

AD-A115 434

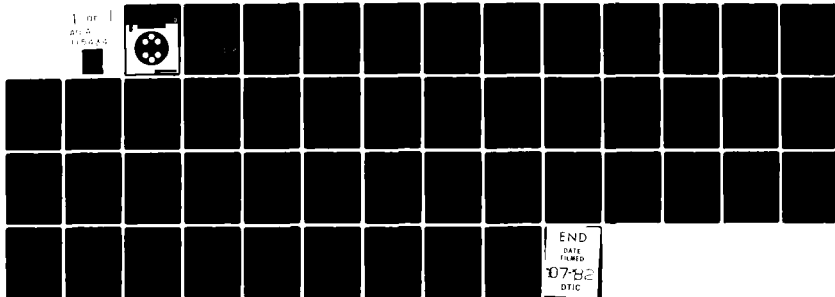
BROWN UNIV PROVIDENCE RI LEFSCHETZ CENTER FOR DYNAMI--ETC F/6 20/1
ROBUSTNESS AND APPROXIMATION OF ESCAPE TIMES AND LARGE DEVIATIO--ETC(U)
MAR 82 H J KUSHNER
LCDS-82-5

N00014-76-C-0279

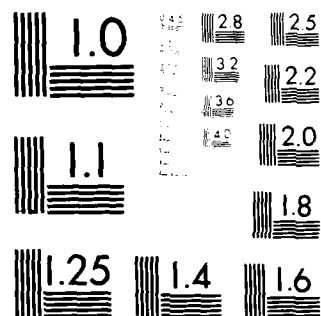
NL

UNCLASSIFIED

1 of 1
AD-A115 434



END
DATE
FILMED
07-82
DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AD A115434

121
DTIC
ELECTE
JUN 11 1982

Lefschetz Center for Dynamical Systems

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

Division of Applied Mathematics

Brown University Providence RI 02912

82 05 10 178

12

ROBUSTNESS AND APPROXIMATION OF ESCAPE
TIMES AND LARGE DEVIATIONS ESTIMATES FOR SYSTEMS
WITH SMALL NOISE EFFECTS

by

Harold J. Kushner

March 1982

LCDS Report #82-5

DTIC
ELECTE
JUN 11 1982
S H D

DISTRIBUTION STATEMENT A
Approved for public release
Distribution Unlimited

ROBUSTNESS AND APPROXIMATION OF ESCAPE TIMES AND LARGE
DEVIATIONS ESTIMATES FOR SYSTEMS WITH SMALL NOISE EFFECTS

Harold J. Kushner

Divisions of Applied Mathematics and Engineering
Lefschetz Center for Dynamical Systems
Brown University
Providence, Rhode Island 02912

March 1982

[†]Work supported in part by the Air Force Office of Scientific Research under AFOSR 81-0116, by the National Science Foundation under NSF-Eng 77-12946-A02 and in part by the Office of Naval Research under N00014-76-C-0279-P0004.

ROBUSTNESS AND APPROXIMATION OF ESCAPE TIMES AND LARGE
DEVIATIONS ESTIMATES FOR SYSTEMS WITH SMALL NOISE EFFECTS

Abstract

→ For the purposes of estimating escape time from a given set, or other statistical properties of systems with small noise effects, it is generally assumed in applications that the system noise is white Gaussian. The Gaussian assumption greatly simplifies the computation, but is not adequate for many important classes of applications to control and communication theory. For example, when the noise is small, the mean escape time from a set can be quite sensitive to the underlying statistics even though in the study of the effects of the noise over any fixed finite time interval, the Gaussian approximation might be a good one. This paper is concerned with the sensitivity of these statistical quantities to the underlying statistical structure, when the noise effects are small, and also with the question of when the Gaussian assumption makes sense. Consider a sequence of systems with small noise effects whose statistics converge in some sense to those of a "limit" system. The techniques developed involve approximation and limit theorems for a sequence of variational problems associated with the minimization of the action functionals which arise when the theory of large deviations is applied to the above mentioned systems. The admissible paths and velocity fields are characterized. Techniques are developed for approximating ϵ -optimal or optimal paths and values of the action functionals with "restricted velocity fields", and these are used to get the desired limit, approximation and robustness theorems. Degenerate and non-degenerate cases with both bounded and Gaussian noise are considered. Several examples and an application to a phase locked loop system which arises in communication theory are discussed. These indicate when the Gaussian assumption might be acceptable in practice. The results are of potential use in computation, for they indicate when the results for a simpler "more computable" noise process might be a good approximation to the results for the true noise process. The results concerning convergence and approximation seem to be of independent interest for treating convergence of the solutions of a sequence of more general variational problems.



Accession For	
NTIS GRA&I	☒
DTIC TAB	☐
Unannounced	☐
Justification	☐
By <i>file</i>	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
<i>A</i>	

1. Introduction

For the purpose of estimating escape time from a given set, or other statistical properties of systems with small noise effects, it is generally assumed in applications that the system noise is white Gaussian. The Gaussian assumption greatly simplifies the computation, but is not adequate for many important classes of applications to control and communication theory. For example, when the noise is small, the mean escape time from a set can be quite sensitive to the underlying statistics even though in the study of the effects of the noise over any fixed finite time interval, the Gaussian approximation might be a good one. This paper is concerned with the sensitivity of these statistical quantities to the underlying statistical structure, when the noise effects are small, and also with the question of when the Gaussian assumption makes sense. Consider a sequence of systems with small noise effects whose statistics converge in some sense to those of a "limit" system. The techniques developed involve approximation and limit theorems for a sequence of variational problems associated with the minimization of the action functionals which arise when the theory of large deviations is applied to the above mentioned systems. The admissible paths and velocity fields are characterized. Techniques are developed for approximating ϵ -optimal or optimal paths and values of the action functionals with "restricted velocity fields", and these are used to get the desired limit, approximation and robustness theorems. Degenerate and non-degenerate cases with both bounded and Gaussian noise are considered. Several examples and an application to a phase locked loop system which arises in communication theory are discussed. These indicate when the Gaussian assumption might be acceptable in practice. The results are of potential use in computation, for they indicate when the results for a simpler "more computable" noise process might be a good approximation to the results for the true noise process.

The results concerning convergence and approximation seem to be of independent interest for treating convergence of the solutions of a sequence of more general variational problems.

We will be concerned with robustness, approximation and applications of large deviations methods [1] - [7] for processes of the type (1.1) - (1.4). The $\{\xi(\cdot)\}$, $\{\xi_n\}$ are bounded and stationary, $w(\cdot)$ is a standard Wiener process, $\{\rho_n\}$ is i.i.d. Gaussian, and $\{\rho_n\}$ and $w(\cdot)$ are independent of $\{\xi_n\}$ and $\xi(\cdot)$, and $E\rho_n = 0$, $\text{covar } \rho_n = I$. The functions $\sigma(\cdot)$, $\bar{b}(\cdot)$ and $b(\cdot, \xi)$ are Lipschitz continuous (uniformly in ξ). In all cases, $x \in R^k$, Euclidean k -space.

$$(1.1) \quad \dot{x}^Y = b(x^Y, \xi(t/\gamma))$$

$$(1.2) \quad dx^Y = b(x^Y, \xi(t/\gamma))dt + \sqrt{\gamma}\sigma(x^Y)dw$$

$$(1.3) \quad x_{k+1}^Y = x_k^Y + \gamma b(x_k^Y, \xi_k)$$

$$(1.4) \quad x_{k+1}^Y = x_k^Y + \gamma b(x_k^Y, \xi_k) + \gamma \sigma(x_k^Y) \rho_k.$$

Several modifications of (1.1) - (1.4) will also be considered. For (1.3) - (1.4), define $x^Y(\cdot)$ to be the piecewise linear interpolation of the function with values x_n^Y at $t = n\gamma$. Define $\bar{b}(\cdot)$ by $T\bar{b}(x) = \lim_N \frac{1}{N} E \sum_{n=0}^{(N-1)T} b(x, \xi_n)$, $Tb(x) = \lim_{\gamma} E_{\gamma} \int_0^{T/\gamma} b(x, \xi(t))dt$, and suppose that $\bar{b}(\cdot)$ is independent of T .

The various assumptions introduced below are not always used together.

Let G be a bounded open set with a piecewise differentiable boundary, define $N_{\epsilon_1}(G) = G_1$ an ϵ_1 -neighborhood of G (henceforth fixed), and assume

(A1.1) $\dot{x} = \bar{b}(x)$ has a unique stable point $\bar{x}_0$ in G_1 and all trajectories originating in G_1 tend to $\bar{x}_0$. Also, these trajectories are never tangent to ∂G .

Define the H -functionals

$$H(x, \alpha) = \frac{1}{T} \lim_{N, M \rightarrow \infty} \frac{1}{(N-M)} \log E \exp \sum_{M}^{N-1} \alpha' (b(x, \xi_n) + \sigma(x) \rho_n) \quad (1.5)$$

$$H(x, \alpha) = \frac{1}{T} \lim_{\gamma \rightarrow 0} \gamma \log E \exp \int_{T_0/\gamma}^{T_1/\gamma} \alpha' [b(x, \xi(s)) ds + \sigma(x) dw(s)].$$

Where $T_1 - T_0 = T$, $N_\gamma = T_1$, $M_\gamma = T_0$, and we assume that the limit does not depend on T or T_1 . When we wish to emphasize the Gaussian component, an affix σ will be used and we write $H^\sigma(x, \alpha) = H^0(x, \alpha) + \tilde{H}^\sigma(x, \alpha)$.

(A1.2) In (1.5), let the convergence be uniform for $x \in G_1$ and also in the initial data, if E is replaced by the expectation given the $\xi(\cdot)$, or $\{\xi_n\}$ data up to time T_0 or T_0/γ (discrete parameter case).

The limits in (1.5) and assumption (A1.2) are phrased as they are because we wish to treat the escape time problem when the noise is not necessarily Markov. If the noise is Markov, it is sufficient to set $T_0 = M = 0$ and let the convergence alluded to in (A1.2) be uniform in x and in ξ_0 , the initial state for the noise process. We will also sometimes use the weaker form

A(1.2') Let $T_0 = M = 0$ and let the convergence in (1.5) be uniform in $x \in G_1$.

Define the Cramer transformation $L(x, \beta) = \sup_{\alpha} [\beta' \alpha - H(x, \alpha)]$, and set $U(x) = \{\beta : L(x, \beta) < \infty\}$. Define $S(T, \phi) = \int_0^T L(\phi(s), \dot{\phi}(s)) ds$, if $\phi(\cdot)$ is absolutely continuous, and set it equal to ∞ otherwise. For $T(\phi) = \inf\{t : \phi(t) \notin G\}$, define $S(\phi) = S(T(\phi), \phi)$, $S_0(x) = \inf\{S(\phi) : \phi(0) = x\}$, $S_0 = S_0(\bar{x}_0)$ and set $\tau_G^Y = \min\{t : x^Y(t) \notin G\}$. The functional $S(T, \phi)$ is called an action functional if for each $a > 0$, $h > 0$ and bounded continuous $\phi(\cdot)$, there is a $\gamma_0 > 0$ such that for $\gamma \leq \gamma_0$,

$$(1.6a) \quad P\{d(x^\gamma, \phi) < \delta\} \geq \exp - (S(T, \phi) + h)/\gamma$$

$$(1.6b) \quad P\{d(x^\gamma, \phi_a) > \delta\} \leq \exp - (a-h)/\gamma,$$

where $d(\cdot, \cdot)$ = sup norm distance, and $\phi_a = \{\text{bounded continuous } \phi(\cdot): S(T, \phi) < a\}$. See [5], Theorem 2.1, where $S(T, \phi)$ is written $S_{0,T}(\phi)$. Also [5], (1.6) implies that for any set A of continuous functions on $[0, T]$,

$$(1.7) \quad - \inf_{\phi \in A^0} S(T, \phi) \leq \lim_{\gamma} \gamma \log P\{x^\gamma(\cdot) \in A\} \\ \leq \overline{\lim}_{\gamma} \gamma \log P\{x^\gamma(\cdot) \in A\} \leq - \inf_{\phi \in \bar{A}} S(T, \phi),$$

where A^0 = interior of A , and $\bar{A}$ = closure of A .

Under broad conditions

$$(1.8) \quad \lim_{\gamma \rightarrow 0} \gamma \log E \tau_G^\gamma = S_0$$

See [3,5]. In [5], $\sigma = 0$ and (A1.2) was implicitly assumed. With $\sigma = 0$, the proof in [5] (Theorem 5.1) is valid for more general $\xi(\cdot)$, provided

(A1.2), (A1.3) hold, and using the convergence of (1.5) uniformly in $x \in G_1$

and in the initial data. It can also be extended to cover $\sigma = \text{constant}$ (see appendix). The case where $\sigma \neq 0$, but $\xi(\cdot)$ does not appear, is in [3].

The proof in [5] was given for the continuous parameter problem, but it also works for the discrete parameter problem. Criteria for (A1.3) appear in theorems 3.8 and 3.9.

(A1.3) For $x \in G_1$, $H(\cdot, \cdot)$ is continuous and $H(x, \cdot)$ is differentiable.

(A.4) The boundary ∂G is piecewise differentiable. In particular, for each $x \in \partial G$, there is a neighborhood $N(x)$ and a finite number of differentiable functions $\theta_1(\cdot)$, $i \leq k$, such that $G \cap N(x) = \{y: \theta_i(y) < 0,$

$i = 1, \dots, k\}$, $\partial G \cap \bar{N}(x) = \{y: \theta_i(y) \leq 0, i = 1, \dots, k, \text{ and some } \theta_i(y) = 0\}$.

