as to be uncorrelated.

Linear and nonlinear oscillators excited by white noise provide the basis for the study of extremal properties of both electrical and mechanical systems.[8–11] Numerical methods for determining the first passage probability densities and mean time to failure for a linear oscillator system have been discussed by Crandall, Chandiramani, and Cook.[7] The relation between this density and the maximum absolute value of the vibrations in a time interval $(0, t)$ for such a linear system using a number of analytic approximations has been reviewed by Crandall.[6] Roberts[8,9] has applied the Fokker–Planck equation approach developed by Stratonovich[5] to the determination of the mean first passage time to a critical value of the energy envelope of a lightly damped nonlinear oscillator excited by white noise. In his predictions he utilized a combination of analysis and numerical computation. He has also examined the first passage time to a critical level of the amplitude envelope for a linear oscillator with weak nonlinear damping.[10] This problem was originally formulated by Stratonovich and investigated rigorously by Khasminskii[21] and by Papanicolaou and Kohler.[22] As pointed out by Roberts[10] there has heretofore been no exact, closed-form solution to the first passage problem for even the simplest engineering systems.

There are, of course, many dynamical processes that are not diffusive, including any process that is not local in time and/or space.[23–25] An example is a viscoelastic material with memory, whose modes of failure can therefore not be obtained from a Fokker–Planck description. There seems to be a complete lack of analysis of the extremal properties of such systems in the engineering literature.

In this paper we deal with four areas of interest in engineering applications of extremal distributions. The first area is that of statistics of extremes of independent events, and is one that finds routine implementation in engineering design problems. We review the standard theory and discuss some applications in Section 2. The second area, having to do with dependent statistical processes described by Fokker–Planck equations, is less well known but has appeared in a number of guises in modeling generic processes assumed to be important in engineering applications. We discuss these processes in Section 3, in particular with the view of providing simple analytic expressions for the extrema properties of a Fokker–Planck system. The third area, that of dependent processes described by a Lévy distribution (in space or in time), seems to be virtually unexplored in engineering applications. We discuss such processes in Section 4. The fourth topic, addressed in Section 5, is implicit in many applications. Therein we distinguish between processes that "fail" with certainty at a precise level (e.g., a false alarm that sounds when a current exceeds a preassigned specific value), and those whose "failure" occurs with a given probability at each level (e.g., a material that fails due to the presence of cracks of varying sizes and spatial locations). We summarize our results in Section 6.

1.2. Definitions[18-20,26]

Let $X(t)$ be a continuous random variable that depends on the time t and let $w(x, t \mid x_0)\,dx$ be the probability that $X(t)$ lies in the interval $(x, x+dx)$ without ever having crossed the boundary $x=\eta$ in the time interval $(0, t)$ given the initial state $X(0)=x_0$. We define a random variable $Y(t)$ to be the largest value achieved by $X(t)$ in this time interval,

$$Y(t) \equiv \max\{X(t), 0 \leqslant \tau \leqslant t\} \tag{1.1}$$

The cumulative distribution function defined by

$$F(\eta, t \mid x_0) \equiv \int_{x_L}^{\eta} dx\, w(x, t \mid x_0) \tag{1.2}$$

where x_L is the lower boundary of the state space occupied by the process, has the probabilistic interpretation

$$F(\eta, t \mid x_0) = \text{Prob}\{Y(t) < \eta \mid X(0) = x_0\} \tag{1.3}$$

Alternatively, we can also interpret the cumulative distribution function as

$$F(\eta, t \mid x_0) = \text{Prob}\{T(\eta) > t \mid X(0) = x_0\} \tag{1.4}$$

where $T(\eta)$ is the time when the random variable first crosses $x = \eta$, i.e.,

$$T(\eta) = \min\{\tau \mid X(\tau) = \eta\} \tag{1.5}$$

Two useful definitions obtained from $F(\eta, t \mid x_0)$ are the maxima and first passage time densities

$$\psi(\eta, t \mid x_0)\, d\eta \equiv d\eta \frac{\partial}{\partial \eta} F(\eta, t \mid x_0) = \text{Prob}\{\eta < Y(t) < \eta + d\eta \mid X(0) = x_0\} \tag{1.6}$$

and

$$\varphi(\eta, t \mid x_0)\, dt \equiv -dt \frac{\partial}{\partial t} F(\eta, t \mid x_0) = \text{Prob}\{t < T(\eta) \leqslant t + dt \mid X(0) = x_0\} \tag{1.7}$$

The cumulative distribution function $F(\eta, t \mid x_0)$ as well as the densities (1.6) and (1.7) are Green's functions for processes with arbitrary initial conditions. Averaging over an initial distribution $\rho(x_0)$ gives

$$F(\eta, t) = \int_{x_L}^{\eta} dx_0\, F(\eta, t \mid x_0)\, \rho(x_0) = \text{Prob}[Y(t) < \eta] = \text{Prob}[T(\eta) > t] \tag{1.8}$$

$$\psi(\eta, t) \equiv \int_{x_L}^{\eta} dx_0\, \psi(\eta, t \mid x_0)\, \rho(x_0) = \frac{\partial}{\partial \eta} F(\eta, t) \tag{1.9}$$

and

$$\phi(\eta, t) = \int_{x_L}^{\eta} dx_0\, \varphi(\eta, t \mid x_0)\, \rho(x_0) = -\frac{\partial}{\partial t} F(\eta, t) \tag{1.10}$$

With these distributions we can calculate the moments of the extremal properties. The conditional nth moment of the first passage time distribution is

$$T_n(\eta \mid x_0) = \int_0^{\infty} dt\, t^n \varphi(\eta, t \mid x_0) = n \int_0^{\infty} dt\, t^{n-1} F(\eta, t \mid x_0) \tag{1.11}$$

The conditional moments of the distribution of maxima are

$$Y_n(t \mid x_0) = \int_{x_L}^{\infty} d\eta\, \eta^n \psi(\eta, t \mid x_0) = x_L^n + n \int_{x_L}^{\infty} d\eta\, \eta^{n-1} [1 - F(\eta, t \mid x_0)] \tag{1.12}$$

The extremal moments averaged over an initial distribution are denoted by $T_n(\eta)$ and $Y_n(t)$.

Another function that is often encountered in applications of extrema statistics to data sets is the hazard function[27,28]

$$h(\eta, t) = \frac{\varphi(\eta, t)}{F(\eta, t)} \tag{1.13}$$

This is the probability density that the process will reach η for the first time in the interval $(t, t+dt)$. From the definition (1.10) of the first passage time density it follows that the survival probability is related to the hazard function by

$$F(\eta, t) = \exp\left[-\int_0^t d\tau h(\eta, \tau)\right] \tag{1.14}$$

In practice it is often impossible to associate a single level η with a survival probability. Rather, to a given value of η there corresponds a probability $p(\eta)$ of survival. The average probability of survival is then given by

$$F(t) \equiv \int_0^\infty d\eta\, p(\eta)\, F(\eta, t) \tag{1.15}$$

2. INDEPENDENT PROCESSES

2.1. Methodology

The theory of extremal behavior has been most completely developed for independent random variables.[1] The underlying reason for this has been the notion that events that are *sufficiently rare* are necessarily statistically independent of one another. Thus, although the underlying process may be continuous in time, the statistics of extrema have been traditionally dealt with in terms of variables defined for discrete times. We will therefore state results in terms of discrete time processes and take the continuous time limit where appropriate. We note that this discrete sampling explains why the theory of extrema has traditionally been viewed as a problem in statistical data analysis rather than one in probability theory.

Let the set $\{X_j\}$, $j = 1,..., m$ denote a sequence of m observations of a process. The X_j are taken to be identically and continuously distributed, mutually independent random variables. We can define the cumulative distribution analogous to (1.3) as

$$F(\eta, m) \equiv \text{Prob}\{T(\eta) > m\Delta t\} \tag{2.2}$$

where Δt is the time between measurements and $T(\eta)$ is defined by (1.5) with $\tau \equiv j\Delta t$, and in analogy with (1.7) we can define

$$\Delta t\varphi(\eta, m) = -[F(\eta, m+1) - F(\eta, m)] = \text{Prob}\left\{m < \frac{T(\eta)}{\Delta t} < m+1\right\} \tag{2.3}$$

Clearly the above quantities do not depend on the "initial" state for these independent observations.

Let $P(x)$ be the cumulative distribution function for each of the random variables X_j and let $p(x)$ be the corresponding probability density:

$$P(x) \equiv \text{Prob}\{X_j < x\}, \qquad j = 1, 2, \ldots, m \tag{2.4}$$

$$p(x) = \frac{d}{dx} P(x) \tag{2.5}$$

The probability $P(x)$ must here be viewed as a phenomenological representation of the unspecified underlying dynamical process. In subsequent sections we relate $P(x)$ to the underlying dynamical process. The phenomenological distribution $P(x)$ contains all the information that one needs to determine the extremal statistics of the set $\{X_j\}$. Thus the cumulative distribution (2.2) is simply the joint probability that each observation does not exceed η, i.e.,

$$F(\eta, m) = [P(\eta)]^m \tag{2.6}$$

from which it follows that

$$\Delta t \varphi(\eta, m) = [P(\eta)]^m [1 - P(\eta)] \tag{2.7}$$

The mean first passage time to η in this discrete representation is given by

$$T_1(\eta) = \Delta t \sum_{m=0}^{\infty} m\varphi(\eta, m) = \frac{\Delta t}{1 - P(\eta)} \tag{2.8}$$

Inverting (2.8) to express $P(\eta)$ in terms of $T_1(\eta)$ allows us to rewrite (2.6) as

$$F(\eta, m) = \left[1 - \frac{\Delta t}{T_1(\eta)}\right]^m \tag{2.9}$$

Setting $m = t/\Delta t$ and taking the limit $\Delta t \to 0$, [i.e., the limit of many observations in the time interval $(0, t)$] yields

$$F(\eta, t) \to e^{-t/T_1(\eta)} \tag{2.10}$$

This exponential form for the cumulative distribution is characteristic of independent events.

The functional dependence of the cumulative distribution on η depends, of course, on the particular process as manifest in the η dependence of $T_1(\eta)$, or, equivalently, of $P(\eta)$. It is possible to provide a general classification of the asymptotic (i.e., large η) behavior of the cumulative distribution (2.10). This can be done because the asymptotic behavior is

dominated by the tails of the phenomenological distribution $P(\eta)$ and is insensitive to other details of its structure. There are three general classes of asymptotic distributions. These classes are called "types -I, -II, and -III extreme values." The type-I extreme value distribution requires that $P(\eta)$ decrease at least exponentially with increasing η in the sense that $P(\eta) \leqslant \exp[-f(\eta)]$, where $f(\eta) \leqslant (\ln \eta)^a$ with $a > 1$. Examples of distributions $P(\eta)$ leading to this class of distributions of extrema are: normal, log-normal, gamma, and Weibull. A type-II extreme value distribution occurs when $P(\eta)$ does not possess finite central moments. For example, power-law distributions and Lévy distributions belong in this class. The type-III distributions arise from $P(\eta)$'s that are bounded in the direction of the extreme value (i.e., bounded above for maxima and bounded below for minima). Thus the distributions appropriate for type-I extrema when truncated yield type-III extrema.

An interesting feature of these three types of extrema distributions is that each one is characterized by a universal form. Let

$$F(y) \equiv \lim_{t \to \infty} F(\eta, t) \tag{2.11}$$

where y is defined below for each type of asymptotic distribution. For type-I extreme values one has

$$F(y) = e^{-e^{-y}} \tag{2.12}$$

where

$$y = \alpha_j(\eta - \beta_j) \tag{2.13}$$

and α_j and β_j are defined by the relations

$$P(\beta_j) = 1 - \frac{1}{j} \tag{2.14a}$$

$$\alpha_j = j \left. \frac{dP(x)}{dx} \right|_{x=\beta_j} \tag{2.14b}$$

For type-II extreme values with $y \geqslant 0$,

$$F(y) = \exp[-1/y^\alpha], \qquad \alpha > 0 \tag{2.15}$$

where

$$y = \eta / v_j \tag{2.16}$$

and v_j is the expected largest value in a sample of size j and is given by the relation

$$P(\eta) = 1 - \frac{1}{j}\left(\frac{v_j}{\eta}\right)^\alpha \tag{2.17}$$

The specific value of α depends on the distribution $P(\eta)$ but is independent of j. The third class of extrema is characterized by the form

$$F(y) = \exp[-(-y)^{\alpha}], \qquad \alpha > 0 \tag{2.18}$$

where

$$y = \left(\frac{w-\eta}{w-v_j}\right) \qquad \text{for} \quad -\infty < \eta \leqslant w \tag{2.19}$$

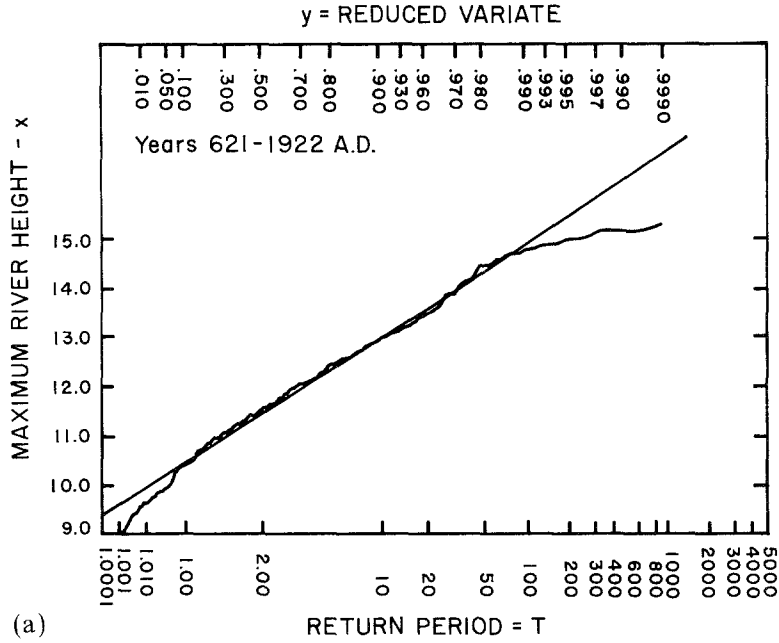
and v_j is again given by (2.17). The j-independent parameter α in (2.18) is defined as the order of the lowest derivative of the probability function that does not vanish at the upper bound $\eta = w$. We note that (2.18) is of the Weibull form.

