The Annals of Probability
1980, Vol. 8, No. 4, 775–792

MODEL PROCESSES IN NONLINEAR PREDICTION WITH APPLICATION TO DETECTION AND ALARM

BY GEORG LINDGREN

University of Lund

A level crossing predictor is a predictor process $Y(t)$, possibly multivariate, which can be used to predict whether a specified process $X(t)$ will cross a predetermined level or not. A natural criterion on how good a predictor is, can be the probability that a crossing is detected a sufficient time ahead, and the number of times the predictor makes a false alarm.

If X is Gaussian and the process Y is designed to detect only level crossings, one is led to consider a multivariate predictor process $Y(t)$ such that a level crossing is predicted for $X(t)$ if $Y(t)$ enters some nonlinear region in R^p. In the present paper we develop the probabilistic methods for evaluation of such an alarm system. The basic tool is a model for the behavior of $X(t)$ near the points where $Y(t)$ enters the alarm region. This model includes the joint distribution of location and direction of $Y(t)$ at the crossing points.

1. Introduction. In some prediction problems the interest focuses on the prediction of certain rare events, such as high level crossings or other rare patterns, occurring in a time varying random function $\{X(t), t \in R\}$. As an example, suppose that we want to predict if $X(t)$ will have a local maximum or minimum within the next interval of some fixed length $m > 0$, or, in other words, that we want to predict the locations of the local maxima and minima of $X(t)$, i.e., the times when $X'(t) = 0$. At time $t - m$ we have at our disposal observations of $X(s), s \leqslant t - m$, together with other information which can be summarized in a multivariate process $Y(t)$. The question now is, how shall the information gathered in $Y(t)$ be used to make a reliable statement about whether or not $X'(\tau)$ will have a zero for some $\tau \in (t - m, t]$.

The naive way of predicting level crossings would be to calculate a predictor of $X'(t)$ based on $Y(t)$, $\hat{X}'(t)$ say, and then state that $X'(t)$ will be zero any time it happens that $\hat{X}'(t) = 0$. However, if the predictor $X'(t)$ is based on a criterion such as minimal mean square prediction error there is no reason why it should also be particularly good at predicting zeros.

Obviously, predictor functions which are based on mean square deviation need not be optimal when used as level crossing predictors, in the sense that they show high ability to detect a crossing without making too many false alarms; some theoretical reasons for this may be found in Lindgren (1975a). Nevertheless, it is often suggested to use a mean square predictor to detect level crossings; the reader may study Sveshnikov (1968), Problems 36.14 and 36.25.

Received June 7, 1978; revised April 9, 1979.

AMS 1970 *subject classifications.* Primary 60G25; secondary 60G15, 60G35.

Key words and phrases. Nonlinear prediction, level crossings, Gaussian processes, point processes.

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Probability.* www.jstor.org

In previous papers, Lindgren (1975b), (1979), de Maré (1980), different aspects on predictors have been developed, based on the idea that the predictor has the restricted task to detect only the occurrence of level crossings, specifically of high levels. It is shown in de Maré (1980) that one is then led to consider a multivariate predictor process $\{Y(t), t \in R\}$ and a region $\Gamma \subseteq R^p, p \geqslant 2$, with nonlinear boundary, and to predict a level crossing for $X(t)$ any time $Y(t)$ enters the alarm region Γ. In the present paper we shall derive the probabilistic tools for evaluation of such alarm systems, in particular the connection between the series of level crossings in $X(t)$ and the series of alarms given by $Y(t)$. Of special interest is the operating characteristic, defined as the long run proportions of level crossings which are accompanied by an alarm, and of alarms accompanied by a level crossing, i.e., of detected crossings and of nonfalse alarms.

Some basic concepts are presented in Section 2, while the behavior of the predictor $Y(t)$ in the neighbourhood of level crossings for $X(t)$ is derived in Section 3; most of these results can be found also in Lindgren (1975b). In Section 4 a model process is derived for the behaviour of $X(t)$ and $Y(t)$ near the points where $Y(t)$ enters the alarm region through a nonlinear boundary. Some proofs are given in Section 5.

2. Notations and basic facts. Suppose $\{X(t),\ t \in R\}$ is a stationary normal process with mean zero and continuously differentiable sample paths, whose covariance function $r_X(\tau) = \mathrm{Cov}(X(t+\tau),\ X(t))$ admits the expansion

$$r_X(\tau) = \lambda_0^x - \lambda_2^x \tau^2/2 + o(\tau^2) \quad \text{as } \tau \to 0,$$

where $\lambda_0^x = r_X(0) = V(X(t))$, $\lambda_2^x = -r_X''(0) = V(X'(t))$. To predict the future of the X-process we have at our disposal a multivariate stationary zero mean process $Y(t) = (Y_1(t), \cdots, Y_p(t))^T$ in which some $Y_k(t)$ may be functions of $X(s)$, $s \leqslant t$; e.g., we may have $Y_1(t) = X(t)$, $Y_2(t) = X(t-h)$, $Y_3(t) = X'(t)$, etc. We suppose that $X(t)$ and $Y(t)$ are jointly normal with matrix covariance function $r_{XY}(\tau) = \mathrm{Cov}(X(t+\tau), Y(t)) = E(X(t+\tau)Y(t)^T)$ of dimension $1 \times p$. We denote by $r_Y(\tau) = E(Y(t+\tau)Y(t)^T)$ the covariance function of $Y(t)$.

Our main object is to predict the times when $X(t)$ crosses a specified level u or, if this can not be achieved, at least, with some degree of certainty, predict whether $X(t)$ will cross u within a specified time or not. If u is a high level we will sometimes call it a catastrophe level, appealing to practical situations of alarms, in which cases we also call $Y(t)$ an alarm function. In the applications we have in mind it is rather the mere occurrence of a level crossing within the near future which is of interest, not the exact time for it. We therefore formulate the prediction problem as a pure two-choice problem; at each time t we make one of two possible statements, either that $X(t+\tau)$ will cross the prescribed level u at least once for some $\tau \in [0, m]$, or that there will be no such crossing. Here m is a fixed constant, the choice of which depends on the application at hand.

We formalize this by defining an *alarm region* $\Gamma \subset R^p$ such that if $Y(t) \in \Gamma$ we believe that $X(t+\tau)$ will cross u for some $\tau \in [0, m]$, while if $Y(t) \in \Gamma^c$ (its complement) we do not believe there will be such a crossing. We call Γ^c the *run region*.

In the special case when the boundary $\partial\Gamma$ is linear there is a linear function

$$Y^*(t) = \Sigma_{k=1}^{p} c_k Y_k(t)$$

and a constant $\hat{u}$ such that $Y(t) \in \Gamma$ if and only if $Y^*(t) > \hat{u}$, and $Y(t)$ enters the alarm region Γ when $Y^*(t)$ has an upcrossing of the *alarm level* $\hat{u}$. Since $X(t)$ and $Y(t)$ are supposed to be jointly normal, $X(t)$ and the level crossing predictor $Y^*(t)$ are also jointly normal, a case which was dealt with in some detail in Lindgren (1975b).

EXAMPLE 2.1. A possible general principle for obtaining an alarm region is to regard the predictor process $Y(t)$ as a test statistic which tests, for some fixed $m > 0$, whether or not $X(t + m) = u$, upcrossing. A formal application of the Neyman-Pearson lemma then yields an alarm region Γ defined by

$$Y(t) \in \Gamma \Leftrightarrow \frac{dP(Y(t)|C(t+m))}{dP(Y(t)|C^c(t+m))} > k,$$

where $C(t + m)$ denotes the event $\{X(t + m) = u, \text{ upcrossing}\}$, and $C^c(t + m)$ is its complement.

This idea is further elaborated in de Maré (1980) where, as an example, it is shown that if $X(t)$ and $Y(t)$ are jointly normal, the alarm region is

$$\Gamma = \{y \in R^p;\ y^T A y + y^T B + \ln\Psi(y^T C) > K\},$$

where A is a $p \times p$ matrix of constants, and B, C constant p-vectors. Furthermore $\Psi(x) = \phi(x) + x\Phi(x)$, ϕ and Φ being the standard normal density and distribution functions. The alarm boundary is approximately a second order surface.

