Journal of Statistical Physics, Vol. 66, Nos. 3/4, 1992

Quasipotentials for Simple Noisy Maps with Complicated Dynamics

Andreas Hamm[1] and Robert Graham[1]

Received May 17, 1991; final August 20, 1991

The theory of nonequilibrium potentials or quasipotentials is a physically motivated approach to small random perturbations of dynamical systems, leading to exponential estimates of invariant probabilities and mean first exit times. In the present article we develop the mathematical foundation of this theory for discrete-time systems, following and extending the work of Freidlin and Wentzell, and Kifer. We discuss strategies for calculating and estimating quasipotentials and show their application to one-dimensional S-unimodal maps. The method proves to be especially suited for describing the noise scaling behavior of invariant probabilities, e.g., for the map occurring as the limit of the Feigenbaum period-doubling sequence. We show that the method allows statements about the scaling behavior in the case of localized noise, too, which does not originally lie within the scope of the quasipotential formalism.

KEY WORDS: Dynamical systems; random perturbations; Feigenbaum attractor; noise scaling.

1. INTRODUCTION

In the study of random perturbations of dynamical systems one approach of special physical relevance is a generalization of a prototype situation one meets with in equilibrium thermodynamics (see, e.g., ref. 1):

Let the macroscopic state of a thermodynamical system near equilibrium be specified by a k-tuple $q=(q^1,..., q^k)$ of macroscopic variables. Phenomenologically, the change of these variables can be described by a first-order differential equation in time:

$$\frac{d}{dt} q = f(q) \tag{1.1}$$

[1] Fachbereich Physik, Universität-GH Essen, 4300 Essen 1, Germany.

0022-4715/92/0200-0689$06.50/0

From the microscopic theory one can derive that the vector field $f(q)$, called the drift, can be split up into a reversible part $r(q)$ and an irreversible part $d(q)$, and that the latter can be written as the gradient of a (coarse-grained) thermodynamic potential $\Phi(q)$:

$$d(q) = -\tfrac{1}{2}\nabla\Phi(q) \tag{1.2}$$

where the gradient has to be taken with respect to a certain metric tensor G which is built up from the so-called transport coefficients. The reversible part $r(q)$ is in this metric orthogonal to $d(q)$.

The same thermodynamic potential $\Phi(q)$ as introduced above has another meaning in the context of thermodynamic fluctuations of the macroscopic variables, which are described by an Ito stochastic differential equation

$$\frac{d}{dt} q = f(q) + \sqrt{\eta}\, \xi_t \tag{1.3}$$

that emerges from (1.1) by adding a k-tuple ξ of mutually independent sources of Gaussian white noise with zero mean and covariance $\langle \xi_t \xi_{t'}^* \rangle = G\delta(t-t')$. The parameter η is Boltzmann's constant here. The thermodynamic potential $\Phi(q)$ enters the equilibrium distribution $w(q)$, which is an invariant probability density for the process defined by Eq. (1.3):

$$w(q) = Z \exp\left(-\frac{\Phi(q)}{\eta}\right) \tag{1.4}$$

with some prefactor Z.

During the past decades much progress has been achieved in the partial adaptation of the above results to systems for which a steady state far from equilibrium supersedes the equilibrium state, or—more abstractly—to general dynamical systems of the form (1.1) which are subject to random perturbations as described by (1.3). For a recent review see ref. 2.

In some cases there exists a continuously differentiable function $\Phi(q)$ such that the drift splits up into two orthogonal parts, one of which satisfies Eq. (1.2). This function $\Phi(q)$ is often referred to as a nonequilibrium potential. In general, the two parts of the drift do not have the interpretation of being the reversible and the irreversible part. What—in a weakend sense—remains is the statistical meaning of the nonequilibrium potential: For an invariant probability density of the process (1.3), Eq. (1.4) holds true not strictly, but as an asymptotic formula in the weak-noise limit, $\eta \to 0$.

In other cases, Eq. (1.4) can still be used as an approximation of the invariant probability density in the weak-noise limit, but the nonequilibrium potential which enters is no longer continuously differentiable.

Apart from its role in the asymptotic behavior of the stationary probability distribution, the nonequilibrium potential can be used to write mean escape times $\langle \tau \rangle$ from a certain region of the configuration space containing the steady state under the influence of the random perturbations in an Arrhenius-type relation:

$$\langle \tau \rangle \sim \exp\left(\frac{\Delta\Phi(q)}{\eta}\right) \tag{1.5}$$

Here $\Delta\Phi$ is the minimal nonequilibrium potential difference between the steady state and the boundary of the region. Again, (1.5) holds true only as an asyptotic formula for $\eta \to 0$.

Independent of the evolution of this subject in a physical context, Wentzell and Freidlin[3] have elaborated a mathematical theory of the weak-noise limit of systems described by (1.3). A detailed account of their work can be found in ref. 4. A main result of this theory as well as of the heuristically based work in the physical literature is the formulation of an extremum principle which presents a tool for finding nonequilibrium potentials, or quasipotentials, as they are called by Wentzell and Freidlin.

The extremum principle has successfully been applied in the search for nonequilibrium potentials of a number of physically motivated examples (e.g., refs. 5–7). However, all these examples have the common feature that the underlying deterministic systems (1.1) are quite simple from the dynamical point of view. Clearly, an extension of the applicability to systems with more complicated properties, such as fractal basin boundaries or strange attractors, is highly desirable.

A reasonable approach to this problem is to begin with an investigation of discrete-time systems which show those complicated properties. The Eq. (1.1) has to be replaced by a difference equation, Eq. (1.3) by a stochastic difference equation. Recently, there has been some work in this direction, both on the mathematical[8–10] and on the physical[7,11–16] sides.

The intuitive physical arguments have resulted, in one-dimensional systems, in working methods for determining the analogue of the nonequilibrium potential, which we choose here to call the quasipotential in accordance with Wentzell and Freidlin. Unfortunately, the nature of these arguments does not allow any rigorous statements on their validity, especially for systems with complicated dynamics. At best one can appeal to numerical evidence *a posteriori.*

Kifer[8] has presented a discrete-time version of the Wentzell–Freidlin theory. However, in his work there is no contact with the physical applications sketched above.

