Evolution of physical objects is often described by ordinary differential equations having the form,

\[ \frac{dX_x(t)}{dt} = B(X_x(t)), \quad X_x(0) = x, \]

where the solution \( X_x(t) \) runs in \( \mathbb{R}^d \) or on a smooth manifold. To ensure existence and uniqueness in the equation (1), the right-hand side \( B \) is usually supposed to be Lipschitz continuous. Then we arrive at a group of diffeomorphisms \( F^t, \) acting by \( F^t x = X_x(t), \) called a dynamical system or a flow to amplify the continuous time here, taking into account that the theory of dynamical systems widely studies also the discrete time case, i.e., iterates of one transformation. A description of a physical evolution in the form (1) can be viewed as an idealization since real motions are subject to various perturbations which can be considered random.

It seems that the problem of studying small random perturbations of dynamical systems was first formulated explicitly by Pontrjagin, Andronov, and Vitt [9] in 1933. As a particular model of random perturbations we can consider a stochastic process \( X^\varepsilon(t) = X^\varepsilon_x(t) \) solving the equation

\[ X^\varepsilon_x(t) = x + \int_0^t B(X^\varepsilon_x(s))ds + \varepsilon w_t, \]

where \( w_t \) is the standard Brownian motion called also the Wiener process. More generally, one of most studied models of random perturbations is represented by a diffusion Markov process \( X^\varepsilon(t), \) \( X^\varepsilon(0) = x, \) whose semigroup of transition operators \( P^\varepsilon_t \) is generated by \( L^\varepsilon = \varepsilon^2 L + (B, \nabla) \) where \( L = \frac{1}{2} (a(x) \nabla, \nabla) + (b(x), \nabla) \) is a second order elliptic (partial) differential operator. Usually, a more convenient way to deal with \( X^\varepsilon_x(t) \) is to represent it as a solution of the Itô stochastic differential equation

\[ dX^\varepsilon_x(t) = \varepsilon \sigma(X^\varepsilon_x(t))dw_t + B(X^\varepsilon_x(t))dt + \varepsilon^2 b(X^\varepsilon_x(t))dt, \]

where \( \sigma(x) \sigma^*(x) = a(x). \) We arrive at the particular case (2) when \( L = \Delta \) is the Laplacian and \( b \equiv 0. \) As in (1), Lipschitz condition on coefficients in (3) ensure existence and uniqueness of solutions there. It is clear that during a fixed bounded time, \( X^\varepsilon_x(t), \) with probability close to one, stays close to the unperturbed motion \( X_x(t) \) but when we are interested in events having very small probabilities or taking place after a very long time, then randomness starts playing a crucial role.

Let \( \varphi_t, t \in [0, T] \) be a continuous curve in the space \( \mathbf{M} \) (either \( \mathbb{R}^d \) or a Riemannian manifold) where both \( X(t) \) and \( X^\varepsilon(t) \) live and we want to estimate the probability that \( X^\varepsilon \) stays in a small neighborhood of \( \varphi, \) i.e., say \( \rho_{0T}(X^\varepsilon, \varphi) \leq \delta \) where \( \rho_{0T}(\varphi, \psi) = \sup_{0 \leq t \leq T} \rho(\varphi_t, \psi_t) \) and \( \rho \) is the distance on \( \mathbf{M}. \) The basic result obtained by Freidlin and Wentzell in their seminal 1970 paper [10] says that the
above probability can be described by means of the functional (which they called
an action functional) defined for absolutely continuous curves by

\[ S_{0T}(\varphi) = \int_0^T L(\varphi_t, \dot{\varphi}_t) dt. \]

Here \( L(x, \beta) = \frac{1}{2}((a(x))^{-1}(\beta - B(x)), (\beta - B(x))) \) and \((\cdot, \cdot)\) denotes the inner
product. If \( \varphi \) is not absolutely continuous, then \( S_{0T}(\varphi) \) is set to be \( \infty \). Now, the
probabilities in question can be described in the following way. For any \( s, \delta, \gamma > 0 \),

\[ P\{\rho_{0T}(X^\varepsilon, \varphi) < \delta\} \geq \exp\{-\varepsilon^{-2}(S_{0T}(\varphi) + \gamma)\}, \]

and for \( \Phi(s) = \{\varphi : S_{0T}(\varphi) \leq s\}, \)

\[ P\{\rho_{0T}(X^\varepsilon, \Phi(s)) \geq \delta\} \leq \exp\{-\varepsilon^{-2}(s - \gamma)\}. \]

These estimates have the form of the lower and upper bounds of the theory of
large deviations. We see from (5) and (6) that on the time interval \([0, T]\), the
perturbed motion \( X^\varepsilon \) stays close to \( X_x \) with overwhelming probability. Still with
exponentially small in \( \varepsilon^{-2} \) probability other paths of \( X^\varepsilon \) are possible, as well.