In the engineering literature one often encounters discussions of lifetimes, failures, and reliability in terms of these extremal types of distribution.[1–17,27–30] The distribution that enters in these discussions is the average survival probability given by (1.15). The variable y in these applications is directly proportional to the time (i.e., to $j \equiv t/\Delta t$). Thus, although one often finds survival probabilities with a type-III distribution (cf. below), one must carefully distinguish whether the distribution is in a variable y as defined in (2.19) or whether the distribution comes about by averaging over η.

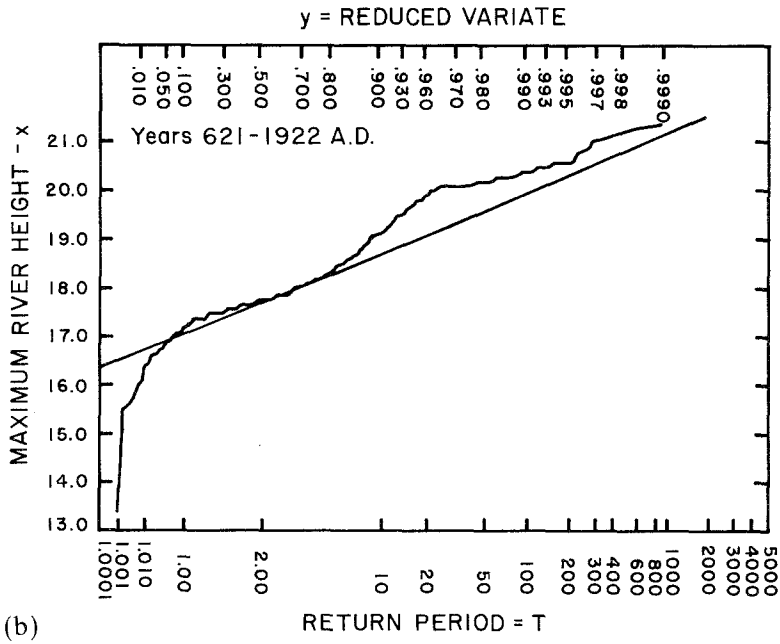
2.2. Some Applications

The utility of the three types of distributions just discussed can be demonstrated by considering a number of data sets from a variety of phenomena. We choose examples that illustrate each of the limiting universal forms.

The pioneer of systematic data analysis from the point of view of extremal types of distributions was Emile Gumbel.[1] He determined that a large variety of meteorological and geophysical phenomena are well described by the limiting form type-I. Examples treated by Gumbel and others[27–31] include daily rainfalls and snowfalls, temperature extremes, wind speed, largest earthquakes, and air pollution levels. Gumbel popularized the notion of constructing specialized graph paper with axes such that $F(y)$ plotted against η yields a straight line. In Fig. 1 we condense over 1300 years of data on the maximum and minimum heights of the Nile River on type-I graph paper.[31] The abscissa is the maximum or minimum height η, while the lower ordinate is the probability of reaching that extreme value η as the number of observations becomes very large. For the minima distribution the straight line corresponds to the distribution



(a)



(b)

Fig. 1. One may use Fig. 1a and 1b to predict the maximum and minimum river heights as follows. Consider a minimum river height of 12 m. By examining the scale in Fig. 1a labeled Return Period, we note that, on the average, every $3\frac{1}{2}$ years the river minimum is 12 m. Correspondingly, the river maximum reaches a value of 20 m every 200 ys. Of course, owing to fluctuations in the river maxima, this latter prediction is less reliable than the former.[31]

function (2.12) with the parameter values $\beta_j = 9.4$ and $\alpha_j = 0.28$, with $j = 1300$. The upper ordinate is the recurrence time between two observations of a given level. Both of these extremal events do yield (at least approximately) a straight line. However, from the figure it is seen that the prediction of droughts (river minimum for each year) is more reliable than that of the flood level (river maximum). It is of course of equal practical importance to be able to accurately predict minimal and maximal levels of water discharge: the former is important for pollution control; both are important for agricultural irrigation.

A more contemporary application of type-I distributions is to the pollution levels in the atmosphere. Figure 2 shows observed SO_2 concentration levels at Long Beach, California for the 19-year period 1956 to 1974.[27] The data were grouped on both a monthly basis and a yearly basis. Both sets of data yield straight lines on Gumbel paper, with the ordinate being the maximum pollution level η and the abscissa the probability of reaching that level. The monthly data has parameters $\alpha_j = 0.115$ and $\beta_j = 14.5$ with $j = 228$; for the yearly data $\alpha_j = 0.081$ and $\beta_j = 31.5$, with $j = 19$. Two interesting features of this analysis should be

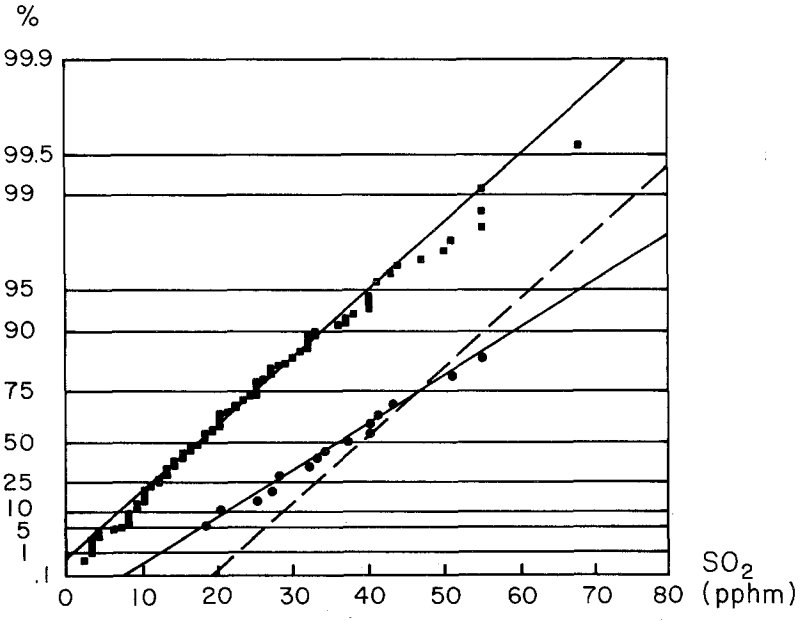


Fig. 2. Observed distribution function of monthly maxima (□) yearly maxima (○) from 1956 to 1975 at Long Beach; fitted double exponential distribution, —— distribution function for yearly maximum extrapolated from monthly maxima.[27]

noted. First, if one used the parameters from the monthly data to estimate the maximal yearly pollution (dashed line), one would overestimate the larger maximum pollution level probabilities and underestimate the low ones. Secondly, we note that the maximum monthly pollution levels are lower than those predicted by the type-I line for small and for large values of η. The effect of this is that the theory "saturates" at lower maximal values than does the data. For example, the theory would predict a very small probability of a 100 pphm concentration of SO_2 because the probability that the maximal pollution is greater than 70 pphm is already in the 0.1 percentile. The data would suggest that these higher concentrations are more likely than this.

As a third and final example of a type-I distribution, we present in Figure 3 the yearly maximum wind speed in London, Ontario for the 23-year period 1939–1961.[27] The straight line has the parameters $\alpha_j = 1/3$ and $\beta_j = 17$, with $j = 23$.

Whereas data fits to type-I distributions are readily available in the engineering literature, it is more difficult to find processes that have been analyzed directly in terms of type-II or type-III distributions. It is not clear whether this situation indicates that the latter types do not give an accurate

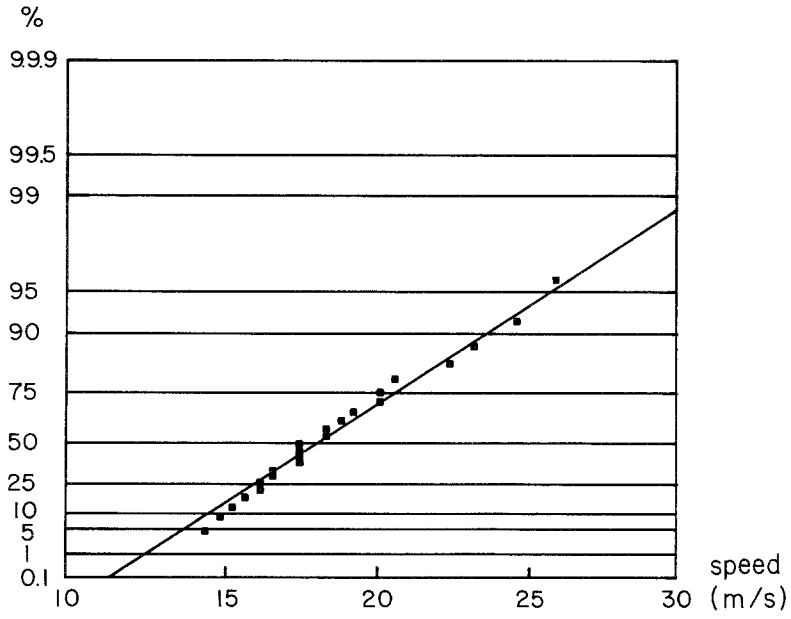


Fig. 3. Observed distribution function of maximum yearly wind speed 1939–1961.[27]

representation of data, or whether it is in fact often not possible to distinguish among fits using the three types of distributions.[32] As an example, consider the data for the probability distribution of the magnitude of earthquakes in the Aleutian Islands region shown in Fig. 4. On the figure it is apparent that both the type-I and the type-III distributions overlay the data, with the second fitting slightly better.

A type-II distribution has been used to fit the annual extreme significant wave-height data for 12 Ocean Station Vessels.[33] This data is shown in Fig. 5, where the *logarithm* of the significant wave height is shown to be described by a type-I distribution. This implies that the significant wave height itself follows a type-II distribution.

One example of the direct application of the type-III (Weibull) distribution occurs in the analysis of the fatigue failure of ductile materials.