In this article we choose an intermediate presentation: We start from rigorous results obtained with reference to the work of Wentzell, Freidlin, and Kifer. Generalizing ideas from the above-cited physical literature, we then systematically derive methods for determining quasipotentials. After that we give a first example of an application. Due to our starting point, the results of this application do not need (but of course can be tested by) numerical verification.

We give a more detailed plan of what follows:

In Section 2 we give the definitions necessary for and the theorems derived from the discrete-time version of the Wentzell–Freidlin theory. This deeply relies on the work of Kifer.[8] Our exposition is more specialized than Kifer's in that it is adjusted to the physical application in mind. On the other hand, it is more general in that we deal with the exit problem, and in that we weaken the fundamental condition for the theorems in order to be able to apply the theory to systems which are limits of a cascade of flip bifurcations. Section 2 does not contain any proof. The reader interested in the mathematical details is referred to the Appendix. It supplies the information necessary to prove those of our statements which differ from ref. 8. We hope that even readers without acquaintance with the Wentzell–Freidlin theory might find the Appendix useful for getting an idea of the arguments involved. Readers who feel no need for mathematical rigor but are familiar with some of the physical work in this area may want to skip Section 2. They should then get used to our notation by reducing statement **O5** in Section 3, setting $D = M$ and $r = 2$, to something they know.

Section 3 is a collection of implications of the general theory which are appropriate tools for an actual determination of quasipotentials. The most general of these implications are placed before those which need special assumptions on the system.

In Section 4 we exemplify a simple application of the tools from Section 3. There we deal with random perturbations of S-unimodal maps on the interval. This class of maps representatively shows all the situations one can meet with in the determination of quasipotentials for one-dimensional maps. A case of special interest is the limit of period-doubling bifurcations. We show how the theory of quasipotentials is able to give a lucid description of the universal noise scaling behavior.

Conclusions are given in Section 5.

2. BASIC DEFINITIONS AND THEOREMS

A discrete-time dynamical system is given by a continuous map $F: M \to M$ on some metric space (M, d). The theory of small random perturbations of such a system deals with Markov sequences $(X_n^\eta(\omega))_{n=0,1,2...}$, $\omega \in \Omega$ (Ω a measurable space of events), with one-step transition probabilities $P^\eta(x, \Gamma) = P_x^\eta\{\omega \in \Omega: X_1^\eta(\omega) \in \Gamma\}$ ($x \in M$, Γ a Borel set of M), which tend to δ-measures centered at $F(x)$ as the small parameter $\eta > 0$ goes to zero. In the following we shall assume that $P^\eta(x, \Gamma)$ has density $p^\eta(x, y)$ with respect to a standard Borel measure on M.

Our main interest lies in an asymptotic description of two objects with obvious physical importance: an invariant measure, which characterizes the stable distribution in the long-time behavior of the system, and the mean first exit time out of some domain D of M (e.g., a basin of attraction in the unperturbed system), which is observable both in simulations and experiments. Recall the following defining equations: for the invariant measure μ^η, provided it exists,

$$\int_M d\mu^\eta(x)\, P^\eta(x, \Gamma) = \mu^\eta(\Gamma) \tag{2.1}$$

for all Borel sets Γ; and for the mean first exit time $\langle \tau_D^\eta \rangle_x$, when starting from $x \in D$,

$$\langle \tau_D^\eta \rangle_x = \int_{\omega \in \Omega} (\inf\{m > 0: X_m^\eta(\omega) \notin D\})\, dP_x^\eta(\omega) \tag{2.2}$$

Kifer's generalization[8] of the Wentzell–Freidlin approach[4] starts with a large-deviation condition on the random perturbations, which we formulate here in a slightly specialized version:

Assumption A. There exists a continuous function $\rho(x, y) \geqslant 0$ on $M \times M$, called the deviation rate, with the property that $y \mapsto \rho(x, y)$ has a unique minimum $\rho(x, F(x)) = 0$, such that uniformly in x and $y \in M$,

$$\lim_{\eta \to 0} \eta \log p^\eta(x, y) = -\rho(x, y)$$

We call such random perturbations ρ-noise.

Although at the moment we do not need to specify the deviation rate, we refer to the best-known example of small random perturbations satisfying Assumption A. This is given by a stochastic difference equation similar to the stochastic differential equation (1.3),

$$X_{n+1}^\eta = F(X_n^\eta) + \sqrt{\eta}\, \xi_n \tag{2.3}$$

where F is a real function and (ξ_n) a sequence of independent real-valued Gaussian random numbers of unit variance. In this case the deviation rate is $\rho(x, y) = \frac{1}{2}|y - F(x)|^2$.

Generalizing this example, we introduce for later use the deviation rates

$$\rho_r(x, y) = \frac{1}{r}\,[d(y, F(x))]^r \tag{2.4}$$

with $r > 0$ and not necessarily integer.

The large-deviation condition allows an exponential estimate of the probability that a realization of the Markov sequence (X_i) stays close to a given sequence (q_i) of length N in M. This probability for small η is approximately $\exp\{-S_N[(q_i)]/\eta\}$ [for a precise formulation see (A.1)]. Here we have introduced the following quantity:

$$\begin{aligned} S_1[(q_0)] &:= 0 \\ S_N[(q_i)_{0 \leqslant i < N-1}] &:= \sum_{i=0}^{N-2} \rho(q_i, q_{i+1}), \qquad N \geqslant 2 \end{aligned} \tag{2.5}$$

which we call the action along the sequence (q_i).

