The Annals of Probability
1987, Vol. 15, No. 2, 646–658

A LARGE DEVIATIONS PRINCIPLE FOR SMALL PERTURBATIONS OF RANDOM EVOLUTION EQUATIONS

BY CAROL BEZUIDENHOUT

University of Minnesota

We prove a result for small random perturbations of random evolution equations analogous to the Ventsel–Freidlin result on small perturbations of dynamical systems. In particular, we derive large deviations estimates and indicate how they can be used to prove an exit result. The processes we study are governed by equations of the form

$$dx^\varepsilon(t) = b(x^\varepsilon(t), y(t))\,dt + \sqrt{\varepsilon}\,\sigma(x^\varepsilon(t))\,dw(t),$$

where x^0 is already a random process. The results include the case where y is an n-state Markov process.

In the special case $\sigma \equiv Id$, the proof of the estimates is a consequence of a generalization of the "contraction principle" for large deviations: We give sufficient conditions on a continuous function F, which ensure that if $\{X_\varepsilon: \varepsilon > 0\}$ satisfies a large deviations principle, then so does $\{F(X_\varepsilon, Y): \varepsilon > 0\}$, where Y is independent of $\{X_\varepsilon: \varepsilon > 0\}$.

Introduction. In their study on small random perturbations of dynamical systems, Ventsel and Freidlin [18] start with a system $dx^0(t) = b(x^0(t))\,dt$ and perturb it by adding a small noise term, leading to the perturbed equation $dx^\varepsilon(t) = b(x^\varepsilon(t))\,dt + \sqrt{\varepsilon}\,\sigma(x^\varepsilon(t))\,dw(t)$. In the special case where $\sigma \equiv Id$, the large deviations estimates, which are the essential part of their theory, follow easily from the corresponding estimates for $\sqrt{\varepsilon}\,w(\cdot)$ (Schilder's theorem [14]) and the "contraction principle" (stated as Lemma 2.3 below), because, in this special case, $x^\varepsilon(\cdot)$ is a continuous function of $\sqrt{\varepsilon}\,w(\cdot)$ (see [16, 6]). Various authors, including Azencott [1], have shown that, in the general case, $x^\varepsilon(\cdot)$ can be regarded as an "almost continuous" function of $\sqrt{\varepsilon}\,w(\cdot)$, and that the large deviations result still follows from Schilder's theorem.

This paper has two main aims. The first is to derive results, analogous to Ventsel and Freidlin's, for processes governed by equations of the form

$$(*) \qquad dx^\varepsilon(t) = b(x^\varepsilon(t), y(t))\,dt + \sqrt{\varepsilon}\,\sigma(x^\varepsilon(t))\,dw(t),$$

where $y(t)$ is a random process, so that $x^0(t)$ is already random. Such processes, known as random evolutions, and generalizations thereof, have been studied by Griego and Hersh [8], Heath [10], Hersh and Papanicolaou [12] and others.

We show that if $y(t)$ is a process which is independent of the Brownian motion $w(t)$, then, assuming that b and σ satisfy certain regularity assumptions, the solution $x^\varepsilon(\cdot)$ of $(*)$ satisfies a large deviations principle. We give an explicit expression for the rate function. In the case where $y(t)$ takes values in a compact

Received May 1985; revised February 1986.
AMS 1980 *subject classifications.* Primary 60F10; secondary 34F05, 60H10, 60H25, 60J60.
Key words and phrases. Large deviations, random evolutions, small random perturbations, Ventsel–Freidlin theory, contraction principle.

set, this rate function has an especially simple form. A precise statement is given in Theorem 1.6, and proved in Section 5. From the large deviations estimates it is possible to use Ventsel and Freidlin's techniques to derive other results, for example about exit from a bounded domain. We state such a result without proof, as Proposition 1.10. This exit result has an interpretation as a statement about systems of partial differential equations. See Proposition 1.13.

The second major purpose of this work is motivated by the first. In the special case where $\sigma \equiv Id$, the solution $x^{\varepsilon}(\cdot)$ of $(*)$ is, for fixed ω, a continuous function (with appropriate topologies) of $\sqrt{\varepsilon}\, w(\cdot)$ and $y(\cdot)$. From this, we abstract the problem formulated in Section 2. The study of this problem is our second main theme. Although the author's interest in studying this question was motivated by its application to $(*)$, it is hoped that the results of Section 2 will be of independent interest. The most important result in this general formulation is Theorem 2.15. This is the result which is specialized to yield the large deviations result for $(*)$ in the special case $\sigma \equiv Id$. This is done in Section 5. For general σ, the result can be proved by a modification of Azencott's technique.

Although the author was not aware of this while studying these problems, similar results had been obtained earlier by Friedlin and Gartner [5]. These authors studied the special case $\sigma \equiv Id$, and assumed that $y(\cdot)$ took values in the space of continuous functions. They used essentially the same arguments as in the original work of Ventsel and Friedlin.