As Example 2.1 shows, one is naturally led to consider nonlinear boundaries when looking for optimal regions for level crossing predictors. In the next two sections we will derive a tool for evaluating the *operating characteristic* of such alarm policies. This concept is defined as follows.

Let

$$[0 <] t_1 < t_2 < \cdots$$

be the locations of the u-upcrossings for $X(t)$, i.e., the catastrophes, and let similarly

$$[0 <] \hat{t}_1 < \hat{t}_2 < \cdots$$

be the times where $Y(t)$ enters the alarm region. Define the operating characteristic as

$$(2.1)\quad 1 - G_m(u; \Gamma)$$

$$= \lim_{T\to\infty} \frac{\#\{\hat{t}_k;\ 0 \leqslant \hat{t}_k \leqslant T,\ X(\hat{t}_k + \tau) = u, \text{ upcrossing, some } \tau \in (0, m)\}}{\#\{\hat{t}_k;\ 0 \leqslant \hat{t}_k \leqslant T\}}$$

and

$$(2.2)\quad 1 - H_m(\Gamma; u) = \lim_{T\to\infty} \frac{\#\{t_k;\, 0 \leqslant t_k \leqslant T,\, Y(t_k - \tau) \text{ enters } \Gamma \text{ for some } \tau \in (0,m)\}}{\#\{t_k;\, 0 \leqslant t_k \leqslant T\}}$$

provided the limits exist, with the interpretation that $G_m(u;\Gamma)$ is the long run proportion of alarms which are not followed by a catastrophe within time m, i.e., $G_m(u;\Gamma)$ is the long run proportion of false alarms. Similarly, $H_m(\Gamma;u)$ is the long run proportion of catastrophes which are not preceded by any alarm within time m, i.e., $H_m(\Gamma;u)$ is the long run proportion of undetected catastrophes. We call $G_m(u;\ \Gamma)$ and $H_m(\Gamma;\ u)$ the *detection errors*.

3. A model process for the predictor near catastrophes. The probability that a catastrophe is detected by the alarm function is the conditional probability that $Y(t-\tau)$ crosses the boundary $\partial\Gamma$ given that $X(t)$ crosses the level u. Such probabilities are perhaps most simply approximated by means of a particular model process, in a different context introduced by Slepian (1962) and further developed by Lindgren (1975b), (1979).

Since $\lambda_2^x = V(X'(t)) < \infty$, the average number of u-upcrossings per time unit by $X(t)$ is given by Rice's formula

$$\frac{1}{2\pi}(\lambda_2^x/\lambda_0^x)^{\frac{1}{2}} \exp(-u^2/2\lambda_0^x) < \infty,$$

which means that there is only a finite number of u-upcrossings in any bounded interval, i.e.,

$$N_T(u) = \#\{t_k; 0 \leqslant t_k \leqslant T\} < \infty.$$

If $A \subseteq R^{pr}$ is a finite-dimensional Borel set and $s = (s_1, \cdots, s_r)$, let

$$N_T(A;\ u) = \#\{t_k; 0 \leqslant t_k \leqslant T, (Y(t_k + s_1), \cdots, Y(t_k + s_r)) \in A\}$$

denote the number of u-upcrossings t_k at which the translated process $Y(t_k + t)$ satisfies the condition defined by A. We shall use the notation $Y(t_k + s) = (Y(t_k + s_1), \cdots, Y(t_k + s_r))$, and write

$$\{Y(t_k + s) \in A\} = \{(Y(t_k + s_1), \cdots, Y(t_k + s_r)) \in A\}$$

if A is a finite-dimensional set. More generally, we write $\{Y(t_k + \cdot\,) \in A\}$ when A is a set of p-variate continuously differentiable functions, in analogy with the notation $\{y(\cdot\,) \in A\}$ when y is a function $R \to R^p$.

Furthermore, let $f_{X(0)}$, $f_{X(0),\,X'(0)}$, $f_{X'(0)|X(0)=u}$, etc., be the density of the random variable $X(0)$, the joint density of $X(0)$, $X'(0)$, the conditional density of $X'(0)$ given that $X(0) = u$, and so on, and write $x^+ = \max(0, x)$. Then

$$\begin{aligned}\gamma_X(u) &= E(N_1(u)) = f_{X(0)}(u)E(X'(0)^+\,|X(0) = u)\\ &= \textstyle\int_{z=0}^{\infty} z f_{X(0),X'(0)}(u,z)\,dz\\ &= \frac{1}{2\pi}(\lambda_2^x/\lambda_0^x)^{\frac{1}{2}} \exp(-u^2/2\lambda_0^x).\end{aligned}$$

The following theorem expresses the long run distribution of $Y(t_k + \cdot)$ as t_k runs through the set of all u-upcrossings for $X(t)$. The theorem is of standard type in crossing theory, and its derivation is postponed to Section 5.

THEOREM 3.1. *If $\{X(t), Y(t), t \in R\}$ is ergodic, $A \subseteq R^{pr}$ is an open set, and $s = (s_1, \cdots, s_r)$, then with probability one,*

$$(3.1)\quad \lim_{T\to\infty} \frac{N_T(A;u)}{N_T(u)} = \frac{E(N_1(A;u))}{E(N_1(u))} = \frac{E\big(1_{Y(s)\in A}\cdot X'(0)^+ \,|X(0) = u\big)}{E\big(X'(0)^+ \,|X(0) = u\big)}$$

$$= \int_{z=0}^{\infty} P(Y(s) \in A | X(0) = u, X'(0) = z)\cdot \gamma_X(u)^{-1} z f_{X(0),X'(0)}(u,z)\, dz.$$

Taking $Y(t) = X'(t)$ and using the fact that for stationary normal processes, $X(0)$ and $X'(0)$ are independent we can obtain as a corollary the well-known Rayleigh distribution of the derivatives $X'(t_k)$.

COROLLARY 3.2. *If $\{X(t), t \in R\}$ is ergodic, the derivative $X'(t_k)$ at the u-upcrossings has a long-run Rayleigh distribution with density $(z/\lambda_2^x)\exp(-z^2/2\lambda_2^x)$, $z > 0$, in the sense that with probability one*

$$\lim_{T\to\infty} \frac{\#\{t_k;\, 0 \leqslant t_k \leqslant T, X'(t_k) < \zeta\}}{\#\{t_k;\, 0 \leqslant t_k \leqslant T\}} = \frac{E\big(1_{X'(0)<\zeta} X'(0)^+ \,|X(0) = u\big)}{E\big(X'(0)^+ \,|X(0) = u\big)}$$

$$= \int_{z=0}^{\zeta} (z/\lambda_2^x)\exp\big(-z^2/2\lambda_2^x\big)\, dz.$$

Since $X(s)$ and $Y(t)$ are supposed to be jointly normal processes, their derivatives are also normal when they exist. This can be used to obtain a probabilistic representation of the density function q_u^s appearing in (3.1) and defined for $y \in R^p$ by

$$q_u^s(y)\, dy = \frac{E\big(1_{Y(s)\in dy}\cdot X'(0)^+ \,|X(0) = u\big)}{E\big(X'(0)^+ \,|X(0) = u\big)}$$

$$= \gamma_X(u)^{-1} \int_{z=0}^{\infty} z f_{X(0),X'(0),Y(s)}(u,z,y)\, dz\, dy.$$

Denote by $r_{YY'}(0)$ the covariance matrix of order $2p \times 2p$ of $Y(t)$ and its derivative $Y'(t)$. The following lemma follows from multivariate normal theory.