It is then a natural question to ask for the least action along all sequences leading from one point x to another point y. We add the condition that the sequences must not leave a domain $D \subset M$ ($\bar{D}$ compact), with the possible exception of the last point y. We split up the minimizing problem into two steps and define

$$V_N^D(x, y) = \min\{S_N[(q_i)]: q_0 = x \in D; q_1, \dots, q_{N-2} \in D; q_{N-1} = y\} \tag{2.6}$$

and

$$V^D(x, y) = \inf\{V_N^D(x, y): N \geqslant 1\} \tag{2.7}$$

$V^D(x, y)$ is continuous on $D \times M$. In the next section we shall be concerned with the practical aspects of the minimizing procedure, but first we discuss the use of the function $V^D(x, y)$.

Consider the following equivalence relation on D, implied by V^D:

$$x \overset{D}{\sim} y \qquad \text{iff} \quad V^D(x, y) = V^D(y, x) = 0 \tag{2.8}$$

The corresponding equivalence classes are denoted by $[x]^D$. They are compact. Due to definition (2.8), the notations $V^D([x]^D, y)$, $V^D(x, [y]^D)$, and $V^D([x]^D, [y]^D)$ make sense, since, e.g., $V^D(x, y)$ has the same value for all equivalent x. We shall omit the superscript D if $D = M$ (M compact).

Among all equivalence classes we concentrate on those which are invariant under F: $F([x]^D) = [x]^D$. They are called basic classes.

Since the action depends on the deviation rate ρ of the random perturbations, so does the notion of equivalence introduced above. Therefore we mention ρ in our notation wherever the results for different deviation rates are compared. Anyway, different deviation rates ρ may lead to the same basic classes. For instance, if, for all $x \in M$,

$$\rho^{(1)}(x, y) = c(x)\, \rho^{(2)}(x, y)$$

with some bounded, positive function c, then $\rho^{(1)}$- and $\rho^{(2)}$-basic classes are identical. This in particular means that introducing a position-dependent diffusion coefficient into the example of Eq. (2.3) does not change the basic classes.

Since the basic classes are decisively involved in the formulation of the conditions under which the main theorems of this section hold true, it may be helpful to add some remarks concerning their meaning and their relation to longer known concepts:

The points of basic classes are recurrent under the ρ-noisy map in the following sense: Starting from each point x of a basic class, there is a sequence of points leading back to x with arbitrarily small action. As a consequence, the Markov sequence (X_i), starting from x, stays near such a recurrent sequence with probability arbitrarily close to 1.

This notion of recurrence is intermediate between the concepts of nonwandering and of chain recurrence (see, e.g., refs. 17 and 18). If x is a nonwandering point of F (i.e., for every neighborhood U of x and $N > 0$ there is $n > N$ such that $F^n U \cap U \neq \varnothing$), then x is contained in some basic class. On the other hand, if x is member of a basic class, then x is chain recurrent [i.e., for every $\varepsilon > 0$ there is a sequence $(q_0 = x,\, q_1, \ldots,\, q_{N-1} = x)$ such that $d(q_{j+1}, F(q_j)) \leqslant \varepsilon$ for $0 \leqslant j < N-1$].

The classification of the ρ-recurrent points into basic classes has its analogies in a spectral decomposition of the nonwandering set into maximal transitive subsets (i.e., maximal subsets containing a dense orbit of F) and in the equivalence classes of chain recurrent points introduced by Ruelle[19] in his study of localized random perturbations.

In those systems where the maximal transitive subsets of a spectral decomposition of the nonwandering set coincide with Ruelle's classes, they coincide with the basic classes, too, which then do not depend on the deviation rate at all.

For a discussion of the r dependence of basic classes for the deviation rates (2.4), see ref. 9. Note that the parameter r can formally be used to mediate between nonwandering ($r \to 0$) and chain recurrence ($r \to \infty$).

The basic classes can be divided up into stable and unstable classes. A basic class $[x]^D$ is called unstable if there is a point $y \notin [x]^D$ such that $V^D(x, y) = 0$.

The stable basic classes are in an informal sense the attractors of the noisy map, but to avoid confusion with the various existing definitions of attractors, we shall not use this name here.

We are now ready to fix the assumption which allows us to find asymptotic estimates of invariant measures and mean exit times, as given by Theorem 1 and Theorem 2 below. The form of this assumption is induced by the method applied in the proofs, for which we give some details in the Appendix.

Roughly speaking, this method is to study in place of the original Markov sequence (X_n^η) a Markov chain on the set of small neighborhoods of basic classes, i.e., to pay attention only to those members of (X_n^η) which are close to ρ-recurrent points. For Markov chains on finite state spaces, easy results on invariant measures and mean exit times are available. Therefore, an obvious standard assumption of Freidlin and Wentzell[4] and of Kifer[8] is that there exists only a finite number of basic classes.

As we shall see in Section 4.2, there are interesting situations where this assumption is not fulfilled, e.g., in the limit situation of the Feigenbaum period-doubling sequence, where besides a stable basic class, which is a Cantor set, there is an infinite number of unstable basic classes, namely the unstable periodic orbits.

Therefore we state a weaker assumption which allows infinite families of unstable basic classes (see category 3 of Assumption B below), provided they satisfy three conditions:

(a) An infinite number of the unstable classes can be absorbed in "coarse-grained" versions of stable classes.

(b) The "coarse-grained" versions of the stable classes cannot be distinguished from the stable classes by the least actions from or to outside points.

(c) In refinements of the "coarse-grained" stable classes the maximal least actions between inside points decrease sufficiently fast.

Under these conditions, made precise below, all necessary estimates can be obtained from Markov chains on the finite number of "coarse-grained" basic classes. This is the basic idea carried through in the Appendix.

The precise form of the assumption is as follows.

Assumption B. In D there are no basic classes other than those which are listed in the following three categories:

1. There is a finite number λ of stable basic classes: $K_0, ..., K_{\lambda-1}$.
2. There may exist a finite number κ of unstable basic classes: $K_\lambda, ..., K_{\lambda+\kappa-1}$.
3. For a number λ_1 of the stable basic classes, say for each ν, $0 \leqslant \nu < \lambda_1$, there may be a family $\{K_\nu^{(j)}, j = 1, 2, ...\}$ of unstable basic classes with the following properties:

(a) There is a decreasing sequence of compact sets $K_\nu^{[j]}$, $K_\nu^{[1]} \supset K_\nu^{[2]} \supset ...$, such that, for every j, $F(K_\nu^{[j]}) \subset K_\nu^{[j]}$, and $K_\nu^{[j]}$ contains K_ν as well as all $K_\nu^{(i)}$ for $i > j$.

