The Annals of Probability
1991, Vol. 19, No. 4, 1629–1639

ON SOME APPLICABLE VERSIONS OF ABSTRACT LARGE DEVIATIONS THEOREMS

BY DIMITRY IOFFE

Technion, Israel Institute of Technology

Two algorithms for calculating rate functions for a family of measures $\{\mu_\varepsilon\}$ on a B-space X are considered. The first one is a relaxed version of the Fenchel transform type theorem for convex rate functions. The second gives conditions under which $\{\mu_\varepsilon\}$ can be replaced by a more convenient family $\{\mu_\varepsilon^x\}$ near admissible points $x \in X$ such that rate functions for both families coincide near x.

As an example, we apply both techniques to investigate large deviation properties of some reaction–diffusion equations with quick random noise.

1. Large deviation theorem in the convex case. Let X be a Hausdorff topological space and $\{\mu_\varepsilon\}$ a family of subprobability measures on X.

DEFINITION 1. $(X, \{\mu_\varepsilon\})$ is called the system of large deviations (S.L.D) with rate function L if:

CONDITION 1. For each open $J \subset X$,

$$\liminf_{\varepsilon \to 0} \varepsilon \ln \mu_\varepsilon(J) \geq -\inf_{x \in J} L(x).$$

CONDITION 2. For each closed $F \subset X$,

$$\limsup_{\varepsilon \to 0} \varepsilon \ln \mu_\varepsilon(F) \leq -\inf_{x \in F} L(x).$$

CONDITION 3. For each $s \geq 0$, the set $\Phi_s = \{x | L(x) \leq s\}$ is compact.

Note that Condition 3 implies the lower-semicontinuity of L.

REMARK. All the measures in the sequel are subprobability measures; for the sake of brevity we refer to them from now on simply as measures.

DEFINITION 2. A family of measures $\{\mu_\varepsilon\}$ is said to be exponentially tight on X if

(A1) $\forall R > 0, \exists$ a compact K_R such that $\limsup_{\varepsilon \to 0} \varepsilon \ln \mu_\varepsilon(K_R^c) \leq -R$.

We work only with exponentially tight families of measures in this article.

Received October 1989; revised September 1990.
AMS 1980 *subject classifications*. Primary 60F10, 35K57.
Key words and phrases. Large deviations, reaction–diffusion equation.

REMARK. The question of whether a particular family of measures satisfying a large deviation principle will also satisfy (A1) is frequently connected to the regularity of this family and we are not going to discuss it here. However, we point out that in a variety of concrete problems, (A1) can be established without reference to L, as in the example below (see also [4], [5]).

Now let us see how much information about the large deviation properties of $\{\mu_\varepsilon\}$ can be achieved by observing this family only on the compact sets appearing in Definition 2.

DEFINITION 3. $(X, \{\mu_\varepsilon\})$ is said to be a γ-truncated S.L.D. with rate function L if it satisfies Conditions 1 and 3 and also the following condition.

CONDITION 2′. For each closed subset $F \subset X$,

$$\limsup_{\varepsilon \to 0} \varepsilon \ln \mu_\varepsilon(F) \le -\Big(\inf_{x \in F} L(x) \wedge \gamma\Big).$$

LEMMA 1.1. *Assume* (A1) *with* $\{K_R\}$ *being a nondecreasing family of compacts. Let* $\alpha\colon \mathbb{R}_+ \to \mathbb{R}_+$ *be such that*

$$\alpha(R) < R \quad \text{and} \quad \lim_{R \to \infty} (R - \alpha(R)) = 0. \tag{1.2}$$

Then the following two statements are equivalent:

(i) $(X, \{\mu_\varepsilon\})$ *is an S.L.D. with rate function L.*

(ii) $\forall R > 0$, $(K_R, \{\mu_\varepsilon^R\})$ *is an* $\alpha(R)$*-truncated S.L.D. with rate function* L_R.

In what is written above, K_R *is understood as a topological space in the induced topology,* μ_ε^R *is the restriction of* μ_ε *to* K_R *and L and* L_R *are related as follows*:

$$L_R(x) = \begin{cases} L(x), & \text{if } L(x) \le \alpha(R), \\ +\infty, & \text{otherwise}, \end{cases} \quad \text{and}$$
$$L(x) = \begin{cases} \lim_{R \to \infty} L_R(x), & \text{if defined}, \\ +\infty, & \text{otherwise}. \end{cases} \tag{1.3}$$

