Journal
of The Franklin Institute
DEVOTED TO SCIENCE AND THE MECHANIC ARTS

Volume 295, Number 3 *March 1973*

Solution of the Stochastic Control Problem in Unbounded Domains

by PRENTISS ROBINSON
Department of Electrical Engineering
University of Maryland, College Park, Maryland
and JOHN MOORE*
Department of Electrical Engineering
University of Newcastle, New South Wales, Australia

ABSTRACT: *Bellman's dynamic programming equation for the optimal index and control law for stochastic control problems is a parabolic or elliptic partial differential equation frequently defined in an unbounded domain. Existing methods of solution require bounded domain approximations, the application of singular perturbation techniques or Monte Carlo simulation procedures.*

In this paper, using the fact that Poisson impulse noise tends to a Gaussian process under certain limiting conditions, a method which achieves an arbitrarily good approximate solution to the stochastic control problem is given. The method uses the two iterative techniques of successive approximation and quasi-linearization and is inherently more efficient than existing methods of solution.

I. Introduction

We consider an important class of nonlinear stochastic control problems. In particular, we consider the optimal control of a system which can be adequately modelled by the stochastic vector differential equation

$$\mathrm{d}x = f(x,t;\, u)\,\mathrm{d}t + \sigma\,\mathrm{d}w, \tag{1}$$

where $x(t)$ is an n-vector random process, $u(t)$ is an r-vector control and $w(t)$ is an n-vector Wiener process with $\dot{w}$ being white Gaussian noise.

In order to calculate an optimal control $u(t)$ for Eq. (1) (to minimize some general performance index), we require the solution of Bellman's dynamic programming equation for the optimal index. This equation is unfortunately

---

* Dr. Moore's work was partially supported by the Australian Research Grants Committee. The computer time was supported by National Aeronautics and Space Administration Grant NsG–398 to the Computer Science Center of the University of Maryland.

a nonlinear parabolic or elliptic partial differential equation, the solution of which is very difficult except in a few special cases such as the case of a linear system with quadratic performance index (**1**, **2**). This difficulty arises since the equation is defined in an unbounded domain and the boundary conditions at infinity are not known *a priori*.

Singular perturbation techniques (**6**, **7**) have been used to achieve approximate solutions. However, accuracy is poor except in the vicinity of the terminal region as indicated by an example studied in the last section of this paper.

In Ref. (**8**), this equation is solved by writing a set of difference equations and solving these using the theory of Markov chains. The difficulty with this approach is that an infinite bounded domain is approximated by a bounded domain, determined using trial and error. There is obviously a compromise between accuracy and calculation time. The Monte Carlo simulation methods (**6**), although tedious, are reasonably reliable but accuracy problems are encountered in simulating white noise.

In this paper we consider a system with a Poisson impulse noise which is a generalization of white Gaussian noise. Now, since the Poisson impulse noise dynamic programming equation is a "first"-order differential-difference equation, its solution at infinity is well defined. This type of noise is a good model for many disturbances but also under suitable conditions, Poisson impulse noise tends to white Gaussian noise and it turns out that the Poisson solution can be used as the boundary condition for the Gaussian noise problem. The method of solution uses the two iterative techniques of successive approximation and quasi-linearization.

In the next section, a description of the Gaussian noise problem is given, and in Section III, the related "Poisson impulse" problem is defined and solved. Examples are given in Section IV, together with a comparison of methods.

II. Equations for Gaussian Noise Case

Consider a dynamical system with white Gaussian noise disturbances described by the stochastic vector differential equation (**1**) repeated here for convenience

$$\mathrm{d}x = f(x,t;\,u)\,\mathrm{d}t + \sigma\,\mathrm{d}w. \tag{1}$$

Without loss of generality, the matrix $\sigma(x,t)$ is assumed to be diagonal with components σ_i, $i = 1, 2, \ldots, n$. The control $u(x,t)$ takes values in a closed bounded set U and the performance index for a particular control is

$$J(x,t;\,u) = E\left\{B(x(t_s)) + \int_t^{t_s} L(x(\tau),\tau)\,\mathrm{d}\tau \,\middle|\, x(t) = x\right\}, \tag{2}$$

where E is the expectation operator, t_s is the random terminal time when $x(t)$ first reaches a terminal region S and B is the terminal or boundary penalty. The cost function L is assumed to be non-negative definite.

