RAPID COMMUNICATIONS

PHYSICAL REVIEW A VOLUME 33, NUMBER 3 MARCH 1986

First-passage times for non-Markovian processes

J. Masoliver* and Katja Lindenberg
Department of Chemistry, B-014, University of California, San Diego, La Jolla, California 92093

Bruce J. West
Center for Studies of Nonlinear Dynamics, La Jolla Institute, 3252 Holiday Court, Suite 208, La Jolla, California 92038

(Received 28 October 1985)

First-passage time statistics for non-Markovian processes have heretofore only been developed for processes driven by dichotomous fluctuations that are themselves Markov. Herein we develop a new method applicable to Markov and non-Markovian dichotomous fluctuations and calculate analytic mean first-passage times for particular examples.

The difficulties encountered in obtaining first-passage time results for even the simplest non-Markovian processes are well known.[1] In great part the subtleties arise from the fact that the first-passage time problem for non-Markovian processes cannot in general be formulated in the traditional way of a Markov process, i.e., as a boundary value problem.[2,3] In this Rapid Communication we present results based on a procedure that avoids this difficulty.

To illustrate our method, we consider the simplest dynamical system driven by external fluctuations with a finite correlation time. Our system is defined by the variable $X(t)$ whose dynamical evolution is specified by the differential equation $\dot{X}=F(t)$. The random variable $F(t)$ is a dichotomous (not necessarily Markov) process, alternately taking on the value of a and $-b$, with $a,b>0$. The times that the variable $F(t)$ retains the value a and $-b$ are, respectively, governed by the distributions $\psi_a(t)$ and $\psi_b(t)$.[4,5] If $F(t)$ is a dichotomous Markov process, then these distributions are exponential, $\psi_j(t)=\lambda_j\exp(-\lambda_j t)$, $j=a,b$, where λ_a^{-1} and λ_b^{-1} are the average residence times in the states $F(t)=a$ and $-b$. Thus, λ_a^{-1} and λ_b^{-1} are average times between switches, and $a\lambda_b=b\lambda_a$. Our first-passage time theory is the first *not* to be restricted to these forms. We assume that successive switches are independent of each other, i.e., that the switching sequence defines a renewal process.[4,5] We note that this assumption in no way restricts the form of the distributions $\psi_a(t)$ and $\psi_b(t)$ that govern the time intervals between switches.[4,5]

The random process $X(t)$ can take on all real values $-\infty\le X(t)\le\infty$, and we wish to calculate the distribution of times for $X(t)$ to first cross the levels $\pm z$. In particular, we are interested in the mean value of this distribution, i.e., in the mean first-passage time to $|X(t)|=z$. Let us begin the process at $X(t=0)=x_0$. Our procedure is based on the fact that the process evolves from this initial state in a series of steps that can be used to construct an actual trajectory by direct integration for any particular realization of $F(t)$. Suppose, for example, that $F(0)=a$. Then we have the following trajectory:

$$\begin{aligned} X(t) &= x_0+at, & 0\le t\le t_1,\\ &= x_0+at_1-b(t-t_1), & t_1\le t\le t_1+t_2,\\ &= x_0+at_1-bt_2+a(t-t_2-t_1), & t_1+t_2\le t\le t_1+t_2+t_3,\\ &= \cdots, & \cdots . \end{aligned} \tag{1}$$

The time intervals t_n are governed by the distributions $\psi_a(t)$ and $\psi_b(t)$. One such trajectory is shown in Fig. 1, where the levels $\pm z$ are also indicated.