PROOF. (i) $\Rightarrow$ (ii): By the construction (1.3), Lebesgue sets of the functional L_R satisfy $\Phi_S^R = \Phi_{S \wedge \alpha(R)} \cap K_R$ and are hence compact. Turning to Condition 1, pick a set J which is open in K_R. If $\inf_J L_R(x) = \infty$, then there is nothing to prove. Otherwise, $\inf_J L_R(x) \le \alpha(R)$. In this case, represent J as an intersection $J = \tilde{J} \cap K_R$, where $\tilde{J}$ is open in X. Note that from the choice of K_R and Condition 1 for $(X, \{\mu_\varepsilon\})$, it follows that $\Phi_{\alpha(R)} \subseteq K_R$. Indeed, assuming the converse, one can find a point x and a neighborhood $\mathscr{U}$ of x such that $\mathscr{U} \subset \Phi_{\alpha(R)}/K_R$. Then by (A1), we obtain

$$\limsup_{\varepsilon \to 0} \varepsilon \ln \mu_\varepsilon(\mathscr{U}) \le \limsup_{\varepsilon \to 0} \varepsilon \ln \mu_\varepsilon(K_R^c) \le -R,$$

and by Condition 1 for the family $\{\mu_\varepsilon\}$,

$$\liminf_{\varepsilon \to 0} \varepsilon \ln \mu_\varepsilon(\mathscr{U}) \geq -\alpha(R) > -R,$$

which is obviously a contradiction. Consequently,

$$\inf_{\tilde{J}} L(x) = \inf_{J} L_R(x) \leq \alpha(R).$$

But $\mu_\varepsilon^R(J) \geq \mu_\varepsilon(\tilde{J}) - \mu_\varepsilon(K_R^c)$ and hence Condition 1 for $(K_R, \{\mu_\varepsilon^R\})$ follows.

Finally, if F is closed in K_R, then it is closed in the original topology of X. Thus Condition 2′ for $(K_R, \{\mu_\varepsilon^R\})$ follows from Condition 2 for $(X, \{\mu_\varepsilon\})$ and the inequality $L \geq L_R \wedge \alpha(R)$.

(ii) ⇒ (i): For each $s > 0$, Φ_S coincides with Φ_S^R if $\alpha(R) > s$ and hence Condition 3 holds. Now if J is open in X and $\inf_J L(x) = \gamma < \infty$, then for any R such that $\alpha(R)$ is greater than γ, $\gamma = \inf_{J \cap K_R} L(x)$. Thus Condition 1 for $(X, \{\mu_\varepsilon\})$ follows from Condition 1 for $\{\mu_\varepsilon^R\}$ and the inequality $\mu_\varepsilon(J) \geq \mu_\varepsilon(J \cap K_R) = \mu_\varepsilon^R(J \cap K_R)$.

In a similar way, for a closed set F, the estimate $\mu_\varepsilon(F) \leq \mu_\varepsilon(F \cap K_R) + \mu_\varepsilon(K_R^c)$ together with Condition 2′ for $(K_R, \{\mu_\varepsilon^R\})$ lead to Condition 2 by letting R tend to infinity.

Now a γ-truncated S.L.D. is invariant under continuous transformations. We will state this fact without proof since it is essentially the same as in the case of S.L.D.'s ([6], Theorem 3.3.1). □

LEMMA 1.4. *Suppose $(X, \{\mu_\varepsilon\})$ is a γ-truncated S.L.D. with rate L_X and F: $X \to Y$ is a continuous map. Then $(Y, \{\nu_\varepsilon\})$ is also a γ-truncated S.L.D. with rate function L_Y, where $\{\nu_\varepsilon\}$ is a family of measures on Y induced by F and $L_Y(y) = \inf_{x \in F^{-1}(y)} L_X(x)$ (as usual* inf *on an empty set equals infinity).*

LEMMA 1.5 (See also [4], Exercise 2.1.20). *Let X, Y be topological Hausdorff spaces, F: $X \to Y$ a continuous injection and $\{\mu_\varepsilon\}$ a family of measures on X satisfying* (A1). *Denote by $\{\nu_\varepsilon\}$ the family of measures induced by F on Y and assume that $(Y, \{\nu_\varepsilon\})$ is an S.L.D. with rate function L_Y.*

Then $(X, \{\mu_\varepsilon\})$ is also an S.L.D. with rate function $L_X = L_Y \circ F$.

PROOF. F has a continuous inverse on each $F(K_R)$, where $\{K_R\}$ is again a family of compacts in X satisfying (A1). Moreover $\{F(K_R)\}$ is an analogous family for $\{\mu_\varepsilon\}$. The claim of the proposition now follows by the successive applications of Lemma 1.5, Lemma 1.4 and again Lemma 1.5. □

REMARK. Note that the above lemma enables us to split the problem of finding a rate function for a family $\{\mu_\varepsilon\}$ on X into two possibly easier ones: First to establish (A1) for $\{\mu_\varepsilon\}$ on X and second to inject X into some Y with a weaker topology, for which the rate function is more or less readily calculable.