The optimal index $V(x,t)$ is defined by

$$V(x,t) = \inf_{u} J(x,t;\, u) \tag{3}$$

and it satisfies Bellman's dynamic programming equation

$$\min_{u\in U}\left\{\frac{1}{2}\sum_{i=1}^{n}\left[\sigma_i^2\frac{\partial^2 V}{\partial x_i^2}+f_i(u)\frac{\partial V}{\partial x_i}\right]+\frac{\partial V}{\partial t}+L(u)\right\} = 0 \tag{4}$$

and $V(x,t) = B(x)$ for $x\in\partial S$, the boundary of S. This equation is a nonlinear parabolic partial differential equation [see (**1, 2**)], and for the cases when $\partial V/\partial t = 0$, this is an elliptic equation—such cases arise when the dynamical system and cost function are time invariant.

Bellman's method of quasi-linearization [approximation in policy space (**9, 10**)] may be used to solve Eq. (4). This method requires the solutions of a series of linear equations of the form

$$\sum_{i=1}^{n}\left[\frac{\sigma_i^2}{2}\frac{\partial^2 V}{\partial x_i^2}+f_i\frac{\partial V}{\partial x_i}\right]+\frac{\partial V}{\partial t}+L = 0 \tag{5}$$

with boundary conditions $V(x,t) = B(x)$ for $x\in\partial S$.

Of course, this Eq. (5) is not really well defined without a second boundary condition for infinite x, but this is not *a priori* known. What we do know, however, is that under certain limiting conditions and constraints a Poisson step process tends to a Wiener process. We now examine this more closely.

III. Relationship between Poisson and Gaussian Cases

Consider a system described by the stochastic differential equation

$$\mathrm{d}x = f(x,t;\, u)\,\mathrm{d}t+\sigma\,\mathrm{d}p, \tag{6}$$

where $p(t)$ is a suitably constructed Poisson step process. For this system —the case of a dynamical system with Poisson impulse disturbances ($\dot{p}$)— consider (for simplicity) that the n-vector Poisson step process $p(t)$ has independent components with steps $\pm h$, each with probability $\frac{1}{2}$, and the mean-rate-of-occurrence of each step is λ. In order that the Poisson process becomes indistinguishable from the Wiener process of the previous section, we introduce the constraint

$$h^2\lambda = 1. \tag{7}$$

The following two theorems prove that the two processes are indistinguishable when the number of steps in $p(t)$ becomes infinite as λ approaches infinity, subject to the constraint (7) (see Ref. (**5**)).

Theorem I. Let $p(t)$ be the Poisson step process defined above. Then $p(t)$ tends to a Wiener process with probability one as $h\to 0, \lambda\to\infty$ such that $h^2\lambda = 1$.

It is straightforward to show that the limit function is strongly stochastically continuous (**13**). The proof of the theorem now follows directly from Theorem 7.1, p. 470 of Doob (**12**).

By Theorem I we know that the Poisson step process tends to a Wiener process. Now the question is: Does the solution of (6) tend to the solution of (1) in some sense? The answer is yes as shown by the next theorem.

Theorem II. The process described by Eq. (6) converges in distribution to $x(t)$ given by Eq. (1) as $h \to 0$, $\lambda \to \infty$ such that $h^2\lambda = 1$.

Denote the solution of (1) by $x_w(t)$ and the solution of (6) by $x_h(t)$. The distributions of $x_w(t)$ and for $x_h(t)$, denoted by $P(x,t\,|\,x(s),s)$ and $P_h(x,t\,|\,x(s),x)$ respectively, are given from Kolmogorov's equations as follows:

$$-\frac{\partial P(x,t\,|\,x(s),s)}{\partial s} = \sum_{i=1}^{n} f_i(x,s)\frac{\partial P(x,t\,|\,x(s),x)}{\partial x_i(s)} + \frac{1}{2}\sum_{i=1}^{n} \sigma_i^2 \frac{\partial^2 P(x,t\,|\,x(s),s)}{\partial^2 x_i(s)^2}$$

$$-\frac{\partial P_h(x,t\,|\,x(s),s)}{\partial s} = \sum_{i=1}^{n} f_i(x,s)\frac{\partial P_h(x,t\,|\,x(s),s)}{\partial x_i(s)}$$

$$+\frac{1}{2}\sum_{i=1}^{n} \sigma_i^2 h^2 \lambda(x,t)\,\Delta_i^2 P_h(x,t\,|\,x(s),s).$$

Here the operation Δ_i^2 is defined from

$$\Delta_i^2 A(x) = 1/h^2[A(x+h_i) - 2A(x) + A(x-h_i)], \tag{8}$$

where h_i is an n-vector with h in the ith entry and zero elsewhere.