Our goal is to calculate the first-passage time probability density $p(t)$ defined as follows: $p(t)\,dt=$ probability that the process $X(\tau)$ crosses z or $-z$ in the time range $t\le\tau\le t+dt$ without ever having crossed either of these levels during the time span $0\le\tau<t$. To calculate $p(t)$, it is useful to denote each time range t_n between switches as an "interval" and to define the auxiliary probability $p_n(t)\,dt=$ probability that the first crossing of z or $-z$ occurs during the nth interval and in the time range $(t,t+dt)$. Clearly,

$$p(t)=\sum_{n=1}^{\infty}p_n(t) \ . \tag{2}$$

The probability densities $p_n(t)$ can be constructed explicitly from the trajectories (1). To illustrate this construction, let us consider a realization that begins with $F(0)=a$, as detailed in Eq. (1). We wish to ensure that no crossings of $\pm z$ occurred in the first $(n-1)$ intervals and that a crossing does occur during the nth interval. During the first interval $X(t)=x_0+at$, and level z is not crossed if the switch to $F(t)=-b$ occurs sufficiently early, i.e., if $X(t_1)=x_0+at_1<z$ or, equivalently, if $t_1<(z-x_0)/a$. The prob-

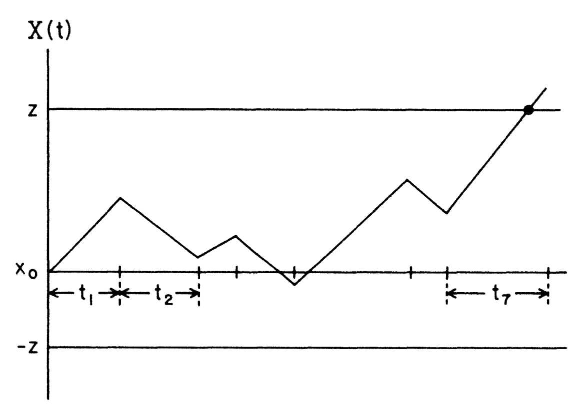

FIG. 1. A typical trajectory with $F(0)=a$. The first crossing of z or $-z$ occurs during the seventh interval.

ability that this inequality holds is

$$\mathrm{Prob}\left\{t_1 < \frac{z-x_0}{a}\right\} = \int_0^{(z-x_0)/a} \psi_a(t_1)\, dt_1 \ . \tag{3}$$

In writing (3) we have assumed that a switch from $F(t)=-b$ to $F(t)=a$ occurred exactly at $t=0$, i.e., that the switching sequence is an "ordinary renewal process."[4] Other initial states can be considered and would require a distribution $\phi_a(t)$ for the first interval that is in general distinct from that of subsequent ones, leading to a so-called "modified renewal process."[4] To ensure that the second interval does not lead to a crossing of level $-z$, we must require that $X(t_2)=x_0+at_1-bt_2>-z$, i.e., that $t_2<(z+x_0+at_1)/b$. The probability that this inequality is satisfied is

$$\mathrm{Prob}\left\{t_2 < \frac{z+x_0+at_1}{b}\right\} = \int_0^{(z+x_0+at_1)/b} \psi_b(t_2)\, dt_2 \ . \tag{4}$$

Similar conditions can be written for the probability that each successive interval up to and including the $(n-1)$st does not lead to a crossing. To proceed with our explicit illustration, we must choose the parity of n: If it is even, then a crossing can only occur at $-z$ during the nth interval, while odd n can only lead to a crossing at z. We select the former and note that the level $-z$ *will* be crossed during the nth interval if $X(t_n)=x_0+at_1-bt_2+\cdots+at_{n-1}-bt_n<-z$. The probability that this inequality is satisfied is

$$\mathrm{Prob}\left\{t_n > \frac{z+x_0+\cdots+at_{n-1}}{b}\right\} = \int_{(z+x_0+\cdots+at_{n-1})/b}^{\infty} \psi_b(t_n)\, dt_n \ . \tag{5}$$

Finally, we must specify *when* during the nth interval the crossing actually occurs. For the crossing to occur at time t, it is necessary that $X(t)=x_0+at_1-bt_2+\cdots+at_{n-1}-b\Delta_{n-1}=-z$, where $\Delta_n=t-(t_1+t_2+t_3+\cdots+t_n)$. The probability density for this crossing event is the delta function

$$p(t:X(t)=-z)=b\delta(z+x_0+at_1-bt_2+\cdots+at_{n-1}-b\Delta_{n-1}) \ . \tag{6}$$