As was mentioned in the beginning of the paper we are able to refine the following large deviation result of Baldi [2]:

Let X be a Banach space, $\{\mu_\varepsilon\}$ a family of measures on X and for each ε define a functional H^ε on X^* (the dual of X) by

$$H^\varepsilon(x^*) = \ln \int_X \exp[x^*(x)]\mu_\varepsilon(dx). \tag{1.6}$$

Assume

(A2) $H(x^*) = \lim_{\varepsilon\to 0} \varepsilon H^\varepsilon(x/\varepsilon)$ exists, is lower-semicontinuous and $\mathrm{Dom}(H) = X^*$.

REMARK. Note that under (A2), H is a convex and proper functional.

Define L to be the Fenchel transform of H (with respect to the pairing between the spaces X and X^*). Set $\mathfrak{L} = \{(x, L(x)) | L$ is subdifferential and strictly convex in $x\}$.

(A3) $\mathfrak{L}$ is dense on the boundary of $\mathrm{epi}(L)$.

THEOREM 1.7. (*From* [2]). *Under* (A1)–(A3), $(X, \{\mu_\varepsilon\})$ *is an S.L.D. with rate function* L.

REMARK. In the original paper [2] all results were stated for locally convex spaces. Moreover, as was recently noted by Baldi [3] in the case of X a reflexive B-space, the claim $\mathrm{Dom}(H) = X$ in (A2) can be replaced by the more familiar finite-dimensional condition that the origin belongs to the set $\mathrm{int}(\mathrm{Dom}(H))$.

Now clearly the finer the topology of X is, the more difficult it becomes to calculate the limit in (A2) or even to establish its existence. Similarly, the only general way we know to establish (A3) is to check the Gateaux differentiability of H, which can be a formidable task if X^* is large.

Suppose that X is a dense subspace of another B-space Y. As usual, let $\{\mu_\varepsilon\}$ be a family of measures on X and $\{\nu_\varepsilon\}$ the induced family on Y. Then Lemma 1.5 enables us to refine Theorem 1.7 as follows.

THEOREM 1.7′. *Assume* (A1) *for* $\{\mu_\varepsilon\}$ *on* X *and* (A2) *and* (A3) *for* $\{\mu_\varepsilon\}$ *on* Y.

Then $(X, \{\mu_\varepsilon\})$ *is an S.L.D. with rate function* $L = L_Y \circ I$, *where* L_Y *is given by* (A2) *for* $\{\nu_\varepsilon\}$ *and* I *is an operator of natural embedding of* X *into* Y.

In particular in the example below, Theorem 1.7′ enables us to make all the calculations in an L_2 space as soon as exponential tightness is established.

2. The nonconvex case. Consider the following reaction–diffusion equation (R.D.E.) which we are going to discuss in greater detail later on.

$$\begin{aligned} &\partial u^\varepsilon/\partial t = \Delta u^\varepsilon + b(u^\varepsilon, \xi(t/\varepsilon, x)), \\ &u^\varepsilon(0,x) = g(x), \qquad (t,x) \in [0,T] \times S^1. \end{aligned} \tag{2.1}$$

If the random field $\xi(\cdot, \cdot)$ has sufficiently good mixing properties in t, then the family of measures $\{\mu_\varepsilon\}$ induced (say on $C([0,T] \times S^1)$) by $\{u^\varepsilon\}$ is likely to obey an L.D. principle. On the other hand, because (2.1) is a nonlinear equation for $\{u^\varepsilon\}$, the appropriate rate function is generally nonconvex so there is no hope of applying Theorem 1.7′ directly.

Let us insert an admissible deterministic function ϕ into the nonlinear term $b(\cdot)$ to obtain an auxiliary family of equations

$$(2.1)_\phi \qquad \begin{aligned} \partial u^{\varepsilon,\phi}/\partial t &= \Delta u^{\varepsilon,\phi} + b(\phi, \xi(t/\varepsilon, x)), \\ u^{\varepsilon,\phi}(0,x) &= g(x), \qquad (t,x) \in [0,T] \times S^1. \end{aligned}$$

$(2.1)_\phi$ is already a linear equation and the Fenchel transform technique of the previous section has a good chance of being successful. Assume that we have managed to calculate the rate function L^ϕ for the induced family $\{\mu_\varepsilon^\phi\}$. We want to find some conditions under which the conclusion that $L(\vartheta) = L^\vartheta(\vartheta)$ is the rate function for the original family $\{\mu_\varepsilon\}$ is justified. The general setting goes as follows.

Let X be a B-space, $\{\mu_\varepsilon\}$ a family of measures on X which is assumed to satisfy (A1) and $\{K_R\}$ an appropriate family of compacts. Suppose further that $\forall\, x \in \bigcup_{R>0} K_R$, one can define an augmented family of measures $\{\mu_\varepsilon^x\}$ which is comparable with $\{\mu_\varepsilon\}$ near x in the following sense.