A study of the above two Kolmogorov equations reveals that the equation for $P_h(x,t\,|\,x(s),s)$ is but a finite difference approximation to the parabolic equation for $P(x,t\,|\,x(s),x)$, at least for the case $h^2\lambda = 1$ and h small. This approximation is valid since $P(x,t\,|\,x(s),s)$ has continuous second partial derivatives, so we have

$$\lim_{\substack{h\to 0\\ h^2\lambda=1}} P_h(x,t\,|\,x(s),s) = P(x,t\,|\,x(s),s)$$

and Theorem II is established.

For system (6), Bellman's dynamic programming equation is (**1, 2**)

$$0 = \min_{u\in U}\left\{\sum_{i=1}^{n}\left[\frac{\sigma_i^2}{2}\Delta_i^2 V_h + f_i(u)\frac{\partial V_h}{\partial x_i}\right] + \frac{\partial V_h}{\partial t} + L(u)\right\} \tag{9}$$

with $V_h(x,t) = B(x)$ for $x \in \partial S$. The subscript h is used to indicate the Poisson solution.

The quasi-linearization method of solution (9) requires the solution of a series of partial differential difference equations similar to (9) but with a fixed control law $u(x,t)$ as now indicated

$$0 = \sum_{i=1}^{n}\left[\frac{\sigma_i^2}{2}\Delta_i^2 V_h + f_i\frac{\partial V_h}{\partial x_i}\right] + \frac{\partial V_h}{\partial t} + L. \tag{10}$$

As will be shown in the next section, Eq. (10) is well defined for all x.

We now introduce an important theorem relating $V_h(x,t)$, the solution of (10) to $V(x,t)$, the solution of (5).

Theorem III. Let $V(x,t)$ and $V_h(x,t)$ denote the function given by (2) where the disturbance is white Gaussian noise and Poisson impulse noise respectively. Then with the constraint (7) holding, the function $V_h(x,t)$ approaches the function $V(x,t)$ as h tends to zero.

The theorem follows from Helly's second theorem and the fact that $x_h(t)$ converges in distribution to $x(t)$ (**11**).

Note that as a consequence of the above theorem we have the result that as $\|x\| \to \infty$ and $h \to 0$, $V_h(x,t) \to V(x,t)$. Clearly then, should we wish to calculate the boundary $\lim_{\|x\| \to \infty} V(x,t)$ we simply evaluate $V_h(x,t)$ for sufficient large $\|x\|$ and sufficiently small h.

IV. Solution of the Poisson Case

Before proceeding to the solution of (10), a basic assumption concerning the deterministic solution of the noise-free problem is necessary: the deterministic solution of (10), denoted by V_0, is unique and the terminal time $T(x,t)$ is finite for all x and t. With this assumption holding, we may write the deterministic solution of (10) as

$$V_0(x,t) = \int_t^T L(x(\tau),\tau)\,\mathrm{d}\tau + B(x(T)), \tag{11}$$

where the integration is along the trajectory

$$\dot{x}(\tau) = f(x(\tau),\tau), \quad x(t) = x. \tag{12}$$

The general solution of (10) may now be written as

$$V_h(x,t) = V_0(x,t) + \int_t^T \sum_{i=1}^n \frac{\sigma_i^2}{2} \Delta_i^2 V_h \,\mathrm{d}\tau \tag{13}$$

along the trajectory (12). (Recall that $\lambda h^2 = 1$.)

This Eq. (13) is a Volterra integral equation. If the terminal time is fixed, this can be solved by the method of successive approximations (**18, 19**) provided that σ_i, f_i and L are exponentially bounded in x for all t, and also provided that these quantities are sufficiently smooth for V and its first four derivatives to be exponentially bounded in x for all t (**3**). This result does not apply here because the terminal time T is a function of (x,t). We were not able to derive a set of sufficient conditions for the more general problem given by (13). However, for the example considered in the next section, the method of successive approximations did converge very rapidly.

Some "comments" on the above results are in order:

(1) An obvious first approximation to use in solving (10) by the method of successive approximations is the deterministic solution V_0 which is usually readily calculated. The choice of an initial control law to use in a quasi-linearization can be achieved using the deterministic optimal law in the vicinity of the terminal region and extrapolating the calculated stochastic optimal law for larger x. Using these ideas very efficient computer programs can be written to solve the optimal control problem.

(2) The convergence rate of the method of successive approximations will clearly depend on σ_i^2. If these numbers are small, the method should converge rapidly so that many iterations will not be required.