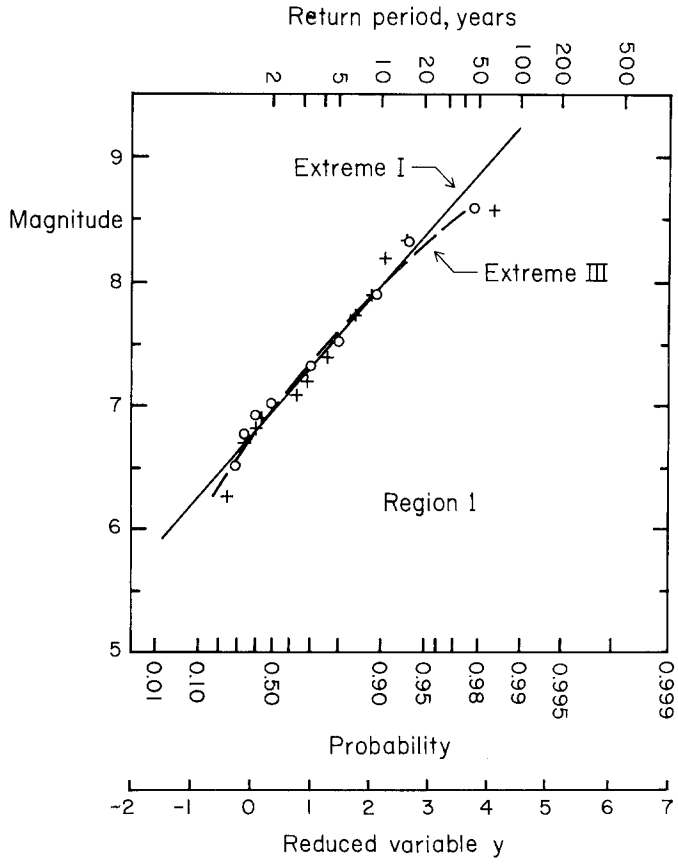


Fig. 4. Magnitude statistics of earthquakes in the Aleutian Islands region.[32]

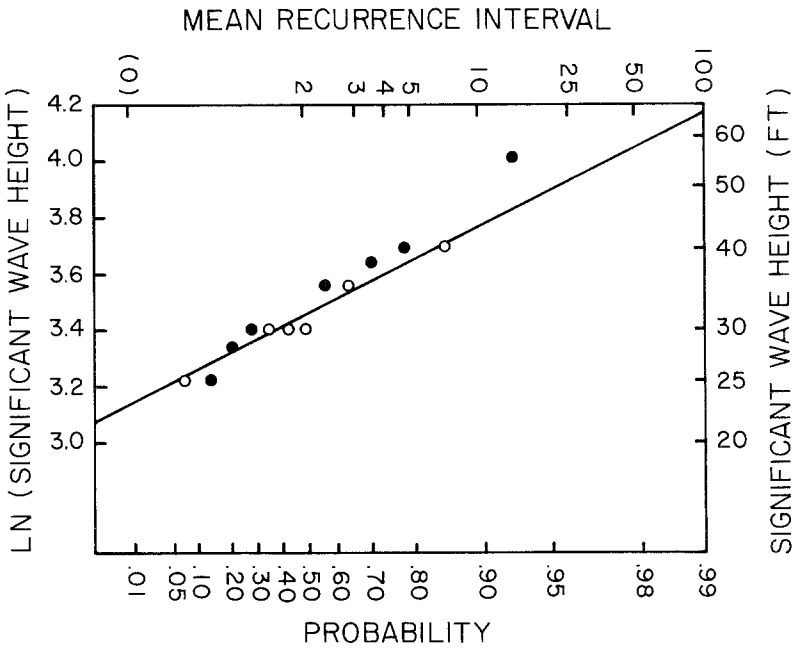


Fig. 5. Probability distribution of significant wave heights. Left ordinate scale: logarithm of significant wave height. Right ordinate scale: significant wave height.[33]

The cumulative probability for failure can be related to the distribution of crack lengths in the material. This distribution is shown in Fig. 6 on type-III probability paper.[34] The abscissa is the square root of the ratio of the crack length to the average crack length.

One often finds the Weibull distribution for averaged survival probabilities as defined in Equation (1.15). For instance, the distribution of failure times of a certain type of thrust bearing under high-temperature conditions is given in terms of the "reliability function"

$$F(t) = e^{-0.00358x^{1.75}}$$

where x is a dimensionless time.[28] A lifetime distribution study of flame nozzles in a particular type of industrial furnace similarly gives

$$F(t) = e^{-0.00178x^{1.8}}$$

where again x is a dimensionless time.[28] Many other such examples can be found in the literature.

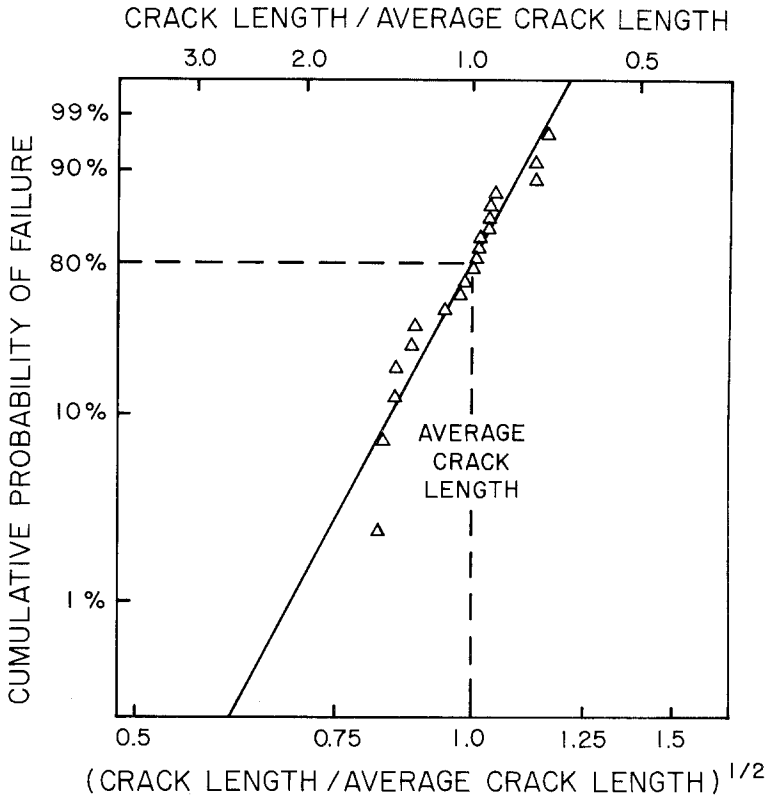


Fig. 6. Size distribution of critical initial defects.[34]

3. FOKKER–PLANCK PROCESSES

In this section we deal with dynamical processes $\mathbf{X}(t)$ that are correlated at different times, i.e., $\mathbf{X}(t)$ and $\mathbf{X}(\tau)$ are not independent even for $t \neq \tau$. We shall consider the "weakest" such dependence, that being a differential Markov process. Such processes can be defined by systems of stochastic rate equations with fluctuations having particular correlation properties to be given subsequently. We further restrict our discussion in this section to processes in which the average change in $\mathbf{X}(t)$ is small when t changes by a small amount Δt, i.e., processes for which

$$E[\mathbf{X}(t+\Delta t)-\mathbf{X}(t)]=0(\Delta t) \tag{3.1a}$$

$$E[\mathbf{X}(t+\Delta t)+\mathbf{X}(t-\Delta t)-2\mathbf{X}(t)]=0(\Delta t) \tag{3.1b}$$

but higher-order differences are of higher order in Δt. Here E denotes the expectation value. Such processes can be described by a differential equation in phase space, i.e., by a Fokker–Planck equation.[35,36]

3.1. Single-Variable Systems

Consider a single-variable dynamical process described by the "Langevin equation"

$$\dot{X} = m_1(X) - \tfrac{1}{2}m_2'(X) + [m_2(X)]^{1/2} f(t) \tag{3.2}$$

Here the dot indicates a time derivative and the prime a derivative with respect to the state variable X. The fluctuating function $f(t)$ is taken to be a zero-centered Gaussian-distributed random variable with correlation function

$$\langle f(t) f(t') \rangle = 2\delta(t - t') \tag{3.3}$$

The brackets $\langle\ \rangle$ denote an average over an ensemble of realizations of the fluctuations $f(t)$. The property (3.3) is a necessary condition for (3.2) to describe a Markov process. The extremal statistics of the process $X(t)$ are, as before, embodied in two types of questions:

1. At what time $T(\eta)$ does $X(t)$ first achieve a preassigned level η?
2. What is the highest value $Y(t)$ achieved by $X(\tau)$ in the preassigned time interval $0 \leqslant \tau \leqslant t$?

One of many possible examples of the application of such models is in the area of population growth. For instance, let $N(t)$ be the instantaneous population whose isolated growth can be described by an equation of the form[37]

$$\frac{dN}{dt} = kNG(N/\vartheta) \tag{3.4}$$

where $G(N/\vartheta)$ is a saturation-inducing growth law which limits the population to a maximum value $N = \vartheta$. A useful form of this function is

$$G(N/\vartheta) = [1 - (N/\vartheta)^{\alpha}]/\alpha \tag{3.5}$$

The choice $\alpha = 0$ leads to the Gompertz equation, while $\alpha = 1$ gives the Verhulst equation, now more commonly referred to as the logistic equation. Equation (3.4) does not account for the variability of the environment as manifest in random fluctuations in, for example, the food

supply, the birth and death rates, immigration and emigration, etc. These effects can be incorporated through a fluctuating growth parameter resulting in the stochastic rate equation[19,37]

$$\frac{dN}{dt} = kNG(N/\vartheta) + \varepsilon N f(t) \tag{3.6}$$

where ε is a measure of the strength of the stochastic rate parameter. The fluctuations in (3.5) allow the population $N(t)$ to exceed ϑ or to go to extinction with finite probabilities. The linear N dependence of the fluctuations insures (not uniquely) that fluctuations do not create a population if none is there and that $N(t)$ remain positive. This behavior is due to the fact that $dN/dt = 0$ when $N(t) = 0$.

The maximum and first passage time statistics are of interest for at least two reasons. First, foreknowledge of the possibility of a large fluctuation in a population and of its most probable magnitude allows for advanced planning. Second, the maxima and first passage times are sensitive probes that allow the determination of the validity of the parameters of a model such as ϑ, α, and ε in Eqs. (3.5) and (3.6).

Equation (3.6) is, of course, of the form (3.2) with the identifications

$$m_1(N) \equiv \frac{kN}{\alpha}\left[1 - \left(\frac{N}{\vartheta}\right)^{\alpha}\right] + \varepsilon^2 N \tag{3.7a}$$

$$m_2(N) = 2\varepsilon^2 N^2 \tag{3.7b}$$

Let us now return to Eq. (3.2). This process has an equivalent phase space description in terms of the conditional probability $W(x, t \mid x_0)\, dx$ that the dynamical variable $X(t)$ is in the interval $(x, x + dx)$ given the initial value $X(0) = x_0$. Because the fluctuations $f(t)$ in (3.2) are Gaussian and delta-correlated, the first and second moments of the process $X(t)$ satisfy (3.1) and the higher-order differences are indeed of higher order in Δt. The evolution of the probability density $W_t \equiv W(x, t \mid x_0)$ is then given [using the Stratonovich interpretation of (3.2)[5,36]] by the Fokker–Planck equation

$$\frac{\partial}{\partial t} W_t = -\frac{\partial}{\partial x}\left[m_1(x)\, W_t\right] + \frac{\partial^2}{\partial x^2}\left[m_2(x)\, W_t\right] \tag{3.8}$$

together with appropriate boundary conditions and the initial condition $W(x, 0 \mid x_0) = \delta(x - x_0)$. In Equation (3.8) the term containing $m_1(x)$ describes the systematic or average evolution of the system, i.e., the "drift," while the second derivative term describes the dispersion about the systematic evolution due to fluctuations. Therefore $m_2(x)$ can be viewed as a "diffusion coefficient" that may be state dependent. In the dynamic

equation (3.2) $m_2(X)$ gives rise to state-dependent fluctuations that have a nonzero average value ($\langle [m_2(X)]^{1/2} f(t) \rangle \neq 0$). The term $[-m_2'(X)/2]$ exactly compensates for this nonvanishing average so that $m_1(X)$ alone is the "drift." In the subsequent analysis we will need the steady state solution of (3.8), given by

$$W_{ss}(x) \equiv \lim_{t \to \infty} W(x, t \mid x_0) = \frac{c}{m_2(x)} \exp\left[\int^x dx' \frac{m_1(x')}{m_2(x')} \right] \tag{3.9}$$

where c is the normalization constant determined by the condition

$$\int_\Omega W_{ss}(x)\, dx = 1 \tag{3.10}$$

and Ω denotes the range of the variable x.

The probability density $w_t \equiv w(x, t \mid x_0)$ appropriate to (3.2) with the additional constraint that $X(\tau)$ has not crossed the value $x = \eta$ for $0 \leqslant \tau \leqslant t$ satisfies Eq. (3.8),

$$\frac{\partial}{\partial t} w_t = -\frac{\partial}{\partial x} [m_1(x)\, w_t] + \frac{\partial^2}{\partial x^2} [m_2(x)\, w_t] \tag{3.11}$$

with the absorbing boundary condition

$$w(\eta, t \mid x_0) = 0 \tag{3.12}$$

and an appropriate lower boundary condition.[38] As for (3.8), the initial condition for (3.11) is $w(x, 0 \mid x_0) = \delta(x - x_0)$.