In Theorems 3.11 and 4.7, the 'continuity' condition A1.5 will be used. For open Q containing $\bar{x}_0$ define $S(Q) = \inf \{S(T, \phi): \phi(0) = \bar{x}_0, \phi(T) \in \partial Q\}$. Then $S(Q)$ is lower semi-continuous in Q , in that if $Q_\alpha \uparrow Q$, then $\lim_{\alpha} S(Q_\alpha) \geq S(Q)$. But it is continuous at 'most' Q in the following sense.

Let $Q_\rho = N_\rho(Q)$ and $\rho < \rho_1$ with $S(Q_{\rho_1}) < \infty$. Then for all but a countable

number of $\rho_0 < \rho_1$, $S(Q_\rho) \rightarrow S(Q_{\rho_0})$ as $\rho \downarrow \rho_0$.

(A1.5) $S(G_\rho)$ is continuous at $\rho = 0$.

The quantity (1.8) is of considerable importance in numerous applications in control and communication theory, and in various applications to stochastic approximation, particularly in estimating escape times from regions in which an algorithm or process has a 'stability' property. Normally such estimates are hard to get unless ϵ is small. Except for the purely Gaussian case, it is now almost impossible to calculate $H(\cdot, \cdot)$, $L(\cdot, \cdot)$ or S_0 , and so the purely Gaussian model is used almost exclusively. However, the value of S_0 can be quite sensitive to the underlying statistical assumptions, and it is not normally satisfactory to use a 'local diffusion or Gaussian' approximation [15]. We study the problem of robustness and approximatability for such problems.

Section 2 contains a brief formal remark on approximatability by a Gaussian system. Section 3 contains various background results concerning smoothness of $H(\cdot, \cdot)$ and the admissible 'velocity fields' for the variational problem associated with getting S_0 or the estimate in (1.7). The main approximation results appear in Section 4. Some examples are discussed in Section 5. Section 6 discusses an application to a phase locked loop, and various problems which arise in connection with that application.

This class of applications seems to be both natural and of increasing popularity for the applications of large deviations or singular perturbation type (partial differential equation based) methods. Since the physical noise in such systems is not white Gaussian, or even Gaussian at all (strictly speaking), that application provides a good example of the role of our results.

In the appendix, there are some remarks concerning extending the proof in [2,5] that $S(T, \phi)$ is an action functional, to the composite cases (1.2) and (1.4), where both Gaussian and non-Gaussian noise appear.

2. A Comment on a Gaussian Approximation

The expression $\alpha' \bar{b}(x) + \frac{1}{2} \alpha' \sigma(x) \sigma'(x) \alpha = H^\sigma(x, \alpha)$ is the H-functional for the system

$$(2.1) \quad x_{n+1}^\gamma = x_n^\gamma + \gamma \bar{b}(x_n^\gamma) + \gamma \sigma(x_n^\gamma) \rho_n.$$

The gradient and Hessian (with respect to α at $\alpha=0$) of $H_N = \frac{1}{N} \log E \exp \alpha' \sum_{n=0}^{N-1} b(x, \xi_n)$ are

$$(2.2a) \quad H_{N,\alpha}(x, 0) = E \sum_{n=0}^{N-1} b(x, \xi_n) / N = \bar{b}_N(x) \rightarrow \bar{b}(x)$$

$$(2.2b) \quad H_{N,\alpha\alpha}(x, 0) = \frac{1}{N} E \sum_{n=0}^{N-1} (b(x, \xi_n) - \bar{b}_N(x)) \sum_{i=0}^{N-1} (b(x, \xi_i) - \bar{b}_N(x))' \rightarrow \sum(x).$$

Both $H_N(x, \cdot)$ and $H(x, \cdot)$ are convex.

Let $\phi(\cdot)$ be absolutely continuous. Then, under broad conditions, the piecewise constant (constant on $[n\gamma, n\gamma + \gamma)$) function which has values

$$\sqrt{\gamma} \sum_{i=0}^{N-1} (b(\phi(i\gamma), \xi_i) - \bar{b}(\phi(i\gamma))) \quad \text{at } t = N\gamma, \quad \text{converges weakly to a Wiener}$$

process with zero mean and covariance $\int_0^t \sum (\phi(s)) ds$. This suggests that a

suitably interpolated (1.3) can be approximated by a 'small noise' diffusion

of the form $dy = \bar{b} dt + \sqrt{\gamma} \sqrt{\Sigma} dw$. But such an approximation is purely

formal, and is not usually valid in the sense of approximation of the large

deviations results. Suppose that $\phi(\cdot)$ is an optimizing (or nearly so) path

for the $S(\cdot)$ of (1.3). If

$$\sup_{\alpha \in Q} [\alpha' \dot{\phi}(t) - H(\phi(t), \alpha)] \approx L(\phi(t), \dot{\phi}(t))$$

for almost all t , where Q is a set where the quadratic approximation $\alpha' \bar{b}(x) + \alpha' \int (x) \alpha/2$ to the H-functional for (1.3) determined by (2.2a,b) is acceptable, then the Gaussian approximation makes sense, but this is normally very difficult to verify.

3. Preliminary Results

Theorems 3.1 to 3.4 give necessary and/or sufficient conditions for β to be in $U(x)$ in terms of the underlying statistics. This is important, since when minimizing $S(\phi)$ or $S(T, \phi)$, we have $\dot{\phi}(t) \in U(\phi(t))$, and the questions of finiteness and approximability of S_0 and $\inf_{\phi \in A} S(T, \phi)$ are related to the properties of $U(x)$. Theorem 3.5 and Corollary 3.6 provide continuity and convergence results which will be useful in the sequel, and Theorems 3.8 and 3.9 provide criteria for (A1.3).

Theorem 3.1. $H(x, \cdot)$, $L(x, \cdot)$ and $U(x)$ are convex. $L(\cdot, \cdot)$ and $U(\cdot)$ (in the Hausdorff topology) are lower semicontinuous.

If $\sigma(x)\sigma'(x)$ is uniformly positive definite on G_1 , then $L(x, \beta) < \infty$, all $x \in G_1$ and all β . Remarks on the degenerate case are given below.

Theorem 3.2 ($\{\xi_n\}$ i.i.d, $\sigma = 0$.) Let $\xi = \xi_n$ have compact support. Then
 $H(x, \alpha) = \log E \exp \alpha' b(x, \xi)$ and

- (a) $L(x, \beta) = \infty$ if $\beta \notin \overline{\text{co range } b(x, \xi)} \equiv C$
- (b) $L(x, \beta) < \infty$ if $\beta \in \text{rel.int. } \overline{\text{co range } b(x, \xi)}$.

Note. $\beta \in \text{range } b(x, \xi)$ if for each $\epsilon > 0$, $P\{b(x, \xi) \in N_\epsilon(\beta)\} > 0$. The

relative interior is relative to the smallest linear manifold which contains the set.

Proof.

(a) Let $\beta_0 \notin C$. Then there are ℓ, c_0 such that $\ell'b < c_0, b \in C$, $\ell'\beta_0 > c$. Also,

$$\begin{aligned} \sup_{\alpha} [\alpha'\beta_0 - H(x, \alpha)] &\geq \sup_{c \geq 0} [c\ell'\beta_0 - \log E \exp c\ell'b(x, \xi)] \\ &\geq \sup_c [c\ell'\beta_0 - cc_0] = \infty. \end{aligned}$$

(b) Let β_0 be such that $P\{b(x, \xi) \in N(\beta_0)\} > 0$ for all neighborhoods $N(\beta_0)$ (in the smallest linear manifold containing C) of β_0 . For convenience, define $M(\beta_0) = \{\xi: b(x, \xi) \in N(\beta_0)\}$, and $P_M(\beta_0) = P\{M(\beta_0)\}$.

Then (Jensen's inequality is used to get the last line)

$$\begin{aligned} \sup_{\alpha} [\alpha'\beta - H(x, \alpha)] &\leq \sup_{\alpha} [\alpha'\beta - \log P_M(\beta_0) \int_{M(\beta_0)} \exp \alpha'b(x, \xi) \frac{dP}{P_M(\beta_0)}] \\ &\leq \sup_{\alpha} [\alpha'\beta - \alpha' \int_{M(\beta_0)} b(x, \xi) \frac{dP}{P_M(\beta_0)}] - \log P_M(\beta_0). \end{aligned}$$

Thus, for $\beta = \int_{M(\beta_0)} b(x, \xi) dP/P_M(\beta_0)$, $L(x, \beta) < \infty$. The assertion (b)

follows from this, Theorem 3.1 and the fact that a convex set is a continuous function (Hausdorff topology) of its extreme points. Q.E.D.

The general case (1.3) or (1.1) for non i.i.d. $\{\xi_n\}$ is more complicated.

We concentrate on (1.3), and first treat the finite Markov chain case.

Theorem 3.3. Let $\{\xi_n\}$ be a finite state Markov chain with state space D , and transition probabilities $\{p_{ij}\}$, and with all states communicating with each other. Then $U(x)$ is the set of β_0 such that there are $N_n \rightarrow \infty$ and z_i such that $p_{z_i z_{i+1}} > 0$ all i and

$$(3.1) \quad \beta_0 = \lim_n \frac{1}{N_n} \sum_{i=0}^{N_n-1} b(x, z_i).$$

Proof. By the ergodicity, such β_0 form a closed convex set. Let $\{z_i\}$, satisfy the hypothesis and define β_0 by (3.1). There is a $q > 0$ such that $p_{z_i z_{i+1}} \geq q$, all i . Then (the limit below exists by the discrete parameter version of Theorem 2.2 of [5]; see also Theorem 3.8 below).

$$\begin{aligned} & \sup_{\alpha} [\alpha' \beta_0 - \lim_N \frac{1}{N} \log E \exp \alpha' \sum_{i=0}^{N-1} b(x, \xi_i)] \\ & \leq \sup_{\alpha} [\alpha' \beta_0 - \lim_n \frac{1}{N_n} \log E \exp \alpha' \sum_{i=0}^{N_n-1} b(x, \xi_i)] \\ & \leq \sup_{\alpha} [\alpha' \beta_0 - \lim_n \frac{1}{N_n} \log (\exp \alpha' \sum_{i=0}^{N_n-1} b(x, z_i)) \prod_{i=1}^{N_n-1} p_{z_{i-1} z_i}] \\ & \leq - \log q. \end{aligned}$$

Thus $\beta_0 \in U(x)$. The reverse case can be proved by a method similar to that of Theorem 3.2 (a). (See also Theorem 3.4.) Q.E.D.

Note that we also proved that $L(x, \xi) \leq -\log q$ on $U(x)$. We now move to the general case (1.3). Put $\beta_0 \in U_0(x)$ if for each $\epsilon > 0$, there are $q_\epsilon > 0$, $N_n \rightarrow \infty$ and $\{z_i\}$ such that (3.1) holds and $P\{b(x, \xi_i) \in N_\epsilon(b(x, z_i))\}$, $i \leq n\} \geq q_\epsilon^n$.

Theorem 3.4. Let the ξ_k each have support in a compact set C_0 . Then
relative interior $\text{co } U_0(x) \subset U(x)$. Define

$$R_k = \text{co range } \left[\sum_{i=0}^{k-1} b(x, \xi_i) / k \right].$$

Suppose that there is a $\delta > 0$ such that the distance between β_0 and R_k
is $\geq \delta$ for an infinite number of k . Then $\beta_0 \notin U(x)$.

Proof. The proof of the first assertion is similar to that of Theorem 3.3. To prove the second part take a subsequence of R_k (also indexed by k) and suppose that $d(\beta_0, R_k) \geq \delta > 0$ for all k (w.l.o.g.). Note that there are $\epsilon_0 > 0$ and bounded $\{c_k\}$ and unit vectors ℓ_k such that $\ell_k' b < c_k$ for $b \in R_k$ and $\ell_k' \beta_0 > c_k + \epsilon_0$. Assume that $\ell_k \rightarrow \ell$ and $c_k \rightarrow c$ (or else work with a convergent subsequence). Then

$$\begin{aligned} \sup_{\alpha} [\alpha' \beta_0 - H(x, \alpha)] &\geq \sup_{c \geq 0} [c \ell' \beta_0 - \lim_k \frac{1}{k} E \exp k c \ell' \sum_{i=0}^{k-1} b(x, \xi_i) / k] \\ &\geq \sup_{c \geq 0} [c \ell' \beta_0 - c c_0] = \infty. \end{aligned} \quad \text{Q.E.D.}$$

There is an obvious continuous parameter analog with an integral replacing the sum and (for measurable $z(\cdot)$) $P\{b(x, \xi(s)) \in N_\epsilon(b(x, z(s)))\}$, $s \leq t\} \geq \exp - t q_\epsilon$, $q_\epsilon > 0$, replacing the analogous condition above.

The continuity of $L(\cdot, \cdot)$.

Theorem 3.5 Let $H(\cdot, \cdot)$ be continuous and let $x_0, \beta_0, N_\epsilon(\beta_0)$ satisfy
 $L(x_0, \beta) < \infty$ for $\beta \in N_\epsilon(\beta_0)$. Then $L(\cdot, \cdot)$ is continuous at (x_0, β_0) .

Proof. By Theorem 3.1, $H(x, \cdot)$ is convex. By the hypothesis and the concavity (in α) of $\alpha'\beta - H(x, \alpha)$, the set of maximizing α (at x_0, β_0) is bounded. Otherwise, for an appropriate but arbitrarily small $\delta\beta$ we would get $L(x_0, \beta_0 + \delta\beta_0) = \infty$. Also $\alpha'\beta - H(x, \alpha) \rightarrow \alpha'\beta_0 - H(x_0, \alpha_0)$ uniformly on bounded α -sets as $(x, \beta) \rightarrow (x_0, \beta_0)$. The concavity and the last three sentences imply that the set of maximizing α in $\sup_\alpha [\alpha'\beta - H(x, \alpha)]$ must also converge to the set of maximizing α for x_0, β_0 . Thus $L(x, \beta) \rightarrow L(x_0, \beta_0)$. Q.E.D.