V. Examples and Comparison of Methods

As a first example to illustrate the ideas of the previous sections, consider the scalar stochastic system described by

$$\mathrm{d}x = -\mathrm{d}t + \mathrm{d}w$$

with the terminal region $S = \{x\colon x \leqslant 0\}$. Let $L = 1$ and $B(x) = 0$ so that $V(x, t)$ is the expected value of the terminal time.

Equation (5) becomes

$$\tfrac{1}{2}\sigma^2 \frac{\mathrm{d}^2 V}{\mathrm{d}x^2} - \frac{\mathrm{d}V}{\mathrm{d}x} + 1 = 0$$

and the solution is

$$V(x) = \begin{cases} x + \dfrac{\alpha\sigma^2}{2}\left[\exp\left(+\dfrac{2x}{\sigma^2}\right) - 1\right], & x \geqslant 0, \\ 0, \quad x < 0, \end{cases}$$

where α is a constant not yet specified.

The Poisson impulse noise equation is

$$\tfrac{1}{2}\sigma^2 \Delta^2 V_h - \frac{\mathrm{d}V_h}{\mathrm{d}x} + 1 = 0.$$

It is now easily verified that the only value for α that enables the solutions V and V_h to be identical at infinite x is the case $\alpha = 0$. For this case $\lim_{x\to\infty} V = \lim_{x\to\infty} V_h = x$.

As a second example, we consider a second-order plant described by the equations,

$$\left.\begin{aligned} \mathrm{d}x_1 &= x_2\,\mathrm{d}t, \\ \mathrm{d}x_2 &= u\,\mathrm{d}t + \sigma\,\mathrm{d}w, \end{aligned}\right\} \tag{14}$$

where $w(t)$ is a Wiener process and the control u is constrained by $|u| \leqslant 1$. The terminal region S is the unit disc

$$S = \{(x_1, x_2)\colon x_1^2 + x_2^2 \leqslant 1\}$$

and the performance index is

$$J = E\left\{\int_t^{t_f} \mathrm{d}\tau \,\middle|\, x(t) = x\right\}.$$

For this problem the optimization equation for $V(x_1, x_2)$ is

$$0 = 1 + \min_{|u| \leqslant 1}\left\{x_2 \frac{\partial V}{\partial x_1} + u\frac{\partial V}{\partial x_2} + \tfrac{1}{2}\sigma^2 \frac{\partial^2 V}{\partial x_2^2}\right\}$$

with the boundary condition $V(x) = 0$ on ∂S.

The solution for $u(x)$ is

$$u = -\operatorname{sgn}\left[\frac{\partial V}{\partial x_2}\right]$$

and therefore u is a bang-bang control with the switching surface determined by the equation

$$\partial V/\partial x_2 = 0.$$

If the disturbance is shot noise, then the optimization equation is

$$0 = 1 + \min_{|u|\leq 1}\left\{x_2\frac{\partial V_h}{\partial x_1} + u\frac{\partial V_h}{\partial x_2} + \frac{\sigma^2}{2}\Delta_2^2 V\right\},$$

where $\lambda h^2 = 1$. This equation may be solved numerically using the method of successive approximations and quasi-linearization. This in fact has been done in another context in Ref. (**16**) and so the results are adapted from Ref. (**16**) to yield the switching curves of the figure. (Computer time less than 1 min.)

Also shown in Fig. 1, for comparison purposes, is the switching curve obtained using the singular perturbation method of Refs. (**6**) and (**7**). It appears that this method is not very accurate for large x, but it appears to be accurate close to the terminal region. This is analogous to the fluid dynamics problem where singular perturbations are used to evaluate boundary layer terms (**17**).