LEMMA 3.3. *The conditional distributions of $Y(t)|X(0) = u, X'(0) = z$ for $t \in R$ are p-variate normal with mean*

$$m_{u,z}^Y(t) = u\frac{r_{XY}(-t)^T}{\lambda_0^x} - z\frac{r'_{XY}(-t)^T}{\lambda_2^x}$$

and covariance matrix function

$$(3.2)\quad r_{Y|X}(s,t) = r_Y(s - t) - \frac{r_{XY}(-s)^T r_{XY}(-t)}{\lambda_0^x} - \frac{r'_{XY}(-s)^T r'_{XY}(-t)}{\lambda_2^x}.$$

In order to handle the conditional distribution of $Y(t)$ given $X(0) = u, X'(0) = z$ in an efficient way, we define an auxiliary nonstationary p-variate normal process

$\{\kappa_{Y|X}(t), t \in R\}$ having mean zero and covariance $r_{Y|X}$ defined by (3.2). Lemma 3.3 can then be reformulated as

$$\mathcal{L}(Y(t)|X(0) = u, X'(0) = z) = \mathcal{L}\big(m^Y_{u,z}(t) + \kappa_{Y|X}(t)\big),$$

where $\mathcal{L}(\cdot)$ and $\mathcal{L}(\cdot|\cdot)$ denote "the law of" and "the conditional law of", respectively. Combining Lemma 3.3 and the process $\kappa_{Y|X}$ with the long-run Rayleigh distribution of the derivatives $X'(t_k)$ at the u-upcrossings, we get the following theorem, where ζ_X is a Rayleigh random variable, independent of $\kappa_{Y|X}$ and with density

$$f_{\zeta_X}(z) = (z/\lambda_2^x)\exp(-z^2/2\lambda_2^x), \qquad z > 0.$$

THEOREM 3.4. *If $\{X(t), Y(t), t \in R\}$ is ergodic, the long-run finite-dimensional distributions of $Y(t_k + \cdot)$ after u-upcrossings t_k are given by*

$$\lim_{T\to\infty} \frac{\#\{t_k;\, 0 \leqslant t_k \leqslant T, Y(t_k + \cdot) \in A\}}{\#\{t_k;\, 0 \leqslant t_k \leqslant T\}}$$
$$= \int_{z=0}^{\infty} f_{\zeta_X}(z) P\big(m^Y_{u,z}(\cdot) + \kappa_{Y|X}(\cdot) \in A\big)\, dz$$
$$= P\big(m^Y_{u,\zeta_X}(\cdot) + \kappa_{Y|X}(\cdot) \in A\big).$$

This fundamental theorem makes it possible to use the process

$$Y_u(t) = m^Y_{u,\zeta_X}(t) + \kappa_{Y|X}(t) \tag{3.3}$$

as a model process for $Y(t_k + t)$ before (if $t < 0$) and after (if $t > 0$) the u-upcrossings t_k, and in particular to express the probability of detecting a level crossing. In fact, the event A in Theorem 3.4 does not need to be finite-dimensional, but can be an event such as

$$\{y(\cdot) \in A\} \Leftrightarrow \{y(t) \text{ enters } A \text{ for some } t \in (-m, 0)\}.$$

THEOREM 3.5. *Under the conditions of Theorem* 3.4

$$1 - H_m(\Gamma;\, u) = P(Y_u(t) \text{ enters } \Gamma \text{ for some } t \in (-m, 0)). \tag{3.4}$$

The detection probability $1 - H_m(\Gamma;\, u)$ can therefore be calculated as the probability that the nonstationary, nonnormal process $Y_u(t)$ crosses the boundary $\partial\Gamma$ in the correct direction for at least one $t \in (-m, 0)$. Such crossing probabilities can be approximated by, indeed, quite accurate bounds, using the expected number of crossings, expressed as suitable integrals. For numerical examples in the context of crossing prediction the reader is referred to Lindgren (1975b), (1979).

4. Crossings and alarms with nonlinear boundaries. In this section we will take a close look at the crossings of a curved boundary $\partial\Gamma$ by the p-variate normal process $\{Y(t),\, t \in R\}$, in previous sections called the alarm process, to see how the presence of crossings influences the process $\{X(t),\, t \in R\}$.

We will use some results by Marcus (1977) to extend a theorem by Belyaev (1968) concerning the expected number of exits across the boundary $\partial\Gamma$ and generalize them to *marked* exists.

Let the boundary $\partial\Gamma$ be defined by means of a real valued function $\Psi(y)$ of $y = (y_1, \cdots, y_p)$ in such a way that

$$y \in \Gamma \Leftrightarrow \Psi(y) > 0,$$
$$y \in \partial\Gamma \Leftrightarrow \Psi(y) = 0.$$

We suppose that $\Psi(y)$ is continuously differentiable for y near $\partial\Gamma$, except possibly for a finite number of points at $\partial\Gamma$, and further that there is a set of new coordinates $\Theta = (\theta^1, \cdots, \theta^p)$ such that $\theta^p = \Psi(y)$. We also suppose that the transformation to new coordinates is continuously differentiable and one to one, at least for y near $\partial\Gamma$. Denote the coordinate transformation by

$$\Theta = h(y),\ y = g(\Theta). \tag{4.1}$$

Further let ν_y be the unit normal perpendicular to the surface $\Psi(y) = C$, C constant, at the point y, i.e., $\nu_y = \dot{\Psi}(y)/|\dot{\Psi}(y)|$, where $\dot{\Psi}(y) = (\partial\Psi/\partial y_1, \cdots, \partial\Psi/\partial y_p)^T$ is supposed to be nonzero near $\partial\Gamma$.

Now, if $\{Y(t),\ t \in R\}$ is a continuously differentiable and p-variate normal process, we can express the exits through $\partial\Gamma$ of $Y(t)$ by means of the zero-crossings of the process $\{\Psi(Y(t)),\ t \in R\}$, which is usually nonnormal, and for which we use the shorthand notation

$$\Psi_t = \Psi(Y(t)).$$

Recalling the notation $\hat{t}_k$, $k = 1, 2, \cdots$ for the alarm points, we then have that Ψ_t has its zero upcrossings at $\hat{t}_1, \hat{t}_2, \cdots$. Note that according to the assumptions on Ψ and $Y(t)$, the process Ψ_t is differentiable, almost surely, with a derivative given by the scalar product

$$\frac{d}{dt}\Psi(Y(t)) = \Psi_t' = \dot{\Psi}(Y(t))^T \cdot Y'(t) = |\dot{\Psi}(Y(t))| \nu_{Y(t)}^T \cdot Y'(t) \tag{4.2}$$

between the directional derivative $Y'(t)$ and the gradient vector $\dot{\Psi}(Y(t))$ perpendicular to the surface $\Psi(y) = \Psi(Y(t))$.

We can now formulate the following analogue of Theorem 3.1. If $B \subseteq R^r$ is a finite-dimensional Borel set, $s = (s_1, \cdots, s_r)$, define

$$N_T(B; \partial\Gamma) = \#\{\hat{t}_k; 0 \leqslant \hat{t}_k \leqslant T,\ X(\hat{t}_k + s) \in B\}$$
$$= \#\{\hat{t}_k; 0 \leqslant \hat{t}_k \leqslant T,\ (X(\hat{t}_k + s_1), \cdots, X(\hat{t}_k + s_r)) \in B\},$$
$$N_T(\partial\Gamma) = \#\{\hat{t}_k; 0 \leqslant \hat{t}_k \leqslant T\},$$

and

$$\gamma_Y(\partial\Gamma) = E(N_1(\partial\Gamma)) = f_{\Psi_0}(0)E((\Psi_0')^+ \mid \Psi_0 = 0) = \int_{z=0}^{\infty} z f_{\Psi_0, \Psi_0'}(0, z)\, dz.$$

THEOREM 4.1. *If $\{X(t), \Phi_t, t \in R\}$ is ergodic, then with probability one*

$$\lim_{T\to\infty} \frac{N_T(B; \partial\Gamma)}{N_T(\partial\Gamma)} = \frac{E(N_1(B; \partial\Gamma))}{E(N_1(\partial\Gamma))} = \frac{E(1_{X(s)\in B} \cdot (\Psi_0')^+ \mid \Psi_0 = 0)}{E((\Psi_0')^+ \mid \Psi_0 = 0)} \tag{4.3}$$
$$= \int_{x \in B} \gamma_Y(\partial\Gamma)^{-1} \int_{z=0}^{\infty} z f_{\Psi_0, \Psi_0', X(s)}(0, z, x)\, dz\, dx.$$

Even if (4.3) is a very simple formula, since Ψ_t and Ψ_t' are in general not normal, it might be difficult to obtain the conditional distribution of $X(\cdot)$ given Ψ_0 and Ψ_0' in a sufficiently explicit form to get a closed form model process. Therefore it can be difficult to evaluate from (4.3) the error probability $G_m(u; \Gamma)$ defined by (2.1). Furthermore, if the aim is to design a good alarm region for a given alarm process $Y(t)$, it is more convenient to retain the explicit dependence on the process $Y(t)$. We shall therefore exploit the normality of this process in order to express (4.3) more explicitly by means of a model process $X(t)$ similar to $Y_u(t)$, which was used in Section 3 to evaluate the error probability $H_m(\Gamma; u)$.