For processes described by the evolution equation (3.11) the moments (2.7) of the first passage time distribution satisfy a hierarchy of equations that follow directly from the backward Kolmogorov equation[18–20,38]

$$\frac{\partial}{\partial t} w_t = m_1(x_0) \frac{\partial}{\partial x_0} w_t + m_2(x_0) \frac{\partial^2}{\partial x_0^2} w_t \tag{3.13}$$

Multiplying (3.13) by nt^{n-1} and integrating over x from the lower boundary x_L to η and over t from 0 to ∞ yields the exact equation

$$m_2(x_0) \frac{\partial^2}{\partial x_0^2} T_n(\eta \mid x_0) + m_1(x_0) \frac{\partial}{\partial x_0} T_n(\eta \mid x_0) = -nT_{n-1}(\eta \mid x_0) \tag{3.14}$$

with $T_0(\eta \mid x_0) \equiv 1$. The boundary conditions for (3.14) are

$$T_n(\eta \mid \eta) = 0 \tag{3.15a}$$

and

$$\frac{\partial}{\partial x_0} T_n(\eta \mid x_0)|_{x_0 = x_L} = 0 \tag{3.15b}$$

In writing (3.15b) we have assumed the lower boundary to be reflecting. Equation (3.14) can be solved exactly for the mean first passage time $T_1(\eta \mid x_0)$:

$$T_1(\eta \mid x_0) = \int_{x_0}^{\eta} dz \, \frac{[\int_{x_L}^{z} w_{ss}(y)\, dy]}{m_2(z)\, w_{ss}(z)} \tag{3.16}$$

For large η one can evaluate the integral (3.16) asymptotically, thereby obtaining a closed-form expression for the mean first passage time.[18–20]

To illustrate the utility of (3.16) let us return to our population example (3.5). With the identifications (3.7) in (3.9) we obtain for $\alpha > 0$ the steady state distribution

$$W_{ss}(N) = \frac{\alpha}{\Gamma(a)} a^a \left(\frac{N}{\vartheta}\right)^{a\alpha - 1} \exp\left[-a\left(\frac{N}{\vartheta}\right)^{\alpha}\right] \tag{3.17}$$

where

$$a \equiv k/\alpha^2 \varepsilon^2 \tag{3.18}$$

and where N is the phase space level of the population. The mean first passage time to $N = \eta$ can then be obtained with (3.17) and (3.7b) in (3.16). For $\eta \gg N_0$ and $\eta \gg \vartheta$ the dominant contribution to the mean first passage time is from the neighborhood of $z \sim \eta$. For these values of z the numerator of the z integrand in (3.16) is approximately unity and[19]

$$\begin{aligned} T_1(\eta) &\sim \Gamma(a)\, a^{1-a} \frac{\alpha^2}{\vartheta} \int^{\eta} dN \exp\left[a\left(\frac{N}{\vartheta}\right)^{\alpha} - (\alpha a + 1) \ln\left(\frac{N}{\vartheta}\right)\right] \\ &\sim \Gamma(a)\, a^{1-a} \alpha^2 \left(\frac{\eta}{\vartheta}\right) \frac{\exp\left[a\left(\frac{\eta}{\vartheta}\right)^{\alpha} - (\alpha a + 1) \ln\left(\frac{\eta}{\vartheta}\right)\right]}{\left[a\alpha\left(\frac{\eta}{\vartheta}\right)^{\alpha} - (\alpha a + 1)\right]} \\ &\quad \times \left[1 - \frac{(\alpha - 1)}{a\alpha\left(\frac{\eta}{\vartheta}\right)^{\alpha}} + \cdots\right] \end{aligned} \tag{3.19}$$

and is independent of N_0. The dominant contribution to the mean first passage time is

$$T_1(\eta) \sim \Gamma(a)\, a^{-a} \alpha \left(\frac{\eta}{\vartheta}\right)^{-a(a+1)} \exp[a(\eta/\vartheta)^{\alpha}] \tag{3.20}$$

For $\alpha = 0$ (Gompertz growth law) the steady state distribution is

$$W_{ss}(N) = \left(\frac{k}{2\pi\varepsilon^2}\right)^{1/2} \left(\frac{\vartheta}{N}\right) \exp\left[-k\left(\ln\frac{N}{\vartheta}\right)^2 \Big/ 2\varepsilon^2\right] \tag{3.21}$$

and the asymptotic mean first passage time then is[19]

$$T_1(\eta) \sim \left(\frac{2\pi k}{\varepsilon^2}\right)^{1/2} \frac{\exp\left[k\left(\ln\frac{\eta}{\vartheta}\right)^2 \Big/ 2\varepsilon^2\right]}{\left[\frac{k}{\varepsilon^2}\ln\left(\frac{\eta}{\vartheta}\right) - 1\right]} \left[1 - \frac{\varepsilon^2}{k\ln\left(\frac{\eta}{\vartheta}\right)} + \cdots\right] \tag{3.22}$$

The extremal properties of the process (3.2) are of course not exhausted by the mean first passage time or even the higher passage time moments [cf. Eq. (3.14)]. Of equal interest is the distribution of the maximum (minimum) value achieved by $X(\tau)$ in a given time interval $0 \leqslant \tau \leqslant t$. To investigate the properties of this distribution we again examine the evolution equation (3.11). Equation (3.11) is a parabolic partial differential equation whose solution can be expressed as the eigenfunction expansion

$$w(x, t \mid x_0) = \sum_{n=0}^{\infty} W_{ss}(x) \frac{U_n(x)\, U_n(x_0)}{N_n} e^{-\lambda_n t} \tag{3.23}$$

where

$$\delta_{mn} N_n = \int_{x_L}^{\eta} dx\; W_{ss}(x)\, U_m(x)\, U_n(x) \tag{3.24}$$

The eigenfunctions $U_n(x)$ and the eigenvalues λ_n satisfy the differential equation

$$m_2(x)\frac{d^2 U_n(x)}{dx^2} + m_1(x)\frac{dU_n(x)}{dx} + \lambda_n U_n(x) = 0 \tag{3.25}$$

subject to the boundary condition

$$W_{ss}(\eta)\, U_n(\eta) = 0 \tag{3.26}$$

and an appropriate boundary condition at x_L. Equation (3.25) can be rewritten in the self-adjoint form

$$\frac{d}{dx}\left[\sigma(x)\frac{dU_n(x)}{dx}\right] + \lambda_n W_{ss}(x)\, U_n(x) = 0 \tag{3.27}$$

where

$$\sigma(x) \equiv m_2(x)\, W_{ss}(x) \tag{3.28}$$

If $\sigma(x)$ and $W_{ss}(x)$ are positive functions with a finite number of poles, then the differential equation (3.27) is of the Sturm–Liouville type and the eigenvalues λ_n are real, nondegenerate, nonnegative, and may be ordered such that $\lambda_0 < \lambda_1 < \lambda_2 < \cdots$.

The eigenfunction–eigenvalue problem (3.27) is in general quite difficult to solve. However, it turns out that if one is only interested in the occurrence of rare events, e.g., in the extremum $Y(t)$ at long times t or the first passage time $T(\eta)$ for large η, then under appropriate initial conditions it is sufficient to evaluate only the lowest eigenvalue $\lambda_0(\eta)$.[18] No other eigenvalues or eigenfunctions are needed. The initial distribution $w(x_0)$ that yields this result is one that is near the steady state, i.e., $w(x_0) \simeq W_{ss}(x_0)$ for $x_L \leqslant x_0 < \eta$. Here η must be sufficiently large that

$$\int_{\eta}^{\infty} dx\, W_{ss}(x) \ll 1 \tag{3.29}$$

The cumulative distribution of extrema (1.5) is then to a good approximation given by

$$F(\eta, t) \simeq e^{-t\lambda_0(\eta)} \tag{3.30}$$

for all times t. It then follows from (1.8) that the moments of the first passage time distribution are given by

$$T_n(\eta) \simeq n!/[\lambda_0(\eta)]^n \tag{3.31}$$

$$= n![T_1(\eta)]^n \tag{3.32}$$

The remarkable outcome of this analysis is that the extrema distribution (3.30) is therefore completely specified by the mean first passage time:

$$F(\eta, t) \simeq e^{-t/T_1(\eta)} \tag{3.33}$$

Since the mean first passage time is completely determined through Eq. (3.16), one need not determine $\lambda_0(\eta)$ by solving the eigenfunction–eigenvalue problem (3.25).

The *form* of (3.33) is noteworthy in that it is identical to that obtained in Eq. (2.15) for independent random variables. This identity reflects the

mutual independence of rare events if the underlying process is correlated over times much shorter than the mean first passage time $T_1(\eta)$. It should be noted that although the same in form, the present theory of *dependent* processes enables us to establish the relation between $T_1(\eta)$ and the underlying dynamic structure of the process. In our earlier discussion of *independent* events $T_1(\eta)$ had to be taken as a functional representation of the data. Thus, provided that the dynamical model (3.8) is applicable, the present results enables us to *interpret* the data.

It is more difficult to make general statements about the asymptotic behavior of maximum moments than about first passage time moments. This situation is a result of our lack of knowledge of the cumulative distribution $F(\eta, t)$ for all η: the form (3.33) is only valid for large η (and all t). It is nevertheless possible to conclude that for $t \geqslant t_c$ (defined below) the mean maximum and higher moments are approximately given by[18–20]

$$Y_n(t) \simeq x_L^n + n \int_{x_L}^{\infty} d\eta\, \eta^{n-1}[1 - e^{-t/T_1(\eta)}], \qquad t > t_c \tag{3.34}$$

The time t_c is given by the relation

$$t_c \simeq T_1(\eta_0) \tag{3.35}$$

where η_0 is the smallest η for which the condition (3.29) is valid. The approximation (3.34) is acceptable because the dominant contributions to the integral come from values of η for which (3.33) is valid. For a fixed value of $t > t_c$, the accuracy of (3.34) decreases with increasing order n. It should also be observed that for large t the mean maximum $Y_1(t)$ is a good measure of the distribution of maxima. The variance of the maximum distribution tends to zero:

$$\lim_{t \to \infty} \frac{Y_2(t) - Y_1^2(t)}{Y_1^2(t)} \to 0 \tag{3.36}$$

This is in marked contrast with the first passage time distribution which has the variance

$$\lim_{\eta \to \infty} \frac{T_2(\eta) - T_1^2(\eta)}{T_1^2(\eta)} \to 1 \tag{3.37}$$

Thus the first passage time distribution is quite broad and $T_1(\eta)$ is not a good measure of it.

Returning to our population example, we calculate the mean maximum population in the time interval $(0, t)$ using (3.20) and (3.22) in

(3.34) with $n=1$. To actually carry out the η integration in (3.34) it is convenient to consider the time derivative[19]

$$\dot{Y}_1(t) \simeq \int_{x_L}^{\infty} d\eta \, \frac{1}{T_1(\eta)} \, e^{-t/T_1(\eta)} \tag{3.38}$$

The integrand of (3.38) for $\alpha > 0$ is sharply peaked about its maximum time-dependent value $\eta_{\max}$ because $T_1(\eta)$ is a rapidly monotonically decreasing function [cf. (3.20)] and therefore we integrate (3.38) using Laplace's method. The value of $\eta_{\max}$ is the solution of the equation $t = T_1(\eta_{\max})$. We then obtain

$$Y_1(t) \sim \frac{(2\pi)^{1/2} \vartheta}{e\alpha a^{2-1/\alpha}} (\ln kt)^{1/\alpha} \tag{3.39}$$

to within a constant. A similar analysis for $\alpha = 0$ gives

$$Y_1(t) \sim \frac{(2\pi)^{1/2} k\vartheta}{e\varepsilon^2} \exp[(2\varepsilon^2 \ln kt/k)^{1/2}] \tag{3.40}$$

It is important to emphasize that the extremal measures $T_n(\eta)$ and $Y_n(t)$ are sensitive to the forms of the drift and diffusion functions in the dynamic equations. Thus these measures can serve as probes into the underlying dynamic structure of the system.

3.2. Two-Variable Systems

The approach reviewed in Section 3.1 is not directly generalizable to systems described by more than one dynamical variable. However, in weakly damped physical systems there are nearly conserved quantities whose variation in time is slower than that of the dynamical variables. If a single such quantity can be identified, e.g., the total energy or the amplitude of an oscillation, then the dynamical description can be reduced to one for this single degree of freedom and the previous approach can be implemented.[5–11,19,20]

As an example for this procedure we consider a lightly damped anharmonic oscillator driven by a stochastic force $f(t)$:

$$\dot{X} = P \tag{3.41}$$

$$\dot{P} = -\lambda Q(X, P) - G(X) + \varepsilon f(t) \tag{3.42}$$

Here $X(t)$ is the displacement of the oscillator, $\lambda Q(X, P)$ and $G(X)$ are damping and restoring forces, respectively, and λ is a damping parameter. The fluctuations $f(t)$ are defined as before.

Such an oscillator is used as a model system in a large variety of engineering applications. Some of these are: wind-generated ocean waves, where $X(t)$ is the vertical displacement of the ocean surface and $f(t)$ represents the incoherent pressure fluctuations from the wind,[39,40] the horizontal structural response $X(t)$ to ground motion during earthquakes in which $f(t)$ models the ground motion,[41–43] the electrical voltage $X(t)$ across a circuit driven by stray voltage fluctuations $f(t)$,[15,44] ocean wave platform and/or pile design where $X(t)$ is the displacement response of the structure to the loading $f(t)$ produced by the water waves.[4]

The dynamic system (3.42) has the equivalent phase space description

$$\frac{\partial}{\partial t} W_t = \left\{ -p \frac{\partial}{\partial x} + \frac{\partial}{\partial p} [\lambda Q(x, p) + G(x)] + \varepsilon^2 \frac{\partial^2}{\partial p^2} \right\} W_t \tag{3.43}$$

where $W_t\, dx dp \equiv W(x, p, t \mid x_0, p_0)\, dx\, dp$ is the probability that $X(t)$ and $P(t)$ lie in the phase space interval $(x, p; x + dx, p + dp)$ given the initial state $X(0) = x_0$ and $P(0) = p_0$.