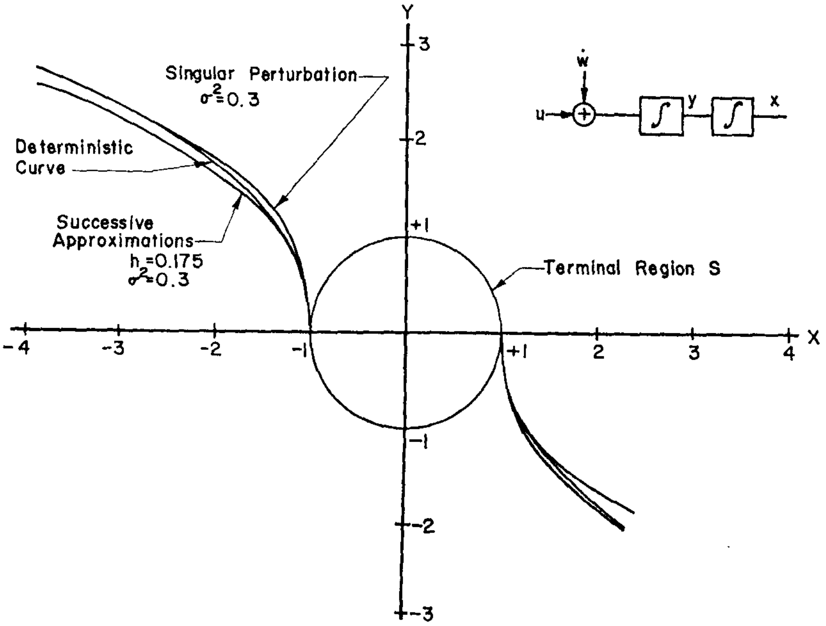

FIG. 1. Switching curves.

Results for the same problem determined using the Monte Carlo simulation procedure (**9**) are close to the singular perturbation results and are therefore not included on the diagram.

The results of Kushner and Kleinman (8) achieve the same accuracy as the method of this paper. It appears, however, that unless a good initial choice for an artificial boundary is made the method is inherently less efficient than the successive approximation and quasi-linearization method of this paper.

References

(1) J. J. Florentine, "Optimal control of continuous-time Markov stochastic systems", *J. Electron. Control*, Vol. 10, pp. 473–488, 1964.

(2) J. J. Kushner, "Stochastic Stability and Control", New York, Academic Press, Chap. I, 1967.

(3) J. B. Moore and P. N. Robinson, "A new method for the solution of the Cauchy problem for parabolic equations", *Communications of ACM*, Vol. 15, No. 2, Dec. 1972.

(4) J. J. Florentine, "Optimal control of systems with generalized Poisson inputs", *J. bas. Eng.*, Vol. 85, No. 2, pp. 217–221, June 1963.

(5) P. N. Robinson, "Analysis of systems with shot noise", *J. Franklin Inst.*, Vol. 293, No. 5, pp. 305–311, May 1972.

(6) P. Dorato, C. M. Hsieh and P. N. Robinson, "Optimal bang-bang control of linear stochastic systems with a small noise parameter", *IEEE Trans. Automatic Control*, Vol. AC-12, pp. 682–690, Dec. 1967.

(7) R. S. Lim, "Asymptotic solutions of an optimal serno problem", *IEEE Trans. Automatic Control*, Vol. AC-13, pp. 45–50, Feb. 1968.

(8) H. J. Kushner and A. J. Kleinman, "Numerical methods for the solution of degenerate nonlinear elliptic equations arising in optimal stochastic control theory", *IEEE Trans. Automatic Control*, Vol. AC-13, pp. 344–353, Aug. 1968.

(9) P. N. Robinson and H. O. Yurtseven, "A Monte-Carlo method for stochastic time-optimal control", *IEEE Trans. Automatic Control* (Corresp.), Vol. AC-14, pp. 574–575, Oct. 1969.

(10) R. Bellman, "Functional equations in the theory of dynamic programming, positivity and quasilinearity", *Proc. natn. Acad. Sci. U.S.A.*, Vol. 41, pp. 743–746, 1955.

(11) B. V. Gnedenko, "The Theory of Probability", p. 283, New York, Chelsea Pub. Co., 1968.

(12) J. L. Doob, "Stochastic Process", p. 420, New York, John Wiley, 1953.

(13) B. V. Gnedenko and A. N. Kolmogorov, "Limit distributions for sums of independent random variables", p. 108, Reading, Mass., Addison-Wesley, 1968.

(14) P. A. Frost and T. Kailath, "Mathematical modeling and transformation theory of white noise processes", in *Stochastic Problems in Control, Symp. Amer. Automat. Contr. Council*, pp. 1–38, 1968.

(15) P. A. Frost and T. Kailath, "An innovations approach to least-squares estimation —Part III: nonlinear estimation in white gaussian noise", *IEEE Trans. Automatic Control*, Vol. AC-16, pp. 217–226, June 1971.

(16) P. N. Robinson, "Time-optimal control in the presence of shot noise", *Proc. Fourth Asilomar Conf.*, pp. 236–242, Nov. 1970.

(17) W. F. Ames, "Nonlinear Partial Differential Equations in Engineering", pp. 211–219, New York, Academic Press, 1965.

(18) L. A. Linsternik and V. J. Sobolev, "Elements of Functional Analysis", pp. 72–32, New York, Ungar, 1961.

(19) A. N. Kolmogorov and S. V. Fomin, "Elements of the Theory of Functional Analysis", pp. 43–51, Vol. 1, Rochester, N.Y., Graylock, 1957.