(b) If, for an arbitrary j, $x \notin K_\nu^{[j]}$, $y \in K_\nu^{[j]}$, then

$$V^D(x, y) = V^D(x, K_\nu) \qquad \text{and} \qquad V^D(y, x) = V^D(K_\nu, x)$$

(c) For each ν, $0 \leqslant \nu < \lambda_1$,

$$\lim_{j \to \infty} j \cdot \rho_{\nu j} = 0$$

where

$$\rho_{\nu j} := \max\{\min_{x \in K_\nu} \rho(x, y), \min_{x \in K_\nu} \rho(y, x) \colon y \in K_\nu^{[j]}\}$$

Now we proceed to give the strict formulations of Eqs. (1.4) and (1.5). Setting $D = M$, we start with the result concerning the invariant measure, expressed in terms of a function $\Phi(x)$ on M which we call the quasi-potential for the reasons indicated in the introduction.

The quasipotential has to be calculated from least actions along sequences beginning at stable basic classes. The general defining formula for $\Phi(x)$ involves some combinatorics, which can be handled most easily in graph-theoretic language.[4] We defer this formula to the Appendix [Eq. (A.11)], but mention here a special case: If there is only one stable basic class K_0, the formula simply reduces to

$$\Phi(x) = V(K_0, x) \tag{2.9}$$

Theorem 1. Consider a dynamical system on M (compact) perturbed by ρ-noise. Suppose that the basic classes in M satisfy Assumption B. Let $w^\eta(x)$ be the density of an invariant measure μ^η and $\Phi(x)$ the quasipotential defined by Eqs. (2.5)–(2.7) and (A.11). Then one has, for all $x \in M$,

$$\lim_{\eta \to 0} \eta \log w^\eta(x) = -\Phi(x) \tag{2.10}$$

The proof of this theorem is briefly sketched in the Appendix.

The following theorem deals with the mean time of first exit from $D \subset M$. It involves a quantity $\Delta\Phi_x^D$, which is, loosely speaking, the quasipotential depth in the region D, but may depend on the starting point $x \in D$. Again we give the general definition of $\Delta\Phi_x^D$ only in the Appendix [Eq. (A.13)]. For the special case in which there is only one stable basic class K_0 in D and $F(D) \subseteq D$ one obtains

$$\Delta\Phi_x^D = \min_{y \in \partial D} V^D(K_0, y) \tag{2.11}$$

independent of $x \in D$.

Theorem 2. Consider a dynamical system on M perturbed by ρ-noise. Suppose that the basic classes in a domain $D \subset M$ ($\bar{D}$ compact) satisfy Assumption B. Let $\langle \tau_D^\eta \rangle_x$ be the mean time of first exit from D, when $x \in D$ is the starting point, and $\Delta\Phi_x^D$ the quantity defined by Eq. (A.13). Then one has

$$\lim_{\eta \to 0} \eta \log\langle \tau_D^\eta \rangle_x = \Delta\Phi_x^D \tag{2.12}$$

For hints on the proof of this theorem, see the Appendix.

We end this section with some remarks on how to establish contact with the continuous-time systems (1.3) mentioned in the introduction, where we put G equal to unity, for simplicity.

One way to compare the results obtained here with the results for continuous time is purely formal: If we let M be a subset of $\mathbb{R}^k$, $\rho(x, y) = \rho_2(x, y) = \frac{1}{2}|y - F(x)|^2$ [cf. (2.4)], and $F(x) = x + \alpha f(x)$ with a real parameter α, then, for $\alpha \to 0$, we can recover the results for the continuous-time systems (1.3) by isolating the lowest order in α.

A second method is rather constructive: For the deterministic system (1.1) there is Poincaré's standard method to reduce to a discrete-time system (see, e.g., 17). Let F^t be the flow of the vector field $f(q)$. After the choice of an appropriate $(k-1)$-dimensional surface $\Sigma \subset \mathbb{R}^k$, this method considers the first return map $F_P: \Sigma \to \Sigma$, $x \mapsto F^{\theta(x)}(x)$, where $\theta(x)$ is the first return time of x to Σ.

The question arises, which noise has to be added to F_P when the perturbed system (1.3) is under investigation? The use of ρ_2-noise is in general a too simple choice. Rather, the results of Freidlin and Wentzell[4] suggest the following deviation rate:

$$\rho_P(x, y) = \inf_{\theta} \inf_{q(t)} \frac{1}{2} \int_0^\theta |f(q(t)) - \dot{q}(t)|^2 \, dt \tag{2.13}$$

where the second infimum has to be taken over all absolutely continuous curves $q(t)$ on M with $q(0) = x \in \Sigma$, $q(\theta) = y \in \Sigma$, and $q(t) \notin \Sigma$ for $0 < t < \theta$.

3. ESTIMATING AND CALCULATING THE LEAST ACTION

The application of the results from the previous section to any specific system requires, above all, the determination of the least action $V^D(x, y)$ between $x \in D$ and $y \in M$. In this section we collect some observations (**O1–O7**) which are useful in this task.

We begin with two simple consequences of the definitions, which hold true under the general Assumption A, but shall later subject the deviation rates and systems to further assumptions.

Since the set of all sequences from x to z via y is a subset of the set of all sequences from x to z, we obtain:

O1 For all $x, y \in D$ and $z \in M$,

$$V^D(x, z) \leqslant V^D(x, y) + V^D(y, z) \tag{3.1}$$

According to the definition (2.5) and Assumption A, the action along orbits of the map F vanishes. Thus, the following condition is sufficient to show the vanishing of the least action.