A remark on the degenerate case.

Suppose that (1.4) has the form $(x = (x_1, x_2), \alpha = (\alpha_1, \alpha_2), \beta = (\beta_1, \beta_2)$

$$(3.2) \quad x_{1,k+1}^Y = x_{1,k}^Y + \gamma \bar{b}_1(x_k^Y), \quad x_1 \in R^{k-l}, \quad x_2 \in R^l,$$

$$x_{2,k+1}^Y = x_{2,k}^Y + \gamma b_2(x_k^Y, \xi_k) + \gamma \sigma_2(x_k^Y) \rho_k.$$

Then

$$(3.3) \quad H(x, \alpha) = \alpha_1' \bar{b}_1(x) + \alpha_2' \sigma_2(x) \sigma_2'(x) \alpha_2 / 2 + \lim_N \frac{1}{N} \log E \exp \alpha_2' \sum_0^{N-1} b_2(x, \xi_n),$$

and $L(x, \beta) = \infty$ if $\beta_1 \neq \bar{b}_1(x)$. But Theorem 3.5 can be used to study continuity with respect to β_2 when $\beta_1 = \bar{b}_1(x)$. If $\sigma_2(x) \sigma_2'(x)$ is uniformly positive definite and $H(\cdot, \cdot)$ is continuous, then $L(x, \beta)$ is continuous in (x, β_2) when $\beta_1 = \bar{b}_1(x)$. Similar remarks hold in the continuous parameter case. Define $U_2(x) = \{\beta_2: L(x; \bar{b}_1(x), \beta_2) < \infty\}$. In the sequel, when we refer to the degenerate case, the form (3.2) is always intended. In the non-degenerate case, we assume that $U(x)$ has a non-empty interior, and in the degenerate case that $U_2(x)$ has a non-empty interior.

In the non-degenerate case, define $\bar{U}^\delta(x)$ by : $\beta \in \bar{U}^\delta(x)$ if $\beta \in U(x)$ and $d(\beta, \partial U(x)) \geq \delta$, where d = Euclidean distance. The set $\bar{U}^\delta(x)$ is called a ' δ -interior' set. Let $\bar{U}_2^\delta(x)$ denote the δ -interior set for $U_2(x)$, and in the degenerate case, define $\bar{U}^\delta(x)$ by $\beta \in \bar{U}^\delta(x)$ if $\beta = (\beta_1, \beta_2)$, $\beta_1 = \bar{b}_1(x)$ and $\beta_2 \in \bar{U}_2^\delta(x)$. The continuity of a set valued function is always in the Hausdoff topology.

An argument similar to that of Theorem 3.5 proves the following.

Corollary 3.6. Let $U(\cdot)$ and $H(\cdot, \cdot)$ be continuous and let $H_n(x, \alpha) \rightarrow H(x, \alpha)$ uniformly on bounded (x, α) sets, where H_n and H are H-functionals.

Then, in the non-degenerate case and for any compact set K and $\delta > 0$,

$L_n(x, \beta) \rightarrow L(x, \beta)$ uniformly on $\{x, \beta : x \in K, \beta \in \bar{U}^\delta(x)\} \equiv \bar{U}^\delta$. In the de-

generate case, let $\bar{b}_n(\cdot)$ denote the 'mean' dynamics for the system yielding

H_n . Then $L_n(\cdot; \bar{b}_{1n}(\cdot), \cdot) \rightarrow L(\cdot; \bar{b}_1(\cdot), \cdot)$ uniformly on $\{x, \beta_2 : x \in K, \beta_2 \in \bar{U}_2^\delta(x)\}$.

If $U(\cdot)$ is not continuous, the convergence holds but might not be uniform.

The proof of the following 'path approximation' theorem is omitted.

$d(\cdot, \cdot)$ is the sup norm distance.

Theorem 3.7. Let $U(\cdot)$ be Lipschitz continuous for $x \in G_1$, and suppose that there is a $\delta_1 > 0$ such that $\bar{U}^{\delta_1}(x)$ is non-empty for all $x \in G_1$. Given $\epsilon > 0$ and $\phi(\cdot)$ such that $\phi(t) \in G_1$ and $\phi(t) \in U(\phi(t))$, $t \leq T$, there are

$\varepsilon_0 > 0$, $\delta > 0$ and absolutely continuous $\phi^\delta(\cdot)$ such that $d(\phi, \phi^\delta) \leq \varepsilon$,
 $d(\dot{\phi}, \dot{\phi}^\delta) \leq \varepsilon_0$, $\dot{\phi}^\delta(t) \in \bar{U}^\delta(\phi^\delta(t))$, $t \leq T$, and $\varepsilon_0 \rightarrow 0$ as $\varepsilon \rightarrow 0$.

Smoothness of $H(\cdot, \cdot)$.

In Theorems 3.8 and 3.9, we stick to the discrete parameter Markov chain case (compact state space D) and use a clever method of Freidlin [5] to slightly extend his results. Analogs of these theorems for the non-Markov case would be quite useful. $C(D)$ denotes the space of continuous functions on D endowed with the sup norm topology. Define the operator $Q(x, \alpha) : C(D) \rightarrow C(D)$ by (use $\xi = \xi_0$)

$$(3.4) \quad Q(x, \alpha)f(\xi) = E_\xi f(\xi_1) \exp \alpha' b(x, \xi_1),$$

where $f(\cdot) \in C(D)$. For $m \geq 1$, let $\|Q^m(x, \alpha)\| = \lambda_m(x, \alpha)$ denote the operator norm. Henceforth $x \in G_1$, and B is a compact (x, α) -set. Theorems 3.8 and 3.9 give conditions under which (A1.3) holds.

Theorem 3.8. Let there be an m such that $Q^m(x, \alpha)$ is compact for each $(x, \alpha) \in B$. Suppose that $Q^m(x, \alpha)f(\xi) > 0$ for all $\xi \in D$ if $0 \neq f(\cdot) \in C(D)$ and $f(\xi) \geq 0$. Then $\lambda_m(x, \alpha)$ is an isolated eigenvalue (with a one dimensional eigenspace) and the corresponding eigenvector $e_m(x, \alpha, \cdot)$ satisfies

$$(3.5) \quad \inf_B \inf_\xi e_m(x, \alpha, \xi) \equiv \delta_0 > 0.$$

Also

$$(3.6) \quad H(x, \alpha) = \frac{1}{m} \log \lambda_m(x, \alpha).$$

The convergence defining $H(x, \alpha)$ is uniform on B and in the initial data and $H(\cdot, \cdot)$ is continuous.

Remark. $Q^m(x, \alpha)f(\xi) = E_\xi f(\xi_m) \exp \alpha' \sum_{k=1}^m b(x, \xi_k).$

Proof.

The continuity of $\lambda_m(x, \alpha)$ is obvious. The rest of the proof is a slight modification of [5, Theorem 2.2]. Write $(x, \alpha) = y$. By Karlin [7], the compactness and strict positivity imply that $\lambda_m(y)$ is an isolated eigenvalue (and has a one dimensional eigenspace), and from this it is not hard to show that the corresponding eigenvector $e_m(y, \cdot)$ is strictly positive. We suppose w.l.o.g., that $\sup_{\xi} e_m(y, \xi) = 1$. Next we prove continuity of $e_m(\cdot, \cdot)$. Let $y_n \rightarrow y$. The set $\{Q^m(y_n)e_m(y_n, \cdot), n \geq 1\}$ lies in a compact set. Take a convergent subsequence, indexed by n , with limit $f(\cdot)$. Also $[Q^m(y_n) - Q^m(y)]e_m(y_n, \cdot)$ and hence $[\lambda_m(y_n)e_m(y_n, \cdot) - f(\cdot)]$ converge uniformly (in ξ) to zero. Since $\lambda_m(y_n) \rightarrow \lambda_m(y)$, $e_m(y_n, \cdot)$ converges uniformly (in ξ) to $f(\cdot)/\lambda_m(y)$, which must be equal to $e_m(y, \cdot)$ by uniqueness of the eigenvector. Thus since $e_m(\cdot, \cdot)$ is continuous in ξ for each y , it is continuous in (y, ξ) . This and the strict positivity of $Q^m(y)$ and of $Q^k(y)e_m(y, \cdot)$, $k < m$, for each $y \in B$ implies (3.5).

Now let $y \in$ compact B . By the above results, there is a $\delta_0 > 0$ such that for each $\xi \in D$

$$\begin{aligned} (3.7) \quad \lambda_m^n(y) Q^k(y) e_m(y, \xi) &= Q^{mn+k}(y) e_m(y, \xi) \leq Q^{mn+k}(y) 1(\xi) \\ &\leq \frac{1}{\delta_0} Q^{mn+k}(y) e_m(y, \xi) = \frac{1}{\delta_0} \lambda_m^n(y) Q^k(y) e_m(y, \xi). \end{aligned}$$

Since

$$(3.8) \quad H(y) = \lim_{\ell} \frac{1}{\ell} \log Q^{\ell}(y) 1(\xi),$$

(3.5) and (3.7) imply (3.6) and that the limit is uniform in $y \in B$, $\xi \in D$.

Q.E.D.

Theorem 3.9. Let $\lambda_m(x, \alpha) = \|Q^m(x, \alpha)\|$ be an isolated eigenvalue of $Q^m(x, \alpha)$ with a one dimensional eigenspace for each x, α . Then $\lambda_m(x, \cdot)$ is differentiable for each x . (We do not use the compactness or positivity here.)

Proof. As noted in [5], this type of result essentially follows from Kato [8]. For each α in some open set A_0 , let $T(\alpha)$ be an operator in $C(D)$, and when $\alpha = \alpha_0$, let $\zeta(\alpha_0)$ be an isolated eigenvalue with a one-dimensional eigenspace. Suppose that $\|T(\alpha) - T(\alpha_0)\| \rightarrow 0$ as $\alpha \rightarrow \alpha_0$. We can then choose eigenvalues $\zeta(\alpha)$ of $T(\alpha)$ such that $\zeta(\alpha) \rightarrow \zeta(\alpha_0)$ as $\alpha \rightarrow \alpha_0$ [8, p. 213]. Let $\Delta > 0$ be such that the distance between $N_{2\Delta}(\zeta(\alpha_0))$ and the {spectrum of $T(\alpha_0)$ minus $\zeta(\alpha_0)$ } is at least Δ . Define $\Gamma = N_{2\Delta}(\zeta(\alpha_0)) - N_{\Delta}(\zeta(\alpha_0))$. Then (Kato, [8], p. 208, Theorem 3.1, remark 3.2 and proof) there is a $C > 0$ such that if $\|T(\alpha) - T(\alpha_0)\| \leq C \min_{\zeta \in \Gamma} \|R(\zeta)\|^{-1}$, where $R(\zeta) = (T(\alpha_0) - \zeta I)^{-1}$, then $T(\alpha)$ has no eigenvalues in Γ , but $\zeta(\alpha) \in N_{\Delta}(\zeta(\alpha_0))$. Since ([9], VII 3.3), $d(\zeta) \geq \|R(\zeta)\|^{-1}$, where $d(\zeta) = \text{distance}(\zeta, \text{spectrum of } T(\alpha_0))$, we find that if $\|T(\alpha) - T(\alpha_0)\| \leq C\Delta$, then $|\zeta(\alpha) - \zeta(\alpha_0)| \leq \Delta$, for small $\Delta > 0$.

Fix x and define the operator $T_i(x, \alpha)$ in $C(D)$ by $T_i(x, \alpha) f(\xi) = E_{\xi} f(\xi_m) \sum_{k=1}^m b_i(x, \xi_k) \exp \alpha' \sum_{k=1}^m b(x, \xi_k)$, where $b_i(\cdot, \cdot)$ is the i^{th} component of $b(\cdot, \cdot)$, and let $\alpha = (\alpha_1, \dots)$. Let $\lambda_m(x, \delta\alpha, \alpha_0)$ denote the eigenvalue of $Q_m(x, \alpha_0) + \sum_i \delta\alpha_i T_i(x, \alpha_0)$ which converges to $\lambda_m(x, \alpha_0)$ as $\delta\alpha \rightarrow 0$. By the first paragraph and a truncated Taylor expansion of $Q_m(x, \alpha_0 + \delta\alpha)$ in $\delta\alpha$, $\lambda_m(x, \alpha_0 + \delta\alpha)$ differs from $\lambda_m(x, \delta\alpha, \alpha_0)$ by $o(|\delta\alpha|)$, where $o(|\delta\alpha|)$ is uniform in α_0 in any bounded set. Thus it is not enough to prove differentiability of $\lambda_m(x, \delta\alpha, \alpha_0)$. But this differentiability follows from the expansion (eqn. (2.17), p. 446 [8]) and the continuity in α of $T_i(x, \cdot)$ and $Q^m(x, \cdot)$ and we omit the details.

Q.E.D.

Theorem 3.10. For each x , let the H-functional $H(x, \cdot)$ be differentiable
at $\alpha = 0$, and let K be compact. For each $\delta > 0$, there is an $\epsilon > 0$
such that $L(x, \beta) \geq \epsilon$ for $|\beta - \bar{b}(x)| \geq \delta$, $x \in K$.