1. A large deviations principle for random evolutions.

Definition of the process. Let $(\Omega, \mathbf{F}, P)$ be a probability space, and $\{\mathbf{F}_t: t > 0\}$ an increasing family of sub-σ-algebras of $\mathbf{F}$. Let $X(t)$ be an $\mathbf{F}_t$-adapted $\mathbf{R}^d$-valued Brownian motion. Let $Y(t)$ be an $\mathbf{F}_t$-adapted process which is independent of $X(\cdot)$ and takes values in $\mathbf{R}^m$. Suppose that b: $\mathbf{R}^d \times \mathbf{R}^m \to \mathbf{R}^d$ has the property that $b(z, y)$ is jointly measurable in (z, y) and suppose that there exists a constant C so that

$$\text{(1.1)}\quad \begin{aligned} &\text{(a)} \quad |b(z, y)| \le C, && \forall\, y \in K, \\ &\text{(b)} \quad |b(z_1, y) - b(z_2, y)| \le C|z_1 - z_2|, && \forall\, y \in \mathbf{R}^m,\ z_1, z_2 \in \mathbf{R}^d, \\ &\text{(c)} \quad |b(z, y_1) - b(z, y_2)| \le C|y_1 - y_2|, && \forall\, z \in \mathbf{R}^d,\ y_1, y_2 \in \mathbf{R}^m. \end{aligned}$$

Fix $x \in \mathbf{R}^d$. Define an $\mathbf{F}_t$-adapted $\mathbf{R}^d$-valued process $Z(t)$ by

$$Z(t) = x + \int_0^t b(Z(s), Y(s))\, ds. \tag{1.2}$$

For fixed $\omega \in \Omega$, (1.2) has a unique solution. This follows, for example, from a theorem of Carathéodory. See Hale [7, Theorem 5.1, page 28].

Suppose that σ: $\mathbf{R}^d \to \mathbf{M}_{d \times d}$, the space of $d \times d$ matrices, has $C_1^b(\mathbf{R}^d)$ entries. For $\varepsilon > 0$, define $Z_\varepsilon(t)$ by

$$Z_\varepsilon(t) = x + \int_0^t b(Z_\varepsilon(s), Y(s))\, ds + \sqrt{\varepsilon} \int_0^t \sigma(Z_\varepsilon(t))\, dX(s). \tag{1.3}$$

The existence of a unique solution of (1.3) which is $\mathbf{F}_t$-adapted and has continuous sample paths is ensured by our assumptions on b and σ, and standard results on existence and uniqueness of solutions of stochastic differential equations with random coefficients. See for example Gikhman and Skorohod [5, Section 5, Theorem 1].

For $g \in C([0,T], \mathbf{R}^d, x)$, define

$$(1.4)\quad I(g) = \tfrac{1}{2}\inf\left\{\int_0^T |\sigma^{-1}(g(t))[\dot{g}(t) - b(g(t), h(t))]|^2\,dt\colon h \in \operatorname{Supp} Y\right\}$$

if g is absolutely continuous and $I(g) = +\infty$ otherwise. Let I^* be the lower semicontinuous regularization of I. See (2.8).

Our aim is to establish that, as a family of random variables with values in $C([0,T], \mathbf{R}^d, x)$, $\{Z_\varepsilon\colon \varepsilon > 0\}$ obeys a large deviations principle with rate I^*.

EXAMPLE 1.5. This includes the case where Y is an n-state Markov process.

Statement of results.

THEOREM 1.6. *Let I be defined by* (1.4) *and let I^* be the lower semicontinuous regularization of I. Then I^* is a rate function, and if Z_ε is as in* (1.3), *then $\{Z_\varepsilon\colon \varepsilon > 0\}$ obeys a large deviations principle with rate I^*. Further, if* $\operatorname{Supp} Y = L^1([0,T], K)$, *where $K \subset \mathbf{R}^m$ is compact, then*

$$(1.7)\qquad I(g) = \tfrac{1}{2}\int_0^T \inf_{y \in K} |\sigma^{-1}(g(t))[\dot{g}(t) - b(g(t), y)]|^2\,dt.$$

The proof of Theorem 1.6 is given in Section 5.

For the following two results, which are given without proof, their proofs being standard, we specialize to the case where Y is as in Example 1.5 and $\sigma \equiv Id$. The first is an exit result which can be proved from Theorem 1.6 using a modification of the Ventsel–Freidlin technique. The details for $n = 2$ were written out in [2]. The second is the interpretation of the first in pde terms.

Suppose that Y is a time-homogeneous n-state Markov process with states $i = 1, 2, \dots, n$ and suppose that there is a positive probability of going from each of the states to each of the others. Assume $\sigma \equiv Id$. Suppose $b\colon \mathbf{R}^d \times \{1, 2, \dots, n\} \to \mathbf{R}^d$ is bounded and Lipschitz continuous in both variables. Suppose further that for $i = 1, 2, \dots, n$,

$$(1.8)\qquad b(0,i) = 0, \qquad x \cdot b(x,i) \le -\delta < 0, \qquad \forall\, x \in \mathbf{R}^d.$$

This hypothesis ensures that energy increases linearly with time for trajectories staying inside a compact set not containing 0. For $T > 0$ and $g \in C([0,T], \mathbf{R}^d, x)$, define $I_T(g)$ by (1.7) with $K = \{1, 2, \dots, n\}$. For $x \in \mathbf{R}^d$, define

$$(1.9)\qquad V(x) = \inf\{I_T(g)\colon g \in C([0,T], \mathbf{R}^d, 0), g(T) = x\}.$$

PROPOSITION 1.10 (An exit result). *Suppose $D \subset \mathbf{R}^d$ is an open bounded set with smooth boundary ∂D, such that $0 \in D$. Let $K \subset \partial D$ be the set of y in ∂D*

for which

$$V(y) = \inf\{V(w) \colon w \in \partial D\}. \tag{1.11}$$

For $\varepsilon > 0$, *define* Z_ε *by* (1.3) *and let* τ_D^ε *be the time of first exit of* Z_ε *from* D. *Then if* $N \subset \partial D$ *is a neighbourhood of* K, *for every* $x \in D$,

$$P_x(\tau_D^\varepsilon \in N) \to 1 \quad \text{as } \varepsilon \to 0. \tag{1.12}$$