In the model process $Y_u(t)$ defined by (3.3) an important role is played by the level u and the Rayleigh distributed derivative ξ_X, which describes the long run random variation of the X-derivative at the u-upcrossings. A similar role will here be played by the distribution of the coordinates $Y(\hat{t}_k)$ and the directional derivative $Y'(\hat{t}_k)$ at the point of exit through the boundary $\partial\Gamma$. These distributions, and consequently also the distribution of $X(\hat{t}_k + \cdot)$, can be obtained as surface densities over the boundary if we express the moments of $N_T(B; \partial\Gamma)$ and $N_T(\partial\Gamma)$ as surface integrals over $\partial\Gamma$. Such formulas have been given by Belyaev (1968), and since they have some intuitive appeal we will digress along that line before turning to the more general results.

Denote by $f_{Y(t)}(y)$ the density of $Y(t)$ at the point $y \in \partial\Gamma$, representing a mass distribution over the surface $\partial\Gamma$. The derivative $Y'(t)$ expresses the direction of a *flow* through $\partial\Gamma$, and the net flow at the point $Y(t) = y$ is equal to the scalar product $\nu_y^T \cdot Y'(t)$ between the unit normal ν_y and the directional derivative $Y'(t)$. The flow from Γ^c into Γ is the positive part $(\nu_y^T \cdot Y'(t))^+$.

The surface integral over an area $A \subseteq \partial\Gamma$,

$$\int_{y \in A} \left(\nu_y^T \cdot Y'(t)\right)^+ f_{Y(t)}(y)\, ds(y)$$

is the total flow at time t from Γ^c into Γ through the region A. Taking expectations, and integrating t over [0, 1] we obtain the mean flow per time unit as

$$\int_{t=0}^{1} \int_{y \in A} E\left(\left(\nu_y^T \cdot Y'(t)\right)^+ \mid Y(t) = y\right) f_{Y(t)}(y)\, ds(y)\, dt. \tag{4.4}$$

This is Belyaev's formula for the mean number of exits across the subset A by the process $Y(t)$.

However, Belyaev states that (4.4) holds if $Y'(t) | Y(t) = y$ has a density, which satisfies certain regularity conditions, and this is too restrictive for our purposes. In particular it does not cover the important example

$$Y(t) = (Y_1(t), Y_2(t))^T = \left(Y_1(t), \frac{d}{dt} Y_1(t)\right)^T$$

for which

$$Y'(t) = \left(Y_2(t), \frac{d}{dt} Y_2(t)\right)^T.$$

In order to use formula (4.4) for the wider class of processes we have in mind, we prefer to couple it to the density $f_{\Psi_t,\Psi_t'}$ of the distance function Ψ_t and its derivative Ψ_t', which was used in (4.3).

THEOREM 4.2. *The mean number of exits per time unit across* $\partial\Gamma$ *from* Γ^c *into* Γ *is given by the formulas*

$$E(N_1(\partial\Gamma)) = f_{\Psi_0}(0)E\big((\Psi_0')^+ \mid \Psi_0 = 0\big)$$

$$(4.5a) \qquad = \int_{z=0}^{\infty} z f_{\Psi_0,\Psi_0'}(0,z)\,dz$$

$$(4.5b) \qquad = \int_{y\in\partial\Gamma} q(y) f_{Y(0)}(y)\,ds(y),$$

where the function q, $R^p \nearrow R$, *is defined by*

$$q(y) = E\big((\nu_y^T\cdot Y'(0))^+ \mid Y(0) = y\big) = |\dot\Psi(y)|^{-1}E\big((\Psi_0')^+ \mid Y(0) = y\big).$$

PROOF. We show here that (4.5a) and (4.5b) give the same result, and defer the proof of (4.5a) to Section 5. Use the coordinate transformation (4.1), $\Theta = h(y) = (\theta^1,\cdots,\theta^{p-1},\theta^p)$ with the inverse $y = g(\Theta)$, and in which $\theta^p = \Psi(y)$, and then define the p-variate process $\xi_t = h(Y(t))$, for which the pth component is $\xi_t^p = \Psi_t$. It has the p-variate density

$$f_{\xi_t}(\Theta) = f_{Y(t)}(g(\Theta))\left|\frac{\partial g(\Theta)}{\partial\Theta}\right|.$$

We then obtain

$$\int_{z=0}^{\infty} z f_{\Psi_0,\Psi_0'}(0,z)\,dz$$

$$= \int_{\theta^1\cdots\theta^{p-1}}\int_{z=0}^{\infty} z f_{\xi_0,\Psi_0'}(\theta^1,\cdots,\theta^{p-1},0,z)\,dz\,d\theta^1\cdots d\theta^{p-1}$$

$$= \int_{\theta^1\cdots\theta^{p-1}} E\big((\Psi_0')^+ \mid \xi_0 = (\theta^1,\cdots,\theta^{p-1},0)\big).$$

$$f_{\xi_0}(\theta^1,\cdots,\theta^{p-1},0)\,d\theta^1\cdots d\theta^{p-1}.$$

Using

$$\Psi_0' = |\dot\Psi(Y(0))|\nu_{Y(0)}^T\cdot Y'(0)$$

and the function q, we see that this is equal to

$$\int_{\theta^1\cdots\theta^{p-1}} q(g(\theta)) f_{Y(0)}(g(\Theta))|\dot\Psi(g(\Theta))|\left|\frac{\partial g(\Theta)}{\partial\Theta}\right| d\theta^1\cdots d\theta^{p-1}$$

which by definition is equal to the surface integral

$$\int_{y\in\partial\Gamma} q(y) f_{Y(0)}(y)\,ds(y),$$

see Apostol (1962), page 293. Thus we have established the equality between (4.5a) and (4.5b). □

Now, for any Borel-set B in R^p, define

$$q(B, y) = E\big(1_{Y'(0)\in B}(\nu_y^T\cdot Y'(0))^+ \mid Y(0) = y\big),$$

so that $q(R^p, y) = q(y)$ in Theorem 4.2. By normalizing with $\gamma_Y(\partial\Gamma) = E(N_1(\partial\Gamma)) = \int_y q(y) f_{Y(0)}(y)\, ds(y) = \int_y \int_z q(dz, y) f_{Y(0)}(y)\, ds(y)$, we can then define a probability measure over $\partial\Gamma \times R^p$ by

$$\gamma_Y(\partial\Gamma)^{-1} q(dz, y) f_{Y(0)}(y)\, ds(y), \qquad y \in \partial\Gamma,\ z \in R^p. \tag{4.6}$$

This measure is fundamental for the rest of this paper. If $Y'(0)$ has a density, conditional on $Y(0) = y$, it can be written

$$\gamma_Y(\partial\Gamma)^{-1}\left(\nu_y^T \cdot z\right)^+ f_{Y'(0)|Y(0)=y}(z) f_{Y(0)}(y)\, dz\, ds(y).$$

We now return to the exits $\hat{t}_k$ and the behaviour of $X(\hat{t}_k + \cdot)$ around these exits. Recalling the notation

$$N_T(B; \partial\Gamma) = \#\left\{\hat{t}_k; 0 \leqslant \hat{t}_k \leqslant T, X\left(\hat{t}_k + s\right) \in B\right\}$$

for the number of B-restricted Y-exits, we can obtain the following theorem, in analogy with Theorem 4.2.