It turns out that (5) and (6) enable us also to study the behavior of \( X^\varepsilon \) on very
long time intervals. Namely, define

\[ V(x, y) = \inf_{\varphi, T \geq 0} \{S_{0T}(\varphi) : \varphi_0 = x, \varphi_T = y\}, \]

which is a Lipschitz continuous function. Suppose that \( D \) is a relatively compact
domain in \( M \) with a smooth boundary \( \partial D \) such that the flow \( F^t \) attracts \( D \cup \partial D \)
to a fixed point \( O \in D \) and at each \( x \in \partial D \) the vector \( B(x) \) points strictly inside \( D \). Suppose that \( \min_{y \in \partial D} V(O, y) = V(O, y_0) \) and that such \( y_0 \) is unique on \( \partial D \).

Let \( \tau_x(D) = \inf\{t \geq 0 : X^\varepsilon_x(t) \in \partial D\} \). Then for any \( x \in D \) and \( \delta > 0 \),

\[ \lim_{\varepsilon \to 0} P\{\rho(X^\varepsilon_x(\tau_x(D)), y_0) > \delta\} = 0, \]

i.e., with overwhelming probability \( X^\varepsilon_x \) exits from \( D \) near \( y_0 \). This result from
[10] has a partial differential equations interpretation. Namely, let \( u^\varepsilon \) be the solution of the Dirichlet problem \( \Delta u^\varepsilon = 0 \) in \( D \) with continuous boundary conditions \( u^\varepsilon|_{\partial D} = g \). Then in the above circumstances \( \lim_{\varepsilon \to 0} u^\varepsilon(x) = g(y_0) \) for any \( x \in D \).

This result was new at the time and it was difficult to approach it by pure partial
differential equations methods, which prompted development of the theory of
viscosity solutions.

Among important questions on random perturbations of dynamical systems
raised in [9] and later repeatedly advertized by Kolmogorov was the problem on
stochastic stability of invariant measures of dynamical systems. Namely, suppose
that both the flow \( F^t \) and its diffusion random perturbation \( X^\varepsilon(t) \) live on compact
Riemannian manifold \( M \). Many interesting classes of dynamical systems, especially
so-called chaotic systems, possess abundance of invariant measures, i.e., probability
measures \( \mu \) such that \( F^t \mu = \mu \), while a nondegenerate diffusion \( X^\varepsilon(t) \) has a unique
invariant measure \( \mu_{\varepsilon} \), i.e., \( \int P_t^\varepsilon f d\mu_{\varepsilon} = \int f d\mu_{\varepsilon} \) for any bounded Borel function \( f \).

It is not difficult to see that all (weak) limit points of \( \mu_{\varepsilon} \) as \( \varepsilon \to 0 \) are invariant
measures of the flow \( F^t \) and, in particular, if the latter has a unique invariant
measure \( \mu \) (i.e., it is uniquely ergodic), then \( \mu_{\varepsilon} \Rightarrow \mu \) as \( \varepsilon \to 0 \). On the other
hand, if \( F^t \) possess a lot of invariant measures and still \( \mu_{\varepsilon} \Rightarrow \mu_0 \) as \( \varepsilon \to 0 \) for
some particular \( \mu_0 \), then it is natural to call the latter the most stable (and thus
having physical meaning) invariant measure of the flow \( F^t \). Initial results in this
direction were obtained by Khasminskii (see [4]). Freidlin and Wentzell showed
in [10] that a part of the above problem can be studied relying on the functions \( V(x, y) \) defined in (7). Namely, let the \( \omega \)-limit set of the flow \( F_t^s \) consist of a finite number of disjoint compacts \( K_1, \ldots, K_\ell \) such that \( V(x, y) = V(y, x) = 0 \) if and only if \( x \) and \( y \) belong to one compact. Then \( V(x, y) = V_{ij} \) for \( x \in K_i \) and \( y \in K_j \) depends only on \( i \) and \( j \). It was shown in [10] that the limiting as \( \varepsilon \to 0 \) behavior of the vector \((\mu_\varepsilon(K_1), \ldots, \mu_\varepsilon(K_\ell))\) can be described via invariant measures of certain Markov chains with \( \ell \) states and transition probabilities determined by \( V_{ij}, i, j = 1, \ldots, \ell \). If the flow \( F_t^s \) is topologically transitive on \( M \), i.e., it has a dense orbit, then \( V(x, y) = 0 \) for any pair \( x, y \in M \), which means that \( \ell = 1 \) and the above approach gives nothing. Even when \( \ell > 1 \), this method cannot give any information about \( \mu_\varepsilon \) inside of each compact \( K_i \). Then other machinery should be applied and, in the case of Anosov flows or, more generally, Axiom A flows, the problem was solved by the reviewer (see [6]) showing that \( \mu_\varepsilon \) exists for any pair \( x, y \in M \) and the above approach gives nothing. Even when \( \ell > 1 \), this method cannot give any information about \( \mu_\varepsilon \) inside of each compact \( K_i \). Then other machinery should be applied and, in the case of Anosov flows or, more generally, Axiom A flows, the problem was solved by the reviewer (see [6]) showing that \( \mu_\varepsilon \) tends weakly as \( \varepsilon \to 0 \) to the so-called Sinai–Ruelle–Bowen invariant measure. A discrete time version of the above Freidlin–Wentzell theory appeared in the reviewer’s paper [7].