In particular, if there is a unique point y_0 *on* ∂D *for which* (1.11) *holds, then* $Z_\varepsilon(\tau_D^\varepsilon) \to y_0$ *in probability as* $\varepsilon \to 0$. *In this case, more can be said. If also there is a unique path* φ *joining* 0 *to* ∂D *which realizes the infimum in* (1.11) (*this path* φ *will in general be defined on a semiinfinite interval* $(-\infty, 0]$) *then with probability approaching* 1 *as* $\varepsilon \to 0$, *the trajectories of the process starting at any point in* D *will exit through a "tube" of fixed radius about* φ.

Proposition 1.10 has an interpretation in term of systems of partial differential equations.

Let Y be as above, $\sigma \equiv Id$, and suppose b satisfies (1.8). Let D be a bounded open subset of $\mathbf{R}^d$ with smooth boundary ∂D and assume that there is a unique point y_0 on ∂D such that (1.11) holds with $y = y_0$.

Let Q be the generator of the Markov process Y. (If $p_{ij}(t) = P^i[Y(t) = j]$ then $q_{ij} = p'_{ij}(0)$ and $Q = (q_{ij})$.) The process

$$\xi_\varepsilon(t) = (Z_\varepsilon(t), Y(t))$$

is a Markov process with state space $\mathbf{R}^d \times \{1, 2, \dots, n\}$ and generator

$$L^\varepsilon u(z, i) = b(z, i) \cdot \nabla_z u(z, i) + \varepsilon \Delta_z u(z, i) + Qu(z, i).$$

Let τ_D^ε be the first time of exit of $Z_\varepsilon(t)$ from D. Let $\tilde{D} = D \times \{1, 2, \dots, n\}$. Then $\tau_D^\varepsilon = \inf\{t \colon \xi_\varepsilon(t) \in \tilde{D}\}$. If g is a function which is continuous on ∂D, and $\tilde{g}$ is defined on $\partial\tilde{D}$ by $g(z, i) = g(z)$, then it follows from general Markov process theory that

$$u(z, y) = E^{z, y}[\tilde{g}(\xi_\varepsilon(\tau_D^\varepsilon))] = E^{z, y}[g(Z_\varepsilon(\tau_D^\varepsilon))]$$

solves the Dirichlet problem

$$\begin{aligned} L^\varepsilon u^\varepsilon &= 0, && \text{in } \tilde{D}, \\ u^\varepsilon(z, y) &= g(z), && \text{on } \partial\tilde{D}. \end{aligned}$$

This can be interpreted as a system of n equations.

The following result follows from (1.12).

Proposition 1.13 (Pde interpretation). *With the above notation,*

$$u^\varepsilon(z, y) \to g(y_0) \quad \text{as } \varepsilon \to 0, \text{ for every } (z, y) \text{ in } \tilde{D}.$$

2. Formulation of a "contraction" problem and statement of results.

Preliminaries. Following Varadhan [17], we define the following:

If E is a complete separable metric space, then a function I defined on E is called a *rate function* if it has the following properties:

$$\text{(2.1)} \quad \begin{array}{ll} \text{(a)} & I \colon E \to [0, \infty],\ I \not\equiv +\infty,\ I \text{ is lower semicontinuous.} \\ \text{(b)} & \text{If } 0 \le a < \infty, \text{ then } C_I(a) \equiv \{x \in E \colon I(x) \le a\} \text{ is compact.} \end{array}$$

If E is a complete separable metric space, B is the Borel σ-field on E, $\{\mu_\varepsilon: \varepsilon > 0\}$ is a family of probability measures on $(E, \mathbf{B})$, and I is a function defined on E and satisfying (2.1), then we say that $\{\mu_\varepsilon\}$ satisfies a *large deviations principle* with rate I if:

(2.2)
(a) For every open subset A of E,

$$\liminf \varepsilon \log \mu_\varepsilon(A) \geq -I(A).$$

(b) For every closed subset A of E,

$$\limsup \varepsilon \log \mu_\varepsilon(A) \leq -I(A).$$

Here, and below, if I is a function defined on a set E and A is a subset of E, then $I(A)$ is defined to be the infimum of I on A. Unless otherwise stated, all lim infs and lim sups are as $\varepsilon \to 0$.

A family $\{X_\varepsilon: \varepsilon > 0\}$ of random variables with values in E is said to obey a large deviations principle with rate I if the corresponding distributions satisfy (2.2).

For future reference, we quote the following well-known result, sometimes known as the "contraction principle." See [17, page 5].