Proof. By using $\alpha'(\beta - \bar{b}(x)) - \tilde{H}(x, \alpha)$, where $\tilde{H}$ is the H-functional for
dynamics $\tilde{b}(x, \xi) = b(x, \xi) - \bar{b}(x)$, we can assume that $\bar{b}(x) \equiv 0$. Fix $\delta > 0$.
Note $\frac{1}{N} \log E \exp \alpha' \sum_{i=0}^{N-1} b(x, \xi_i) \geq \frac{1}{N} \log \exp N\alpha' \frac{E}{N} \sum_{i=0}^{N-1} b(x, \xi_i) \rightarrow 0$ as $N \rightarrow \infty$
since $\bar{b}(x) = 0$. Hence $H(x, \alpha) \geq 0$. Suppose there are $x_n \rightarrow x_0$, $\beta_n \rightarrow \beta_0$,
 $x_n \in K$ such that $L(x_n, \beta_n) \rightarrow 0$ and $|\beta_n| \geq \delta$. By lower semicontinuity,
 $\lim_n L(x_n, \beta_n) \geq L(x_0, \beta_0)$, and $|\beta_0| \geq \delta$. By the convexity and non-negativity
of $L(x, \cdot)$, $L(x_0, \beta) = 0$ for $\beta \in [0, \beta_0]$. Thus $H(x_0, \alpha) \geq \beta' \alpha$ for $\beta \in [0, \beta_0]$.
This, $H(x_0, 0) = 0$ and $H(x, \alpha) \geq 0$ contradict the differentiability at
 $\alpha = 0$.

Q.E.D.

Theorem 3.11. Let $S_0 < \infty$. Then under (A1.1) to (A1.5),

$$\lim_{\gamma \rightarrow 0} \gamma \log E \tau_G^\gamma \leq S_0.$$

(For (1.2) and (1.4)) set $\sigma(\cdot) = \text{constant.}$ Let (A1.2) hold when T_0 is
replaced by a stopping time τ and T_1 by $\tau + T$. Then

$$\lim_{\gamma \rightarrow 0} \gamma \log E \tau_G^\gamma \geq S_0.$$

With the use of the assumptions concerning uniform convergence of the H-
functional, the proof is essentially the same as that of Lemma 1 in [3].

The uniform convergence is important in order to get estimates (of the probability of the events used in Lemma 1 of [3]) uniformly in the conditioning data, since we do not necessarily have a Markovian set up. The proof in [3] implicitly assumes the continuity of $S_0(x)$ at $\bar{x}_0$. But, under our conditions this holds by essentially the same proof as used in Theorem 5.1 of [5] (with our (A1.5) and $S_0 < \infty$ replacing (5.1) of [5]). Condition (A1.5) can be replaced by the controllability condition (A4.7) and (bounded $U(x)$ case), the existence of an ϵ -optimal path satisfying the requirements of Theorem 4.7. In fact, these conditions imply A1.5. Allowing for degenerate (see (3.2)) $b(\cdot, \cdot)$ and non-Markov noise is important in applications.

4. Approximating $U(x), S(T, A)$ and S_0 .

Lemma 4.1 and Theorems 4.2, 3 show that if $H_n \rightarrow H, \phi_n \rightarrow \phi$, then $\lim_n S_n(\phi_n) \geq S(\phi)$, a basic result for the general approximation results. Theorems 4.4 to 4.8 show that $S_n \rightarrow S$ if $H_n \rightarrow H$ and some other conditions hold. Theorem 4.9 gives approximation results for inequality (1.7), when $H_n \rightarrow H$. Many of the auxiliary and intermediate approximations and techniques seem to be of independent interest.

One or more of the following conditions will be used throughout the section, and will occasionally be weakened. Until Theorem 4.9, x is always assumed to be in G_1 .

(A4.1) The H-functionals $H_n(\cdot, \cdot)$ converge to H uniformly on bounded (x, α) sets.

(A4.2) $U(\cdot)$ is continuous in the Hausdorff topology.

(A4.3) $U(x)$ and $\{\xi_n\}$ or $\xi(\cdot)$ are uniformly bounded. (We will also treat the unbounded case.)

For simplicity we consider 2 cases, the non-degenerate and the degenerate of (3.2).

(A.4.4) There is an $\epsilon_0 > 0$ such that for all x either (non-degenerate case) $N_{\epsilon_0}(\bar{b}(x)) \in U(x)$ or (degenerate case) $N_{\epsilon_0}(\bar{b}_2(x)) \in U_2(x)$.

Lemma 4.1. Under (A4.1), $\lim_n L_n(x_n, \beta_n) \geq L(x, \beta)$, if $x_n \rightarrow x, \beta_n \rightarrow \beta$.

Proof. Let $R_N = \{\alpha: |\alpha| \leq N\}$ and define $L^N(x, \beta) = \sup_{\alpha \in R_N} (\alpha' \beta - H(x, \alpha))$. Then

$$L^N(x, \beta) \uparrow L(x, \beta) \text{ as } N \rightarrow \infty. \text{ Also } L_n(x_n, \beta_n) \geq \sup_{\alpha \in R_N} (\alpha' \beta_n - H(x_n, \alpha)) \rightarrow L^N(x, \beta)$$

As $n \rightarrow \infty$. The assertion follows from this, and the arbitrariness of N . Q.E.D.

Let $S_n(T, \phi)$, $S_n(\phi)$ denote the action functionals corresponding to the H -functional H_n . The next theorem is basic for the subsequent approximation results.

Theorem 4.2. Let $\phi_n(\cdot) \rightarrow \phi(\cdot)$ uniformly and $\lim_n T(\phi_n) = T < \infty$. Then, under (A4.1) - (A4.4) and (A1.3),

$$\lim_n S_n(\phi_n) \geq S(\phi).$$

Remark. The case where $T(\phi_n) \equiv T_n \rightarrow \infty$ does not have much significance:

If $T(\phi) < \infty$, then use the fact that $S_n(t, \phi)$ is non-decreasing in t and

the theorem follows by working on $[0, T(\phi)]$ in the proof. If $T(\phi) = \infty$ and

$\lim_n S_n < \infty$, then $\phi(t) \rightarrow \bar{x}_0$, $\phi(\cdot)$ never escapes from G and $S(\phi)$ is not defined. In any

case, if each $T(\phi_n) < \infty$ and $\sup_n S_n(\phi_n) < \infty$, for each $\epsilon > 0$ we can show that

there is a sequence $\phi_n^\epsilon(\cdot)$ such that $S_n(\phi_n^\epsilon) \leq S_n(\phi_n) + \epsilon$ and $\sup_n T(\phi_n^\epsilon) < \infty$.

Proof. Assume w.l.o.g. (choose a subsequence if necessary), that

$\lim_n S_n(\phi_n) < \infty$, and that $T(\phi_n) \rightarrow T \geq T(\phi)$, $T < \infty$, and let $m(\cdot)$ denote

Lebesgue measure. For any $\epsilon > 0$, $m\{t: \dot{\phi}_n(t) \notin N_\epsilon(U(\phi_n(t))), t \leq T_n\} \rightarrow 0$

as $n \rightarrow \infty$, since $L_n(x, \beta) \rightarrow \infty$ uniformly in (β, x) in any compact subset of

$\{\beta, x: \beta \notin N_\epsilon(U(x))\}$.

To see this, suppose that there are $\{x_n, \beta_n\}$ and $K < \infty$ such that $L_n(x_n, \beta_n) \leq K$, where $\beta_n \notin N_\varepsilon(U(x_n))$ and $x_n \rightarrow x$, $\beta_n \rightarrow \beta$. But $\beta \notin N_\varepsilon(U(x))$ and

$\lim_n L_n(x_n, \beta_n) \geq L(x, \beta) = \infty$ by Lemma 4.1, a contradiction. By

this, the convexity of $U(x)$, continuity of $U(\cdot)$, and weak convergence of $\dot{\phi}_n(\cdot)$ to $\dot{\phi}(\cdot)$, we have $m\{t: \dot{\phi}(t) \notin \overline{U(\phi(t))}, t \leq T\} = 0$; in fact it can be shown that $\overline{U(\phi(t))}$ can be replaced by $U(\phi(t))$ there.

Now recall the definition of the δ -interior set $\overline{U}^\delta(x)$ and define $U_\varepsilon(x) = N_\varepsilon(U(x))$ and let $I_\varepsilon^n(\cdot)$ be the indicator of the set on which $\dot{\phi}_n(s) \in U_\varepsilon(\phi_n(s))$. We have

$$(4.1) \quad B_1 \equiv \lim_n \int_0^T L_n(\phi_n(s), \dot{\phi}_n(s)) ds \geq \lim_n \int_0^T L_n(\phi_n(s), \dot{\phi}_n(s)) I_\varepsilon^n(s) ds \equiv B_2.$$

Let $\delta > \varepsilon$. For large n and small δ and ε , there is a measurable function $\dot{\Delta}_n(\cdot)$ with $|\dot{\Delta}_n(t)| \leq 2\delta$ and such that $\dot{\phi}_n(t) + \dot{\Delta}_n(t) \in \overline{U}^\delta(\phi_n(t))$ for all t such that $\dot{\phi}_n(t) \in U_\varepsilon(\phi_n(t))$ and such that for these t and small δ and ε ,

$$(*) \quad L_n(\phi_n(t), \dot{\phi}_n(t) + \dot{\Delta}_n(t)) \leq L_n(\phi_n(t), \dot{\phi}_n(t)) + \delta_1,$$

where $\delta_1 \rightarrow 0$ as $\delta \rightarrow 0$.

To prove the last assertion, define $h_\delta(x, \beta)$ as follows, for $\beta \in U_\varepsilon(x) - \overline{U}^\delta(x)$. (We do the non-degenerate case, the proof in the degenerate case is almost the same.) Let $h_\delta(x, \beta)$ be the unique intersection

on $\partial \bar{U}^\delta(x)$ of the line segment $\{z: z = s\beta + (1-s)\bar{b}(x), 0 \leq s \leq 1\}$ connecting β and $\bar{b}(x)$. If $\beta \in \bar{U}^\delta(x)$, set $h_\delta(x, \beta) = \beta$. Then $[h_\delta(\phi_n(\cdot), \dot{\phi}_n(\cdot)) - \dot{\phi}_n(\cdot)] I_\epsilon^n(\cdot) \equiv \dot{\Delta}_n(\cdot)$ is measurable. Now we prove (*). Suppose that (*) is false. In particular, suppose that for each small $\delta_0 > 0$ there are $x, \beta, x_n \rightarrow x, \beta_n \rightarrow \beta, \delta_n \rightarrow 0, \epsilon_n \rightarrow 0$, with $\beta_n \in U_{\epsilon_n}(x_n)$ and such that $L_n(x_n, h_{\delta_n}(x_n, \beta_n)) - L_n(x_n, \beta_n) \geq \delta_0$. Then the convexity of $L_n(x, \cdot)$ and the fact that $h_{\delta_n}(x_n, \beta_n) - \beta_n \rightarrow 0$ as $n \rightarrow \infty$ implies that the derivative (in the direction of increasing s) at some $s = s_n \rightarrow 0$ along the line segment $\{s\beta_n + (1-s)\bar{b}(x_n)\}$ increases to ∞ as $n \rightarrow \infty$. By convexity, the derivative is non-decreasing as s increases. This and the uniform convergence $L_n(\cdot, \cdot) \rightarrow L(\cdot, \cdot)$ on $\{x, \beta: x \in \text{compact } K, \beta \in \bar{U}^\delta(x)\}$ for each K and $\delta > 0$ lead to a contradiction to (A4.2), (A4.4). In particular, we get $L(x, \bar{b}(x)) = \infty$, contradicting $L(x, \bar{b}(x)) = 0$. Thus (*) holds.

By (4.1) and (*)

$$B_2 \geq \lim_n \int_0^T L_n(\phi_n(t), \dot{\phi}_n(t) + \dot{\Delta}_n(t)) I_\epsilon^n(s) ds - \delta_1 T.$$

By Corollary 3.6,

$$\lim_\epsilon \lim_n \sup_{\substack{|x-y| \leq \epsilon \\ \beta \in \bar{U}^\delta(x)}} |L_n(y, \beta) - L_n(x, \beta)| = 0, \quad y, x \in \text{Compact in } G_1.$$

Thus, by the uniform convergence $\phi_n(\cdot) \rightarrow \phi(\cdot)$, for each $\delta_0 > 0$ there is an $\epsilon_0 > 0$ such that $|t-\tau| \leq \epsilon_0$ implies that for large n

$$|L_n(\phi_n(\tau), \beta) - L_n(\phi_n(t), \beta)| \leq \delta_0, \quad \beta \in \bar{U}^\delta(\phi_n(t)).$$

Define a finite sequence $\{t_i, i = 1, \dots, q\}$ such that

$t_{i+1} > t_i, t_{i+1} - t_i \leq \varepsilon_0, t_0 = 0, t_q = T + \varepsilon_0$, and set $L_n(\phi_n(t), \dot{\phi}_n(t)) = 0$ for $t \geq T_n$. Then (the last inequality below uses Jensen's inequality and the convexity of $L_n(x, \cdot)$)

$$(4.2) \quad B_2 \geq -\delta_1 T - \delta_0 T \\ + \lim_n \sum_i \int_{t_i}^{t_{i+1}} L_n(\phi_n(t_i), \dot{\phi}_n(s) + \dot{\Delta}_n(s)) I_\varepsilon^n(s) ds \\ \geq -(\delta_1 + \delta_0)T + \lim_n \sum_i (t_{i+1} - t_i) L_n(\phi_n(t_i), f_i^{n, \varepsilon}),$$

where

$$f_i^{n, \varepsilon} \equiv \frac{1}{(t_{i+1} - t_i)} \int_{t_i}^{t_{i+1}} (\dot{\phi}_n(s) + \dot{\Delta}_n(s)) I_\varepsilon^n(s) ds.$$

Assume (or take a suitable subsequence) that $\Delta_n(\cdot)$ converge to a function $\Delta(\cdot)$. Define

$$f_i = \left[\frac{\phi(t_{i+1}) - \phi(t_i)}{t_{i+1} - t_i} + \frac{\Delta(t_{i+1}) - \Delta(t_i)}{t_{i+1} - t_i} \right]$$

Then $f_i^{n, \varepsilon} \rightarrow f_i$ as $n \rightarrow \infty$, for each $\varepsilon > 0$. By Lemma 4.1, (4.2) and the lower semicontinuity of $L(\cdot, \cdot)$ and its continuity on $\{x, \beta : \beta \in \overline{U}^\delta(x)\}$

$$(4.3) \quad B_1 \geq -T(\delta_1 + \delta_0) + \sum_i (t_{i+1} - t_i) L(\phi(t_i), f_i) \\ \xrightarrow{\varepsilon_0 \rightarrow 0} -T\delta_1 + \int_0^T L(\phi(s), \dot{\phi}(s) + \dot{\Delta}(s)) ds.$$

Finally, letting $\varepsilon \rightarrow 0, \delta \rightarrow 0$ and again using the lower semicontinuity of

$L(\cdot, \cdot)$, yields the theorem.