O2 If in each neighborhood of $F(x)$ $(x \in D)$ there is a point x', and in each neighborhood of $y \in M$ a point y', such that $y' = F^n(x')$ for some $n \geqslant 0$, and $F^j(x') \in D$ for $0 \leqslant j < n$, then

$$V^D(x, y) = 0$$

This in particular implies that, if F acts transitively on an invariant set $K \subset D$, all points of K are equivalent.

An immediate consequence of the first two observations is the fact that least actions do not increase along orbits of the deterministic map F. Due to (A.11) [or (2.9)], this behavior is true for the quasipotential $\Phi(x)$, too.

Henceforth we assume that the function $y \mapsto \rho(x, y)$ introduced in Assumption A increases monotonically with $d(y, F(x))$.

O3 Consider $D' \subset D$ with $F(D') \subset D'$. Then for all $x \in D'$ and $z \notin D'$ it follows that

$$V^D(x, z) \geqslant \min_{y \in \partial D'} V^D(x, y)$$

If there is, furthermore, a subset $B' \subset D$ such that $\partial D' \subset F(B')$, then

$$V^D(x, z) \geqslant \min_{y \in B'} V^D(y, z)$$

The first statement is due to the fact that for any sequence (q_i) from x to z one can find a sequence from x to the boundary of D' with smaller or equal action: Let q_k be the last point of (q_i) in D' and y' the point in $\partial D'$ nearest to $F(q_k)$. Since $\rho(q_k, y') \leqslant \rho(q_k, q_{k+1})$, the action along $(x, q_1,..., q_k, y')$ does not exceed the action along (q_i).

In an analogous way, the second statement makes use of the points following q_k for obtaining a lower bound of the action.

By a similar argument one obtains the following result.

O4 For all $x \in D$ and $z \notin F(D)$, it follows that

$$V^D(x, z) \geqslant \min_{y \in \partial(F(D))} V^D(x, y)$$

Whereas so far we have just gathered some direct implications of the definitions from Section 2, we now want to investigate analytical methods for determining least actions. To this end, we confine ourselves to systems with:

1. (M, d) a subset of Euclidean $(\mathbb{R}^k, |\cdot|)$
2. F continuously differentiable on $\bar{D}$
3. Deviation rate (cf. Assumption A) ρ_r as defined by (2.4), where $r > 1$

If we were to deal with time-continuous systems, the problem of finding the least action would be a variational problem, which would be solved by standard methods known from classical mechanics. This procedure was pursued in refs. 5 and 20–22.

In discrete-time systems the problem reduces to a minimization problem. The following observation and the correlated remarks systematize and generalize results of refs. 7, 10, and 12–16.

O5 Let $(q_i)_{0 \leqslant i < N}$ be a sequence such that $q_0 = x \in D$, $q_i \in D$ for all $i < N-1$, and $q_{N-1} = y \in M$. We call this sequence a minimal N-sequence if $S_N[(q_i)] = V_N^D(x, y)$. If for $0 < i < N-1$ the transposed maps of the derivatives of F at q_i, $DF|_{q_i}^T$, are invertible, then a minimal N-sequence (q_i) satisfies the following two-step recursion $(0 < i < N-1)$:

$$q_{i+1} = F(q_i) + \left\{ \frac{|q_i - F(q_{i-1})|}{|(DF|_{q_i}^T)^{-1} [q_i - F(q_{i-1})]|} \right\}^{(r-2)/(r-1)} \times (DF|_{q_i}^T)^{-1} [q_i - F(q_{i-1})] \tag{3.2}$$

This two-step recursion can equivalently, and more clearly, be arranged by introducing an auxiliary sequence $(p_i)_{1 \leqslant i < N}$, which obeys $(1 \leqslant i < N-1)$

$$\begin{aligned} p_1 &= |q_1 - F(q_0)|^{r-2} [q_i - F(q_0)] \\ p_{i+1} &= (DF|_{q_i}^T)^{-1} p_i \end{aligned} \tag{3.3}$$

Then (q_i) satisfies the following recursion $(0 \leqslant i < N-1)$:

$$q_{i+1} = F(q_i) + |p_{i+1}|^{-(r-2)/(r-1)}\, p_{i+1} \tag{3.4}$$

Equation (3.2) follows from the condition that the gradient of the action along sequences of length N from x to y, considered as a function on $D^{(N-2)}$, must vanish at the minimum. It can be interpreted as a discrete Lagrange equation.

Equations (3.3) and (3.4) are the corresponding discrete Hamilton equations of a $2k$-dimensional Hamiltonian system.

If a minimal N-sequence (q_i) from x to y or the related sequence (p_i) is known, the following formula for the corresponding action is useful:

$$\begin{aligned} V_N^D(x, y) &= \frac{1}{r} \sum_{i=1}^{N-1} |p_i|^{r/(r-1)} \\ &= \frac{1}{r} \sum_{i=1}^{N-1} \left| \prod_{j=1}^{i-1} (DF|_{q_j}^T)^{-1}\, p_1 \right|^{r/(r-1)} \end{aligned} \tag{3.5}$$

Here we use the notational convention that $\prod_{j=n}^{n-1} (\cdots) := 1$ for any integer n.

Interpreting minimal sequences as projections of orbits of a Hamiltonian map on the $(p=0)$-submanifold allows some useful conclusions. We here only make some brief remarks in this connection, since there is barely a difference (at least for $r=2$) from the continuous-time situation which has been described in ref. 21.