LEMMA 2.3. *Suppose E and E' are Polish spaces, I is a rate function defined on E, $\{X_\varepsilon: \varepsilon > 0\}$ is a family of E-valued random variables satisfying a large deviations principle with rate I and f: $E \to E'$ is continuous. For $\varepsilon > 0$, define $Y_\varepsilon = f(X_\varepsilon)$. Then the family $\{Y_\varepsilon: \varepsilon > 0\}$ of E'-valued random variables satisfies a large deviations principle with rate I' defined for $y \in E'$ by*

$$I'(y) = \inf\{I(x): f(x) = y\}.$$

Here and below we shall use the convention

$$\inf \varnothing = +\infty. \tag{2.4}$$

Formulation of the problem. To formulate the problem to be studied here, let E_X, E_Y and E_Z denote complete separable metric spaces. Suppose $\{X_\varepsilon: \varepsilon > 0\}$ is a family of random variables with values in E_X satisfying a large deviations principle with rate function I_X, and Y a random variable with values in E_Y. We shall always assume

$$X_\varepsilon \text{ is independent of } Y \text{ for } \varepsilon > 0. \tag{2.5}$$

F will denote a function F: $E_X \times E_Y \to E_Z$ which is continuous if $E_X \times E_Y$ is given the product topology, and for $\varepsilon > 0$, we shall define

$$Z_\varepsilon = F(X_\varepsilon, Y). \tag{2.6}$$

We investigate the problem of whether the family $\{Z_\varepsilon: \varepsilon > 0\}$ defined by (2.6) satisfies a large deviations principle.

Heuristic considerations lead one to conjecture that a possible rate function for the family $\{Z_\varepsilon: \varepsilon > 0\}$ is

$$I_F(Z) = \inf\{I_X(x): \exists\, y \in \operatorname{Supp} Y \ni F(x, y) = z\}. \tag{2.7}$$

I_F is not always lower semicontinuous (see second remark after (5.2) below); in case it is not, we introduce the lower semicontinuous regularization I_F^* of I_F, defined by

$$I_F^*(Z) = \lim_{r\to 0} I_F(B_Z(z,r)), \tag{2.8}$$

where $B_Z(z, r)$ is the ball of radius r in E_Z centred at z. Since I_F is a decreasing function on sets, the limit in (2.8) exists.

Statement of results. We answer the question of whether the family $\{Z_\varepsilon: \varepsilon > 0\}$ defined by (2.6) satisfies a large deviations principle negatively by giving a counterexample. This is the content of Section 3. However, under certain conditions on the function F, a large deviations principle does apply. The principal results, whose proofs appear in Section 4, are listed below. For the application to the motivating problem of obtaining a large deviations theorem for random evolutions, the most important result is Theorem 2.15.

Let E_X, E_Y, E_Z, $\{X_\varepsilon: \varepsilon > 0\}$ and Y be as above. Suppose (2.5) holds and that $F: E_X \times E_Y \to E_Z$ is continuous. For $\varepsilon > 0$, define Z_ε by (2.6).

LEMMA 2.9 (Lower bound). *If $A \subset E_Z$ is open,*

$$\liminf \varepsilon \log P(Z_\varepsilon \in A) \geq -I_F(A). \tag{2.10}$$

LEMMA 2.11 (Upper bound). *If $A \subset E_Z$ is closed, then*

$$\limsup \varepsilon \log P(Z_\varepsilon \in A) \leq -I_X\big(\overline{\Pi_X(A)}\big), \tag{2.12}$$

where

$$\Pi_X(A) = \{x \in E_X: \exists\, y \in \operatorname{Supp} Y \ni F(x,y) \in A\} \tag{2.13}$$

PROPOSITION 2.14. *If Y has compact support, then $\{Z_\varepsilon: \varepsilon > 0\}$ satisfies a large deviations principle with rate function I_F.*

THEOREM 2.15. *Let I_F^* be as in* (2.8). *Suppose I_F^* is a rate function, i.e., satisfies* (2.1). *Suppose F has the following two properties*:

(2.16)
(a) $\{F(\cdot, y)\}_{y\in \operatorname{Supp} Y}$ *is an equicontinuous family of functions.*
(b) *If $K \subset E_x$ is precompact, then $F(K \times \operatorname{Supp} Y)$ is precompact.*

Then $\{Z_\varepsilon: \varepsilon > 0\}$ obeys a large deviations principle with rate I_F^*.

It is natural to ask what happens if Y is also allowed to depend on ε. In answer to this, we have the following result.

PROPOSITION 2.17. *Let E_X, E_Y, E_Z, $F: E_X \times E_Y \to E_Z$ and $\{X_\varepsilon: \varepsilon > 0\}$ be as before and let $\{Y_\varepsilon: \varepsilon > 0\}$ be a family of E_Y-valued random variables satisfying a large deviations principle with rate function I_Y. Suppose for each*

$\varepsilon > 0$, X_ε is independent of Y_ε. For $\varepsilon > 0$, let $Z_\varepsilon = F(X_\varepsilon, Y_\varepsilon)$. Define

$$I_F(z) = \inf\{I_X(x) + I_Y(y): F(x, y) = z\} \tag{2.18}$$

for $z \in E_Z$. Then $\{Z_\varepsilon: \varepsilon > 0\}$ satisfies a large deviations principle with rate function I_F.

3. A counterexample. In this section, we give an example in which a large deviations principle does not apply. In fact we shall construct closed subsets A of E_Z for which $P(Z_\varepsilon \in A)$ tends to zero arbitrarily slowly, although $I_F^*(A) > 0$.

In this example, $E_X = C([0,1], \mathbf{R}, 0)$ is the space of continuous real-valued functions defined on $[0,1]$ and taking the value 0 at time $t = 0$, with the topology induced by the uniform norm

$$\| f \| = \sup\{ | f(t) |: 0 \le t \le 1\}.$$

$X_\varepsilon = \sqrt{\varepsilon} X$, where X is a one-dimensional Brownian motion on $[0,1]$. It is known (Schilder [14]) that $\{X_\varepsilon\}$ satisfies a large deviations principle with rate function given by

$$I_X(f) = \int_0^T | \dot{f}(t) |^2 \, dt \tag{3.1}$$

if f is absolutely continuous, and $I_X(f) = +\infty$ otherwise. Here $T = 1$, but we shall need formula (3.1) later for arbitrary $T > 0$ and $f \in C([0, T], \mathbf{R}^d, 0)$ for $d \ge 1$.