$\forall\, \rho > 0$ and $\forall$ $\mathscr{U}$-neighborhoods of x $\exists$ $\mathscr{U}_1^\rho$- and $\mathscr{U}_2^\rho$-neighborhoods of x such that for ε sufficiently small,

$$(2.2) \qquad \mu_\varepsilon(\mathscr{U}) \ge \mu_\varepsilon^x(\mathscr{U}_1^\rho) - \exp(-\rho/\varepsilon),$$

$$(2.3) \qquad \mu_\varepsilon^x(\mathscr{U}) \ge \mu_\varepsilon(\mathscr{U}_2^\rho) - \exp(-\rho/\varepsilon).$$

LEMMA 2.4. *In the above notation, assume that $(X, \{\mu_\varepsilon^x\})$ is a S.L.D. with rate function L^x for each $x \in \bigcup_{R>0} K_R$. Set*

$$(2.5) \qquad L(x) = \begin{cases} L^x(x), & \text{if } x \in \bigcup_{R>0} K_R, \\ +\infty, & \text{otherwise}. \end{cases}$$

Then the lower-semicontinuity of L implies a L.D. principle for $\{\mu_\varepsilon\}$ on X with L as a rate function.

PROOF. Lower bound: Obviously we have to prove Condition 1 only for the case $J \cap \bigcup_{R>0} K_R \ne \varnothing$. But for each x which lies in $J \cap \bigcup_{R>0} K_R$ and for each $\rho > 0$, one can pick a neighborhood $\mathscr{U}^\rho$ such that by (2.2) and by the L.D. assumption on $\{\mu_\varepsilon^x\}$,

$$\begin{aligned} (-\rho) \vee \liminf_{\varepsilon \to 0} \varepsilon \ln \mu_\varepsilon(J) &\ge \liminf_{\varepsilon \to 0} \varepsilon \ln \mu_\varepsilon^x(\mathscr{U}^\rho) \\ &\ge -\inf_{\mathscr{U}^\rho} L^x \ge -L^x(x) = -L(x). \end{aligned}$$

Because the above is true for all $\rho > 0$ and $x \in J$, the claim follows.

Now Condition 3 for L is an immediate consequence of the lower bound, exponential tightness and lower-semicontinuity.

Upper bound: Let $F \subseteq X$ be closed. Set $\gamma = \inf_F L$. If $\gamma = 0$, Condition 2 is trivially satisfied. Otherwise, pick $0 < \alpha < \gamma$. Note that

$$\limsup_{\varepsilon \to 0} \varepsilon \ln \mu_\varepsilon(F) \le (-2\alpha) \vee \limsup_{\varepsilon \to 0} \varepsilon \ln \mu_\varepsilon(F \cap K_{2\alpha}). \tag{2.6}$$

Next we construct an open covering of the compact $F \cap K_{2\alpha}$ as follows: For each $x \in F \cap K_{2\alpha}$, choose first a neighborhood $\mathscr{U}_x$ such that $\overline{\mathscr{U}}_x \cap \{y | L^x(y) \le \alpha\} = \varnothing$. Then x will be covered by $\mathscr{V}_x^\rho$, where $\mathscr{V}_x^\rho$ satisfies (2.3) with $\mathscr{U}_x$ on the left side. Let $\mathscr{V}_{x_1}^\rho, \dots, \mathscr{V}_{x_n}^\rho$ be a finite subcovering. We then have

$$\begin{aligned}\limsup_{\varepsilon \to 0} \varepsilon \ln \mu_\varepsilon(F \cap K_{2\alpha}) &\le \sup_i \limsup_{\varepsilon \to 0} \varepsilon \ln \mu_\varepsilon\left(\mathscr{V}_{x_i}^\rho\right) \\ &\le (-\rho) \vee \sup_i \limsup_{\varepsilon \to 0} \varepsilon \ln\{\mu_\varepsilon^{x_i}\}, \left(\overline{\mathscr{U}}_{x_i}\right) \le (-\rho) \vee \alpha,\end{aligned}$$

where the second inequality is exactly (2.3) and the last one follows by the L.D. principle for the family $\{\mu_\varepsilon^{x_i}\}$, $i = 1, \dots, n$. Combining this with (2.6) and letting ρ tend to infinity we obtain that for any $\alpha \le \gamma$,

$$\limsup_{\varepsilon \to 0} \varepsilon \ln \mu_\varepsilon(F) \le -\alpha.$$

Therefore Condition 2 follows.