THEOREM 4.3. *The mean number of B-restricted exits per time unit is given by*

$$E(N_1(B; \partial\Gamma)) = f_{\Phi_0}(0) E\left(1_{X(s)\in B}(\Psi_0')^+ \mid \Psi_0 = 0\right) \tag{4.7a}$$

$$= \int_{y\in\partial\Gamma} E\left(1_{X(s)\in B}\left(\nu_y^T \cdot Y'(0)\right)^+ \mid Y(0) = y\right) f_{Y(0)}(y)\, ds(y) \tag{4.7b}$$

(4.7c)

$$= \int_{y\in\partial\Gamma}\int_{z\in R^p} P(X(s) \in B \mid Y(0) = y, Y'(0) = z) q(dz, y) f_{Y(0)}(y)\, ds(y).$$

PROOF. The proof of (4.7a) is similar to that of (4.5a) and is deferred to Section 5. To see that (4.7a) and (4.7c) agree, proceed as in the proof of Theorem 4.2 but replace

$$\int_{z=0}^{\infty} z f_{\Psi_0, \Psi_0'}(0, z)\, dz$$

by

$$\int_{z=0}^{\infty} P(X(s) \in B \mid \Psi_0 = 0, \Psi_0' = z) z f_{\Psi_0, \Psi_0'}(0, z)\, dz.$$

To obtain the intermediate form (4.7b), start from (4.7c) and write $p(\cdot\,, y)$ for the conditional distribution on R^p for $Y'(0)$ given $Y(0) = y$. Then $q(\cdot\,, y)$ is absolutely continuous with respect to $p(\cdot\,, y)$ with density

$$\frac{dq(\cdot\,, y)}{dp(\cdot\,, y)}(z) = \left(\nu_y^T \cdot z\right)^+,$$

and it follows that

$$\begin{aligned} E&\left(1_{X(s)\in B}\left(\nu_y^T \cdot Y'(0)\right)^+ \mid Y(0) = y\right) \\ &= \int_{z\in R^p} E\left(1_{X(s)\in B} \mid Y(0) = y, Y'(0) = z\right)\left(\nu_y^T \cdot z\right)^+ p(dz, y) \\ &= \int_{z\in R^p} E\left(1_{X(s)\in B} \mid Y(0) = y, Y'(0) = z\right) q(dz, y), \end{aligned}$$

which shows the equality between (4.7b) and (4.7c). □

We have now obtained the instrument to construct a model process $X_{\partial\Gamma}(t)$ in analogy with $Y_u(t)$ in which the assumed normality of X and Y is fully exploited. From Theorem 4.3, formula (4.7c) we first obtain the following theorem.

THEOREM 4.4. *If $\{X(t), Y(t), t \in R\}$ is ergodic and $B \subseteq R^r$ is open, then with probability one*

$$\lim_{T\to\infty} \frac{N_T(B;\, \partial\Gamma)}{N_T(\partial\Gamma)} = \int_{y\in\partial\Gamma}\int_{z\in R^p} P(X(s) \in B \mid Y(0) = y,\ Y'(0) = z)$$

$$\cdot\, \gamma_Y(\partial\Gamma)^{-1} q(dz,\, y) f_{Y(0)}(y)\, ds(y).$$

The following theorem is the multivariate analogue to Corollary 3.2.

THEOREM 4.5. *If $\{Y(t), t \in R\}$ is ergodic, the long-run joint distribution of the location $Y(\hat{t}_k)$ and the directional derivative $Y'(\hat{t}_k)$ at the exits of $Y(t)$ across $\partial\Gamma$, is given by*

$$\gamma_Y(\partial\Gamma)^{-1}\left(\nu_y^T\cdot z\right)^+ p(dz,y) f_{Y(0)}(y)\, ds(y), \qquad y \in \partial\Gamma, z \in R^p,$$

where $p(dz,y)$ is the conditional distribution of $Y'(0)$ given that $Y(0) = y$.

PROOF. Take $X(t) = (Y(t), Y'(t))$ and use Theorem 4.4 with $q(dz,y) = (\nu_y^T\cdot z)^+ p(dz,y)$. A strict proof will be given in Section 5 in the framework of marked exits, since Theorem 4.4 is only formulated for one-dimensional $X(t)$. □

Theorem 4.5 gives the p-variate version of the long-run Rayleigh distribution of the derivative at an upcrossing, and it enables us to weight the conditional probability in Theorem 4.4 to give a long-run interpretable total. The following lemma parallels Lemma 3.3.

LEMMA 4.6. *If $Y(t)$ is differentiable, the conditional distributions of*

$$X(t) \mid Y(0) = y = (y_1,\cdots,y_p)^T, \qquad Y'(0) = z = (z_1,\cdots,z_p)^T$$

for $t \in R$ are univariate normal with mean

$$m_{y,z}^X(t) = (r_{XY}(t), -r'_{XY}(t)) r_{YY'}(0)^{-1}\binom{y}{z}$$

and covariance function

$$r_{X|Y}(s,t) = r_X(s-t) - (r_{XY}(s), -r'_{XY}(s)) r_{YY'}(0)^{-1}\begin{bmatrix} r_{XY}(t)^T \\ -r'_{XY}(t)^T \end{bmatrix}.$$

Now define a nonstationary normal process $\kappa_{X|Y}(t)$ with mean zero and covariance function $r_{X|Y}(s,t)$ defined above. Also define a $2p$-variate random variable (η_Y, ζ_Y) with values in $\partial\Gamma \times R^p$, independent of $\kappa_{X|Y}$ and with a distribution defined by (4.6), which we from now on denote

$$f_{\eta_Y,\zeta_Y}(y, dz)\, ds(y) = \gamma_Y(\partial\Gamma)^{-1} q(dz,y) f_{Y(0)}(y)\, ds(y).$$

This variable will replace ζ_X in Theorem 3.4 and shall illustrate the random variations of $Y(\hat{t}_k)$ and $Y'(\hat{t}_k)$. If inserted instead of y, z in the mean value function $m^X_{y,z}$ in Lemma 4.6 it yields a random function

$$m^X_{\eta_Y,\zeta_Y}(t) = \left(r_{XY}(t), -r'_{XY}(t)\right) r_{YY'}(0)^{-1}\begin{pmatrix}\eta_Y\\ \zeta_Y\end{pmatrix}.$$

THEOREM 4.7. *If $\{X(t), Y(t), t \in R\}$ is ergodic, the long-run finite-dimensional distributions of $X(\hat{t}_k + \cdot)$ after exits across $\partial\Gamma$ by $Y(\hat{t}_k)$ are given by*

$$\lim_{T\to\infty} \frac{\#\left\{\hat{t}_k;\, 0 \leqslant \hat{t}_k \leqslant T,\ X\left(\hat{t}_k + \cdot\right) \in B\right\}}{\#\left\{\hat{t}_k;\, 0 \leqslant \hat{t}_k \leqslant T\right\}}$$
$$= \int_{y\in\partial\Gamma}\int_{z\in R^p} f_{\eta_Y,\zeta_Y}(y, dz) P\left(m^X_{y,z}(\cdot) + \kappa_{X|Y}(\cdot) \in B\right) ds(y)$$
$$= P\left(m^X_{\eta_Y,\zeta_Y}(\cdot) + \kappa_{X|Y}(\cdot) \in B\right).$$

PROOF. The result follows simply from Theorem 4.4 and Lemma 4.6. □

The theorem motivates that we use the process

$$X_{\partial\Gamma}(t) = m^X_{\eta_Y,\zeta_Y}(t) + \kappa_{X|Y}(t) \tag{4.8}$$

as a model process for $X(\hat{t}_k + \cdot)$ and use it to calculate the error probability for a false alarm by means of

$$G_m(u; \Gamma) = P\left(X_{\partial\Gamma}(t) \text{ has no } u\text{-upcrossing for } t \in (0, m)\right).$$

EXAMPLE 4.8. Let $Y(t) = (Y_1(t), \cdots, Y_p(t))^T$ have independent normal components, each with mean zero and variance one, and let the alarm region be the complement of the sphere

$$\Gamma = \left\{y \in R^p;\ \|y\| = \left(\Sigma_{i=1}^p y_i^2\right)^{\frac{1}{2}} \leqslant r\right\},$$

and take $\Phi(y) = \|y\|^2 = \Sigma_1^p y_i^2$. In engineering sciences this is used as a reasonable model for symmetric loadings on circular structures; see Sharpe (1978).