E_Y is the set $\{1, 2, 3, \dots\}$ with the discrete topology. Y is a fixed random variable taking values in E_Y and independent of X. Define

$$p_n = P\{Y = n\} \tag{3.2}$$

and assume that $p_n > 0$ infinitely often. (Otherwise Y has compact support and, as stated earlier [(2.14)] a large deviations result does hold.) E_Z is the space $E_X \times E_Y$ with the produce topology and F is the identity map.

We shall need the following fact. See Orey [13, Theorem 2].

LEMMA 3.3. *There exists a constant m, $0 < m < \infty$, such that*

$$\rho^2 \inf\{I_X(f): \|f - X\| < \rho\} \to m$$

almost surely as $\rho \to 0$.

For $0 \le a < \infty$, let $C(a) = C_{I_X}(a)$ where the latter set is as in (2.1)(b). We claim:

LEMMA 3.4. *Let $\{p_n\}$ be a sequence of nonnegative numbers with $p_n > 0$ infinitely often and $\sum p_n = 1$. Let H: $[0, \infty) \to [0, \infty)$ be a continuous nondecreasing function with $H(0) = 0$. Then there exists a nonincreasing sequence of positive numbers $\{q_n\}$ so that $q_n \downarrow 0$ as $n \to \infty$, and a number $\varepsilon_0 > 0$ so that for $\varepsilon \le \varepsilon_0$,*

$$\sum p_n P[d(X_\varepsilon, C(1)) \ge q_n] \ge H(\varepsilon). \tag{3.5}$$

PROOF. It follows from Lemma 3.3 that there exists $\rho_0 > 0$ so that if $\rho < \rho_0$, then

$$P\left[\rho^2 \inf\{I_X(f): \|X - f\| < \rho\} > m/2\right] > \tfrac{1}{2}. \tag{3.6}$$

If $0 < \varepsilon < \varepsilon_0 \equiv 2\rho_0^2/m$ and $0 < \delta \le c\varepsilon$, where $c = \sqrt{(m/2)}$, then $\delta/\sqrt{\varepsilon} < \rho_0$ and $\delta^2/\varepsilon^2 \le m/2$. Therefore, from (3.6)

$$\begin{aligned} P\left[d(X_\varepsilon, C(1)) \ge \delta\right] &= P\left[d(X, C(1/\varepsilon)) \ge \delta/\sqrt{\varepsilon}\right] \\ &= P\left[\delta^2/\varepsilon \inf\{I_X(f): \|X - f\| < \delta/\sqrt{\varepsilon}\} \ge \delta^2/\varepsilon^2\right] \\ &> \tfrac{1}{2}. \end{aligned} \tag{3.7}$$

Define

$$P_n = \sum_{k \ge n} p_k,$$

$$q_n = c \inf\{\eta: P_{n+1} \le 2H(\eta) \le P_n\},$$

where $c = \sqrt{(m/2)}$. Since $\{P_n\}$ is nonincreasing and H is nondecreasing, $\{q_n\}$ is nonincreasing. Assume that $H(\varepsilon) > 0$ for $\varepsilon > 0$. [Otherwise (3.5) holds trivially.] Then $q_n > 0$ for every n. If $0 < \varepsilon \le \varepsilon_0$, there exists n so that $P_{n+1} < 2H(\eta) \le P_n$. For $k \ge n$, we have $q_k \le c\varepsilon$. From this and (3.7), it follows that the left-hand side of (3.5) is greater than

$$\frac{1}{2} \sum_{\{k:\, q_k \le c\varepsilon\}} p_n \ge \tfrac{1}{2} P_n \ge H(\varepsilon). \qquad \square$$

EXAMPLE 3.8. Suppose the setup is as described at the beginning of this section. Let H be a prescribed function which has the properties described in Lemma 3.4. Define subsets V_n of E_x by

$$V_n = \{f \in E_X: d(f, C(1)) \ge q_n\}$$

and let $A \subset E_Z$ be defined by

$$A = \bigcup_{n=1}^{\infty} V_n \times \{n\}.$$

Then A is closed, $\lim_{\varepsilon \to 0} P(Z_\varepsilon \in A) = 0$ and by (3.5), $P(Z_\varepsilon \in A) \ge H(\varepsilon)$. However, $I_F^*(A) = I_F(A) = 1 > 0$.

4. Proofs of the large deviations bounds. In this section, we give the proofs of the large deviations results for the family $\{Z_\varepsilon: \varepsilon > 0\}$ of (2.6).