At first glance, the main obstacle in the use of the above lemma is the a priori lower-semicontinuity of L. In fact it can be verified even without actual calculation of L^x, using only the L.D. assumption on $\{\mu_\varepsilon^x\}$ and some regularity conditions. Namely, let us forget for a moment about $\{\mu_\varepsilon\}$ and consider a B-space X with $\{K_R\}$ a nondecreasing family of compacts on X. Suppose again that for any $x \in M = \bigcup_{R>0} K_R$, the family of measures $\{\mu_\varepsilon^x\}$ is defined. Let the following two conditions hold:

$$\text{If } x \in M/K_R, \text{ then } \exists\ \mathscr{U}\text{-neighborhood of } x \text{ such that} \quad \limsup_{\varepsilon \to 0} \varepsilon \ln \mu_\varepsilon^x(\mathscr{U}) \le -R. \tag{2.7}$$

Denote by $B(y; d)$ the ball with radius d and center in y. Then

$$\begin{aligned}&\forall\, x \in M, r > 0 \text{ and } \rho > 0, \exists\, \delta > 0 \text{ and } \bar{r} > 0 \text{ such that} \\ &\text{whenever } \|y - x\| < \delta, y \in M \text{ and } \varepsilon \text{ is small enough,} \\ &\mu_\varepsilon^x(B(x; r)) \ge \mu_\varepsilon^y(B(x; \bar{r})) - \exp(-\rho/\varepsilon).\end{aligned} \tag{2.8}$$

□

LEMMA 2.9. *Assume that* (2.7), (2.8) *are satisfied and* $(X, \{\mu_\varepsilon^x\})$ *is an S.L.D. with rate function* L^x *for each* $x \in M$. *Then the functional* L *given by* (2.5) *is lower-semicontinuous.*

PROOF. Assume that $\{x_n\}$ is a sequence of points in X and $x_n \to x$. We have to show that $L(x) \le \liminf_{n \to \infty} L(x_n)$. Obviously it is enough to consider only $x_n \in M$. Let us distinguish between the two cases $x \in M$ and $x \notin M$.

$x \in M$: Pick a ball $B(x; r)$. Then condition (2.8) implies the existence of $\bar{r} > 0$ such that for n sufficiently large,

$$(-\rho) \vee \liminf_{\varepsilon \to 0} \varepsilon \ln \mu_\varepsilon^x(B(x, r)) \ge \limsup_{\varepsilon \to 0} \varepsilon \ln \mu_\varepsilon^{x_n}(\overline{B}(x, \bar{r})).$$

But

$$\limsup_{\varepsilon\to 0} \varepsilon \ln \mu_\varepsilon^{x_n}\big(\bar{B}(x_n,\bar{r})\big) \ge -L^{x_n}(x_n) \quad \text{if } x_n \in \bar{B}(x,\bar{r}),$$

which will be the case for sufficiently large n. Thus we have $\forall\, \rho > 0$ and $r > 0$,

$$(2.10) \qquad (-\rho) \vee \Big(-\inf_{B(x;r)} L^x\Big) \ge -\limsup_{n\to\infty} L^{x_n}(x_n).$$

By the lower-semicontinuity of L^x,

$$\lim_{r\to 0}\Big(\inf_{B(x;r)} L^x\Big) = L^x(x).$$

Therefore shrinking r to zero and letting ρ tend to infinity in (2.10), we arrive at the desired result.

$x \notin M$: In this case, the sequence $\{x_n\}$ is eventually in K_R^c for each R. Hence by (2.7) and the L.D. property, $L^{x_n}(x_n) \ge R$ for arbitrarily large R and n large enough. □

3. Example: Random perturbation of R.D.E. We turn now to the system (2.1). Set $U = C([0,T] \times S^1)$ and denote by $\mu_\varepsilon, \mu_\varepsilon^\phi$ the measures induced on U by u^ε and $u^{\varepsilon,\phi}$, respectively. For brevity we make the following assumptions (for a discussion of the more general case see [7]):

ASSUMPTION 1. ξ is a step random field such that $\xi(t,x) = \eta_k(x)$ on each of the time intervals $[k-1,k]$ and $\{\eta_k(\cdot)\}$ is a family of i.i.d. Gaussian processes on S^1 with zero mean and continuous correlation function.

ASSUMPTION 2. $b(\cdot)$ is bounded along with its derivative, $\|b\|, \|b'\| \le m$ and $g \in W^2(S^1)$. Now for each $t \in [0,T]$, $\phi \in W^{1,2}([0,T]\times S^1)$ and $\phi(0,\cdot) = g(\cdot)$ define B_ϕ^t as the correlation operator of the process $b(\phi(t,\cdot))\eta_1(\cdot)$. Then B_ϕ^t acts on $L_2(S^1)$. Define the inverse $\mathscr{B}_\phi^t = (B_\phi^t)^{-1/2}$ in the usual way (see [6]). Finally define the functional L^ϕ by

$$(3.1) \quad L^\phi(\vartheta) = \begin{cases} \dfrac{1}{2}\displaystyle\int_0^T \left\| \mathscr{B}_\phi^t\left(\frac{\partial\vartheta}{\partial t} - \Delta\vartheta\right)\right\|_{L_2}^2 dt, & \text{if defined (in the distributional sense) and } \phi(0,\cdot) = g(\cdot), \\ +\infty, & \text{otherwise.} \end{cases}$$

Then the large deviation result is:

LEMMA 3.2. *$(U, \{\mu_\varepsilon\})$ is an S.L.D. with rate function L given by*

$$(3.3) \qquad L(\phi) = \begin{cases} L^\phi(\phi), & \textit{if defined}, \\ +\infty, & \textit{otherwise}. \end{cases}$$

First of all we establish:

LEMMA 3.4. *The families* $\{\mu_\varepsilon\}$, $\{\mu_\varepsilon^\phi\}$ *are exponentially tight on* U.