Freidlin showed in [11] that a machinery similar to the above large deviations can be applied to some problems in the theory of averaging. The averaging setup can be viewed as a perturbation of a family of dynamical systems. Namely, consider a family of flows \( F_t^s y = Y_{x,y}(t) \) determined by ordinary differential equations

\[
\frac{dY_{x,y}(t)}{dt} = b(x, Y_{x,y}(t)), \quad Y_{x,y}(0) = y
\]

parametrized by \( x \in \mathbb{R}^d \). We may think about the flow \( \Phi_t^s(x, y) = (x, F_t^s y) \) as describing an idealized physical motion where parameters \( x = (x_1, \ldots, x_d) \) play the role of constants of motion. In a real system these parameters evolve slowly and we arrive at a “fully coupled” system

\[
\frac{dZ_{x,y}(t)}{dt} = \varepsilon B(Z_{x,y}(t), Y_{x,y}(t)), \quad \frac{dY_{x,y}(t)}{dt} = b(Z_{x,y}(t), Y_{x,y}(t)),
\]

where \( Z_{x,y}(0) = x, Y_{x,y}(0) = y \). For obvious reasons \( Z_{x,y} \) and \( Y_{x,y} \) are called the slow and the fast motions, respectively. Sometimes, by a change of variables we can “decouple” (10), arriving at the equation

\[
\frac{dZ_{x,y}(t)}{dt} = \varepsilon B(Z_{x,y}(t), F_t^s y), \quad Z_{x,y}(0) = x,
\]

where now the fast motion does not depend anymore on the slow one. In place of the flow \( F_t^s \) in (11) we can consider a stochastic process \( \xi_t \), which is the case studied both in [11] and in Chapter 7 of the book under review. So we proceed with the description of results in this “uncoupled” situation

\[
\frac{dZ_{x}^\varepsilon(t)}{dt} = \varepsilon B(Z_{x}^\varepsilon(t), \xi_t), \quad Z_{x}^\varepsilon(0) = x,
\]

referring the reader to the study of the more difficult fully coupled case both in deterministic and stochastic setups to the reviewer’s work [8]. The classical averaging principle yields that if \( B(z, y) \) is Lipschitz continuous in \( z \) and the limit

\[
\bar{B}(z) = \lim_{t \to \infty} \frac{1}{t} \int_0^t B(z, \xi_s) ds
\]

exists, then the averaged motion \( \bar{X}_x \) solving the equation

\[
\frac{d\bar{X}_x(t)}{dt} = \bar{B}(\bar{X}_x(t)), \quad \bar{X}_x(0) = x
\]
approximates the time changed slow motion \( X^\varepsilon(t) = Z^\varepsilon(t/\varepsilon) \) on any bounded time interval \([0, T]\), i.e., \( \lim_{\varepsilon \to 0} \sup_{0 \leq t \leq T} |X^\varepsilon(t) - \tilde{X}_x(t)| = 0 \). Suppose that for any piecewise constant \( \mathbb{R}^d \)-valued functions \( \varphi_s \) and \( \alpha_s \) on \([0, T]\),

\[
(13) \quad \lim_{\varepsilon \to 0} \varepsilon \ln E \exp \left( \frac{1}{\varepsilon} \int_0^T (\alpha_s, b(\varphi_s, \xi_s/\varepsilon)) ds \right) = \int_0^T H(\varphi_s, \alpha_s) ds
\]

for some function \( H \) on \( \mathbb{R}^d \times \mathbb{R}^d \). Then \( H(x, \alpha) \) is continuous and convex in \( \alpha \) and its Legendre transform