Q.E.D.

We next treat an unbounded $U(x)$ case.

- (A4.5) Let (a) $\inf_{x \in G_1} \frac{L(x, \beta)}{|\beta|} \rightarrow \infty$ as $|\beta| \rightarrow \infty$,
- (b) (nondegenerate) $\sup_{|\beta| \leq B} \sup_{x \in G_1} L(x, \beta) < \infty$, all $B < \infty$.
- (degenerate), let $\beta_1 = \bar{b}_1(x)$, and take sup only over
- $|\beta_2| \leq B$.

The conditions hold for (1.2), (1.4) if (non-degenerate case) $\sigma(x)\sigma'(x)$ is uniformly positive definite, and (degenerate case) if $\sigma_2(x)\sigma_2'(x)$ is uniformly positive definite.

Theorem 4.3. Under (A1.3), (A4.1), (A4.5), the conclusions of Theorem 4.2 hold.

Proof. Let $\phi_n(\cdot) \rightarrow \phi(\cdot)$ uniformly and w.l.o.g. let $\overline{\lim}_n S_n(\phi_n) < \infty$ and

$T(\phi_n) = T_n \rightarrow T \geq T(\phi)$, $T < \infty$. For notational simplicity, we do the nondegenerate case only. The proof for the degenerate case requires only minor modifications.

Since $U(x) = \text{entire space}$, Corollary 3.6 implies that

$$(4.4) \quad \lim_{|\beta| \rightarrow \infty} \frac{\lim_n \inf_{x \in G_1} \frac{L_n(x, \beta)}{|\beta|}}{n} = \infty.$$

Also, (4.4) and $\overline{\lim}_n S_n(\phi_n) < \infty$ imply that

$$(4.5) \quad \lim_{K \rightarrow \infty} \overline{\lim}_n m\{t: t \leq T_n, |\dot{\phi}_n(t)| \geq K\} = 0.$$

Define $I_K^n(\cdot)$ = indicator of set where $|\dot{\phi}_n(t)| \leq K$. Then (4.1) holds with I_K^n replacing I_ϵ^n . By the uniform convergence of $L_n(\cdot, \cdot)$ to $L(\cdot, \cdot)$ on bounded sets and the continuity of $L(\cdot, \cdot)$, for each $\delta_0 > 0$, there are $\epsilon_0 > 0$, $\{t_i\}$, $t_0 = 0$, $0 < t_{i+1} - t_i \leq \epsilon_0$, as in Theorem 4.2, and such that

$$\begin{aligned} \frac{\lim}{n} \int_0^{T_n} L_n(\phi_n(s), \dot{\phi}_n(s)) I_K^n(s) ds &\geq -\delta_0 T + \\ &\frac{\lim}{n} \sum_i \int_{t_i}^{t_{i+1}} L_n(\phi_n(t_i), \dot{\phi}_n(s)) I_K^n(s) ds. \end{aligned}$$

The proof is completed in essentially the same way that the proof of Theorem 4.2 was completed, except that $K \rightarrow \infty$ replaces $\epsilon \rightarrow 0$, there is no need to introduce $\Delta_n(\cdot)$, and

$$(4.6) \quad \lim_K \overline{\lim}_n \int_0^{T_n} |\dot{\phi}_n(s)| (1 - I_K^n(s)) ds = 0$$

is used to get $\int_{t_i}^{t_{i+1}} \dot{\phi}_n(s) I_K^n(s) ds \rightarrow \phi(t_{i+1}) - \phi(t_i)$, as $n \rightarrow \infty$, then $K \rightarrow \infty$.

Q.E.D.

Limits of $\{S_n\}$. The functional H_n corresponds to a system of one of the types (1.1) to (1.4) with dynamical terms $b, \bar{b}, \sigma$ subscripted by n and ξ_k^n replacing ξ_k , where the 'mean' dynamical term is $\bar{b}_n(\cdot)$. As $n \rightarrow \infty$, $\bar{b}_n(x) \rightarrow \bar{b}(x)$ and many types of assumptions on the behavior of $\dot{x} = \bar{b}_n(x)$ can be dealt with. Here we simply assume (A4.6).

(A4.6) The system corresponding to H-functional H_n satisfies (A1.1), but where $\bar{x}_n$ replaces $\bar{x}_0$ and $\bar{x}_n \rightarrow \bar{x}_0$ as $n \rightarrow \infty$.

For the degenerate case, we need the 'controllability' condition (A4.7). In the non-degenerate case, with the unbounded $U(x)$, (A4.7) always holds if the conditions $|\dot{\phi}_2| \leq M$ and $\dot{\phi}_1 = \bar{b}_1(\phi)$ are replaced by $|\dot{\phi}| \leq M$. In the non-degenerate case with bounded $U(x)$, (A4.7) always holds if the condition $\dot{\phi}_2(t) \in \bar{U}_2^\delta(\phi(t))$ is replaced by $\dot{\phi}(t) \in \bar{U}^\delta(\phi(t))$.

(A4.7) (Unbounded $U(x)$ case.) There is an $M < \infty$ such that for each small $\varepsilon_2 > 0$ and each $y \in N_{\varepsilon_2}(\bar{x}_0)$, there is a function

$\phi(\cdot) = (\phi_1(\cdot), \phi_2(\cdot))$ such that $\phi(0) = \bar{x}_0$, $\phi(t_y) = y$ for some $t_y \leq T$, where $T \rightarrow 0$ as $\varepsilon_2 \rightarrow 0$, and $\dot{\phi}_1 = \bar{b}_1(\phi)$, $|\dot{\phi}_2| \leq M$.

(Bounded $U(x)$ case.) Simply replace M and $|\dot{\phi}_2| \leq M$ by $\dot{\phi}_2(t) \in \bar{U}_2^\delta(\phi(t))$, for some $\delta > 0$.

Theorem 4.4. (Unbounded $U(x)$ case) Assume (A4.1), (A4.5), (A4.6) and (A4.7) (for the degenerate case) and (A1.1), (A1.3), (A1.4). Then

$$S_n \rightarrow S_0.$$

Note (A1.2) is not used here. The theorem makes no direct claim concerning escape times and the H-functionals are defined by (1.5).

Proof. Fix $\varepsilon > 0$, let $S_0 < \infty$ and let $\phi^\varepsilon(\cdot)$ be an ε -optimal path for $S(\cdot)$ with $\phi^\varepsilon(0) = \bar{x}_0$. Write $T_\varepsilon = T(\phi^\varepsilon)$. Below, we show that for small $\varepsilon_3 > 0$, $\phi^\varepsilon(\cdot)$ can be selected such that it is defined until T'_ε , the exit time from $N_{\varepsilon_3}(G)$, and $S(T'_\varepsilon, \phi^\varepsilon) \leq S_0 + 3\varepsilon$ and for some

$K < \infty$, $|\dot{\phi}^\varepsilon(t)| \leq K$, and $\phi^\varepsilon(\cdot)$ is not tangent to any of the boundary

curves at the exit point from G . Assume this for the moment. In this part of the proof we do only the (more difficult) degenerate problem.

Define $\phi_n^\varepsilon = (\phi_{1n}^\varepsilon, \phi_{2n}^\varepsilon)$ by (in the non-degenerate case we would set $\phi_n^\varepsilon(0) = \bar{x}_n, \dot{\phi}_n^\varepsilon(t) = \dot{\phi}^\varepsilon(t)$)

$$\phi_{1n}^\varepsilon(t) = \bar{x}_{1n} + \int_0^t \bar{b}_n(\phi_n^\varepsilon(s)) ds$$

$$\phi_{2n}^\varepsilon(t) = \bar{x}_{2n} + \int_0^t \dot{\phi}_2^\varepsilon(s) ds,$$

where $\bar{x}_n$ is defined in (A4.6), and $\bar{b}_n = (\bar{b}_{1n}, \bar{b}_{2n})$. Recall that $\bar{b}(\cdot)$ is Lipschitz continuous. Then, by the properties of ϕ^ε assumed in the last paragraph, $T_n^\varepsilon = T(\phi_n^\varepsilon) < \infty$

for large n and $T_n^\varepsilon \rightarrow T$ as $n \rightarrow \infty$. By the boundedness of $\dot{\phi}^\varepsilon(\cdot)$ and the uniform convergence of $L_n(x; \bar{b}_{1n}(x), \beta_2)$ to $L(x; \bar{b}_1(x), \beta_2)$ on bounded (x, β_2) sets,

$$(4.7a) \quad S_n \equiv S_n(\bar{x}_n) = S_n(T_n^\varepsilon, \phi_n^\varepsilon) = \int_0^{T_n^\varepsilon} L_n(\phi_n^\varepsilon(s); \bar{b}_{1n}(\phi_n^\varepsilon(s)), \dot{\phi}_2^\varepsilon(s)) ds \\ \longrightarrow \int_0^T L(\phi^\varepsilon(s); \bar{b}_1(\phi^\varepsilon(s)), \dot{\phi}_2^\varepsilon(s)) ds \leq S_0 + 3\varepsilon$$

Thus

$$(4.7b) \quad \lim_n S_n \leq S_0.$$

We now show that there is a $\phi^\varepsilon(\cdot)$ of the desired form.

Let $\phi^\varepsilon(\cdot)$ be an ε -optimal path for $S(\phi)$ with $\phi^\varepsilon(0) = \bar{x}_0$. We do the non-degenerate case only, for the sake of notational simplicity. A very similar construction yields $\phi^\varepsilon(\cdot)$ of the desired form for the degenerate case.

Let $I_\varepsilon^K(\cdot)$ denote the indicator of the set where $|\dot{\phi}^\varepsilon(s)| \leq K$. By (A4.5) and $S_0 < \infty$,

$$(4.8) \quad \lim_{K \rightarrow \infty} \int_0^T |\dot{\phi}^\varepsilon(s)| (1 - I_\varepsilon^K(s)) ds = 0.$$

For $y \in G$ and any $K < \infty$ define $\tilde{\phi}_y^\varepsilon(\cdot)$ by

$$\tilde{\phi}_y^\varepsilon(t) = y + \int_0^t \dot{\phi}^\varepsilon(s) I_\varepsilon^K(s) ds.$$

There is an $M < \infty$ and $\varepsilon_2 > 0$ such that for each y satisfying

$|y - \bar{x}_0| \leq \varepsilon_2$ there is a $\hat{\phi}_y^\varepsilon(\cdot)$ satisfying $\hat{\phi}_y^\varepsilon(0) = \bar{x}_0$, $\hat{\phi}_y^\varepsilon(t_y) = y$, with

$|\dot{\hat{\phi}}_y^\varepsilon(t)| \leq M$ and $S(t_y, \hat{\phi}_y^\varepsilon) \leq \varepsilon$, where $t_y \rightarrow 0$ as $\varepsilon_2 \rightarrow 0$. Define

$\bar{\phi}_y^\varepsilon(\cdot)$ by

$$\begin{aligned} \bar{\phi}_y^\varepsilon(t) &= \hat{\phi}_y^\varepsilon(t), \quad t \leq t_y \\ &= \tilde{\phi}_y^\varepsilon(t - t_y), \quad t > t_y. \end{aligned}$$

By (4.8) and the continuity of $L(\cdot, \cdot)$, we can find a sequence $\{y_\alpha, K_\alpha\}$

where $K_\alpha \rightarrow \infty$ as $\alpha \rightarrow \infty$ and such that for large α , $\bar{\phi}_{y_\alpha}^\varepsilon(\cdot)$ satisfies the

conditions required on $\phi^\varepsilon(\cdot)$ in the first paragraph of the proof (where

ε_3 now depends on the chosen y_α, K_α). Recall $S_n = \inf \{S_n(\phi) : \phi(0) = \bar{x}_n\} = S_n(\bar{x}_n)$.