We defer the proof of this lemma until the next section but note that this already enables us to use Theorem 1.7′. Next we calculate the rate function for the family $\{\mu_\varepsilon^\phi\}$.

LEMMA 3.5. $(U, \{\mu_\varepsilon^\phi\})$ *is an S.L.D. with rate function* L^ϕ.

PROOF. Denote by $G_t(x)$ the heat kernel for the operator $((\partial/\partial t) - \Delta)$, $t \in \mathbb{R}_+$ and $x \in S^1$. Denote further by $*$ the usual convolution on S^1 and by $\circ$ the following operation:

$$\alpha \circ \beta(t,x) = \int_0^t \int_{S^1} \alpha(t-s, x-y)\beta(s,y)\,dy\,ds, \qquad \alpha, \beta \in L_2([0,T] \times S^1).$$

Then $u^{\varepsilon,\phi} = G * g + G \circ [b(\phi)\xi(\cdot/\varepsilon, \cdot)]$. Because $\{\mu_\varepsilon^\phi\}$ is exponentially tight on U, it suffices to prove the proposition only in $L_2([0,T] \times S^1)$. Pick a function $\beta \in L_2([0,T] \times S^1)$ and compute

$$\begin{aligned} H^\varepsilon(\beta) &= \ln E \exp\left(\int_0^T \langle \beta(t,\cdot), u^{\varepsilon,\phi}(t,\cdot)\rangle\,dt\right) \\ &= \int_0^T \langle \beta(t,\cdot), G * g\rangle\,dt \\ &\quad + \ln E \exp\left(\int_0^T \langle [G \circ \hat{\beta}] b(\hat{\phi})(t,\cdot), \hat{\xi}(t/\varepsilon, \cdot)\rangle\,dt\right), \end{aligned}$$

where $\langle \cdot, \cdot \rangle$ denotes the scalar product in $L_2(S^1)$ and $\hat{\alpha}(t,\cdot) = \alpha(T-t,\cdot)$. Hence the generalized Cramér transform H^ϕ is given by

$$\begin{aligned} H^\phi(\beta) &= \int_0^T \langle \beta(t,\cdot), G * g\rangle\,dt \\ &\quad + \lim_{\varepsilon \to 0} \varepsilon \ln E \exp\left(\frac{1}{\varepsilon}\int_0^T \langle [G \circ \hat{\beta}] b(\hat{\phi})(t,\cdot), \hat{\xi}(t/\varepsilon, \cdot)\rangle\,dt\right). \end{aligned}$$

To calculate the second term above, denote by $\omega(\cdot)$ the uniform modulus of continuity (in t) of the function $[G \circ \hat{\beta}]b(\hat{\phi}) = F$. If t_k^ε is some point in the interval $[\varepsilon(k-1), \varepsilon k]$, then

$$\begin{aligned} \frac{1}{\varepsilon}\int_0^T \left\langle F(t,\cdot), \hat{\xi}\left(\frac{t}{\varepsilon}, \cdot\right)\right\rangle dt &= \frac{1}{\varepsilon}\sum_1^{T/\varepsilon} (F(t_k^\varepsilon, \cdot), \eta_k) \\ &\quad + \frac{1}{\varepsilon}\sum_1^{T/\varepsilon} \langle 0_k^\varepsilon(\omega(\varepsilon)), \eta_k\rangle, \end{aligned} \tag{3.6}$$

where $0_k^\varepsilon(\cdot)$ satisfies $\|0_k^\varepsilon(r)\|_{L^\infty(S^1)} \le r$ for $r \ge 0$. Now for any two random variables $\mathscr{x}$ and $\mathscr{y}$ with finite exponential moments,

$$\begin{aligned}(3.7)\qquad & p \ln E \exp(\mathscr{x}/p) - (p/q)\ln E \exp(-q\mathscr{y}/p) \\ & \leq \ln E \exp(\mathscr{x} + \mathscr{y}) \leq (1/p)\ln E \exp(p\mathscr{x}) + (1/q)\ln E \exp(q\mathscr{y}),\end{aligned}$$

where p and q are Hölder conjugate numbers. Therefore, picking $\mathscr{x}$ and $\mathscr{y}$ as the first and second terms, respectively, in (3.6) and letting q tend to infinity, we obtain

$$(3.8)\qquad H^{\phi}(\beta) = \int_0^T \langle \beta(t,\cdot), G * g\rangle\, dt + \tfrac{1}{2}\int_0^T \langle \beta_{\phi}^{t}(G \circ \hat{\beta}), G \circ \hat{\beta}\rangle\, dt.$$