PROOF OF LEMMA 2.9 (Lower bound). To establish (2.10) for open $A \subset E_Z$, it is enough to show that if M and N are open subsets of E_X and E_Y, respectively, $x \in M$, $y \in N$, $I_X(x) < \infty$, $y \in \operatorname{Supp} Y$ and $F(M \times N) \subset A$, then $\liminf \varepsilon \log P((X_\varepsilon, Y) \in M \times N) \ge -I_X(x)$. But this follows immediately from our assumptions, using the independence of X_ε and Y. $\square$

PROOF OF LEMMA 2.11 (Upper bound). Obvious. □

PROOF OF PROPOSITION 2.14 (Large deviations principle when Y has compact support). The lower bound follows from Lemma 2.9. The upper bound follows from Lemma 2.11 if we observe first that it is always true that $I_F(A) = I_X(\Pi_X(A))$, and then that if Y has compact support, $\Pi_X(A)$ is closed whenever A is closed. The fact that I_F is a rate function in this case follows easily from the compactness of Supp Y and the fact that I_X is a rate function. □

PROOF OF THEOREM 2.15 (Large deviations principle when F satisfies (2.16)). It is easy to see that if A is an open subset of E, then $I_F(A) = I_F^*(A)$. So the lower bound follows from (2.9). For the upper bound, it suffices, by (2.11), to show that under the above hypotheses on F,

$$(4.1) \qquad I_F^*(A) \le I_X\big(\overline{\Pi_X(A)}\big) \quad \text{for every closed } A \subset E_Z,$$

where $\Pi_X(A)$ is as in (2.13).

We now show that if (2.16) holds, then so does (4.1). If x is in the closure of $\Pi_X(A)$, then there exist $x_n \in \Pi_X(A)$ and $y_n \in \operatorname{Supp} Y$ so that $z_n = F(x_n, y_n) \in A$ and $x_n \to x$ in E_X. In particular, $\{x_n\}$ is precompact in E_X. Therefore, by (2.16)(b), $\{z_n\}$ is precompact in E_Z, and so we may assume that $\{z_n\}$ converges in E_Z, say to z. Since $z_n \in A$, and A is closed, $z \in A$. Define $\tilde{z}_n = F(x, y_n)$. By (2.16)(a), $d_Z(z_n, \tilde{z}_n) \to 0$ as $n \to \infty$, so $\tilde{z}_n \to z$ also; so $I_F^*(A) \le I_F^*(z) \le \liminf_{n\to\infty} I_F(\tilde{z}_n) \le I_X(x)$. □

NOTE 4.2. A sufficient condition for I_F^* to be a rate function is that $\{z: I_F(z) \le a\}$ be precompact for $0 \le a < \infty$.

PROOF OF PROPOSITION 2.17 (Large deviations principle when $Z_\varepsilon = F(X_\varepsilon, Y_\varepsilon)$). By the "contraction principle" (2.3), it is enough to consider the case $E_Z = E_X \times E_Y$, and $F = id$ [i.e., $F(z) = z$].

We prove first that I_{id} is a rate function. I_{id}: $E_Z \to [0, \infty]$ is defined by

$$(4.3) \qquad I_{id}(x, y) = I_X(x) + I_Y(y),$$

which is clearly lower semicontinuous if I_X and I_Y are. Hence the set $\{z: I_{id}(z) \le a\}$ is closed for $0 \le a < \infty$. Since it is contained in the compact set $\{x: I_X(z) \le a\} \times \{y: I_Y(z) \le a\}$, the result follows.

NOTE. Since $\{z: I_{id}(z) \le a\} \supset \{x: I_X(z) \le a/2\} \times \{y: I_Y(z) \le a/2\}$, I_{id} is a rate function if and only if both I_X and I_Y are.

The lower bound follows easily from the assumption that for each $\varepsilon > 0$, X_ε is independent of Y_ε and the lower bounds for X_ε and Y_ε. The proof is very similar to that of (2.9).

For the upper bound note first that it is obvious that if $A \subset E_Z$ is closed, then

$$(4.4) \qquad \limsup \varepsilon \log P(Z_\varepsilon \in A) \le -\sup \min_{i=1,2,\dots,n} \{I_X(M_i) + I_Y(N_i)\},$$

where the supremum is taken over all finite collections of pairs of closed subsets M_i of E_X and N_i of E_Y such that

$$A \subset \bigcup_{i=1}^{n} M_i \times N_i.$$

Suppose $a < I_{id}(A)$. Let $C_X = C_X(a)$, $C_Y = C_Y(a)$ and $C_Z = C_Z(a)$ be the sets $\{I \le a\}$ for $I = I_X$, I_Y and I_{id}, respectively. By hypothesis and since I_F is a rate function, these sets are all compact. Let d_X and d_Y denote the metrics on E_X and E_Y, respectively. Define d_Z on E_Z by

$$d_Z((x_1, y_1), (x_2, y_2)) = d_X(x_1, x_2) + d_Y(y_1, y_2).$$

Since $a < I_{id}(A)$, A and C_Z are disjoint and so, since the latter is compact, $d_Z(A, C_Z) = \delta > 0$. By compactness, $C_X \subset \bigcup B_X(x_i, \delta/16)$ and $C_Y \subset \bigcup B_Y(y_j, \delta/16)$, where i runs from 1 to m and j from 1 to n for some finite m and n.

Suppose that $d_Z((x, y), C_Z) > \delta$, $d_X(x, C_X) \le \delta/16$ and $d_Y(y, C_Y) \le \delta/16$. Then $x \in B_X(x_i, \delta/8)$ and $y \in B_Y(y_j, \delta/8)$ for some i and j, and if $(x', y') \in B_X(x_i, \delta/8) \times B_Y(y_j, \delta/8)$ then $d_Z((x', y'), C_Z) \ge \delta/2$, and hence

$$I_X(B_X(x_i, \delta/8)) + I_Y\big(B_Y(y_j, \delta/8)\big) \ge a. \tag{4.5}$$

Therefore,

$$\begin{aligned} A \subset {} & \{x\colon d_X(x, C_X) \ge \delta/16\} \times E_Y \\ & \cup E_X \times \{y\colon d_Y(y, C_Y) \ge \delta/16\} \\ & \cup \bigcup_{(i,j) \in I} B_X(x_i, \delta/8) \times B_Y(y_j, \delta/8), \end{aligned} \tag{4.6}$$

where I is a finite set of pairs of indices (i, j) such that if $(i, j) \in I$, then (4.5) holds. From (4.6) and (4.4), it follows that the left-hand side of (4.4) is no greater than $-a$. Letting $a \uparrow I_{id}(A)$, we get the result. □