In many applications, such as those mentioned above, one is interested in the earliest time t when the absolute displacement $|X(t)|$ of the oscillator attains a threshold value x_c, i.e., the first passage time to a specified boundary. The difficulties associated with obtaining the first passage time statistics of $X(t)$ for an oscillator described by (3.41) and (3.42) are well known.[7–11,19,20,45] It has so far not been possible to obtain first passage time statistics since one cannot construct a well-posed boundary value problem for Eq. (3.43). Traditionally this difficulty is circumvented by constructing an approximate equation for a single "slow" variable such as the energy or the amplitude envelope of the oscillator.[5] The first passage time to attain a critical energy E_c or critical amplitude A_c is then studied by solving this approximate single variable Fokker–Planck equation. The first passage time to E_c or to A_c provides a lower bound for the first passage time to the absolute displacement x_c.

Several authors[5,7–11,19,20] have obtained approximate Fokker–Planck equations for the energy and amplitude envelopes of lightly damped nonlinear oscillators. The equation for the energy envelope probability density $W(E, t \mid E_0) \equiv W(E, t)$ is

$$\frac{\partial}{\partial t} W(E, t) = \left\{ \frac{\partial}{\partial E} \left[\lambda \frac{\psi(E)}{\varphi'(E)} - \varepsilon^2 \right] + \varepsilon^2 \frac{\partial^2}{\partial E^2} \frac{\varphi(E)}{\varphi'(E)} \right\} W(E, t) \tag{3.44}$$

where the energy E is defined through the relation

$$E = \tfrac{1}{2}p^2 + u(x) \tag{3.45}$$

$u(x)$ is the potential energy of the oscillator,

$$u(x) = \int^x G(x')\, dx' \tag{3.46}$$

and the other functions appearing in (3.44) are

$$\varphi'(E) = \frac{1}{2}\int_{E>u(x)} \frac{dx}{[E-u(x)]^{1/2}} \tag{3.47}$$

$$\varphi(E) = \int_{E>u(x)} dx[E-u(x)]^{1/2} \tag{3.48}$$

$$\psi(E) = \frac{1}{\sqrt{2}}\int_{E>u(x)} dx\, Q(x, \{2[E-u(x)]\}^{1/2}) \tag{3.49}$$

The integrations in (3.47)–(3.49) are carried out over all x for which the square root $[E-u(x)]^{1/2}$ is real. The Fokker–Planck equation for the amplitude envelope is often used when $u(x)$ is a harmonic potential, i.e., $G(x) = \omega_0^2 x$, and is given by

$$\frac{\partial}{\partial t} W(A, t) = \left\{\frac{\partial}{\partial A}\left[\lambda \frac{C(A)}{\omega_0} - \frac{\varepsilon^2}{2A\omega_0^2}\right] + \frac{\varepsilon^2}{2\omega_0^2}\frac{\partial^2}{\partial A^2}\right\} W(A, t) \tag{3.50a}$$

where

$$A = (x^2 + p^2/\omega_0^2)^{1/2} \tag{3.50b}$$

and $C(A)$ is defined in terms of the nonlinear damping as

$$C(A) = -\frac{1}{2\pi}\int_0^{2\pi} Q(A\cos\vartheta, -\omega_0 A \sin\vartheta)\sin\vartheta d\vartheta \tag{3.51}$$

Of course $E = \omega_0^2 A^2$ and hence $(2\omega_0^2 A)^{-1} W(A, t) = W(E, t)$.

The theory developed in the previous section can now be brought to bear by properly identifying the functions m_1 and m_2 appearing in (3.8). For instance, let us take the restoring potential to have a single stable fixed point at $x = 0$ and to have the form

$$G(x) = k\,|x|^m \operatorname{sgn} x \tag{3.52a}$$

and let us take the damping to be linear,

$$Q(x, p) = p \tag{3.52b}$$

Then

$$m_1(E) = \varepsilon^2 - 2\frac{(m+1)}{(m+3)}\lambda E \tag{3.53}$$

$$m_2(E) = 4\frac{(m+1)}{(m+3)}\varepsilon^2 E \tag{3.54}$$

Using these forms in (3.9) and (3.16) with $x_L = 0$ and performing an asymptotic expansion for large E_c we find[20]

$$T_1(E_c) \sim \frac{1}{\lambda}\Gamma\left(\frac{3m+5}{2m+2}\right)\left(\frac{\varepsilon^2}{\lambda E_c}\right)^{(m+3)/(2m+2)} e^{\lambda E_c/\varepsilon^2} \times \left[1 + \frac{(m+3)\,\varepsilon^2}{(2m+2)\,\lambda E_c} + \cdots\right] \tag{3.55}$$

The first passage time to energy E_c provides a lower bound for the first passage time to the absolute displacement x_c. These threshold values are related by Eq. (3.45) with $p = 0$, i.e.,

$$E_c = \frac{k}{m+1}x_c^{m+1} \tag{3.56}$$

so that

$$T_1^L(x_c) \sim \frac{1}{\lambda}\Gamma\left(\frac{3m+5}{2m+2}\right)\left[\frac{\varepsilon^2(m+1)}{\lambda k}\right]^{(m+3)/(2m+2)} \frac{e^{[\lambda k/\varepsilon^2(m+1)]x_c^{m+1}}}{x_c^{(m+3)/2}} \times \left[1 + \frac{\varepsilon^2}{2\lambda}\frac{(m+3)(m+1)}{k^2 x_c} + \cdots\right] \tag{3.57}$$

Let us now present the leading terms for the mean first passage time results just obtained for specific restoring forces. For a linear oscillator ($m = 1$) (3.57) reduces to

$$T_1^L(x_c) \sim \frac{2\varepsilon^2}{\lambda^2 k}\frac{e^{\lambda k x_c^2/2\varepsilon^2}}{x_c^2} \tag{3.58}$$

where $k^{1/2} \equiv \omega_0$ is the natural frequency of the linear oscillator. Equation (3.58) is a well-known linear oscillator result and blends smoothly into exact numerical results when $k^{1/2}x_c^2/\varepsilon^2 \gtrsim 3.5$ for $\lambda/k^{1/2} = 0.02$ and 0.16. For an anharmonic oscillator with $m = 3$ we obtain

$$T_1^L(x_c) \sim \frac{\Gamma(7/4)}{\lambda^2}\left(\frac{4\varepsilon^2}{\lambda k}\right)^{3/4}\frac{e^{\lambda k x_c^4/4\varepsilon^2}}{x_c^3} \tag{3.59}$$

The heavy dots in Figure 7 indicate the numerical simulations of Roberts for $\ln \lambda T_1^L(x_c)$ versus $\tilde{x}_c^4$ where

$$\tilde{x}_c = \left(\frac{\lambda k}{4\varepsilon^2}\right)^{1/4} \left(\frac{\Gamma(1/4)}{\Gamma(3/4)}\right)^{1/2} x_c \tag{3.60}$$

The curve is calculated from Eq. (3.59). Our curve begins at the value $\tilde{x}_c = 2.40$ ($\tilde{x}_c^4 = 33.18$) because the contributions of the corrections to Eq. (3.59) are smaller than 2% beyond this value and decrease as $\tilde{x}_c^{-4}$. The agreement with Roberts' numerical simulations is clear, ours lying slightly below his. Note that in the range of values of $\tilde{x}_c$ in this figure all terms explicitly shown in (3.59) must be retained. Thus this expression provides an analytic representation for the mean first passage time in the region where computer costs for doing numerical simulations become prohibitive.

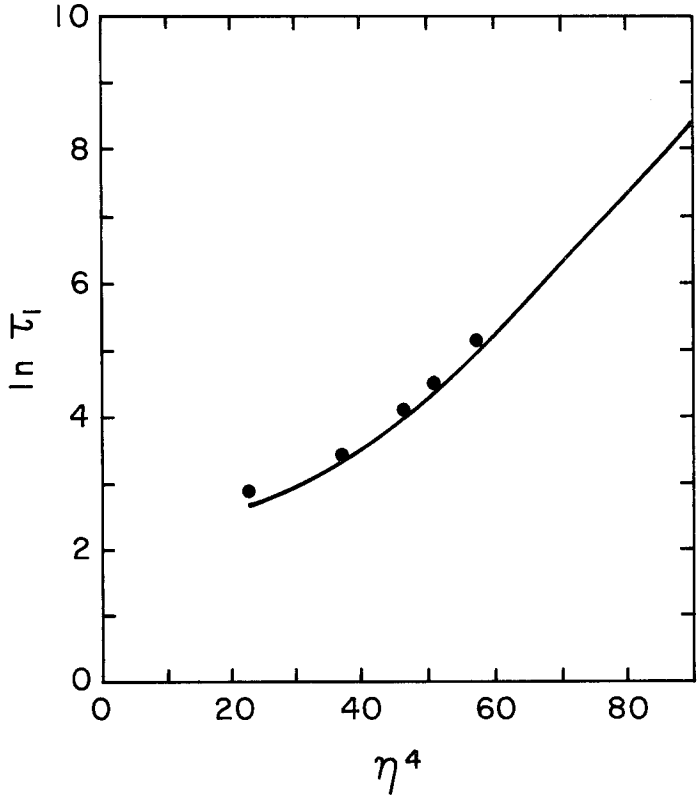


Fig. 7. Mean first passage time for power law oscillator with $m = 3$. Solid curve: Eq. (3.59); circles: Roberts' numerical simulation.[20]

We can now use the mean first passage time result (3.55) to calculate the maximum energy envelope moments from (3.34). Using Laplace's method for evaluating the integral for $\dot{e}_1(t)$ [cf. (3.38)] we obtain for the mean maximum $e_1(t)$ the asymptotic result[20]

$$e_1(t) \sim \frac{\varepsilon^2}{\lambda} \ln \lambda t \tag{3.61}$$

A similar calculation of the mean square maximum energy $e_2(t)$ yields

$$e_2(t) \sim \left(\frac{\varepsilon^2}{\lambda}\right)^2 (\ln \lambda t)^2 \tag{3.62}$$

indicating that the variance vanishes as $t \to \infty$. The distribution of the maximum energy is therefore narrow at long times, consistent with the results of the previous section. Equation (3.61) provides an upper bound for the mean maximum displacement $\hat{x}_1(t)$:

$$\hat{x}_1(t) \lesssim \left(\frac{\lambda k}{(m+1)\,\varepsilon^2} \ln(\lambda t)\right)^{1/(m+1)} \tag{3.63}$$

We have carried out a similar analysis for the linearly damped Duffing oscillator, i.e.,[20]

$$G(x) = \omega_0^2 x + k x^3 \tag{3.64}$$

In this case we obtain the mean first passage time for the energy envelope

$$T_1(E_c) \sim \frac{1}{\lambda} \frac{\Gamma(\mu+1)}{(\lambda E_c/\varepsilon^2)^\mu} \frac{e^{\lambda E_c/\varepsilon^2}}{(1-\varepsilon^2\mu/\lambda E_c)} \tag{3.65}$$

where μ is a parameter that lies in the interval $1/2 < \mu \leqslant 1$ whose precise value is determined by k and ω_0. This expression is of the same form as (3.55) but here μ replaces the m-dependent numerical factors. The comparison of (3.65) with Roberts' numerical simulations are shown in Fig. 8.