To see that the Fenchel transform of H^{ϕ} is given exactly by (3.1), note that for $\vartheta \in W^{1,2}([0,T] \times S^1)$,

$$(3.9)\qquad \begin{aligned}\int_0^T \langle \beta(t,\cdot), \vartheta\rangle\, dt &= \int_0^T \langle \beta(t,\cdot), G * \vartheta(0,\cdot)\rangle\, dt \\ &\quad - \int_0^T \langle G \circ \hat{\beta}, \frac{\partial \hat{\vartheta}}{\partial t} + \Delta \hat{\vartheta}\rangle\, dt.\end{aligned}$$

(3.8) clearly implies (A2), so to complete the proof we have to verify (A3). But the latter condition readily follows from the fact that H^{ϕ} is Gateaux differentiable (hence L^{ϕ} is strictly convex in any point where it is subdifferentiable) and the Brondsed–Rockafellar theorem [9], which ensures that the set $\{(\vartheta, L^{\phi}(\vartheta)) | L^{\vartheta}$ is subdifferentiable in $\vartheta\}$ is dense in $\partial\, \mathrm{epi}(L^{\phi})$. □

We conclude the proof of Lemma 3.2 by showing that the families $\{\mu_{\varepsilon}\}$, $\{\mu_{\varepsilon}^{\phi}\}$ actually satisfy the conditions of Lemma 2.4 and Lemma 2.9.

LEMMA 3.10. *Denote by* $\|\cdot\|_U$ *the norm in the space* U $(= C([0,T] \times S^1))$. *Then there exists an exponentially bounded family of random variables* $\{D_{\varepsilon}\}$ *such that whenever* $\{u^{\varepsilon,\phi}\}$, $\{u^{\varepsilon,\vartheta}\}$ *are defined the following inequalities hold*:

$$(3.11)\qquad \|u^{\varepsilon} - \phi\|_U \leq [mD_{\varepsilon} \exp(mD_{\varepsilon}) + 1]\|u^{\varepsilon,\phi} - \phi\|_U,$$

$$(3.12)\qquad \|u^{\varepsilon,\phi} - \phi\|_U \leq [1 + mTD_{\varepsilon}]\|u^{\varepsilon} - \phi\|_U,$$

$$(3.13)\qquad \|u^{\varepsilon,\phi} - \phi\|_U \leq \|u^{\varepsilon,\vartheta} - \phi\|_U + [1 + mD_{\varepsilon}]\|\phi - \vartheta\|_U.$$

PROOF OF LEMMA 3.2. The inequalities (3.11)–(3.13) are readily translated into the conditions (2.2), (2.3) and (2.8), respectively. On the other hand, in the course of proving Lemma 3.4, we will see that all of the families $\{\mu^{\varepsilon,\phi}\}$ are exponentially tight with the same family of compacts $\{K_R\}$. This implies the missing condition (2.7). Therefore, Lemma 2.4 applies. □

PROOF OF LEMMA 3.10. Set $z^\varepsilon = u^\varepsilon - u^{\varepsilon,\phi}$ and $v^\varepsilon = u^{\varepsilon,\phi} - u^{\varepsilon,\vartheta}$. From (2.1), $(2.1)_\phi$ and $(2.1)_\vartheta$, we obtain the following equation for z^ε and v^ε:

$$\partial z^\varepsilon/\partial t = \Delta z^\varepsilon + [b(u^\varepsilon) - b(\phi)]\xi(t/\varepsilon, x), \qquad z^\varepsilon(0,\cdot) = 0; \tag{3.14}$$

$$\begin{aligned} \partial z^\varepsilon/\partial t = \Delta z^\varepsilon &+ [b(u^{\varepsilon,\phi}) - b(\phi)]\xi(t/\varepsilon, x) \\ &+ b'(u^{\varepsilon,\phi} - \theta(u^\varepsilon - u^{\varepsilon,\phi}))\xi(t/\varepsilon, x)z^\varepsilon, \\ & z^\varepsilon(0,\cdot) = 0 \quad \text{and} \quad 0 \le \theta \le 1; \end{aligned} \tag{3.15}$$

$$\partial v^\varepsilon/\partial t = \Delta v^\varepsilon + [b(\phi) - b(\vartheta)]\xi(t/\varepsilon, x), \qquad v^\varepsilon(0,\cdot) = 0. \tag{3.16}$$

Set $D_\varepsilon = \int_0^T \|\xi(t/\varepsilon,\cdot)\|_{C(S^1)}\,dt$. From the proof of Lemma 3.3, it becomes apparent that the family $\{D_\varepsilon\}$ is exponentially bounded. Then from (3.14), we obtain $\|z^\varepsilon\|_U \le mTD_\varepsilon\|u^\varepsilon - \phi\|_U$ and hence (3.12). In a similar way, (3.16) implies that $\|v^\varepsilon\|_U \le mTD_\varepsilon\|\phi - \vartheta\|_U$ and this gives us (3.13). On the other hand, from (3.15) and the Feynman–Kac formula, it follows that $\|z^\varepsilon\|_U \le mTD_\varepsilon \exp(mTD_\varepsilon)\|u^{\varepsilon,\phi} - \phi\|_U$ and therefore (3.11) holds. □