To obtain the distribution of the location and direction of $Y(t)$ at the crossing points, we use Theorem 4.5 and simplify the distribution

$$\gamma_Y(\partial\Gamma)^{-1}\left(\nu_y^T \cdot z\right)^+ p(dz, y) f_{Y(0)}(y)$$

where $p(dz, y)$ denotes the conditional distribution of $Y'(0)$ given that $Y(0) = y$. In this case $Y(t)$ and $Y'(t)$ are independent, and since furthermore the components of $Y'(t) = (Y_1'(t), \cdots, Y_p'(t))^T$ are independent and normal with mean zero and variances $V(Y_i'(t)) = \lambda_i$, say, we have

$$p(dz, y) = \Pi_{i=1}^p (2\pi\lambda_i)^{-\frac{1}{2}} \exp\left(-z_i^2/2\lambda_i\right) dz.$$

Further, since the unit normal ν_y at y is equal to $r^{-1}y$, we obtain the following density for the location $Y(\hat{t}_k)$ and direction $Y'(\hat{t}_k)$ at the exit points:

$$\gamma_Y(\partial\Gamma)^{-1} r^{-1}(2\pi)^{-p/2} \exp(-r^2/2)(y^T \cdot z)^+ \exp(-\|z\|_\lambda^2/2), \tag{4.9}$$

where we have written $\|z\|_\lambda^2 = \Sigma_1^p z_i^2/\lambda_i$.

To obtain the model process $X_{\partial\Gamma}$ we now define the variables η_Y and ζ_Y with values in $\partial\Gamma$ and R^p, and with density $f_{\eta_Y,\zeta_Y}(y, z)$ given by (4.9). Due to the simple structure of $Y(t)$, the function $m^X_{y,z}(t)$ and the covariance function $r_{X|Y}(s, t)$ are quite simple. Writing $\Lambda = \text{diag}(\lambda_i)$ we have

$$r_{YY'}(0) = \begin{pmatrix} I & 0 \\ 0 & \Lambda \end{pmatrix}, \qquad r_{YY'}(0)^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & \Lambda^{-1} \end{pmatrix},$$

so that

$$m^X_{y,z}(t) = r_{XY}(t)\cdot y - r'_{XY}(t)\Lambda^{-1}\cdot z$$

and

$$r_{X|Y}(s, t) = r_X(s - t) - r_{XY}(s)r_{XY}(t)^T - r'_{XY}(s)\Lambda^{-1}r'_{XY}(t)^T.$$

The model process is

$$X_{\partial\Gamma}(t) = r_{XY}(t)\cdot\eta_Y - r'_{XY}(t)\Lambda^{-1}\cdot\zeta_Y + \kappa_{X|Y}(t).$$

EXAMPLE 4.9. Let, as in the previous example, $Y(t) = (Y_1(t), \cdots, Y_p(t))^T$ consist of independent processes with $V(Y_i(t)) = 1$, $V(Y_i'(t)) = \lambda_i$, and define the alarm region by

$$\Gamma = \{y \in R^p; \max_i y_i > \hat{u}\}.$$

This region could be possible if p independent processes can give alarm independently of each other.

Since $\partial\Gamma = \{y \in R^p; \max_i y_i = \hat{u}\} = \cup_{i=1}^p \partial\Gamma_i$ where $\partial\Gamma_i = \{y \in R^p; y_i = \hat{u}, y_j \leqslant \hat{u} \text{ for } j \neq i\}$ we can define the distribution of (η_Y, ζ_Y) over each $\partial\Gamma_i$ separately. If $y \in \partial\Gamma_i$, the unit normal is $\nu_y = \delta_i$, the ith unit vector, so we get for $y \in \partial\Gamma_i$

$$f_{\eta_Y,\zeta_Y}(y,z) = c\left(\nu_y^T\cdot z\right)^+ e^{-\|z\|^2_\lambda/2}\cdot e^{-\|y\|^2/2} \tag{4.10}$$

$$= cz_i^+ e^{-z_i^2/2\lambda_i}\cdot e^{-\Sigma_{j\neq i}z_j^2/2\lambda_j}\cdot e^{-\Sigma_{j=1}^p y_j^2/2}.$$

The constant c is determined by integrating $f_{\eta_Y,\zeta_Y}(y,z)$ over $\partial\Gamma$, and we obtain

$$1 = \int_{y\in\partial\Gamma}\int_{z\in R^p} f_{\eta_Y,\zeta_Y}(y,z)\,dz\,ds(y)$$

$$= \Sigma_{i=1}^p \int_{y\in\partial\Gamma_i}\int_{z\in R^p} cz_i^+ e^{-z_i^2/2\lambda_i - \Sigma_{j\neq i}z_j^2/2\lambda_j - \hat{u}^2/2 - \Sigma_{j\neq i}y_j^2/2}\,dz\,dy^i$$

where $dy^i = dy_1 \cdots dy_{i-1}\,dy_{i+1} \cdots dy_p$.

We immediately find that this is equal to

$$ce^{-\hat{u}^2/2}\Sigma_{i=1}^p \int_{y_j<\hat{u}} e^{-\Sigma_{j\neq i}y_j^2/2}\,dy^i \int_{z_i>0} z_i e^{-z_i^2/2\lambda_i}\int e^{-\Sigma_{j\neq i}z_j^2/2\lambda_j}\,dz^i$$

$$= ce^{-\hat{u}^2/2}\cdot(\Phi(\hat{u}))^{p-1}(2\pi)^{p-1}\Sigma_{i=1}^p\lambda_i\Pi_{j\neq i}(\lambda_j)^{\frac{1}{2}}$$

which gives

$$c = \left\{e^{-\hat{u}^2/2}(\Phi(\hat{u}))^{p-1}(2\pi)^{p-1}\cdot\Pi_{j=1}^p(\lambda_j)^{\frac{1}{2}}\cdot\Sigma_{j=1}^p(\lambda_j)^{\frac{1}{2}}\right\}^{-1}.$$

Proceeding as in Example 4.8 we obtain a model process of the same type,

$$X_{\partial\Gamma}(t) = r_{XY}(t)\cdot\eta_Y - r'_{XY}(t)\Lambda^{-1}\cdot\zeta_Y + \kappa_{X|Y}(t),$$

where (η_Y, ζ_Y) has the density (4.10) over $\partial\Gamma_i = \{y \in R^p; y_i = \hat{u}, y_j \leqslant \hat{u} \text{ for } j \neq i\}, i = 1, \cdots, p$.

Note that if this alarm system gives alarm at the boundary $\partial\Gamma_i$, so that $Y_i(t)$ has an upcrossing of $\hat{u}$, the situation is not the same as when $Y_i(t)$ works as a single predictor, since we here know, in fact, that $Y_j(t) < \hat{u}, j \neq i$. This will probably imply that the probability of detection, $1 - H$, is not improved by a factor p, but by something less than p, compared to prediction from $Y_i(t)$ alone.

5. A theorem on marked crossing processes. In Section 4 we left out the justification for the formulas for the mean number of exits across a boundary $\partial\Gamma$, at which $X(t)$ or $Y(t)$ satisfies some special requirement. In fact, Theorem 3.1, Corollary 3.2, Theorem 3.5, Theorem 4.1, formulas (4.5a) and (4.7a), and Theorem 4.5 will all follow from a general theorem on the average number of marked exits, to be given here.