Collecting the results (3)–(6) immediately gives us the following integral form for the density $p_n(t)$:

$$p_n(t)|_{\substack{F(0)=a\\ n\ \mathrm{even}}} = b\int_0^{(z-x_0)/a} dt_1\,\psi_a(t_1)\int_0^{(z+x_0+at_1)/b} dt_2\,\psi_b(t_2)\cdots\int_0^{(z-x_0-at_1+bt_2+\cdots+bt_{n-2})/a} dt_{n-1}\,\psi_a(t_{n-1})$$
$$\times\int_{(z+x_0+at_1-bt_2+\cdots+at_{n-1})/b}^{\infty} dt_n\,\psi_b(t_n)\,\delta(z+x_0+at_1-bt_2+\cdots+at_{n-1}-b\Delta_{n-1}) \ , \tag{7a}$$

for $n\geq 2$, where we have explicitly indicated the initial value $F(0)=a$ and the parity of n. For odd n, the density $p_n(t)$ is found by similar arguments to be given by

$$p_n(t)|_{\substack{F(0)=a\\ n\ \mathrm{odd}}} = a\int_0^{(z-x_0)/a} dt_1\psi_a(t_1)\int_0^{(z+x_0+at_1)/b} dt_2\,\psi_b(t_2)\cdots\int_0^{(z+x_0+at_1-bt_2+\cdots+at_{n-2})/b} dt_{n-1}\,\psi_b(t_{n-1})$$
$$\times\int_{(z-x_0-at_1+bt_2+\cdots+bt_{n-1})/a}^{\infty} dt_n\,\psi_a(t_n)\,\delta(z-x_0-at_1+bt_2+\cdots+bt_{n-1}-a\Delta_{n-1}) \ , \tag{7b}$$

for $n\geq 2$, and

$$p_1(t)|_{F(0)=a} = a\int_{(z-x_0)/a}^{\infty} dt_1\,\psi_a(t_1)\,\delta(z-x_0-at_1) = \delta\left\{t-\frac{z-x_0}{a}\right\}\int_{(z-x_0)/a}^{\infty} dt_1\,\psi_a(t_1) \ . \tag{7c}$$

Similar expressions can clearly be obtained for $F(0)=-b$.

The next step in our procedure is to Laplace transform Eq. (7) and to establish an integral recursion relation to connect the nth and $(n+2)$nd densities. The recursion relations for even n and for odd n must thus be constructed separately. Upon summing the resulting relations over n we obtain an integral relation for each of the functions,

$$K_l(s;x_0,z;a) \equiv \sum_{n=0}^{\infty} \tilde{p}_{2n+l}(s)|_{F(0)=a} \ , \tag{8}$$

where $l=1,2$, and where the dependences on the initial values of $F(t)$ and $X(t)$ as well as the barrier height z have been indicated explicitly. Since the integral relations for K_1 and for K_2 turn out to have the same form, we can sum them and exhibit a single relation for the combination $K(s;x_0,z;a)=K_1(s;x_0,z;a)+K_2(s;x_0,z;a)$. We obtain

$$K(s;x_0,z;a)=\tilde{p}_1(s)+\tilde{p}_2(s)+\int_0^{(z-x_0)/a} dt_1\int_0^{(z+x_0+at_1)/b} dt_2\,\psi_a(t_1)\psi_b(t_2)e^{-s(t_1+t_2)}K(s;x_0+at_1-bt_2,z;a) \ . \tag{9}$$

Finally, in terms of these functions and the probabilities $w_0(a|x_0)$ and $w_0(b|x_0)$ that $F(t=0)=a$ and $F(t=0)=-b$ given that $X(0)=x_0$, the Laplace transform of the first-passage time probability density $p(t)$ then is

$$\tilde{p}(s)=K(s;x_0,z;a)w_0(a|x_0)+K(s;x_0,z;b)w_0(b|x_0) \ . \tag{10}$$

Thus the entire problem has been reduced to the solution of integral equations of the form (9). These equations can in general not be solved exactly for arbitrary forms of $\psi_a(t)$ and $\psi_b(t)$, but they lend themselves to approximation schemes

RAPID COMMUNICATIONS

appropriate to specific forms of these functions. There are, however, situations when the integral equations *can* be solved exactly for certain forms of $\psi_a(t)$ and $\psi_b(t)$, and we here give two such examples.