Now, to get the reverse inequality to (4.7b) for either the degenerate or the non-degenerate case, let $\sup_n S_n < \infty$ and let $\phi_n^\varepsilon(\cdot)$ be the ε -optimal path for $S_n(\phi)$. We can select $\phi_n^\varepsilon(\cdot)$

such that $T_n^\epsilon = T(\phi_n^\epsilon) \rightarrow T < \infty$. Let $I_\epsilon^{K,n}(\cdot)$ denote the indicator of the set where $|\dot{\phi}_n^\epsilon(t)| \geq K$. By (A4.5), the convexity of $L_n(x, \cdot)$ and $L(x, \cdot)$ and the uniform convergence on bounded sets, for each large $N < \infty$ there is a $K_N < \infty$ such that

$$(4.9) \quad S_n(\phi_n^\epsilon) \geq N \int_0^{T_n^\epsilon} |\dot{\phi}_n^\epsilon(t)| I_\epsilon^{K_N, n}(t) dt$$

for large n . Thus, the set $\{\phi_n^\epsilon(\cdot), n \text{ large}, \epsilon > 0\}$ is uniformly absolutely continuous. Extract a convergent subsequence, indexed by n , and with limit $\bar{\phi}^\epsilon(\cdot)$, where $\bar{\phi}^\epsilon(0) = \bar{x}_0$. By Theorem 4.3,

$$(4.10) \quad \epsilon + \lim_n S_n(\bar{x}_n) \geq \lim_n S_n(\phi_n^\epsilon) \geq S(\bar{\phi}^\epsilon) \geq S_0,$$

$$\lim_n S_n \geq S_0.$$

Thus, $S_n \rightarrow S_0$.

Q.E.D.

A useful special case is given by Theorem 4.5. See also Theorem 4.6.

Theorem 4.5. Let the H-functionals satisfy $H_n(x, \alpha) \rightarrow H(x, \alpha)$, each x, α .

Then $S_n \leq S_0$ and under the conditions of Theorems 4.2 or 4.3, $S_n \rightarrow S_0$ as $n \rightarrow \infty$.

The theorem is obvious, since $L_n(x, \beta) \leq L(x, \beta)$. A case of particular interest is where $b(x, \xi_n) = \tilde{b}(x, \xi_n) + \hat{b}_n(x, \xi_n)$, and $\{\tilde{\xi}_n\}$ and $\{\hat{\xi}_n\}$ are independent of one another and $\hat{H}_n(x, \alpha) \rightarrow 0$, uniformly on bounded (x, α) sets (where $\hat{H}$ and $\tilde{H}$ are the H-functionals corresponding to $\hat{b}$ and $\tilde{b}$,

respectively). Then if the system corresponding to $\tilde{b}$ satisfies the conditions of Theorems 4.2 or 4.3, $S_n \rightarrow \tilde{S}_0$ as $n \rightarrow \infty$.

The H-function for (1.2) or (1.4) takes the form (where H^0 is the H-functional for $\sigma = 0$)

$$H^\sigma(x, \alpha) = H^0(x, \alpha) + \alpha' \sigma(x) \sigma'(x) \alpha / 2.$$

Theorem 4.6. Let $H_n^\sigma(x, \alpha) = H_n^0(x, \alpha) + \alpha' \sigma(x) \sigma'(x) \alpha / 2$, where we assume the conditions of Theorem 4.4 with H^σ and H_n^σ replacing H and H_n resp.

Then $S_n^\sigma \rightarrow S_0^\sigma$ as $n \rightarrow \infty$. Furthermore, if H^0 satisfies (A4.1 to 4), (A1.3) in the bounded $U(x)$ case or (A1.3), (A4.1), (A4.3) in the unbounded $U(x)$ case, then $S_0^\sigma \rightarrow S_0^0$ as $\sigma \rightarrow 0$.

The theorem follows from Theorems 4.2 to 4.5. Thus, when the system contains (independent) Gaussian noise, the exit times are robust with respect to changes in the other system noises. Also, the addition of small Gaussian noise changes the exit times only slightly under broad conditions.

Theorem 4.7. (Bounded $U(x)$) Assume (A1.1, 3, 4) and (A4.1, 2, 3, 4, 6, 7).

Suppose that for each $\epsilon < 0$, there is a $\delta > 0$ such that there is an ϵ -optimal path $\phi^\epsilon(\cdot)$ (with $\phi^\epsilon(0) = \bar{x}_0$) for $S(\cdot)$, with $\dot{\phi}^\epsilon(t) \in \bar{U}^\delta(\phi^\epsilon(t))$. Then $S_n \rightarrow S_0$ as $n \rightarrow \infty$.

The proof uses arguments developed in the theorems of this section and only a few comments will be made. To get (4.7b) we roughly follow the proof of that result in Theorem 4.4. The controllability (A4.7), the continuity of $U(\cdot)$, and a piecing together argument (such as used for the construction of $\bar{\phi}_y^\epsilon(\cdot)$ in Theorem 4.4) are used to get an ϵ -optimal path (starting from $\bar{x}_0$) which satisfies the requirements of

the third sentence of the proof of Theorem 4.4, except that

$|\dot{\phi}_2^\varepsilon(t)| \leq M$ is replaced by (degenerate case) $\dot{\phi}_2^\varepsilon(t) \in \bar{U}_2^\delta(\phi^\varepsilon(t)), \dot{\phi}_1^\varepsilon(t) = \bar{b}_1(\phi^\varepsilon(t))$, or

(non-degenerate case) $\dot{\phi}^\varepsilon(t) \in \bar{U}^\delta(\phi^\varepsilon(t))$, for some $\delta > 0$. Then define

$\phi_n^\varepsilon(\cdot)$ as in the second paragraph of the proof of Theorem 4.4, and in the analogous way for the non-degenerate case. There is a

$\delta' > 0$ such that for large n , (degenerate case) $\dot{\phi}_{2n}^\varepsilon(t) \in \bar{U}_{2n}^{\delta'}(\phi_n^\varepsilon(t))$

(and $\dot{\phi}_n^\varepsilon(t) \in \bar{U}^{\delta'}(\phi_n^\varepsilon(t))$ for the non-degenerate case.

Then use (4.7a) (or the analogous formula for the non-degenerate case) and the convergence $L_n(\cdot, \cdot) \rightarrow L(\cdot, \cdot)$ uniformly on $\{x, \beta: x \in \text{compact } K, \beta \in \bar{U}^{\delta'}(x)\}$ to get (4.7b). The proof of (4.10) is very similar to the proof used in Theorem 4.4, whether or not the $U_n(x)$ are bounded. The appropriate convergent subsequence of $\{\phi_n^\varepsilon(\cdot)\}$ is extracted by using the nature of the convergence of $L_n(\cdot, \cdot) \rightarrow L(\cdot, \cdot)$ and the boundedness of the $U(x)$.

In the next theorem we show that the $\bar{U}^\delta(x)$ - approximation required by the last theorem exists under reasonable conditions. We actually show the existence of a slightly modified set, called $\tilde{U}^\delta(x)$, which can be used in place of $\bar{U}^\delta(x)$. For $0 < \delta \leq 1$, define $\{x_k^{\gamma, \delta}\}$ by

$$x_{k+1}^{\gamma, \delta} = x_k^{\gamma, \delta} + \gamma \bar{b}(x_k^{\gamma, \delta}) + \gamma(1-\delta) \tilde{b}(x_k^{\gamma, \delta}, \varepsilon_k), \quad x_0^{\gamma, \delta} = x_0^\gamma = \bar{x}_0,$$

where $\tilde{b} = b - \bar{b}$. Let L^δ denote the L -functional for $\{x_k^{\gamma, \delta}\}$ and let $\tilde{H}$ denote the H -functional for $\tilde{b}$. Then

$$\begin{aligned} L^\delta(x, \beta) &= \sup_{\alpha} [\alpha'(\beta - \bar{b}(x)) - \tilde{H}(x, (1-\delta)\alpha)] \\ &= L(x, \frac{v}{1-\delta} + \bar{b}(x)), \end{aligned}$$

where $v = \beta - \bar{b}(x)$. Define $\tilde{U}^\delta(x)$ by: $\beta \in \tilde{U}^\delta(x)$ if $\beta = \bar{b}(x) + (1-\delta)v$, where

$\bar{b}(x) + v \in U(x)$. Clearly, under (A4.4), $\tilde{U}^\delta(x)$ can be used instead of $\bar{U}^\delta(x)$ in the previous theorems (analogously for $\tilde{U}_2^\delta(x)$ and $\bar{U}_2^\delta(x)$ in the degenerate case).

Define $\tilde{L}^\delta(\cdot, \cdot)$ by $\tilde{L}^\delta(x, \beta) = L(x, \beta)$ for $\beta \in \tilde{U}^\delta(x)$, and equal to infinity otherwise. Let $\tilde{S}^\delta(\cdot)$ denote the action functional corresponding to $\tilde{L}^\delta(\cdot)$. Let $x^{\gamma, \delta}(\cdot)$ denote the piecewise linear interpolation of $\{x_k^{\gamma, \delta}\}$ with interpolation interval γ .

Theorem 4.8. Under (A4.1 to 4) and (A1.1 to 5), $\tilde{S}_0^\delta \rightarrow S_0$ as $\delta \rightarrow 0$. If (A4.6,7) also holds, then $S_n \rightarrow S_0$.

Remark. The first sentence of the theorem implies that $\tilde{U}^\delta(x)$ satisfies the requirements put on $\bar{U}^\delta(x)$ in Theorem 4.7.

Proof. For notational reasons only, we work with the non-degenerate case.

First we show that for each compact x -set K there is a $c(\delta)$ which goes to 0 as $\delta \rightarrow 0$ and such that (if $\beta \notin \tilde{U}^\delta(x)$, then both sides are infinite)

$$(4.11) \quad \tilde{L}^\delta(x, \beta) \leq L^\delta(x, \beta) + c(\delta), \quad x \in K.$$

Suppose (4.11) is false. Then there are $c > 0$, $\delta_n \rightarrow 0$, $x_n \in K$ and $\bar{b}(x_n) + v_n = \beta_n \in \tilde{U}^{\delta_n}(x_n)$ and $\delta_n \rightarrow 0$ such that (recall the form of L^δ given above the theorem)

$$(*) \quad L(x_n, \bar{b}(x_n) + v_n) - L(x_n, \frac{v_n}{1-\delta} + \bar{b}(x_n)) \geq c.$$

This relation is impossible unless $d(\bar{b}(x_n) + v_n, \partial U(x_n)) \rightarrow 0$ as $n \rightarrow \infty$. Using this and (*) and the convexity of the $L(x, \cdot)$, we get that $L(x_n, \bar{b}(x_n)) \rightarrow \infty$ as $n \rightarrow \infty$, a contradiction. Thus (4.11) holds. We can show that

$$(4.12) \quad \lim_{\delta \rightarrow 0} \tilde{S}_0^\delta \geq S_0$$

by a proof similar to that in the last part of Theorem 4.4. (See also the comment after Theorem 4.7).

We now adapt a device used in [5, Lemma 5.1]. Let $S_0 < \infty$. For $T < \infty$ and small $\rho > 0$, define the sets $A^{\gamma, \delta} =$ event that $x^{\gamma, \delta}(\cdot)$ leaves G by time T , and $A_\rho^\gamma =$ event that $x^\gamma(\cdot)$ leaves $N_\rho(G)$ by time T . For small $\delta > 0$, $P\{A_\rho^{\gamma, \delta}\} \geq P\{A_\rho^\gamma\}$. For each small $h > 0$, there is a $\rho > 0$, a $T^{h, \rho}$ (which we can suppose is bounded uniformly in h, ρ) and a function $\phi^{h, \rho}(\cdot)$ such that $\phi^{h, \rho}(0) = \bar{x}_0$, $\phi^{h, \rho}(T^{h, \rho}) \in \partial N_\rho(G)$ and (recall (A1.5)) $S_0 \leq S(T^{h, \rho}, \phi^{h, \rho}) \leq S_0 + h$. Then, by (1.6), $P\{A_\rho^\gamma\} \geq \exp - [S_0 + 2h]/\gamma$ for small γ . Let $\tilde{\Lambda} =$ set of continuous paths $\phi(\cdot)$ which leave G by time T and have $\phi(0) = \bar{x}_0$. Then, for each h there is a $\gamma_0 > 0$ such that for $\gamma \leq \gamma_0$,

$$P\{A^{\gamma, \delta}\} \leq \exp - \left[\inf_{\phi \in \tilde{\Lambda}} S^\delta(T, \phi) - h \right] / \gamma$$

Combining (4.11) with the estimates in the last paragraph yields that for some $T_0 < \infty$

$$(4.13) \quad \tilde{S}_0^\delta - c(\delta)T_0 \leq S_0^\delta \leq \inf_{\phi \in \tilde{\Lambda}} S^\delta(T, \phi) \leq S_0 + 3h,$$

where T_0 can be taken to be an upper bound (over small $\delta > 0$) for $\{T(\phi^\delta)\}$, where ϕ^δ are such that $S^\delta(\phi^\delta) \leq S^\delta + \delta$. Combining (4.12) and (4.13) yields the first assertion of the theorem.

Using $\tilde{U}^\delta(x)$ for the $\bar{U}^\delta(x)$ in Theorem 4.7 yields the last assertion of the Theorem. Q.E.D.

There are also approximation Theorems for the inequalities (1.7).

Let A be a set of continuous functions on $[0, T]$. Define $S_n(T, A) =$

$$\inf_{\phi \in A} S_n(T, \phi).$$

Then

$$-S_n(T, A^0) \leq \liminf_{\gamma} \gamma \log P\{x_n^\gamma(\cdot) \in A\} \leq \overline{\lim}_{\gamma} \gamma \log P\{x_n^\gamma(\cdot) \in A\} \leq -S_n(T, \bar{A}),$$

where H_n is the H-functional arising from the processes $\{x_n^\gamma(\cdot), \gamma > 0\}$,

$n = 1, 2, \dots$, each of which is of the form (1.1), (1.2) or the interpolation of forms (1.3), (1.4) for suitable $b_n, \sigma_n, \bar{b}_n, \xi_k^n$ replacing $b, \sigma, \bar{b}, \xi_k$, and S_n is the Cramer transformation of H_n . Let $A, \varepsilon_0 > 0$ and compact G_0 be such that for $\phi(\cdot) \in N_{\varepsilon_0}(A)$, $\phi(T) \in G_0$, $t \leq T$.