Let $\{\xi_t, t \in R\}$ be a stationary, real valued, differentiable stochastic process defined on some probability space $\{\Omega, \mathfrak{F}_1, P\}$, and let $\{\eta_t, t \in R\}$ be a family of random elements, also defined on $\{\Omega, \mathfrak{F}, P\}$ but with values in some topological space $\{S, \mathfrak{S}\}$. Suppose further that η_t is jointly stationary with ξ_t.

In the applications we have in mind the state space S is R^p, R^{pr}, or C, the set of continuous functions with the topology of uniform convergence on compact sets. For Theorem 4.5 we may, e.g., take $\xi_t = \Psi_t = \Psi(Y(t))$ and $\eta_t = (Y(t), Y'(t))$, while for formula (4.7a) we can take $\eta_t = X(t + s) \in R^r$.

We will consider η_t as a mark attached to the process ξ at time t. Define, for $A \in \mathfrak{S}$, the number of A-restricted zeros of ξ_t as

$$N(A) = \#\{t \in [0, 1]; \xi_t = 0, \eta_t \in A\},$$

which is the number of times that $\xi_t = 0$ and simultaneously $\eta_t \in A$. (That this is, indeed, a random variable will follow from the construction below). The following theorem about the average number of marked exits will be stated and proved along similar lines as in Marcus (1977), with a method dating back to Kac (1943).

Define, for $n = 1, 2, \cdots$, the polynomial approximation $\xi_{t,n}$ by taking $\xi_{t,n} = \xi_t$ for $t = k/2^n, k = 0, 1, \cdots, 2^n$, and linear in between, and let $\eta_{t,n}$ be the piecewise constant approximation of η_t defined as $\eta_{t,n} = \eta_{k/2^n}$ if $k/2^n \leqslant t < (k+1)/2^n$. Finally, write for any fixed set $A \in \mathfrak{S}$,

$$g_t(x) = f_{\xi_t}(x) E\big(1_{\eta_t \in A} |\xi'_t| \,\|\, \xi_t = x\big)$$

and

$$g_{t,n}(x) = f_{\xi_{t,n}}(x) E\big(1_{\eta_{t,n} \in A} |\xi'_{t,n}| \,\|\, \xi_{t,n} = x\big).$$

Note that since conditional expectations are defined only up to equivalence, we need some sort of continuity condition to ensure identifiability. In particular, we require

$$f_{\xi_t}(0)E\big(1_{\eta_t\in A}|\xi_t'||\xi_t=0\big) = g_t(0) = \lim_{\Delta\to 0}\frac{1}{2\Delta}\int_{-\Delta}^{\Delta}g_t(x)\,dx.$$

The following conditions are needed in order to justify the approximation of the pair (ξ_t, η_t) by $(\xi_{t,n}, \eta_{t,n})$:

(5.1) for all n, there exist $M_n, \delta > 0$ such that $g_{t,n}(x) \leqslant M_n$ for all $|x| < \delta, t \in [0,1]$;

(5.2) $$\lim_{\Delta\to 0}\frac{1}{2\Delta}\int_{-\Delta}^{\Delta}g_{t,n}(x)\,dx = g_{t,n}(0) \text{ for all } n \text{ and } t \in [0,1];$$

(5.3) there exists M such that $g_{t,n}(0) \leqslant M$ for all n and $t \in [0,1]$;

(5.4) $$\lim_{n\to\infty} g_{t,n}(0) = g_t(0).$$

Note that in many applications

$$f_{\xi_{t,n}}(x)E(|\xi_{t,n}'||\xi_{t,n} = x)$$

can be chosen to be uniformly bounded in n, t and x, and then (5.1) and (5.3) are trivially satisfied.

We further have to require that the set $A \subseteq S$ is open in the weak topology, i.e., we assume that

(5.5) the set $\eta^{-1}A(\omega) = \{t \in (0,1);\ \eta_t(\omega) \in A\}$ is an open set, for almost all outcomes ω in Ω.

In the following theorem we write $N = N(S) = \#\{t \in [0,1];\ \xi_t = 0\}$.

THEOREM 5.1. *Suppose* $\{\xi_t, \eta_t, t \in R\}$ *satisfies* (5.1)–(5.4), *A is open in the sense of* (5.5) *and that* $E(N) < \infty$. *Then the mean number of A-restricted zeros of ξ_t is*

$$E(N(A)) = f_{\xi_0}(0)E\big(1_{\eta_0\in A}|\xi_0'||\xi_0 = 0\big)$$

provided $f_{\xi_0}(x)E(1_{\eta_0\in A}|\xi_0'||\xi_0 = x)$ *is continuous at* $x = 0$.

PROOF. The proof is similar to that of Theorem 2.1 in Marcus (1977). Define, for any $s \in S$, $x \in R$, $\Delta > 0$,

$$\chi_\Delta(x,s) = \begin{matrix} 1 & \text{if } |x| < \Delta, s \in A, \\ 0 & \text{otherwise,} \end{matrix}$$

and let

$$N_n(A) = \#\{t \in [0,1];\ \xi_{t,n} = 0,\ \eta_{t,n} \in A\}.$$

Then with probability one,

$$N_n(A) = \lim_{\Delta\to 0}\frac{1}{2\Delta}\int_0^1\chi_\Delta(\xi_{t,n},\eta_{t,n})|\xi_{t,n}'|\,dt$$

and it holds

(5.6) $$\frac{1}{2\Delta}\int_0^1\chi_\Delta(\xi_{t,n},\eta_{t,n})|\xi_{t,n}'|\,dt \leqslant 2^n.$$

This follows exactly as Lemma 3.2 in the paper by Marcus, if we notice that $\eta_{t,n}$ is constant in each of the intervals $[k/2^n, (k+1)/2^n)$, and that for almost all realizations $\xi_{t,n} \neq 0$ for $t = k/2^n$, $k = 0, \cdots, 2^n$. (This is the classical zero counting device of Kac (1943).)

Since A is open, we can also conclude that

$$N_n(A) \to N(A) \quad \text{as } n \to \infty,$$

since suppose $\xi_t = 0$ for a finite number of times $t_1, \cdots, t_{m'}$, and that at m of these $\eta_t \in A$, say at $t_1, \cdots, t_m$. Due to (5.5), $t_1, \cdots, t_m$ are interior points in $\eta^{-1}A$ so that $\eta_t \in A$ for all t in a neighbourhood of $t_1, \cdots, t_m$. As $n \to \infty$ these m time points will be counted by $N_n(A)$ and no else if $N(A) = m < \infty$, so that $N_n(A) \to N(A)$.

Incidentally, we have the promised argument that $N(A)$ is a random variable, since $N_n(A)$ is.

Since $N_n(A) \leqslant N$ and $E(N) < \infty$, we have that $E(N(A)) = \lim_{n\to\infty} E(N_n(A))$, so our task is to show that $\lim_{n\to\infty} E(N_n(A)) = g_0(0)$. To do this we evaluate $\lim E(N_n(A))$ using (5.1)–(5.4), and first observe that (5.6) implies that

$$\lim_{n\to\infty} E(N_n(A)) = \lim_{n\to\infty} \lim_{\Delta\to 0} E\left\{ \frac{1}{2\Delta} \int_0^1 \chi_\Delta(\xi_{t,n}, \eta_{t,n}) |\xi'_{t,n}| \, dt \right\}.$$