5. Application of Section 2 to the proof of Theorem 1.6. We present in this section the proof of the large deviations theorem for random evolutions stated as Theorem 1.6 We give the details only in the case where $\sigma \equiv Id$, the $d \times d$ identity matrix. The result for general $\sigma \in C_1^b(\mathbf{R}^d)$ can be proved in much the same way as Azencott [1] proves the general case of the Ventsel–Freidlin result. In the special case $\sigma \equiv Id$, the large deviations estimates required for Theorem 1.6 follow directly from Theorem 2.15.

PROOF OF THEOREM 1.6 WHEN $\sigma \equiv Id$. To express this in the notation of Section 2, let E_X be the space $C([0, T], \mathbf{R}^d, 0)$, $E_Y = L^1([0, T], \mathbf{R}^m)$ and $E_Z = C([0, T], \mathbf{R}^d, x)$.

REMARK. In Example 1.5 it is more usual to think of Y as an element of the space D of right continuous functions with the Skorohod topology. Note however that the Skorohod topology is stronger than the L^1 topology.

Let $X_\varepsilon = \sqrt{\varepsilon} X$. By Schilder's theorem [14] $\{X_\varepsilon\}$ satisfies a large deviations principle with rate given by (3.1). Define F: $E_X \times E_Y \to E_Z$ as follows. $F(f, h) = g$ if

$$g(t) = x + \int_0^t b(g(s), h(s))\, ds + f(t). \tag{5.1}$$

Clearly the solution Z_ε of (1.3) satisfies $F(X_\varepsilon, Y) = Z_\varepsilon$. It follows from a standard argument using Gronwall's inequality that

Lemma 5.2. *F: $E_X \times E_Y \to E_Z$ defined by* (5.1) *is continuous.*

Remark. Since convergence in D implies convergence in L^1, this also establishes that in Example 1.5 F is continuous as a map from $E_X \times D$ to E_Z.

Using (3.1), we see that the function I_F defined by (2.7) is the same as I in (1.4) (with $\sigma \equiv Id$). Define $I^* = I_F^*$ by (2.8).

Remark. To demonstrate the necessity for introducing I^*, we remark that $I = I_F$ is not necessarily lower semicontinuous. As an example, take Y to be a two-state Markov process with states 0 and 1. Let $T = 1$, $b(y, i) = (-1)^i$ for $y \in \mathbf{R}$, $i = 0, 1$. Let g_n be absolutely continuous with derivative $(-1)^k$ on $(k/2^n, (k+1)/2^n)$, $k = 0, 1, \dots, 2^n - 1$. Let $h_n(t) = 0$ when $\dot{g}_n(t) = 1$ and $h_n(t) = 1$ when $\dot{g}_n(t) = -1$. Then for almost every $t \in [0, 1]$, $\dot{g}_n(t) = b(g_n(t), h_n(t))$ and so $g_n = F(0_X, h_n)$ where 0_X is the function identically zero on $[0, 1]$. So $I_F(g_n) \le I_X(0_X) = 0$. But $g_n \to 0_Z$ in E_Z and $I_F(0_Z) = \frac{1}{2}$.

To show that $\{Z_\varepsilon: \varepsilon > 0\}$ obeys a large deviations principle with rate I^*, we shall use Theorem 2.15. In order to apply Theorem 2.15, we have to show that $I^* = I_F^*$ is a rate function and check the condition (2.16).

Lemma 5.3. *I_F^* is a rate function.*

Proof. As observed in (4.2), it is enough to check that $\{g: I_F(g) \le a\}$ is precompact for $0 \le a < \infty$. From the definition (1.4) of $I(g)$, and the boundedness of b, it follows that

$$\{g: I_F(g) \le a\} \subset \left\{g: \int_0^T |\dot{g}(t)|^2\, dt \le A\right\}$$

for some $A < \infty$, and the latter set is precompact by Arzela-Ascoli. □

Lemma 5.4. *The function F defined in* (5.1) *satisfies the condition* (2.16).

Proof. It follows from the standard argument using Gronwall's inequality that in fact if F is defined as in (5.1), then F is Lipschitz continuous. This clearly implies that $\{F(\cdot, h)\}_{h \in L^1}$ is equicontinuous, whence (2.16)(a) follows.

Since the family $\{\int_0 b(g(t), h(t))\,dt\}_{g\in C, h\in L^1}$, where C is the space E_Z is equicontinuous and uniformly bounded, (2.16)(b) follows from the definition of F. □