The third nonlinear oscillator that we consider using these techniques has a harmonic restoring potential,[20]

$$G(x) = \omega_0^2 x \tag{3.66a}$$

but a nonlinear dissipation

$$Q(x, p) = p[1 + K|p|^m] \tag{3.66b}$$

The analysis is done using the amplitude envelope (3.50b) rather than the energy envelope (3.45) so that direct comparisons with Roberts' numerical

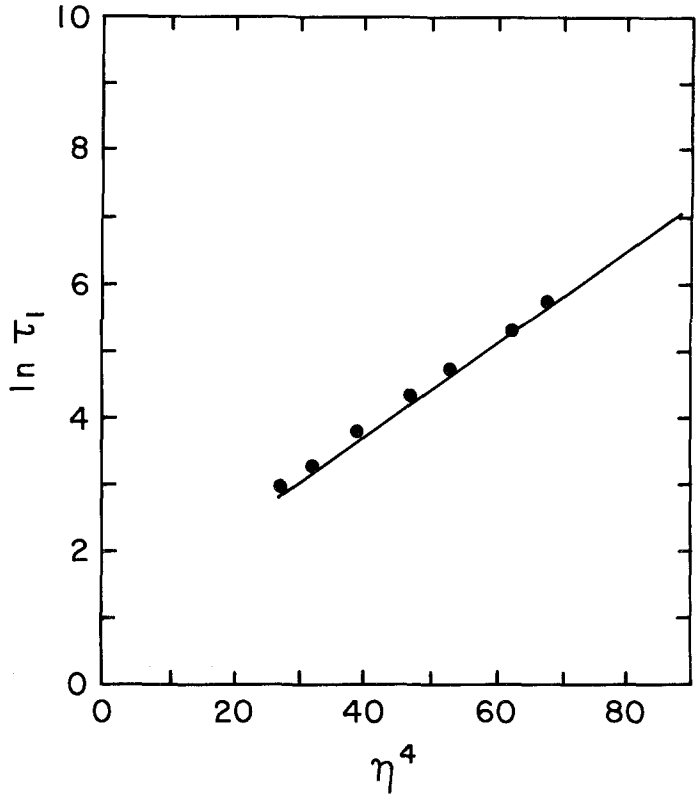


Fig. 8. Mean first passage time for the Duffing oscillator with $kD/\lambda\omega_0^4 = 0.2$ and $\mu = 1$. Solid curve: Eq. (3.65); circles: Roberts' numerical simulation.[20]

calculations can be made. The function $C(A)$ defined in (3.51) is then given by

$$C(A) = \frac{\omega_0 A}{2} + \frac{2^{m+2}}{\pi} K(\omega_0 A)^{m+1} \left[\Gamma\left(\frac{m+3}{2}\right) \right]^2 \Big/ \Gamma(m+3) \tag{3.67}$$

The steady state solution to the Fokker–Planck equation (3.50a) is

$$W_{ss}(A) = \gamma_m A \exp\left[-\frac{1}{2}\frac{\lambda\omega_0^2}{\varepsilon^2} A^2 - \beta_m A^{m+2} \right] \tag{3.68}$$

where γ_m is the normalization constant and

$$\beta_m = \lambda^{\varepsilon^2} \omega_0^{m+2} \frac{2^{m+3}}{\pi(m+2)} K \left[\Gamma\left(\frac{m+3}{2}\right) \right]^2 \Big/ \Gamma(m+3) \tag{3.69}$$

The mean first passage time to a large amplitude A_c is then found to be[20]

$$T_1(A_c) \sim \frac{2}{\lambda \gamma_m}\left[\left(\frac{\lambda\omega_0^2}{\varepsilon^2}\right)^{1/2} A_c^2 - 1 + \beta_m(m+2)\, A_c^{m+2}\right]^{-1} \times \exp\left[\frac{\lambda}{2\varepsilon^2}\omega_0^2 A_c^2 + \beta_m A_c^{m+2}\right] \tag{3.70}$$

For $m=1$ we compare the result (3.70) with the numerical simulations of Roberts in Fig. 9. The two results are indistinguishable and hence only a single line is shown.

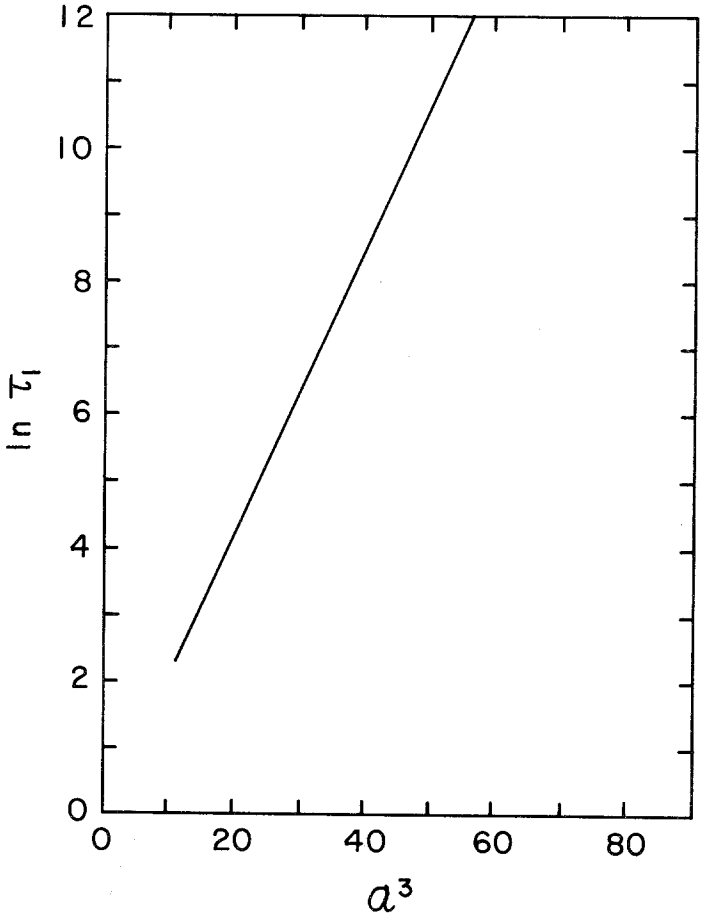


Fig. 9. Mean first passage time for the nonlinearly damped oscillator, Eq. (3.70), with $m=1$, $\gamma_1 = 1.414$, and $\beta_1 = 0.1414$.[20]

The mean first passage time results can, as before, be used to calculate maxima moments. In particular, the mean maximum amplitude $a_1(t)$ and the mean square maximum $a_2(t)$ are found to be

$$a_1(t) \sim \left(\frac{\varepsilon^2}{\lambda\omega_0^2}\right)^{1/2} (\ln \lambda t)^{1/(m+2)} \tag{3.71}$$

and

$$a_2(t) \sim \left(\frac{\varepsilon^2}{\lambda\omega_0^2}\right) (\ln \lambda t)^{2/(m+2)} \tag{3.72}$$

Again the variance $[a_2(t) - a_1^2(t)]/a_1^2(t)$ vanishes asymptotically.

3.3. Limiting Distribution

In engineering applications which utilize the assumption of independence of observations there is an implicit ambiguity in the specification of the minimum length of the time interval between observations that will insure independence. The *minimum* length depends on the underlying dynamics. For the Fokker–Planck processes considered in this section we have in effect specified this measure: it is given by the time $t_c = T_1(\eta_0)$ defined in Eq. (3.35).

A second question that is of importance in engineering applications is the asymptotic form of the distribution of extrema. We have seen in Section 2 that one obtains one of the three classes of asymptotes, type-I, -II, or -III, depending on the functional form of $T_1(\eta)$. For the Markov processes considered in this section one finds that the requirement of normalizability of the probability density W_{ss} constrains the dependence of T_1 on η through (3.16). In particular, for large η we have argued that

$$T_1(\eta) \sim \int^{\eta} dz \, \frac{W_{ss}^{-1}(z)}{m_2(z)} = c \int^{\eta} dz \, e^{\int^z dx \, m_1(x)/m_2(x)} \tag{3.73}$$

Since normalizability of $W_{ss}(x)$ requires that $m_1(x)/m_2(x) \gtrsim 0[(\ln x)^a]$ with $a > 0$, (3.73) implies that $T_1(\eta) \sim \exp[f(\eta)]$ where $f(\eta) \gtrsim (\ln \eta)^a$. This is precisely the condition that leads to a type-I asymptotic maxima distribution. We therefore conclude that (2.12) is the appropriate limiting distribution for Fokker–Planck processes with the $P(\eta)$ that enters (2.14) related to $T_1(\eta)$ calculated in this section via (2.8).

4. LÉVY PROCESSES

We have so far restricted our attention to dependent processes that are local in space and in time, i.e., to "continuous stochastic processes." In this section we consider the extrema statistics of a class of processes that are "discontinuous," i.e., nonlocal in space or in time. This class was first extensively studied by P. Lévy and so we refer to its members as Lévy processes.[46]

The evolution of the continuous process of the previous section was describable by means of a partial differential equation for the probability density $W(x, t|x_0)$ [cf. Eq. (3.8)]. It is not possible to construct a differential representation for a Lévy process. Rather, an integral representation for it becomes necessary. To specify the form of the processes to be considered, let us define a random variable Y that depends on an index z. The process is assumed to satisfy the chain condition[25]

$$W(y, z) = \int_{-\infty}^{\infty} dy'\, W(y - y', z - z')\, W(y', z') \tag{4.1}$$

where $W(y, z)$ is the probability that $Y(z) - Y(0) = y$. Processes that satisfy the chain condition (4.1) are Markovian. The characteristic function $\varphi(u, z)$ is defined as the Fourier transform of the probability density, i.e.,

$$\varphi(u, z) = \int_{-\infty}^{\infty} dy\, e^{iuy} W(y, z) \tag{4.2}$$

For processes satisfying the chain rule (4.1), $\varphi(u, z)$ obeys the product rule

$$\varphi(u, z) = \varphi(u, z - z')\, \varphi(u, z') \tag{4.3}$$

and is therefore an infinitely divisible stable distribution. The most general form of $\varphi(u, z)$ for such distributions was obtained by Lévy,[46] and Khinchine and Lévy.[47] One can verify by direct substitution that the form

$$\varphi(u, z) = e^{-bz|u|^{\mu}} \tag{4.4}$$

satisfies the product rule (4.3) and is the form of the Lévy distribution for a symmetric process if the parameters μ and b obey the restrictions $b \geqslant 0$ and $0 \leqslant \mu \leqslant 2$.

One of the important properties of Lévy distributions is that, except for $\mu = 2$, the distributions $W(y, z)$ do not possess finite y moments of all orders. This is most directly seen from the fact that for $z > 0$ and $\mu < 2$[25]

$$\lim_{y \to \infty} W(y, z) \sim \mu b z \Gamma(\mu) \sin(\pi\mu/2)/\pi |y|^{\mu + 1} \tag{4.5}$$

Thus, all v moments defined by

$$\langle |y|^v \rangle = \int_{-\infty}^{\infty} dy\, |y|^v\, W(y, z) \tag{4.6}$$

are finite for $v < \mu$ and are infinite for $v \geqslant \mu$. In particular, the variance is infinite. This behavior is in marked contrast to that of the solutions of the Fokker–Planck equations discussed in the preceding section. It is worthy of note that power-law distributions of the form (4.5) lead to type-II extremal distributions for independent random variables [cf. (2.17) and (2.15)].

Even though Lévy distributions are characterized by an apparently simple equation [cf. (4.4)], there are a number of difficulties in understanding the system-specific implications of these distributions. Firstly, the probability density $W(y, z)$ cannot be evaluated in closed form except for special choices of the parameter μ.[25] Apart from the familiar Gaussian ($\mu = 2$) and the Cauchy ($\mu = 1$) cases, only a very few other one-dimensional distributions have been constructed. The second difficulty is that it is not possible to construct simple evolution equations of the Fokker–Planck-type for the probability density, i.e., equations involving $\partial W/\partial z$. Such first derivative equations in general contain integral operators in the y variable. For example, when $\mu \neq 2$ the evolution equation can be shown to be of the form[23]

$$\frac{\partial}{\partial z} W(y, z) = \frac{b}{\pi} \sin(\pi\mu/2)\, \Gamma(\mu + 1) \int_{-\infty}^{\infty} dy' \frac{W(y', z)}{|y - y'|^{\mu+1}} \tag{4.7}$$

This equation has the form of a master equation with a long-range interaction. Expansion of (4.7) in a Kramers–Moyal series leads to derivative terms of all orders that cannot be approximated by a truncation at second order.

A number of additional interesting properties of Lévy distributions are worth mentioning. From the inverse Fourier transform of (4.4),

$$W(y, z) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-bz|u|^\mu} e^{-iuy}\, du \tag{4.8}$$

one immediately verifies the scaling relation

$$W(\beta^{1/\mu} y, \beta z) = \beta^{-1/\mu} W(y, z) \tag{4.9}$$

This relation implies that if the process $Y(z)$ is a random variable with probability density $W(y, z)$ then the two random variables $Y(\beta z)$ and $\beta^{1/\mu} Y(z)$ have the same distribution. This scaling relation establishes that

irregularities are generated at each scale in a statistically identical manner. This scaling is a type of self-similarity and is one of the defining properties of a *fractal* process.[48]

Another property of interest is the probability that $Y(z)$ is outside some interval $(-y, y)$ at fixed z:

$$\text{Prob}(|Y(z)| > y) \sim \frac{\text{const}}{y^{\mu}} \qquad \text{as} \quad y \to \infty \tag{4.10}$$

which is obtained by integration of (4.5) and is a hyperbolic distribution. Such distributions preserve self-similarity and have trajectories with fractal dimensionality D. Thus the Lévy process with exponent μ has a trajectory $Y(z)$ versus z of fractal dimensionality $D = 2 - 1/\mu$ (if $\mu < 1/2$ then $D \equiv 0$). Fractal processes in Euclidean dimension 1 and $D < 1$ (i.e., $\mu < 1$) are not space-filling, i.e., in order to maintain the scaling property such a fractal process can only occupy y space in clustered or localized patches.[23,25,48]

4.1. Fractal "Spatial" Processes

Let us consider processes for which we identify the variable z with physical time ($z = t$). We are then interested in calculating the first passage time distribution of the random variable $Y(z) \equiv X(t)$ to a prescribed level η and the distribution of maximum values attained by $X(\tau)$ in a time interval $0 \leqslant \tau \leqslant t$. A dynamical description of this process is given by the stochastic differential equation

$$\dot{X}(t) = f(t) \tag{4.11}$$

where $f(t)$ is a δ-correlated Lévy process. Unfortunately, the techniques developed for Fokker–Planck processes are not applicable except for $\mu = 2$, where (4.8) describes a diffusive process [i.e., one with $m_1(x) = 0$ and $m_2(x) = \text{const}$]. (Note that the homogeneity of the process in x implies the absence of a "potential" so that there is no drift.) The reason for this failure is that the processes considered here with $\mu < 2$ involve long-range transitions and therefore may bypass any localized absorbing barrier without ever encountering it. The first passage time distribution to η can therefore not be represented as a solution to a boundary value problem.