The gradient of $y \mapsto V_N^D(x, y)$ can be obtained in the Hamiltonian formulation as

$$\nabla_y V_N^D(x, y) = p_{N-1} \tag{3.6}$$

Note that the dynamics of the Hamiltonian system (3.3), (3.4), restricted to the invariant $(p=0)$-plane, retrieves the unperturbed dynamics of F. Thus, for instance, to each fixed point x of F the point $(0, x)$ in the (p, q)-space is a fixed point of the Hamiltonian system. If x is stable for F, $(0, x)$ is a saddle of the Hamiltonian system, the $(p=0)$-plane being its stable manifold. For $r<2$ the unstable manifold is tangential to the $(q=x)$-plane-plane. For $r=2$ the unstable manifold is transverse to the stable manifold with an inclination determined by $DF|_x^T$ (see Fig. 1b). For $r>2$ the Hamiltonian map bears the deficience of nondifferentiability at $p=0$. The unstable set of $(0, x)$ asymptotically approaches the stable manifold near $(0, x)$ in this case.

For the determination of $V(x, y)$, where x is the above fixed point, one has to look for orbits of the Hamiltonian map starting arbitrarily close to

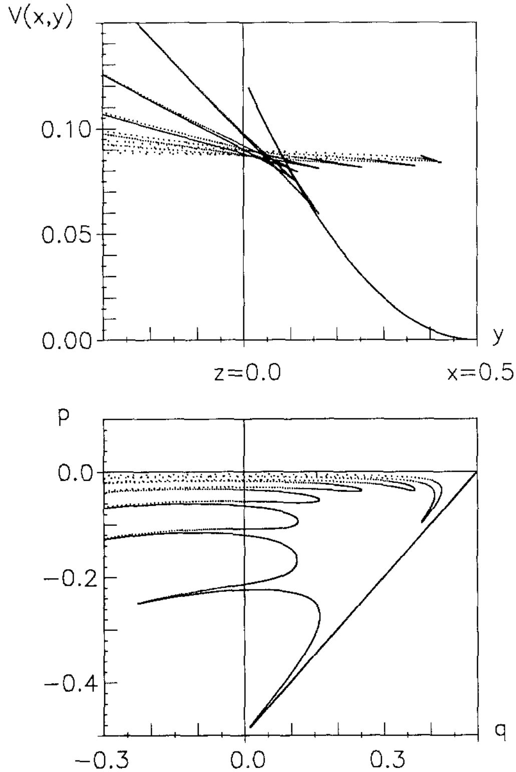

Fig. 1. (a) Least action $V(x, y)$ (for $r = 2$) obtained as the lower envelope of the actions along minimal N-sequences starting near the stable fixed point $x = 0.5$ of $F(y) = y + 0.15 \sin(2\pi y)$. The discontinuities in the slope of the least action can be traced back to the heteroclinic tangles in (b) the unstable manifold of the point $(0, x)$ in the (p, q)-plane for the Hamiltonian system defined by (3.3), (3.4), which accumulate near the unstable fixed point $z = 0$ of F.

$(0, x)$, therefore lying on the unstable set, and hitting the $(q = y)$-plane. Consider the case where, besides the stable fixed point x of F, there is an unstable fixed point z (see Fig. 1a). The least action $V(x, z)$ is obtained by a heteroclinic orbit from $(0, x)$ to $(0, z)$. If the heteroclinic points are transverse, they imply a heteroclinic tangle of the unstable set of $(0, x)$ near $(0, z)$ (see Fig. 1b). In this case there will be several intersections of the unstable set with the $(q = y)$-plane if y is close to z. The proper choice among these for evaluating the action is guided by the minimizing condition (2.7). For nearby points y_1, y_2, the intersections so chosen with the planes $q = y_1$ and $q = y_2$ need not both lie on the same lap of the unstable set; this jumping from lap to lap results in jumps in the gradient of $V(x, y)$

because of (3.6). The conclusion is that in general near unstable fixed points (and with analogous arguments near all types of unstable basic classes) discontinuities in the gradient of the quasipotential accumulate. Since the heteroclinic orbit still asymptotically approaches $(0, z)$, (3.6) implies also that the gradient of the quasipotential tends to zero near z.

For the sake of completeness we add the analogue of the Hamilton–Jacobi approach now:

O6 Define for some $x \in D$ the function $D \ni y \mapsto \phi(y) = V^D(x, y)$. On some regions of D, this function satisfies the following functional equation:

$$\begin{aligned}\phi(y) - \phi(F(y) + |(DF|_y^T)^{-1} \nabla\phi(y)|^{-(r-2)/(r-1)} (DF|_y^T)^{-1} \nabla\phi(y))\\ + \frac{1}{r} |(DF|_y^T)^{-1} \nabla\phi(y)|^{r/(r-1)} = 0 \qquad (3.7)\end{aligned}$$

Equation (3.7) can be derived in close analogy to the ordinary Hamilton–Jacobi equation by looking for an appropriate canonical transformation of the (p, q) coordinates of Eqs. (3.3), (3.4). A different derivation of Eq. (3.7) (for $r = 2$, $k = 1$) has been given in ref. 11.

While for continuous-time systems the Hamilton–Jacobi equation is a major tool for the determination of quasipotentials (see, e.g., refs. 21–25), for discrete-time systems Eq. (3.7) is of little practical use, for two reasons: From a technical point of view, the solution of a functional equation like (3.7) is by far more problematic than the solution of a partial differential equation of the Hamilton–Jacobi type. Moreover, we had to make the restriction in **O6** that Eq. (3.7) is only satisfied in certain regions of D. This is connected with the facts that, in general, there are discontinuities in $\nabla\phi$, and that—in place of the time derivative in the Hamilton–Jacobi equation—Eq. (3.7) contains a difference of the values of ϕ at possibly distant points. There is no way to determine the regions of validity for Eq. (3.7) *a priori*.

We now return to Eqs. (3.3), (3.4). Finding a minimal N-sequence between a given pair of points is a boundary value problem which in general can only be solved numerically (cf. the remarks at the end of this section). Therefore, an evaluation of (3.5) also requires numerical calculation. There is, however, an important method of estimating $V_N^D(x, y)$ if a minimal N-sequence from x to y is ε-shadowed by the orbit of x. Recall that a sequence $(q_i)_{0 \leqslant i < N}$ is said to be ε-shadowed by the orbit of x if $|q_i - F^i(x)| < \varepsilon$ for $0 \leqslant i < N$.