Theorem 4.9. Let $S(T, A^0) = S(T, \bar{A})$ if the $U(x)$ are bounded. Assume (A1.2', 3)

where G_1 is replaced by G_0 in (A1.2'). Assume (A4.1, 2) and also

(A4.3, 4) for the bounded $U(x)$ case, and (A4.5) for the unbounded $U(x)$

case. Then in the bounded $U(x)$ case, $S_n(T, A) \rightarrow S(T, A^0)$, and in the unbounded $U(x)$ case, $\lim_n S_n(T, \bar{A}) \geq S(T, \bar{A})$, $\lim_n S_n(T, A^0) \leq S(T, A^0)$.

Proof. Only an outline will be given. The techniques are similar to those used in the previous theorems of this section. Fix $\varepsilon > 0$. Let $\phi_n^\varepsilon(\cdot)$ be such that $S_n(T, \phi_n^\varepsilon) \leq S_n(T, \bar{A}) + \varepsilon$. We can always choose such a sequence for which there is a convergent subsequence. Let n index the subsequence

and denote the limit by $\bar{\phi}^\epsilon(\cdot)$. Then $\bar{\phi}^\epsilon(\cdot) \in \bar{A}$ and by Theorem 4.2 and the arbitrariness of $\epsilon > 0$, we have

$$(4.14) \quad \liminf_n S_n(T, \bar{A}) \geq S(T, \bar{A}).$$

To get the reverse inequality, first consider the 'unbounded $U(x)$ ' case, and let $\phi^\epsilon(\cdot)$ be an ϵ -optimal path for $S(T, A^0)$ such that $|\dot{\phi}^\epsilon(t)|$ is bounded. Then use an argument similar to that used in connection with (4.7a) to get

$$(4.15) \quad \limsup_n S_n(T, A^0) \leq S(T, A^0),$$

and the Theorem is proved for the unbounded $U(x)$ case. We need not concern ourselves with 'exit times' in this Theorem.

Now, to complete the proof for the bounded $U(x)$ case, we use the technique and terminology of Theorem 4.8. Define $x^{\gamma, \delta}(\cdot)$ as above Theorem 4.8. Suppose that A^0 is non empty and for small $\rho > 0$ define the open set $A^\rho = \{\phi: \phi \in A, d(\phi, \partial A) > \rho\}$. For small $\delta > 0$

$$(4.16) \quad P\{x^{\gamma, \delta}(\cdot) \in A\} \geq P\{x^\gamma(\cdot) \in A^\rho\},$$

and for each $h > 0$ there is a $\gamma(h) > 0$ such that for $\gamma \leq \gamma(h)$, the right side of (4.16) is $\geq \exp - [S(T, A^\rho) + h]/\gamma$ and the left side is $\leq$

$\exp - [S^\delta(T, \bar{A}) - h]/\gamma$. Using the terminology and result of Theorem 4.8,

$\tilde{S}^\delta(T, \bar{A}) - c(\delta)T \leq S^\delta(T, \bar{A})$. Now, by the hypothesis $S(T, A^0) = S(T, \bar{A})$,

we have that $S(T, A^\rho) + S(\bar{A}) = S(A^0)$ as $\rho \rightarrow 0$.

Now

$$(4.17) \quad h + S(T, A^0) \xleftarrow{\rho \rightarrow 0} S(T, A^\rho) + h \geq S^\delta(T, \bar{A}) - h \geq \tilde{S}^\delta(T, \bar{A}) - c(\delta)T, - h.$$

Also, as in Theorem 4.8,

$$(4.18) \quad \lim_{\delta} \tilde{S}^\delta(T, \bar{A}) \geq S(T, \bar{A}) = S(T, A^0).$$

Relations (4.17), (4.18) imply that $\lim_{\delta \rightarrow 0} \tilde{S}^\delta(T, A^0) = S(T, A^0)$. Thus, for

small $\delta > 0$, we can select an ε -optimal path ϕ^ε for $S(T, A^0)$ with $\dot{\phi}^\varepsilon(t) \in \tilde{U}^\delta(\phi^\varepsilon(t))$ (or, equivalently, in $\bar{U}^\delta(\phi^\varepsilon(t))$, $t \leq T$, if we wish).

The proof of (4.15) follows from this, the convergence of $L_n(\cdot, \cdot)$ to $L(\cdot, \cdot)$ uniformly on $\{x, \beta: x \in G_0, \beta \in \bar{U}^\varepsilon(x)\}$ (or the analogous result for the degenerate case) and the boundedness of $U(x)$. Q.E.D.

5. Examples of convergence of H_n to Gaussian H-functional.

5.1 Let $\bar{b}_n(\cdot)$ and $b_n(\cdot, \xi)$ be Lipschitz continuous, uniformly in ξ . Let

$N_n \rightarrow \infty$ as $n \rightarrow \infty$, and let $\{\xi_{ki}, i \geq 0, k \geq 0\}$ be i.i.d. with

$$E \tilde{b}_n(x, \xi_{ki}) = 0, \text{ and define } \tilde{b}_n(x, \xi_k^n) = \sum_{i=1}^{N_n} \tilde{b}_n(x, \xi_{ki}) / \sqrt{N_n}. \text{ Define } \{x_k^Y\}$$

by (suppress the n index on x_k^Y) by

$$(5.1) \quad x_{k+1}^Y = x_k^Y + \gamma \bar{b}_n(x_k^Y) + \gamma \tilde{b}_n(x_k^Y, \xi_k^n).$$

Let H_n denote the H-functional when $\bar{b}_n(x) = 0$. For convergence to the

Gaussian H-functional $H(x, \alpha) = \alpha' \bar{b}(x) + \alpha' \Sigma(x) \alpha / 2$ we need $\bar{b}_n(x) \rightarrow \bar{b}(x)$ and

$$\tilde{H}_n(x, \alpha) = N_n \log E \exp \alpha' \tilde{b}_n(x, \xi_{ki}) / \sqrt{N_n} \rightarrow \alpha' \Sigma(x) \alpha / 2,$$

uniformly in $x \in G_1$ for some smooth $\Sigma(x)$. If the ξ_{ki} are bounded, then clearly $\Sigma(x) = \lim_{n \rightarrow \infty} E \tilde{b}_n(x, \xi_{ki}) \tilde{b}_n'(x, \xi_{ki})$. But in general, the convergence or lack of it depends on the higher moments of $\tilde{b}_n(x, \xi_{ki})$.

5.2. Now let $\tilde{b}_n(x, \xi_k^n) = \sum_{i=1}^{N_n} \tilde{b}_n(x, \xi_{ki}^n)$, where N_n is Poisson with parameter λ_n , and for each n , $E \tilde{b}_n(x, \xi_{ki}^n) = 0$ and $\{\xi_{ki}^n, k \geq 0, i \geq 0\}$ are i.i.d. for each n .

Then

$$\tilde{H}_n(x, \alpha) = \lambda_n [E \exp \alpha' \tilde{b}_n(x, \xi_{ki}^n) - 1].$$

Let $\lambda_n \rightarrow \infty$ and $\lambda_n E \tilde{b}_n(x, \xi_{ki}^n) \tilde{b}_n'(x, \xi_{ki}^n) \rightarrow \Sigma(x)$ uniformly in $x \in G_1$, as $n \rightarrow \infty$.

Then, for H_n to converge to the Gaussian H-functional, it is sufficient

that $\tilde{b}_n(x) \rightarrow \bar{b}(x)$ and that $\lambda_n \sum_{\ell=3}^{\infty} |\alpha|^\ell E |\tilde{b}_n(x, \xi_{ki}^n)|^\ell / \ell! \rightarrow 0$ uniformly

in bounded α -sets, as $n \rightarrow \infty$. This depends on the higher moments of

$\tilde{b}_n(x, \xi_{ki}^n)$. If $|\tilde{b}_n(x, \xi_{ki}^n)| \leq \delta_n \rightarrow 0$ as $n \rightarrow \infty$ and $\overline{\lim}_n \lambda_n \delta_n^2 < \infty$, then

the sum converges to zero as desired.

5.3. Consider the continuous parameter case

$$(5.2) \quad dx^\gamma = \bar{b}(x^\gamma) dt + \sigma(x^\gamma) dJ^n(t/\gamma),$$

where $J^n(\cdot)$ is a centered Poisson jump process with rate λ_n and jump

random variables $\{\xi_i^n\}$. Then

$$\begin{aligned} H_n(x, \alpha) &= \alpha' \bar{b}(x) + \lim_{\gamma \rightarrow 0} \gamma \log E \exp \alpha' \int_0^{1/\gamma} \sigma(x) dJ^n(t) \\ (5.3) \quad &= \alpha' \bar{b}(x) + \lambda_n [E \exp \alpha' \sigma(x) \xi_i^n - 1], \end{aligned}$$

and the comments made in the discrete parameter case also apply here.

5.4 Let $J(\cdot)$ be a jump process with jump intervals $c > 0$ and bounded i.i.d. jump random variables $\{\psi_i\}$ with $E\psi_i = 0$ and consider the system

$$(5.4) \quad \dot{x}^\gamma = \bar{b}(x^\gamma) + v(x^\gamma) \xi(t/\gamma),$$

where $\xi(\cdot)$ is the filtered noise

$$(5.5) \quad \xi(t) = \int_0^t h(t-s) dJ(s).$$

For computational simplicity, let $h(s) = \exp -as$, $a > 0$ and set

$K_\gamma = 1/c\gamma$. Then

$$\begin{aligned} \int_0^{1/\gamma} \xi(t) dt &= \int_0^{1/\gamma} dt \int_0^t h(t-s) dJ(s) \\ &= \int_0^{1/\gamma} dJ(s) \int_s^{1/\gamma} h(t-s) dt = \sum_{i \leq 1/c\gamma} \frac{\psi_i}{a} [1 - \exp -ac(\frac{1}{c\gamma} - i)] \end{aligned}$$

Thus

$$\begin{aligned} &\lim_{\gamma \rightarrow 0} \gamma \log E \exp \alpha v(x) \int_0^{1/\gamma} \xi(s) ds \\ (5.6) \quad &= \lim_{\gamma \rightarrow 0} \gamma \log (E \exp \frac{\alpha v(x)}{a} \psi)^{K_\gamma} \\ &= \frac{1}{c} \log E \exp \frac{\alpha v(x)}{a} \psi. \end{aligned}$$

Now, replace (c, a, ψ_i) by (c_n, a_n, ψ_i^n) , let $|\psi_i^n| \leq \delta_n$, where $\delta_n/a_n \rightarrow 0$ as $n \rightarrow \infty$. Let $\lim E (\psi^n)^2 / a_n^2 c_n^2 = u^2 > 0$. Then as $n \rightarrow \infty$, (5.6) converges to the Gaussian form $\alpha^2 v^2(x) u^2 / 2$. If the deterministic intervals c were replaced by i.i.d. and exponentially distributed intervals, the (5.4), (5.5) would be close to actual physical noise models. We expect that the same conclusions would hold in this case, suggesting that the Gaussian approximation is indeed useful for a large class of physical noise models.

6. A Phase Locked Loop (PLL) Example.

This example does not completely fit the previous theorems, but it represents an important and interesting class of applications where further work is required, but where approximation theorems such as those here are essential if the results are to be physically meaningful. Let $\{z_i(\cdot), \psi_i, i = 1, 2\}$ be mutually independent with $z_i(\cdot)$ scalar valued Gaussian with mean zero and integrable covariance function $\rho(\cdot)$, and ψ_i uniformly distributed on $[0, 2\pi]$. A standard method of representing wide bandwidth but 'band pass' noise $n^Y(\cdot)$ in communication systems is by using the form

$$n^Y(t) = [z_1^Y(t) \cos(\omega_0^Y t + \psi_1) + z_2^Y(t) \sin(\omega_0^Y t + \psi_2)],$$

where $\omega_0^Y = \omega_0/\gamma$, $z_i^Y(t) = z_i(t/q_\gamma)$, $\gamma/q_\gamma \rightarrow 0$ as $\gamma \rightarrow 0$. For notational

simplicity, we set $\psi_i = 0$. The bandwidth of $n^Y(\cdot)$ is $O(1/q_\gamma)$ and the center

frequency $O(1/\gamma)$. Let the input $y^\gamma(\cdot)$ to the system be the sum of a signal plus noise

$$y^\gamma(t) = A_0^\gamma(t) \sin(\omega_0^\gamma t + \theta) + n^\gamma(t),$$

where $A_0^\gamma(t) = A_0(t/q_\gamma)$ is a deterministic signal. Suppose that there is a constant $\bar{A} \neq 0$ such that the convergence

$$\lim_{\gamma} q_\gamma \int_{T_0/q_\gamma}^{T_1/q_\gamma} A_0(t) dt = \bar{A}(T_1 - T_0)$$

is uniform in $(T_1 - T_0)$. As noted in more detail at the end of this section, the function of the PLL is to track changes in $\theta(\cdot)$, a job of fundamental importance in many modern low error communications systems. [10], [11]. The scaling used here for the input signal and noise allows us to exploit the asymptotic method, but the general type of scaling used is consistent with that required by many applications where the center frequencies and bandwidths are large but the bandwidth is small relative to the center frequency. In fact to use asymptotic methods on such problems (i.e., to be able to replace the actual system noises by simple stochastic processes), such a relation between the bandwidth and center frequency seems to be required.