No one has yet devised an exact method for determining the extremal properties of Lévy processes. Since $X(t)$ is (by construction) a symmetric variable we will consider absolute extremal properties, e.g., the first passage time to $|x| = \eta$ and the absolute maximum value of $X(\tau)$ in the time interval $0 \leqslant \tau \leqslant t$. The definitions of Section 1.2 are then modified by the replacement of the lower limits x_L by $-\eta$. One can estimate such proper-

ties via a scaling argument and a sequence of approximations that improve the coefficients of the scaled quantity. The simplest approximation procedure for calculating the mean first passage time $T_1(\eta)$ rests on the assumption that one can replace the (conditional) probability $w(x, t)$ appearing in (1.2) by $W(x, t)$ for a Lévy process. With this approximation the mean first passage time to η is given by[23]

$$T_1(\eta) = \int_0^\infty F(\eta, t)\, dt \simeq \int_0^\infty dt \int_{-\eta}^{\eta} W(x, t)\, dx \tag{4.12}$$

Introducing the scaled variables $\hat{x} = x/\eta$, $\hat{k} = u\eta$ and $\hat{t} = bt/\eta^\mu$ and substituting (4.8) into (4.12) yields

$$T_1(\eta) \simeq \frac{\eta^\mu}{b}\left[\frac{1}{2\pi}\int_0^\infty d\hat{t} \int_{-1}^{1} d\hat{x} e^{-i\hat{k}\hat{x} - \hat{t}|\hat{k}|^\mu}\right] \tag{4.13a}$$

$$\sim C_\mu \eta^\mu \tag{4.13b}$$

Thus $T_1(\eta)$ scales as η^μ according to this approximation. However, the coefficient C_μ obtained from (4.13a) is finite only when $\mu < 1$ as can be seen by examining the singularities in the integrand of (4.13a) after integrating over $\hat{t}$. The integral is thus finite only for "transient" processes (i.e., those which have a zero probability of return to the origin) because the probability of return after leaving the interval $(-\eta, \eta)$ is small in this case. For $\mu > 1$ the process is "persistent" (i.e., has a finite probability of return to the origin). The estimate provided by (4.12) then leads to a divergent coefficient C_μ. The most stringent test of this method in the range $0 \leqslant \mu \leqslant 1$ is the boundary $\mu = 1$ (the least transient processes in this range). At this boundary one finds $C_1 = \pi$, i.e.,

$$T_1(\eta) = \frac{\pi}{b}\eta \tag{4.14}$$

The next level of approximation is to replace $F(\eta, t)$ in (4.12) by a probability

$$F(\eta, t) \simeq \int_{-\eta}^{\eta} w(x, t) \tag{4.15}$$

where $w(x, t)$ is the probability density in the presence of absorbing points within an interval dx around $+\eta$ and around $-\eta$. This eliminates at least those realizations that actually impinge on these intervals from returning into the range $|x| < \eta$. For $0 \leqslant \mu \leqslant 1$ we expect the mean first passage time obtained from (4.15) to be close to that obtained from (4.13). For per-

sistent processes ($1 < \mu \leqslant 2$) where (4.13) fails we expect that (4.15) will provide a finite estimate of C_μ. The replacement (4.15) is of course exact for $\mu = 2$.

The technical problem which remains is to determine the probability density in the presence of boundaries, $w(x, t)$, when a differential description is not appropriate. A technique which has proven to be successful is the method of images.[24,25,49] In this method "image densities" are introduced in such a way that $w(\eta, t) = w(-\eta, t) = 0$ at all times. One requires an infinite number of images and is therefore led to a solution in the form of an infinite series. When this series is substituted into (4.15) one obtains[23,25]

$$F(\eta, t) \simeq \frac{4}{\pi} \sum_{l=0}^{\infty} \frac{(-1)^l}{(2l+1)} \exp\left\{ -\left[\frac{(2l+1)\pi}{2\eta} \right]^\mu bt \right\} \tag{4.16}$$

The mean first passage time is then given by (4.13b) with

$$C_\mu = \frac{2^{2+\mu}}{2\pi^{1+\mu}} \sum_{l=0}^{\infty} \frac{(-1)^l}{(2l+1)^{1+\mu}} \tag{4.17}$$

For $\mu = 1$ and for $\mu = 2$ this series can be evaluated exactly. For the former one obtains[23]

$$T_1(\eta) = \eta \left[\frac{8G}{b\pi^2} \right] \tag{4.18}$$

where G is the Catalan constant 0.915956 .... The coefficient of η in (4.18) is about one quarter of that in (4.14). The *exact* mean first passage time for $\mu = 1$ has been calculated by Kac and Pollard[50] and is given by

$$T_1(\eta) = \eta / b \tag{4.19}$$

For $\mu = 2$ (4.17) yields

$$T_1(\eta) = \eta^2 / 2b \tag{4.20}$$

which is of course the exact result for a diffusion process.

The approximations given above suggest that further improvement in the estimation of the first passage time can be obtained by considering larger absorbing intervals at the edge of the boundary. This absorbing interval would further suppress the reentry of the process into the interval $|x| < \eta$.

As in the previous section, we can use the mean first passage time to calculate the asymptotic form of the absolute maxima moments by sub-

stituting either (4.16) or the even cruder approximation made in (4.12) into (1.12) with x_L replaced by zero. With (4.16) one obtains the fractional moment of order α as

$$Y_\alpha(t) = A_\alpha t^{\alpha/\mu} \tag{4.21}$$

provided $\alpha < \mu$. The constant A_α is given by the series

$$A_\alpha = \frac{4\alpha\pi^{\mu-1}b^{\alpha/\mu}}{\Gamma(\alpha/\nu)} \int_0^\infty d\hat{\eta}\, \hat{\eta}^{\alpha-1} \sum_{l=0}^{\infty} \frac{(-1)^l(2l+1)^{\mu-1}}{(2\hat{\eta})^\mu + [\mu(2l+1)]^\mu} \tag{4.22}$$

where $\hat{\eta} = \eta/(tb)^{1/\mu}$. If $\alpha \geqslant \mu$ this integral diverges and so does the moment $Y_\alpha(t)$. In particular, all integer moments beyond the first diverge for a Lévy process, and the absolute maximum is finite only for $\mu > 1$.

4.2. Fractal "Temporal" Processes

Now we consider processes for which we identify the variable y with times ($y = t$). Whereas in the previous section the process unfolded continuously in time but could make discontinuous or long-range jumps in "space," here the process unfolds continuously in space but erratically or intermittently in time. This intermittency is often introduced by considering the state space ($z = x$) to be discrete and by defining a waiting distribution $\psi(t)$ for the time t between transitions among the discrete states. If $\psi(t)$ is a distribution whose first moment diverges, then the process is temporally fractal. For instance, consider a spatially diffusive process (as an appropriate limit of the discrete process) with intermittent transitions described by a waiting time distribution whose Laplace transform is

$$\psi(s) \equiv \int_0^\infty e^{-st}\psi(t)\, dt = \left[1 + \left(\frac{s}{\lambda}\right)^{2\mu}\right]^{-1} \tag{4.23}$$

where λ is a rate parameter and $0 < \mu < 1/2$. It has been shown[24,52] that the characteristic function $\hat{\varphi}(\eta, \omega)$ for the first passage time distribution to a level $|x| = \eta$ for this process is given by $\varphi(\eta, t)$ [cf. (1.10)]

$$\hat{\varphi}(\eta, \omega) = e^{-b\eta|\omega|^\mu} \tag{4.24}$$

where b is given by $\cos(\pi\mu/2)/D_\mu^{1/2}$ and D_μ is a generalized diffusion coefficient. This characteristic function has the form (4.4) and therefore the first passage time distribution obeys the chain condition (4.1):

$$\varphi(\eta, t) = \int_0^\infty dt'\, \varphi(\eta - \eta', t - t')\, \varphi(\eta', t') \tag{4.25}$$

The equation that describes the evolution of $W(x, t)$ corresponding to the process with this first passage time distribution is[24]

$$W(x, t) = \delta(x) + \frac{D_\mu}{\Gamma(2\mu)} \frac{\partial^2}{\partial x^2} \int_0^t dt' \frac{W(x, t')}{|t - t'|^{1-2\mu}} \tag{4.26}$$

with $0 < \mu < 1/2$.

The mean first passage time to $|x| = \eta$ is evaluated using (4.24). From (4.11) one sees that $T_1(\eta)$ is the derivative of (4.24) with respect to ω evaluated at $\omega = 0$. Since $\mu < 1/2$, this derivative diverges and therefore the mean first passage time is infinite. This divergence is a consequence of the temporal clustering of the process. The intermittency of the process implies long periods of quiescence that dominate the extremal moments.

5. DISTRIBUTED FAILURE (OR SUCCESS) LEVELS

In Sections 2–4 we have established that for processes that are Markov in time the asymptotic cumulative distribution or survival probability $F(\eta, t)$ has the form

$$F(\eta, t) = \sum_{l=0}^{\infty} c_l e^{-\lambda_l(\eta) t} \tag{5.1}$$

In many applications the failure (or success) can indeed be characterized by a single value of η. Thus, for instance, a false alarm will be sounded at this preassigned level; a river will flood when its height exceeds the embankment and a wave will break when its slope exceeds a critical value. In other applications a distribution of values for η may be more representative of the behavior of an ensemble of systems. Thus, for example, a mechanical structure may have a range of possible failure modes each with an appropriate failure probability. The response of individuals to levels of toxins in the environment is similarly distributed. In these latter examples one is interested in the average survival probability

$$F(t) \equiv \int_0^\infty d\eta\, p(\eta)\, F(\eta, t) \tag{5.2}$$

and the corresponding net mean first passage time

$$T_1 \equiv \int_0^\infty d\eta\, p(\eta)\, T_1(\eta) \tag{5.3}$$

where $p(\eta)$ is the probability of failure (or success) when the process

reaches the interval $(\eta, \eta + d\eta)$. A commonly chosen $p(\eta)$ is the Poisson form

$$p(\eta) = 2\Gamma e^{-2\Gamma\eta} \tag{5.4}$$

Another example is the (restricted) Gaussian

$$p(\eta) = 4(\gamma/\pi)^{1/2} e^{-4\gamma\eta^2} \tag{5.5}$$

Let us first consider the average survival probability for the spatially fractal process of Section 4.1. There we showed that the η dependence of the eigenvalue $\lambda_l(\eta)$ is of the form [cf. (4.16)]

$$\lambda_l(\eta) = A_l \eta^{-\mu} \tag{5.6}$$

where μ is the Lévy exponent. The integrals that must be performed to calculate $F(t)$ are of the form

$$I_l(t) = C_l \int_0^\infty d\eta\, p(\eta)\, e^{-tA_l\eta^{-\mu}} \tag{5.7}$$

and thus

$$F(t) = \sum_{l=0}^{\infty} I_l(t) \tag{5.8}$$

One can evaluate (5.7) asymptotically using Laplace's method. For a Poisson distribution of intervals [cf. (5.4)] one obtains the survival probability

$$F(t) \sim ((2\Gamma)^\mu \mu bt)^{1/(2\mu+2)} \left(\frac{\pi}{2(\mu+1)}\right)^{1/2} \sum_{l=0}^{\infty} C_l A_l^{1/(2\mu+2)} \times \exp\left\{-(\mu+1)\left(\frac{2\Gamma A_l^{1/\mu}}{\mu}\right)^{\mu/(\mu+1)} (bt)^{1/(\mu+1)}\right\} \tag{5.9}$$

Note that each term in the sum has the same time dependence. At long times we can approximate (5.9) by the $l=0$ term:

$$F(t) \sim \left(\frac{8}{\mu+1}\right)^{1/2} \left(\frac{\Gamma^\mu}{\pi}\right)^{1/(2\mu+2)} (bt)^{1/(2\mu+2)} \times \exp\left[-(\mu+1)\left(\frac{\Gamma\pi}{\mu}\right)^{\mu/(\mu+1)} (bt)^{1/(\mu+1)}\right] = \gamma_\mu t^{1/(2\mu+2)} e^{-\beta_\mu t^{1/(1+\mu)}} \tag{5.10}$$

For $\mu = 2$, as pointed out earlier, the single trapping points at $\pm\eta$ completely prevent returns from outside the interval and therefore (5.10) is then exact (aside from the quadrature evaluation):

$$F(t)|_{\mu=2} \sim \left(\frac{8}{3\pi^{1/3}}\right)^{1/2} (\Gamma^2 bt)^{1/6} \exp[-3(\Gamma\pi/2)^{2/3} t^{1/3}] \tag{5.11}$$

Other analyses using approximate approaches relying on the distribution of spans in an unbounded diffusive process lead to survival probabilities $F(t) \sim ct^{1/2} \exp(-bt^{1/3})$. The coefficient b in those analyses is smaller by a factor of $(2)^{1/3}$ than that in (5.11). In addition, whereas our prefactor increases as $t^{1/6}$, theirs increases at $t^{1/2}$. The physical reason why (5.11) gives a lower survival probability than previously obtained is that our initial value is confined to a given interval whereas in some applications one initializes the system in any of an ensemble of intervals.