The estimate is much easier to obtain in the one-dimensional case. We therefore restrict now consideration to $k = 1$. In addition, we assume that F is twice continuously differentiable on $\bar{D} \subset \mathbb{R}$.

The property of ε-shadowing then implies

$$F(q_i) - F(F^i(x)) = [1 + \mathcal{O}(\varepsilon)]\, F'(q_i)[q_i - F^i(x)]$$

Writing

$$[q_i - F^i(x)] = q_i - F(q_{i-1}) + F(q_{i-1}) - F(F^{i-1}(x))$$

we can iterate this estimate to obtain

$$[q_{N-1} - F^{N-1}(x)] = [1 + \mathcal{O}(\varepsilon)] \sum_{j=1}^{N-1} \left(\prod_{i=j}^{N-2} F'(q_i) \right) [q_j - F(q_{j-1})] \tag{3.8}$$

With the aid of Eqs. (3.3) and (3.4), this leads to

$$|q_{N-1} - F^{N-1}(x)| = [1 + \mathcal{O}(\varepsilon)] \prod_{i=1}^{N-2} |F'(q_i)| \times \sum_{j=1}^{N-1} \prod_{k=1}^{j-1} |F'(q_k)|^{-r/(r-1)}\, |p_1|^{1/(r-1)} \tag{3.9}$$

As a combination of Eqs. (3.5) and (3.9), we obtain the following result, where instead of calculating the derivatives at the unknown points of the minimal sequence, we evaluate them at the easier-to-obtain nearby points of the deterministic orbit:

O7 If a minimal N-sequence from x to y is ε-shadowed by the orbit of x, the following estimate holds true:

$$V_N^D(x, y) = [1 + \mathcal{O}(\varepsilon)] \frac{1}{r} |y - F^{N-1}(x)|^r \left[\sum_{j=1}^{N-1} \prod_{k=j}^{N-2} |F'(F^k(x))|^{r/(r-1)} \right]^{1-r} \tag{3.10}$$

As a first application of **O7**, we consider the quasipotential in the vicinity of a stable periodic orbit (period m) of a one-dimensional map F. Let this stable orbit be the only stable basic class of the system. The quasipotential in a point y near the point x of the periodic orbit has to be calculated as the least action along sequences from x to y [see (2.9)]. Obviously, a minimal $(l \cdot m + 1)$-sequence (for each $l = 1, 2,...$) from x to y is $|y - x|$-shadowed by the periodic orbit (l revolutions).

We estimate $V_{lm+1}^D(x, y)$ by Eq. (3.10) and notice that the action decreases with increasing l. After some calculation, we obtain in the limit $l \to \infty$ for sufficiently small $|y - x|$

$$\Phi(y) = \frac{C_{(x)}}{r} |y - x|^r + \mathcal{O}(|y - x|^{r+1}) \tag{3.11}$$

where the following coefficient appears:

$$C_{(x)} = \left(\frac{1 - |(F^m)'(x)|^{r/(r-1)}}{\sum_{j=1}^{m} |(F^{m-j})'[F^j(x)]|^{r/(r-1)}} \right)^{r-1} \tag{3.12}$$

(Note that the $j = m$ term of the sum in the denominator equals 1.)

With very similar arguments one can show that the increase of the quasipotential toward a point z in an unstable basic class which is an unstable periodic orbit of period m (which generally shows accumulating points of nondifferentiability) can be enveloped by a curve of the form

$$\Phi(y) \approx \Phi(z) - \frac{c_{(z)}}{r} |y - z|^r \tag{3.13}$$

where

$$c_{(z)} = \left(\frac{1 - |(F^m)'(z)|^{-r/(r-1)}}{\sum_{j=1}^{m} |(F^j)'(z)|^{-r/(r-1)}} \right)^{r-1} \tag{3.14}$$

Further application of the above observations will be demonstrated in Section 4.

Although nonnumerical arguments are in the foreground of the present article, we stress that the detailed computation of quasipotentials even for simple systems can only be performed numerically. [Strictly speaking, the functions which are to be computed are the least actions $V^D(K_\nu, x)$ for all stable basic classes K_ν. If there is more than one such class, the numerical results have to be used to take the minimum according to Eq. (A.11).] We sketch the two computational methods which are suggested by the above observations.

The first way makes use of **O5**. The boundary value problem of finding a minimal N-sequence from a point of the stable class to a point outside of that stable class is by shooting method arguments turned into an initial value problem. The recursion (3.3), (3.4) is repeatedly started with various suitable initial values—with the intention to reach all relevant points in M. The action along the computed sequences is obtained by Eq. (3.5). In general one will find several minimal sequences of different lengths which lead to the same point. According to (2.7), among the different values of the action which then result from Eq. (3.5), the lowest one has to be chosen.

The most problematic detail of this method is the distribution of the initial values in order to obtain a sufficiently uniform distribution of the subsequent points. Besides, it is not a good idea to put q_0 on the stable basic class itself, because then the infimum in (2.7) is generally not attained for any sequence of finite length [as can be exemplified with the limiting procedure which was necessary to obtain (3.11)]. Instead, one should start with (p_1, q_1) close to a point of the stable class on its unstable set under the Hamiltonian map (3.3), (3.4).

The second, more general way, which was introduced by Reimann and Talkner,[12,13] is in the present context a direct application of the definitions and is related to **O1**. Based on

$$V(K_v, z) = \inf_{y \in M} [V(K_v, y) + \rho(y, z)] \tag{3.15}$$

one can, from a guess of $y \mapsto V(K_v, y)$, iteratively obtain improvements by inserting the guess in the right-hand side of (3.15) and reading off the improvement on the left-hand side. The convergence of this procedure depends on starting with a reasonable guess. In the one-dimensional case with periodic orbits, Eq. (3.11) suggests itself as initial guess.

The second method is free from the problem of a possibly inhomogeneous distribution of those points for which the quasipotential is known. On the other hand, the possibility of selected attention to special points or regions can be an advantage of the first method.