I. DICHOTOMOUS MARKOV PROCESS $F(t)$

A dichotomous Markov process $F(t)$ is characterized by exponential distributions of switching times as given earlier and depicted in Fig. 2. In this case the integral equation (9) can readily be converted to a second-order differential equation. For the particular example considered here, we obtain

$$\left[\frac{d^2}{dx_0^2}+\left(\frac{\lambda_b+s}{b}-\frac{\lambda_a+s}{a}\right)\frac{d}{dx_0}-\frac{s}{ab}(s+\lambda_a+\lambda_b)\right]K(s;z,x_0;a)=\left[\frac{d^2}{dx_0^2}+\left(\frac{\lambda_b+s}{b}-\frac{\lambda_a+s}{a}\right)\frac{d}{dx_0}-\frac{(\lambda_a+s)(\lambda_b+s)}{ab}\right][\tilde{p}_1(s)+\tilde{p}_2(s)] \ . \tag{11}$$

The boundary conditions for (11) are deduced directly from (9) and are given by $K(s;z,z;a)=1$ (ensuring that a process that begins at $x_0=z$ with positive velocity escapes with certainty), and the integral relation

$$\frac{d}{dx_0}K(s;z,x_0;a)|_{x_0=z}=\frac{d}{dx_0}[\tilde{p}_1(s)+\tilde{p}_2(s)]_{x_0=z}-\frac{\lambda_a\lambda_b}{ab}\int_{-z}^{z}du\, e^{-(\lambda_b+s)(z-u)/b}K(s;z,u;a) \ . \tag{12}$$

We note that the constants of integration in the solution of (11) can be found either from these boundary conditions or from a substitution of the solution back into the integral relation (9). The solution of (11) when substituted into (10) with the initial choice $F(0)=a$ [i.e., with $w(a|x_0)=1$ and $w(b|x_0)=0$] gives the result

$$\tilde{p}(s)=[(\alpha+r)e^{2rz}-(\alpha-r)e^{-2rz}]^{-1}\left\{e^{-\beta(z-x_0)}[(\alpha+r)e^{r(z+x_0)}-(\alpha-r)e^{-r(z+x_0)}]+\frac{\lambda_a}{a}e^{\beta(z+x_0)}(e^{r(z-x_0)}-e^{-r(z-x_0)})\right\} , \tag{13}$$

where

$$\alpha=\frac{a\lambda_b+b\lambda_a+(a+b)s}{2ab},\quad \beta=\frac{1}{2}\left(\frac{s+\lambda_a}{a}-\frac{s+\lambda_b}{b}\right) , \tag{14}$$

$$r=[4\beta^2+4s(s+\lambda_a+\lambda_b)/ab]^{1/2} \ . \tag{15}$$

The distribution (13) yields analytic expressions for arbitrary first-passage time moments. In particular, the mean first-passage time is given by

$$T_1(z,x_0)=\int_0^\infty dt\, tp(t)=-\frac{d}{ds}\tilde{p}(s)|_{s=0} \ . \tag{16}$$

Using Eq. (13) in Eq. (16), we obtain

$$T_1(z,x_0)=\frac{z^2-x_0^2}{2D}+\frac{z-x_0}{a}\frac{a/\lambda_a+z(a+b)/b}{a/\lambda_a+2z} , \tag{17}$$

where $D=a^2b/\lambda_a(a+b)$. This result was also obtained by Hänggi and Talkner[1] using an entirely different procedure restricted to dichotomous Markov processes $F(t)$. For the special case $a=b$ and $\lambda_a=\lambda_b$, the mean first-passage time