4. Proof of Lemma 3.3. Denote by $\|\cdot\|_{1,2}$ and $\|\cdot\|_2$, respectively, the norms in the Sobolev spaces $W^{1,2}([0,T]\times S^1)$ and $W^2(S^1)$. The following two facts play the principal role in our verification of (A1) for the family $\{\mu_\varepsilon\}$.

1. Consider the equation

$$\frac{\partial u}{\partial t} = \Delta u + f, \qquad u(0,\cdot) = g(\cdot). \tag{4.1}$$

   Suppose that $f \in L^2([0,T]\times S^1)$ and $g \in W^2(S^1)$. Then there exists a unique weak solution $\bar{u}$ of (4.1) which belongs to $W^{1,2}([0,T]\times S^1)$. Moreover

$$\|\bar{u}\|_{1,2} \le \text{const.}\left[\|g\|_2 + \left(\int_0^T\int_{S^1} f^2\,dx\,dt\right)^{1/2}\right] \qquad (\text{from [1]}).$$

2. $W^{1,2}([0,T]\times S^1)$ is compactly embedded into $C([0,T]\times S^1)$. This follows from the results of [10].

Combining 1 and 2 we see that the proof reduces to the verification of the exponential boundness of the family $\{C_\varepsilon = \int_0^T\int_{S^1}\xi(t/\varepsilon,x)^2\,dx\,dt\}$. We will do even more and establish the exponential boundness of the family $\{D_\varepsilon\}$ defined above. The latter will follow if we show that $\forall\ R > 0$, $\exists\ N_R$ such that

$$\lim_{n\to\infty}\frac{1}{n}\ln P\left\{\frac{1}{n}\sum_1^n \|\eta_k\|_{C(S^1)} > N_R\right\} \le -R. \tag{4.2}$$

We set $l_k = \|\eta_k\|_{C(S^1)}$. The proof of (4.2) is based on the estimate of Fernique and Shepp for the tail of the distribution of the suprema of a gaussian process [8]: $\forall\ \Delta > d\ \exists\ m(\Delta)$: $P\{l_1 > z\} \le \exp(-z^2/2\Delta)$ as soon as $z > m(\Delta)$, where $d = \sup_{x\in S^1} E\eta_1^2(x)$. What we actually need is the finiteness of the exponential moment $E\exp(l_1) = A < \infty$. Now for any N, we have

$$(4.3)\quad P\left\{\frac{1}{n}\sum_1^n l_k > N\right\} \le \left[E\exp(l_1)\right]^n e^{-nN} = \exp(-n(N - \ln A)).$$

Setting $N_R = R + \ln A$ in (4.3) we obtain (4.2), thereby proving the lemma.

Acknowledgments. I am grateful to M. I. Freidlin who suggested my studying (2.1) and made some valuable comments concerning this equation. I also want to thank R. G. Pinsky and A. D. Ioffe for helpful discussions.

REFERENCES

[1] AGRANOVICH, M. G. and VYSHYK, M. I. (1964). Elliptic problems with parameter and parabolic problems of general form. *Uspehi Mat. Nauk* **19** 53–161.

[2] BALDI, P. (1988). Large deviations and stochastic homogenization. *Ann. Mat. Pura Appl.* (*4*) 161–177.

[3] BALDI, P. (1990). Private communication.

[4] DEUSCHEL, J.-D. and STROOCK, D. W. (1989). *Large Deviations*. Academic, New York.

[5] FREIDLIN, M. I. (1988). Random perturbations of reaction–diffusion equations. *Trans. Amer. Math. Soc.* **305** 665–697.

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

[7] IOFFE, D. (1991). Large deviations for reaction–diffusion equations with rapidly oscillating random noise. In *Proceedings of the Third Trento Conference on SPDE*. Springer, New York. To appear.

[8] MARCUS, M. B. and SHEPP L. A. (1972). Sample behavior of Gaussian processes. *Proc. Sixth Berkeley Symp. Math. Statist. Probab.* **2** 423–441.

[9] PHELPS, R. R. (1989). *Convex Functions, Monotone Operators and Differentiability. Lecture Notes in Math.* **1364**. Springer, New York.

[10] VOLEVICH, L. and PONEYAH, B (1965). Some spaces of generalized functions and embedding theorems. *Russian Math. Surveys* **20** 3–74.

DEPARTMENT OF MATHEMATICS
TECHNION
ISRAEL INSTITUTE OF TECHNOLOGY
HAIFA 32000
ISRAEL