The dynamical equations of the two forms of PLL of Fig. 1 are given by (6.1a) and (6.1b), respectively.

$$\begin{aligned} \dot{v}^\gamma &= Dv^\gamma + Hw^\gamma, \quad \dot{\hat{\theta}}^\gamma = cv^\gamma, \\ (6.1a) \quad w^\gamma(t) &= y^\gamma(t) \cos(\omega_0^\gamma t + \hat{\theta}^\gamma(t)). \end{aligned}$$

$$(6.1b) \quad \dot{\hat{\theta}}^\gamma = Kw^\gamma.$$

Here $\hat{\theta}^\gamma(\cdot)$ is the systems estimate of $\theta(\cdot)$ and the $\cos(\omega_0^\gamma t + \hat{\theta}^\gamma)$ term is generated by the systems 'voltage controlled oscillator'. Also $v^\gamma(\cdot)$ is the state of a stable filter which is used in the 'forward' path in Figure 1a. A trigonometric expansion of $w^\gamma(t)$ yields terms involving sin or cos of $2\omega_0^\gamma t$. If we retain these terms, then their effects would drop out below when $\lim_{\gamma \rightarrow 0}$ is taken. So, for convenience, we expand $w^\gamma(t)$, drop these 'high frequency' terms and replace $w^\gamma(t)$ by

$$u^\gamma(t) = A_0^\gamma(t) \sin(\theta - \hat{\theta}^\gamma)/2 + [z_1^\gamma(t) \cos \hat{\theta}^\gamma - z_2^\gamma(t) \sin \hat{\theta}^\gamma]/2.$$

For (6.1a) define $x = (v, \hat{\theta})$, $\alpha = (\alpha_1, \alpha_2)$ and set

$$\bar{b}(x) = \begin{Bmatrix} Dv + H\bar{A} \sin(\theta - \hat{\theta})/2 \\ Cv \end{Bmatrix} = \begin{Bmatrix} \bar{b}_1(x) \\ \bar{b}_2(x) \end{Bmatrix}.$$

Then the H-functional for (6.1a) is (note that the appropriate scaling is q_γ not γ here)

$$\begin{aligned} H(x, \alpha) &= \alpha' \bar{b}(x) + \lim_{\gamma \rightarrow 0} q_\gamma \log E \exp \alpha_1' H/2 \int_0^{1/q_\gamma} [z_1(t) \cos \hat{\theta} - z_2(t) \sin \hat{\theta}] dt \\ (6.2) \quad &= \alpha' \bar{b}(x) + \frac{(\alpha_1' H)^2}{2} \bar{\sigma}^2 \end{aligned}$$

$$\bar{\sigma}^2 = \frac{1}{4} \int_{-\infty}^{\infty} \rho(s) ds.$$

This is also the H-functional for the system

$$\begin{aligned} dv &= \bar{b}_1(v) dt + H \bar{\sigma} \sqrt{q_\gamma} dw, \\ \dot{\hat{\theta}} &= Cv \end{aligned}$$

where $w(\cdot)$ is a standard Wiener process. Since $\bar{w}^\gamma(t) =$

$\int_0^t [z_1^\gamma(s) \cos \hat{\theta}^\gamma - z_2^\gamma(s) \sin \hat{\theta}^\gamma] ds / 2\sqrt{q_\gamma}$ converges weakly to a Wiener process

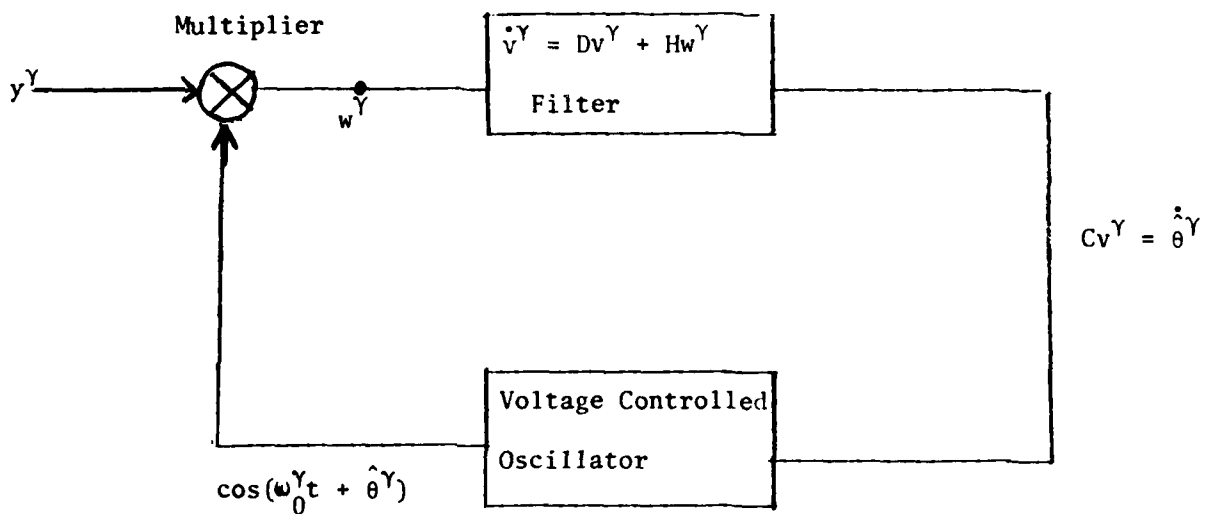
$\bar{w}(\cdot)$ with variance $\sigma^2 t$, the small 'white noise' approximation to (6.1) makes sense here. But we are not aware of a proof that $H(x, \alpha)$ actually gives an action functional and the exit time formulas (1.6) to (1.8), for the systems of (6.1a or b) where the normalizing factor γ is replaced by q_γ . Possibly such a proof can be based on Azencott [1] for this purely Gaussian process. In any case, it is not adequate to simply proceed from there, without some sort of limit or approximation argument.

Although $\bar{w}^\gamma(\cdot)$ converges weakly to $\bar{w}(\cdot)$, if the $z_i(\cdot)$ are only (sufficiently) strongly mixing but not Gaussian, the H-functionals are not usually of the form (6.2).

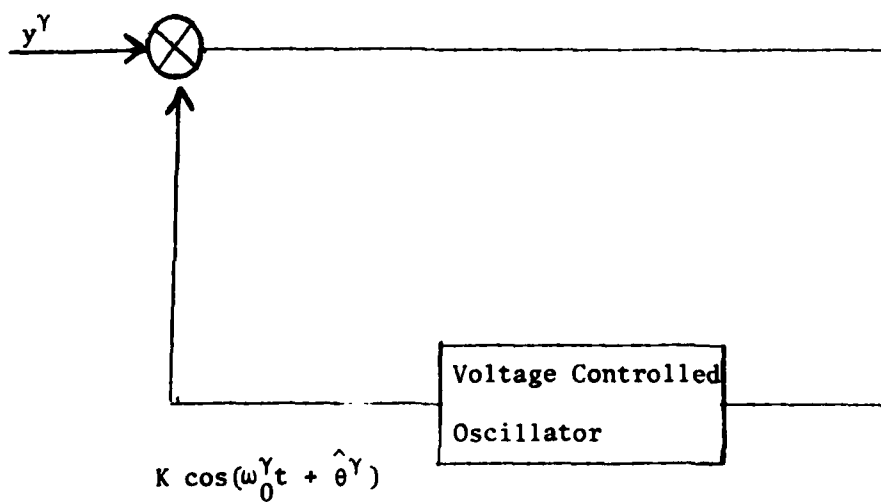
Suppose that $n^\gamma(\cdot)$ was obtained from an impulsive or shot noise process which was suitably filtered in order to guarantee that the actual noise entering the system have bandwidth $O(1/q_\gamma)$ and center frequency $O(1/\gamma)$. Rough calculations similar to those in Section 5.4 suggest that the limit would take the form (6.2) under reasonable conditions. Such a result would be quite useful in applications; in many cases, such shot noise based processes are closer to the true physical noise than is the Gaussian noise. It would also be interesting to work with $n^\gamma(t)/q_\gamma^\delta$, for some $\delta < 1/2$ and use Freidlin's idea of moderate deviations [5].

The PLL systems considered above are an important class of applications to which large deviations or singular perturbation methods have been applied [12], [16], although it is now common practice to ignore the limit and approximation questions, and even the (usual) 'pass-band' nature of the PLL in order to write down a 'small' noise Itô equation model directly, and allow the analysis to start from there.

Let $\theta(t) = \theta_0$. The mean equation is $\dot{x} = \bar{b}(x)$, and for the usual filters, $(0, \theta_0) = \bar{x}_0$ is a locally asymptotically stable point of this equation. For the simple 'first order' PLL of Fig. 1, there is no filter and the limit equation is $\dot{\hat{\theta}} = K\bar{A} \sin(\theta - \hat{\theta})/2$, where $K > 0$ is a scalar. An important communications theory problem is to estimate the minimum time for $(v(t), \hat{\theta}(t))$ or $\hat{\theta}(t)$ to leave the stability set of the limit equation. Owing to the difficulty of the problem, and to the fact that the noise is often 'rapidly fluctuating' and with 'small' effects, 'small noise' methods are appealing. Above, we have given an outline of the role of the theory of large deviations. But for the actual physical non-white noise model, a number of questions concerning modelling and approximation of the noise, and proof of the escapetime formula (1.8) still remain open.



(a) Higher order phase locked loop



(b) First order phase locked loop

Fig. 1. Phase Locked Loops

Appendix

The proof of (1.6) in [5] is not quite valid for (1.2), (1.4), since $\{\rho_n\}$ is unbounded. The proofs in [3] do not account for the ξ_n or $\xi(\cdot)$ terms. If $\sigma(x) = \sigma$, the proofs given or referenced in [5] remains valid, with a few modifications. Here, we remark on the required changes, without proofs. For concreteness, the discrete parameter case only will be dealt with.

The set $\{\phi: \tilde{S}_{[0,T]}(\phi) \leq a, \phi(0) = x\}$ on top of p.136 [5] is still compact in the unbounded $U(x)$ case, by (A4.5). Lemma 3.1 of [5] requires a few modifications, since $\{\rho_n\}$ is unbounded. The inequality below (3.2), p. 138 [5] is no longer true, but it does hold modulo the probability of a set $A_{\epsilon, \Delta}$ where $P\{A_{\epsilon, \Delta}\} \leq \exp -N/\epsilon$, where N can be made as large as we wish by choosing ϵ, Δ small enough. Similarly, the set inclusion below [5, p.138] holds modulo a set of probability $\leq \exp -N/\epsilon$, where $N \rightarrow \infty$ as $\Delta \rightarrow 0, \epsilon \rightarrow 0$. With these changes Lemma 3.1 of [5] holds.

Proof of Theorem 2.1 [5]. If $\sigma(x) = \text{constant}$, the last set inclusion on [5, p. 141] holds by the uniform Lipschitz condition on $b(\cdot, \xi)$. Concerning the argument on p. 142 [5], the trajectories $\tilde{x}^{\epsilon, \psi}(\cdot), x^{\epsilon}(\cdot)$, do not belong to a compact set. But for any large N , there is a set of probability $\geq 1 - \exp -N/\epsilon$ such that on this set the trajectories do belong to a compact set of continuous functions- they satisfy a common holder condition.

REFERENCES

- [1] R. Azencott, "Grandes deviations et applications," Lecture Notes in Math., No. 774, 1980, Springer-Verlag, New York.
- [2] R. Azencott, G. Ruget, "Melanges d'equations differentielles et grands écartes à la loi des grand nombres," Z. Wahr. 38, 1977, pp. 1-54.
- [3] A. D. Ventsel, M. I. Freidlin, "Some problems concerning stability under small random perturbations," Theory of Prob. and Applic., 17, 1972, pp. 269-283.
- [4] A. D. Ventsel, "Rough limit theorems on large deviations for Markov processes," I, II, Theory of Prob. and Applic., 21, 1976, pp. 227-242.
- [5] M. I. Freidlin, "The averaging principle and theorems on large deviations," Russian Math. Surveys, 33, July-Dec., 1978, pp. 117-176.
- [6] M. Donsker, S. R. S. Varadhan, "Asymptotic evaluation of certain Markov process expectations for large time," I, II, III, Comm. Pure Appl. Math 28, 1975, pp. 1-47, 229-301, 389-461.
- [7] S. Karlin, "Positive operators," J. Math and Mech., 8, 1959, pp. 907-937.
- [8] T. Kato, Perturbation Theory for Linear Operators, 2nd ed., Springer, Berlin, 1976.
- [9] N. Dunford, J. T. Schwartz, Linear Operations, Part I, Wiley, New York, 1966.
- [10] A. J. Viterbi, Principles of Coherent Communication, McGraw-Hill, New York, 1966.
- [11] F. L. Gardner, Phaselock Techniques, 2nd ed., 1979, Wiley, New York.
- [12] Z. Schuss, Theory and Applications of Stochastic Differential Equations, 1980, Wiley, New York.
- [13] D. Ludwig, "Persistence of dynamical systems under random perturbations," SIAM Rev. 17, 1975, pp. 605-640.
- [14] G. Ruget, "Some tools for the study of channel-sharing algorithms," CISM publication, ed. Longo, Springer, 1980.
- [15] H. J. Kushner, "A cautionary note on the use of singular perturbation methods for 'small noise' models," Stochastics, 6, 1982, pp. 117-120.
- [16] B. Z. Bobrovsky, Z. Schuss, "A singular perturbation method for the computation of the mean first passage time in a non-linear filter," SIAM J. on Appl. Math., 42, 1982, p. 174-187.

D

E

ED

B2

C