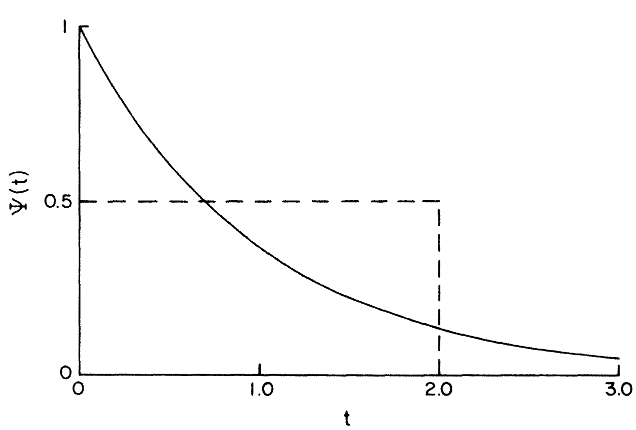

FIG. 2. Distribution $\psi(t)$ of time intervals between switches of $F(t)$ from one value to another. The mean time between switches is unity ($\lambda=1$). Solid curve: exponential $\psi(t)$ corresponding to a dichotomous Markov process $F(t)$. Dashed curve: a $\psi(t)$ corresponding to a non-Markovian $F(t)$.

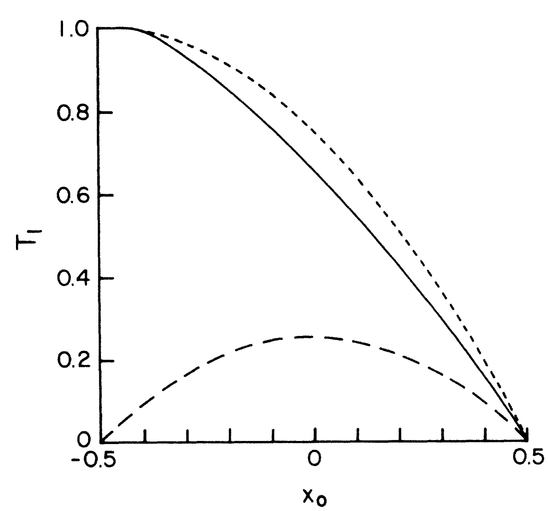

FIG. 3. Mean first-passage time T_1 to ± 0.99 vs initial position x_0 for a process that begins with positive slope. Dotted curve: dichotomous Markov fluctuations with $\lambda=1$ and $a=1$. Solid curve: dichotomous non-Markovian fluctuations with $\lambda=1$ and $a=1$. Dashed curve: diffusive process with $2D=1$.

(17) reduces to the simpler form

$$T_1(z,x_0)=\frac{z^2-x_0^2}{2D}+\frac{z-x_0}{a} , \tag{18}$$

where now $D=a^2/2\lambda_a$. The first term in (17) or (18) is the mean first-passage time in the limit of Gaussian white noise in which $b,a\to\infty$, $\lambda_b,\lambda_a\to\infty$, and $D=\text{constant}$. The remaining contribution in (17) and (18) gives the deviation from this limit and leads to a jump in the mean first-passage time at $x_0=-z$. This finite value of $T_1(z,-z)$ is the reason why the first-passage time problem in the presence of colored noise cannot be described in terms of a standard boundary value problem. Equation (18) is shown in Fig. 3 as is the Gaussian white noise result for comparison.

II. A NON-MARKOVIAN PROCESS $F(t)$

As an example of a non-Markovian dichotomous process $F(t)$, we choose the distributions $\psi_a(t)$ and $\psi_b(t)$ to have the form shown in Fig. 2, i.e., for $j=a,b$ we take $\psi_j(t)=\lambda_j/2$ if $0\leq t\leq 2/\lambda_j$, and $\psi_j(t)=0$ otherwise. The mean times between switches are λ_a^{-1} and λ_b^{-1}, and again $a\lambda_b=b\lambda_a$. With this choice of switching distributions, it is again possible to convert the integral equation (9) to a second-order differential equation:

$$\left[\frac{d^2}{dx_0^2}+\frac{\lambda^2-4s^2}{4a^2}\right]K(s;x_0,z;a)=\left[\frac{d^2}{dx_0^2}-\frac{s^2}{a^2}\right][\tilde p_1(s)+\tilde p_2(s)] , \tag{19}$$

where we have taken $a/\lambda_a=b/\lambda_b>z$. The boundary conditions can again be found directly from (9), or alternatively, the solution of (19) can be substituted into Eq. (9) for the determination of the constants of integration. For the case $a=b$ and $\lambda_a=\lambda_b=\lambda$ and with the initial choice $F(0)=a$, we obtain for the Laplace transform of the first-passage time distribution

$$\tilde p(s)=\left[\left(\frac{s}{a}+\beta\right)e^{2\beta z}-\left(\frac{s}{a}-\beta\right)e^{-2\beta z}\right]^{-1}\left(1-4\frac{s}{\lambda}-e^{-2sz/a}\right)\times\left[\frac{\lambda}{2a}\left(e^{\beta(z-x_0)}-e^{-\beta(z-x_0)}\right)+\left(\frac{s}{a}+\beta\right)e^{\beta(z+x_0)}-\left(\frac{s}{a}-\beta\right)e^{-\beta(z-x_0)}\right]+e^{-s(z+x_0)/a}+4\frac{s}{\lambda}e^{-s(z-x_0)/a} , \tag{20}$$

where $\beta=i(\lambda^2-4s^2)^{1/2}/2a$. Once again all first-passage time moments can be obtained analytically from (20). In particular, we obtain for the mean first-passage time

$$T_1(z,x_0)=\frac{2}{a\lambda}\frac{(2a-\lambda z)}{\cos(\lambda z/a)}\left(\sin\frac{\lambda(z-x_0)}{2a}+\cos\frac{\lambda(z+x_0)}{2a}\right)-\frac{1}{a}\left(\frac{4a}{\lambda}-z-x_0\right) . \tag{21}$$

Again we note that, whereas $T_1(z,z)=0$, $T_1(z,-z)$ has a finite value. Further comparisons between the results (21) and the Markov results (18) are shown in Fig. 3.

We find that in general for a given value of λ (i.e., for a given mean time between switches) the mean first-passage time for the process to reach a boundary is longer (shorter) for a process driven by the non-Markovian $F(t)$ if the initial state is very near the lower (upper) boundary and the process begins with a positive slope. The differences in the mean first-passage times of the processes driven by Markov and non-Markovian dichotomous fluctuations are most pronounced when the mean distance a/λ covered in one interval is comparable to the distance $2z$ between the boundaries. When a/λ is very small, the process becomes diffusive, while a larger value of a/λ leads in either case to a mean first-passage time $T_1=(z-x_0)/a$, representing a direct arrival at the boundary during the first interval.

ACKNOWLEDGMENTS

This research was supported in part by National Science Foundation Grants No. ATM-8507820, No. ATM-8507821, and No. DMR-8315950, and by Defense Advanced Research Projects Agency Contract No. DAAG 29-85-K-0246. One of us (J.M.) would like to thank the Fulbright Foundation for financial support. The Center for Studies of Nonlinear Dynamics (B.J.W.) is affiliated with the University of California, San Diego, La Jolla, CA.

*Permanent address: Department de Matematiques, E.T.S. Enginyers Telecomunicacio, Universitat Politecnica De Catalunya, Jordi Girona Salgado s/n, 08034 Barcelona, Spain.

[1]P. Hänggi and P. Talkner, Phys. Rev. A **32**, 1934 (1985), and references therein.

[2]K. Lindenberg and V. Seshadri, J. Chem. Phys. **71**, 4075 (1979).

[3]B. J. Matkowsky and Z. Schuss, SIAM J. Appl. Math. **33**, 365 (1977).

[4]D. R. Cox and H. D. Miller, *The Theory of Stochastic Processes* (Wiley, New York, 1965), pp. 340 and 341.

[5]A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 2nd ed. (McGraw-Hill, New York, 1984), pp. 365–367.