For $\mu < 2$ we again point out that the expression (5.1) for the survival probability for a fixed η is approximate, but that the approximation seems to be a good one (at least for $1 \leqslant \mu < 2$). In this sense, (5.10) is then also an approximation that one expects to be quite accurate. We note that the form (5.10) differs from the Weibull distribution $F(t) \sim \exp(-t^{\alpha})$ usually assumed in failure studies.

Let us next consider the average survival probability when the distribution of intervals is Gaussian [cf. (5.5)]. Now we obtain

$$F(t) \sim \frac{1}{(\mu+2)^{1/2}} \sum_{l=0}^{\infty} c_l \times \exp\left\{-[(4\gamma)^{\mu/2} A_l]^{2/(\mu+2)} \left[\left(\frac{\mu}{2}\right)^{2/(\mu+2)} + \left(\frac{2}{\mu}\right)^{\mu/(\mu+2)}\right] (bt)^{2/(\mu+2)}\right\} \tag{5.12}$$

The leading term of (5.12) at long times is again the $l = 0$ term:

$$F(t) \sim \frac{4}{\pi(\mu+2)^{1/2}} \times \exp\left\{-(\gamma\pi^2)^{\mu/(\mu+2)} \left[\left(\frac{\mu}{2}\right)^{2/(\mu+2)} + \left(\frac{2}{\mu}\right)^{\mu/(\mu+2)}\right] (bt)^{2/(\mu+2)}\right\}$$

$$\sim \alpha'_{\mu} \exp(-\beta'_{\mu} t^{2/(\mu+2)}) \tag{5.13}$$

For $\mu = 2$ the entire series (5.12) can in fact be summed exactly, with the result

$$F(t) = \frac{1}{4(\gamma\pi^3)^{1/2}} \operatorname{sech}[2\pi(\gamma bt)^{1/2}] \tanh[2\pi(\gamma bt)^{1/2}] \tag{5.14a}$$

$$\underset{t \to \infty}{\sim} \frac{1}{2(\gamma\pi^3)^{1/2}} e^{-2\pi(\gamma bt)^{1/2}} \tag{5.14b}$$

which agrees with (5.13) when $\mu = 2$. Note that (5.13) has the form of a Weibull distribution.

Now let us consider a process describable by a Fokker–Planck equation. For some choices of the drift $m_1(x)$ and diffusion $m_2(x)$, the mean first passage time has the form $T_1(\eta) \sim \eta^\mu$ [cf. (5.6)] but unlike a fractal process μ here is greater than 2. For example, if $m_2(x)$ is a constant and $m_1(x) \sim x/(a^2 + x^2)$ then $T_1(\eta) \sim \eta^3$. For these cases the results (5.10) and (5.13) for Poisson and Gaussian interval distributions, respectively, are valid with the appropriate μ. We stress again that for such diffusive processes the survival time formulation in terms of absorbing boundaries is exact. For most Fokker–Planck processes the mean first passage time increases more rapidly than algebraic and, correspondingly, the survival probability at a given time is increased. As an example let us consider the average survival probability for a linear dissipative process, i.e., one with $m_2(x) = D$ and $m_1(x) = -\lambda x$. The mean first passage time to $\pm\eta$ has the asymptotic form

$$T_1(\eta) \sim \frac{\sqrt{\pi}}{2} \left(\frac{2D}{\lambda}\right)^{1/2} \frac{e^{\lambda\eta^2/2D}}{\eta} \tag{5.15}$$

With $F(\eta, t) \sim \exp[-t/T_1(\eta)]$ in (5.2) we obtain the following averge survival probability when $p(\eta)$ has the Poisson form (5.4):

$$F(t) \sim \frac{\exp\{-2\Gamma'[\ln(2\lambda t/\sqrt{\pi}\,\Gamma')]^{1/2}\}}{[\ln(2\lambda t/\sqrt{\pi}\,\Gamma')]^{1/4}} \tag{5.16}$$

where $\Gamma' \equiv \Gamma(2D/\lambda)^{1/2}$. When $p(\eta)$ is Gaussian as in (5.5), then

$$F(t) \sim \frac{t^{-4\gamma'}}{[\ln(\lambda t/4\gamma'\sqrt{\pi})]^{1/2}} \tag{5.17}$$

where $\gamma' \equiv 2D\gamma/\lambda$. We thus see that the survival probabilities of Fokker–Planck processes decay much more slowly (essentially algebraically) than do those for the Lévy processes.

6. SUMMARY

Let us summarize the major points of the preceding sections:

1. For statistically independent processes, extrema can be categorized by three types of distributions. Each of these extrema distributions is characterized by parameters determined from phenonomenological probability distributions deduced from the data. For many purposes this formulation of extremal statistics is adequate.

2. One may wish to understand the relation of the phenomenological probability distribution to the underlying dynamical process. For diffusive processes this relation is well established and provides analytic asymptotic expressions for the survival probability. These expressions rely entirely on knowledge of the steady-state properties of the system.

3. For nondiffusive processes the relation between the underlying dynamics and the extrema distribution is less well known. We provide approximate results that appear to be satisfactory in those cases where a basis for comparison is available. For nondiffusive processes one needs to know the dynamical behavior (and not just the steady state properties) of the system in order to arrive at these results.

4. One must distinguish between survival probabilities when failure occurs at a preset level from those that involve a probability of failure at each level. In the former case the survival probability depends on the level, whereas in the latter an average over levels is performed. This averaging procedure can profoundly influence the time dependence of the survival probability.

ACKNOWLEDGMENT

This research was supported in part by NSF Grant Nos. ATM83-10672 and ATM83-10673 and in part by DARPA contract DAAG 29-85-K-0246.

REFERENCES

1. E. J. Gumbel, *Statistics of Extremes* (Columbia University Press, New York, 1958).
2. W. Weibull, *Ingenioervetenskapsacad. Handl.* **151**:15–60 (1939).
3. V. V. Bolotin, *Statistical Methods in Structural Mechanics* (translated by S. Aroni), (Holden–Day, San Francisco, 1969).
4. G. Moe and S. H. Crandall, Journal of Mechanical Design, Transactions of the American Society of Mechanical Engineers Paper No. 77-DET-82 (1977).
5. R. L. Stratonovich, *Topics in the Theory of Random Noise* Vol. 1 (Gordon and Breach, New York, 1967), pp. 115–119.
6. S. H. Crandall, *J. Sound Vib.* **12**:285–299 (1970).
7. S. H. Crandall, K. L. Chandiramani, and R. G. Cook, *J. Appl. Mech.* **33**:532–538 (1966).

8. J. B. Roberts, *J. Eng. Mech. Div. Am. Soc. Civil Engineers* **102**:851–866 (1976).
9. J. B. Roberts, *Stochastic Problems in Dynamics*, B. L. Clarkson, ed. (Pitman, London, 1977).
10. J. B. Roberts, *J. Appl. Mech.* **45**:175–180, (1978).
11. J. B. Roberts, *J. Sound Vib.* **46**:1–14, (1976).
12. S. H. Crandall and W. D. Mark, *Random Vibrations in Mechanical Systems* (Academic Press, New York, 1963).
13. D. Middleton, *An Introduction to Statistical Communication Theory* (McGraw-Hill, New York, 1960).
14. S. O. Rice, in *Noise and Stochastic Processes*, N. Wax, ed. (Dover, New York, 1954).
15. W. D. Mark, *Proc. Inst. Elect. Electron. Eng.* **54**:316–317 (1966).
16. R. E. Barlow, F. Proschan, and L. C. Hunter, *Mathematical Theory of Reliability*, SIAM Series in Applied Mathematics (John Wiley and Sons, New York, 1965).
17. D. R. Cox, *Renewal Theory* (Methuen, London, 1962).
18. K. Lindenberg, K. E. Shuler, J. Freeman, and T. J. Lie, *J. Stat. Phys.* **12**:217–251 (1975).
19. K. Lindenberg, and V. Seshadri, *J. Chem. Phys.* **71**:4075–4084 (1979).
20. V. Seshadri, B. J. West, and K. Lindenberg, *J. Sound Vib.* **68**:553 (1980).
21. R. Z. Khasminskii, *Theory Prob. Applications* **11**:390 (1966).
22. G. C. Papanicolaou and W. Kohler, *Commun. Pure Appl. Math.* **27**:641–668, (1974).
23. V. Seshadri, and B. J. West, *Proc. Natl. Acad. Sci. USA* **79**:4501 (1982).
24. V. Balakrishnan and M. Khantha, *Pramána* **21**:187 (1983); V. Balakrishnan, preprint.
25. E. W. Montroll and B. J. West, in *Fluctuation Phenomena*, E. W. Montroll and J. L. Lebowitz, eds. (North-Holland, Amsterdam, 1979).
26. A. T. Bharucha-Reid, *Elements of the Theory of Markov Processes and Their Applications* (McGraw-Hill, New York, 1960).
27. M. R. Leadbetter, G. Lindgren, and H. Rootzeń, *Extremes* (Springer-Verlag, New York, 1980).
28. K. V. Bury, *Statistical Models in Applied Science* (Wiley, New York, 1975).
29. C. Lomnitz and E. Rosenbleuth, eds., *Seismic Risk and Engineering Decisions* (Elsevier, Amsterdam, 1976).
30. J. M. Bloom and J. C. Ekvall, eds., *Probabilistic Fracture Mechanics and Fatigue Methods* (ASTM, Philadelphia, 1981).
31. A. Budgor and B. J. West, in *Mathematical Models as a Tool for the Social Sciences*, B. J. West, ed. (Gordon and Breach, New York, 1980).
32. L. Esteva, in *Seismic Risk and Engineering Decisions*, C. Lomnitz and E. Rosenbleuth, eds. (Elsevier, Amsterdam, 1976), Chap. 6.
33. H. C. S. Thom, *J. Marine Res.* **29**:19 (1971).
34. G. G. Trantina and C. A. Johnson, in *Probabalistic Fracture Mechanics and Fatigue Methods*, J. M. Bloom and J. C. Ekvall, eds. (ASTM, Philadelphia, 1981).
35. M. C. Wang and G. E. Uhlenbeck, *Rev. Mod. Phys.* **17**:323–342, (1945).
36. K. Lindenberg, K. E. Shuler, V. Seshadri, and B. J. West, in *Probabilistic Analysis and Related Topics*, Vol. 3, A. T. Bharucha-Reid, ed. (Academic Press, New York, 1983).
37. N. Goel, S. Maitra and E. Montroll, *Rev. Mod. Phys.* **43**:231 (1971).
38. L. Pontryagin, A. Adronow, and A. Witt, *Zh. Eksp. Theor. Fiz.* **3**:172–178 (1933).
39. O. M. Phillips, *The Dynamics of the Upper Ocean*, 2nd ed. (Cambridge University Press, Cambridge, 1977).
40. B. J. West, *Weak Interaction Theory of Deep Water Gravity Waves* (Springer-Verlag, Berlin, 1981).
41. M. Kaul, *Earthq. Eng. Struct. Dyn.* **6**:497 (1978).
42. G. P. Solomos and P-TD. Spanos, *Soil Dyn. Earthq. Eng.* **2**:110 (1983).

43. C. W. Helstrom, *IRE Trans. Inf. Theory* **IT-5**:139–140 (1959).
44. S. H. Crandall, S. S. Lee, and J. H. Williams Jr., *J. Appl. Mech.* **41**:1094, (1974).
45. A. M. Hasofer and P. Petocz, *Adv. Appl. Prob.* **10**:711–716, (1978).
46. P. Lévy, *Théorie de l'Addition des Variables Aléatoires* (Guthier-Villars, Paris, 1937).
47. A. Ya. Khintchine and P. Lévy, *C. R. Acad. Sci* **202**:274 (1936).
48. B. B. Mandelbrot, *Fractals, Form, Chance, and Dimension* (W. H. Freeman, San Francisco, 1977); *The Fractal Geometry of Nature* (W. H. Freeman, San Francisco, 1982).
49. G. H. Weiss and R. J. Rubin, in *Advances in Chemical Physics*, Vol. 52, I. Prigogine and S. A. Rice, eds. (Wiley, New York, 1983).
50. M. Kac and H. Pollard, *Can. J. Math.* **2**:375 (1950).
51. E. W. Montroll and G. H. Weiss, *J. Math. Phys.* **6**:167 (1965).
52. E. W. Montroll and M. F. Shlesinger, in *Nonequilibrium Phenomena II: From Stochastics to Hydrodynamics*, J. L. Lebowitz and E. W. Montroll, eds. (North-Holland, Amsterdam, 1984).