\[
(14) \quad L(z, \beta) = \sup_{\alpha} \left( (\alpha, \beta) - H(z, \alpha) \right)
\]

is convex and lower semicontinuous. Now introduce the “action” functional by \( L \) only with \( L \) given by \( (14) \). If \( H(x, \alpha) \) is differentiable in \( \alpha \), then according to \([1]\) the slow motion \( X^\varepsilon \) satisfies the large deviations estimates \( (5) \) and \( (6) \) with this new “action” functional. Condition \( (13) \) can be verified (and the function \( H \) identified) for some stochastic processes \( \xi_t \) such as nondegenerate finite states Markov chains with continuous time (the case considered in \([1]\)), nondegenerate diffusions on compact manifolds (more generally, for random evolutions), and for some classes of dynamical systems (see \([8]\)). Assuming special structure of the flow \( \tilde{F}_t \) generated by the averaged vector field \( \tilde{B} \), we can obtain results about very long time behavior of \( X^\varepsilon(t) \) similar to the random perturbations case described above. This concerns, in particular, exits of \( X^\varepsilon(t) \) from a domain \( D \) attracted by the flow \( \tilde{F}_t \) to a fixed point \( O \in D \). A word of caution is due here: these results described in Section 6 of Chapter 7 in the book under review hold true if the function \( V \) defined by \( (7) \) (with the “action” functional obtained by \( L \) from \( (14) \)) is lower semicontinuous. This is not so obvious here as in the random perturbations case and requires some extra assumptions (see \([8]\)). Furthermore, at the end of the same section there is a discussion on invariant measures of \( X^\varepsilon(t) \), which is not quite clear since it is not a Markov process, in general. Thus this section should be read with a “grain of salt”.

Up to now we discussed the theory developed in \([10]\) and \([1]\), which was the base for the whole first edition of the book as well as for Chapters 1–7 of the second and the third editions. The main addition in the latter editions is the study of random perturbations of Hamiltonian systems. Thus we return to the setup at the beginning of this review where now \( B \) in \([1]\) is a vector field of the form \( B = \nabla H = (\partial H/\partial p, -\partial H/\partial q) \) with a Hamiltonian \( H = H(p, q) \). The unperturbed motion \( X^0(t) = X(t) \) stays all the time on one surface of constant Hamiltonian (energy level) but the perturbed process \( X^\varepsilon(t) \) can drift slowly across these surfaces. In the two dimensional case, i.e., for systems with one degree of freedom, when energy levels in a region form a family of closed orbits, the problem was studied by Khasminskii in \([5]\), who showed that the drift across these orbits by the time rescaled process \( Z^\varepsilon(t) = X^\varepsilon(t/\varepsilon^2) \) can be asymptotically described by a one dimensional diffusion process with characteristics obtained by averaging corresponding characteristics of \( Z^\varepsilon \) along closed orbits above. The motion across energy levels can be conveniently represented by the process \( H(Z^\varepsilon(t)) \), and the above result yields that the latter converges in distribution to a one dimensional diffusion. Suppose now that the Hamiltonian has several critical points, then the energy levels will have more complicated structures. Let, for instance, the Hamiltonian has two local minima and one local maximum so that the Hamiltonian flow has three families
of closed orbits separated by an \( \infty \)-shape curve (consisting of a saddle fixed point and two orbits). It turns out that in this and more general two dimensional cases the motion across constant energy curves can be asymptotically described by a diffusion on a graph obtained by identifying points on each such curve so that each edge will correspond to one connected component filled entirely by closed orbits. This diffusion is well defined in the interior of each edge of the graph which should be supplemented by appropriate boundary conditions at vertices to ensure correct transition laws from edge to edge. In dimension higher than two the global geometry of constant energy surfaces may become quite complicated and the description of the slow evolution of \( X^\varepsilon(t) \) across them was obtained so far only in special cases described in Chapter 9 of the book under review. This part of the book is based on several more recent papers by Freidlin and Wentzell, in particular, [2], [3] and others.

The last two chapters are devoted to brief description of other related asymptotic problems, such as stability under random perturbations, asymptotics of transition densities for processes with small diffusion, reaction-diffusion equations with small parameters, and random perturbations of infinite dimensional semiflows generated by reaction-diffusion equations.

The book under review is neither a textbook nor a reference book devoted to a classical topic, but rather a research monograph that comprises the work of the authors (sometimes together with others) for about 40 years. It is quite rare that such a book would go already through its third edition, which demonstrates the high quality, originality, and importance of the theory developed there.

REFERENCES


Yuri Kifer

Institute of Mathematics, Hebrew University, Jerusalem 91904, Israel

E-mail address: kifer@math.huji.ac.il