REMARK. In fact (although this is not needed) there is equality in (4.1) in this case. To see this, suppose $g \in A$ and $I_F^*(g) < \infty$. Then $I_F^*(g)$ equals $\lim_{n\to\infty} I_F(g_n)$, where $g_n \to g$. There are $f_n \in E_X$, and $h_n \in \operatorname{Supp} Y$ so that $g_n = F(f_n, h_n)$ and $\lim_{n\to\infty} I_X(f_n) = I_F^*(g)$. In particular, $\{I_X(f_n)\}$ is bounded, and so since I_X is a rate function, $\{f_n\}$ is precompact, and we may assume that $f_n \to f$, for some $f \in E_X$. By definition, for $t \in [0, T]$,

$$\begin{aligned} g_n(t) &= x + \int_0^t b(g_n(s), h_n(s))\,ds + f_n(t) \\ &= x + k_n(t) + f_n(t), \end{aligned}$$

where the second equality defines k_n. It follows from the boundedness of b that $\{k_n\}$ is precompact, and so we may assume $k_n \to k$, for some $k \in C([0, T], \mathbf{R}^d, 0)$, uniformly on $[0, T]$. Since $g_n \to g$, we have that $g = x + k + f$. Hence $g = F(l_n, h_n)$, where

$$l_n(t) = \int_0^t [b(g_n(s), h_n(s)) - b(g(s), h_n(s))]\,ds - k_n(t) + k(t) + f(t).$$

Since b is Lipschitz in its first variable, $g_n \to g$ and $k_n \to k$ as $n \to \infty$, it follows that $l_n \to f$ in E_X. Since $g \in A$ and $g = F(l_n, h_n)$, $l_n \in \Pi_X(A)$. So f is in the closure of $\Pi_X(A)$ and hence

$$I_X\big(\overline{\Pi_X(A)}\big) \le I_X(f) \le \liminf_{n\to\infty} I_X(f_n) = I_F^*(g).$$

It remains only to show that:

LEMMA 5.5. *If the support of Y is all of $L^1([0, T], K)$, for some compact $K \subset \mathbf{R}^m$, the rate function has the simpler form given* (1.7).

PROOF. We may assume $I_F(g) < \infty$. Denote the expression on the right side of (1.7) by $I_F'(g)$. Clearly $I_F'(g) \le I_F(g)$. For fixed t, since K is compact and $b(g(t), \cdot)$ is continuous, the set

$$\begin{aligned} A(t) = \Big\{ y \in K \colon &|\sigma(g(t))[\dot g(t) - b(g(t), y)]| \\ &= \inf_{w\in K} |\sigma(g(t))[\dot g(t) - b(g(t), w)]| \Big\} \end{aligned}$$

is compact and nonempty. Standard results (see for example [3]) ensure that we can select $h(t) \in A(t)$ in such a way that $h(t)$, $t \in [0, T]$, is measurable. Since $A(t) \subset K$, $h(\cdot) \in L^1([0, T], K) = \operatorname{Supp} Y$, and so

$$I_F(g) \le \tfrac{1}{2}\int_0^T |\dot g(t) - b(g(t), h(t))|^2\,dt = I_F'(g).$$ □

REMARK. In Example 1.5, a sufficient condition for the hypothesis of Lemma 5.4 to hold is that there should be a positive probability of going from each of the n states to every other state.

REFERENCES

[1] AZENCOTT, R. (1980). Grandes deviations et applications. *Ecole d'Été de Probabilités de Saint-Flour, VIII—1978. Lecture Notes in Math.* **774** 1–176. Springer, New York.

[2] BEZUIDENHOUT, C. (1985). Small random perturbations of stochastic systems. Ph.D. Thesis, Univ. of Minnesota.

[3] CHRISTENSEN, J. (1974). *Topology and Borel Structures.* North-Holland. Amsterdam.

[4] FREIDLIN, M. (1985). *Functional Integration and Partial Differential Equations.* Princeton Univ. Press.

[5] FREIDLIN, M. and GARTNER, J. (1978). A new contribution to the question of large deviations for random processes. *Vestnik Moskov. Univ. Ser. I Mat. Meh.* **5** 52–59.

[6] FREIDLIN, M. I. and WENTZELL, A. D. (1984). *Random Perturbations of Dynamical Systems.* Springer, New York.

[7] GIKHMAN, I. I. and SKOROHOD, A. V. (1972). *Stochastic Differential Equations.* Springer, New York.

[8] GRIEGO, R. J. and HERSH, R. (1971). Random evolutions, Markov chains and systems of partial differential equations. *Trans. Amer. Math. Soc.* **156** 405–418.

[9] HALE, J. K. (1969). *Ordinary Differential Equations.* Wiley, New York.

[10] HEATH, D. (1969). Probabilistic analysis of hyperbolic systems of partial differential equations. Ph.D. Thesis, Univ. of Illinois.

[11] HERSH, R. (1974). Theory of random evolutions with applications to partial differential equations. *Rocky Mountain J. Math.* **4** 443–476.

[12] HERSH, R. and PAPANICOLAOU, G. (1972). Noncommuting random evolutions, and an operator-valued Feynman–Kac formula. *Comm. Pure Appl. Math.* **25** 337–367.

[13] OREY, S. (1983). Two strong laws for shrinking Brownian tubes. *Z. Wahrsch. verw. Gebiete* **63** 281–288.

[14] SCHILDER, M. (1966). Some asymptotic formulas for Wiener integrals. *Trans. Amer. Math. Soc.* **125** 63–85.

[15] STROOCK, D. (1984). *An Introduction to the Theory of Large Deviations.* Springer, New York.

[16] VARADHAN, S. R. S. (1980). *Lectures on Diffusion Problems and Differential Equations.* Tata Institute of Fundamental Research, Bombay.

[17] VARADHAN, S. R. S. (1984). *Large Deviations and Applications.* SIAM, Philadelphia.

[18] VENTSEL, A. D. and FREIDLIN, M. I. (1970). Small random perturbations of dynamical systems. *Russian Math. Surveys* **25** 1–55.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF WISCONSIN
MADISON, WISCONSIN 53706