Expanding the expectation we obtain the expression

$$\lim_{n\to\infty} \lim_{\Delta\to 0} \frac{1}{2\Delta} \int_0^1 \int_{x=-\Delta}^{\Delta} f_{\xi_{t,n}}(x) E\left(1_{\eta_{t,n}\in A} |\xi'_{t,n}| \,|\, \xi_{t,n} = x\right) dx\, dt$$

$$= \lim_{n\to\infty} \lim_{\Delta\to 0} \int_0^1 \frac{1}{2\Delta} \int_{x=-\Delta}^{\Delta} g_{t,n}(x)\, dx\, dt = \lim_{n\to\infty} \int_0^1 g_{t,n}(0)\, dt,$$

using (5.1) and (5.2). Finally (5.3) and (5.4) imply that

$$\lim_{n\to\infty} E(N_n(A)) = \int_0^1 g_t(0)\, dt = g_0(0) = f_{\xi_0}(0) E\left(1_{\eta_0 \in A} |\xi'_0| \,|\, \xi_0 = 0\right),$$

which was to be proved. □

PROOF OF THEOREM 3.1 AND COROLLARY 3.2. For Theorem 3.1 take $\xi_t = X(t) - u$ and $\eta_t = (X'(t), Y(t+s)) = (X'(t), Y(t+s_1), \cdots, Y(t+s_r))$, $s = (s_1, \cdots, s_r)$, let A be an open pr-dimensional set, and define

$$A' = \{(y,z) \in R \times R^{\mathrm{pr}}; y > 0, z \in A\}.$$

Then, writing $p_{t,n}(x,y)$ for the joint density of $\xi_{t,n}$ and $\xi'_{t,n}$, we have

$$g_{t,n}(x) = \int_{-\infty}^{\infty} |y| p_{t,n}(x,y) P(\eta_{t,n} \in A' | \xi_{t,n} = x, \xi'_{t,n} = y)\, dy$$

$$\leqslant \int_{-\infty}^{\infty} |y| p_{t,n}(x,y)\, dy = f_{\xi_{t,n}}(x) E(|\xi'_{t,n}| \,|\, \xi_{t,n} = x),$$

and it is standard normal theory to show that (5.1) and (5.3) are satisfied. Since A' is open, the probability $P(\eta_{t,n} \in A' | \xi_{t,n} = x, \xi'_{t,n} = y)$ is a continuous function of x for all t and n, which implies that $g_{t,n}(x)$ is also continuous. Convergence of $g_{t,n}(0)$ as $n \to \infty$ follows by normal theory, so that (5.1)–(5.5) are actually satisfied.

Theorem 5.1 then gives that

$$E(N_1(A;\, u)) = f_{X(0)}(u) E\big(1_{\eta_0 \in A'} | X'(0) \| X(0) = u\big)$$
$$= f_{X(0)}(u) E\big(1_{Y(s)\in A}(X'(0))^+ \,| X(0) = u\big),$$

which implies Theorem 3.1.

For Corollary 3.2, just take $\eta_t = X'(t)$ and proceed as above. ▯

PROOF OF THEOREM 3.5. We have to show that

$$\frac{E(\#t_k \in [0,1];\ Y(t_k+\tau) \text{ enters } \Gamma \text{ for some } \tau \in (-m, 0))}{E(\#t_k \in [0,1])} \tag{5.7}$$
$$= P(Y_u(\tau) \text{ enters } \Gamma \text{ for some } \tau \in (-m,0)).$$

Take S to be the space C^p of p-variate continuously differentiable functions $y(\cdot) = (y_1(\cdot), \cdots, y_p(\cdot))^T$ with the topology of uniform convergence on compact sets for each $y_i(\cdot)$, and define the set

$$A = \{y \in S;\ y(\tau) \text{ enters } \Gamma \text{ for some } \tau \in (-m, 0)\}.$$

If we, for any outcome ω, take $\eta_t = Y(t + \cdot) \in S$ to be the Y-process shifted the time t, we have that, with probability one,

$$\eta^{-1}A(\omega) = \{t \in (0,1);\ \eta_t(\omega) \in A\}$$
$$= \{t \in (0,1);\ Y(t+\tau) \text{ enters } \Gamma \text{ for some } \tau \in (-m,0)\}$$

is an open subset of the interval (0, 1) so that Theorem 5.1 implies that

$$E(\#t_k \in [0,1];\ Y(t_k + \tau) \text{ enters } \Gamma \text{ for some } \tau \in (-m, 0))$$
$$= f_{X(0)}(u) E\big(1_{\eta_0 \in A}(X'(0))^+ \,| X(0) = u\big)$$
$$= f_{X(0)}(u) E\big(1_{Y(\cdot)\in A}(X'(0))^+ \,| X(0) = u\big).$$

The ratio (5.7) is therefore equal to

$$\frac{E\big(1_{Y(\cdot)\in A}(X'(0))^+ \,| X(0) = u\big)}{E((X'(0))^+ \,| X(0) = u)},$$

which, by the definition of $Y_u(\cdot)$ equals $P(Y_u(\cdot) \in A)$. This concludes the proof of the assertion. ▯

PROOF OF (4.5a) AND (4.7a), LEADING TO THEOREM 4.1. Take $\xi_t = \Psi_t$ and $\eta_t = (X(t+s), \Psi_t') \in R^r \times R$, and proceed as in the proof of Theorem 3.1. Whether (5.1)–(5.5) are satisfied or not depends on the smoothness of the function $\Psi(y)$, and we do not embark upon a general treatment here. ▯

PROOF OF THEOREM 4.5. Take $\eta_t = (Y(t), Y'(t), \Psi_t') \in R^p \times R^p \times R$, and let $A_1 \subseteq \partial\Gamma$ be relatively open and $A_2 \subseteq R^p$ open. Take $A \subseteq R^p \times R^p$ such that $A \cap (\partial\Gamma \times R^p) = A_1 \times A_2$, and define

$$A' = \{(x, y, z) \in R \times R^p \times R^p;\ x > 0, (y, z) \in A\}.$$

Then

$$E\left(1_{\eta_0 \in A'} | \xi_0' | | \xi_0 = 0\right) = E\left(1_{Y(0) \in A_1,\ Y'(0) \in A_2} (\Psi_0')^+ | \Psi_0 = 0\right),$$

and we get the mean number of exits across $\partial\Gamma$ such that $Y(\hat{t}_k) \in A_1$ and $Y'(\hat{t}_k) \in A_2$ as

$$f_{\Phi_0}(0) E\left(1_{Y(0) \in A_1,\ Y'(0) \in A_2} (\Psi_0')^+ | \Psi_0 = 0\right)$$

$$= \int_{y \in A_1} \int_{z \in A_2} \left(\nu_y^T \cdot z\right)^+ p(dz, y) f_{Y(0)}(y)\, ds(y),$$

where the last equality follows as in Theorem 4.2. □

Acknowledgment. I am grateful to the referee for much valuable criticism.

REFERENCES

[1] Apostol, T. M. (1962). *Calculus, Vol.* 2. Blaisdell, New York.
[2] Belyaev, Yu. K. (1968), On the number of exits across the boundary of a region by a vector stochastic process. *Theor. Probability Appl.* **13** 320–324.
[3] Kac, M. (1943). On the average number of real roots of a random algebraic equation. *Bull. Amer. Math. Soc.* **49** 314–320.
[4] Leadbetter, M. R. (1966). On crossings of levels and curves by a wide class of stochastic processes. *Ann. Math. Statist.* **37** 260–267.
[5] Lindgren, G. (1975a). Prediction from a random time point. *Ann. Probability* **3** 412–423.
[6] Lindgren, G. (1975b). Prediction of catastrophes and high level crossings. *Bull. Internat. Statist. Inst.* **46** Book 2, 225–240.
[7] Lindgren, G. (1979). Prediction of level crossings for normal processes containing deterministic components. *Adv. Appl. Probability* **11** 93–117.
[8] Marcus, M. B. (1977). Level crossings of a stochastic process with absolutely continuous sample paths. *Ann. Probability* **5** 52–71.
[9] de Maré, J. (1980). Optimal prediction of catastrophes with applications to Gaussian processes. *Ann. Probability* 841–850.
[10] Sharpe, K. (1978). Some properties of the crossings process generated by a stationary χ^2 process. *Adv. Appl. Probability* **10** 373–391.
[11] Slepian, D. (1962). On the zeros of Gaussian noise. In *Time Series Analysis*. (M. Rosenblatt, Ed.) 104–115. Wiley, New York.
[12] Sveshnikov, A. A. (1968). *Problems in Probability Theory, Mathematical Statistics and Theory of Random Functions*. Saunders, Philadelphia.

Department of Mathematical Statistics
University of Lund
Box 725
S-220 07 Lund
Sweden