4. QUASIPOTENTIAL OF S-UNIMODAL MAPS

4.1. Maps with a Finite Number of Basic Classes

We now apply the results from the previous section to a particularly well-known class $\mathscr{C}$ of one-dimensional maps, the even, S-unimodal maps (the S indicates negative Schwarzian derivative) of the interval $I = [-a, a] \subset \mathbb{R}$. For the classical results on these maps see ref. 26. The facts which we use here are concisely summarized in Chapter 2 of ref. 27.

Recall that the members of the family of logistic maps

$$F_\mu(x) = 1 - \mu x^2 \qquad \left(a = \frac{1 + (1 + 4\mu)^{1/2}}{2\mu}\right) \tag{4.1}$$

belong to $\mathscr{C}$; varying the parameter μ, $0 < \mu \leqslant 2$, one can study examples for all the situations discussed below.

Choosing $M = D = I$, we can use Theorem 1 for asymptotic statements about the stationary probability density for small random perturbations in terms of the quasipotential. Most of the various former investigations of the influence of noise on maps of the interval (see ref. 28 for a review of the early work) start with a definition of the noisy system by a stochastic difference equation [see (2.3)]. In order to show the relevance of the present approach for those models, we give an example for a stochastic difference equation

$$X_{n+1}^{\eta} = F(X_n^{\eta}) + \xi_n^{\eta}(F(X_n^{\eta})) \tag{4.2}$$

($F \in \mathscr{C}$) that defines a Markov sequence on $M = I$ satisfying Assumption A with deviation rate ρ_r. Here a state dependence of the random variables $\xi_n^{\eta}(z)$, $z \in I$, is introduced to prevent the noise from driving the system out of $M = I$.[11,29] The unmodified choice of Gaussian random variables as in (2.3) clearly would not work. However, if one restricts the values of $\xi_n^{\eta}(z)$ to the interval $I_{(z)} := [-a-z, a-z]$, then (X_n^{η}) does not escape from I. Thus we may choose the following probability density $\psi_{(z)}^{\eta}(\xi)$ for $\xi_n^{\eta}(z)$:

$$\psi_{(z)}^{\eta}(\xi) = \begin{cases} \mathcal{N}_{(z)}^{\eta} \exp(-|\xi|^r/r\eta) & \text{for} \quad \xi \in I_{(z)} \\ 0 & \text{else} \end{cases} \tag{4.3}$$

where the factor $\mathcal{N}_{(z)}^{\eta}$ normalizes the density. With this choice, (4.2) defines a Markov sequence on I with the transition probability density

$$p^{\eta}(x, y) = \mathcal{N}_{(F(x))}^{\eta} \exp\left(-\frac{|y - F(x)|^r}{r\eta}\right)$$

Since the factor $\mathcal{N}_{(F(x))}^{\eta}$ shows just an algebraic dependence on η, Assumption A holds true with $\rho_r(x, y)$ from (2.4).

Note that a further possible application, which we do not take up here, would be to employ Theorem 2 with the choice $M = \mathbb{R}$ and $D = I$ to estimate mean exit times when the noise is allowed to throw the system out of the interval I.

Theorem 2.4 of ref. 27, which is largely due to Jonker and Rand,[30] gives a decomposition of the nonwandering set of any map in $\mathscr{C}$, which—in the present context of ρ_r-noise—simultaneously supplies a specification of the basic classes of the map. The decomposition may be finite or infinite. We defer the latter case to the next subsection and assume until then that there is only a finite number $\kappa + \lambda$ of basic classes (see Assumption B). As an implication of the assumed negative Schwarzian derivative, there is exactly one stable basic class: $\lambda = 1$. Equation (2.9) applies. The above-mentioned decomposition theorem says that, apart from degenerate special cases, there are two alternatives.

1. The stable class K_0 is a stable periodic orbit of period m. The behavior of the quasipotential near K_0 follows from Eqs. (3.11), (3.12).

2. The stable class K_0 is a disjoint union of m closed intervals Δ_j, $0 < j \leqslant m$, such that F maps Δ_j homeomorphically onto Δ_{j+1} for $j < m$

and F^m is on Δ_m conjugate to a tent map. Since F shows on K_0 a sensitive dependence on initial conditions, this set is a strange attractor.

In order to describe the behavior of the quasipotential (see Fig. 2) near the strange attractor, we arrange the endpoints x_K^j and x_K^{j+m} of the intervals Δ_j such that $F(x_K^i) = x_K^{i+1}$ for $0 < i < 2m$. Thus, for any outside point y near x_K^i there is an orbit along endpoints shadowing a minimal $(i+1)$-sequence from K_0 to y. We conclude by **O7** from Section 3 that

$$\Phi(y) \approx \frac{1}{r} \left(\sum_{j=1}^{i} \prod_{k=j}^{i-1} |F'(x_K^k)|^{r/(r-1)} \right)^{1-r} |y - x_K^i|^r \tag{4.4}$$

There are two main types of unstable basic classes as well.

1. The unstable class K_ν $(1 \leqslant \nu \leqslant \kappa)$ is an unstable periodic orbit of period m. The quasipotential decreases on both sides of the periodic points with the behavior described by Eqs. (3.13), (3.14).

2. The unstable class K_ν $(1 \leqslant \nu \leqslant \kappa)$ is a Cantor set on which F is conjugate to a transitive subshift of finite type. [Conjugacy to a subshift of finite type means that there exist a finite number, say l, of closed intervals $I_1, ..., I_l$ and a transition matrix $(a_{ij})_{1 \leqslant i,j \leqslant l}$ whose coefficients are all 0 or 1 such that for all $x \in K_\nu$ and all $k \geqslant 0$

$$F^k(x) \in I_i \Rightarrow F^{k+1}(x) \in \bigcup_{a_{ij} \neq 0} I_j$$

Transitivity is guaranteed if for each pair (i, j), $1 \leqslant i, j \leqslant l$, there is an integer power of the transition matrix whose (i, j)-coefficient is greater than 0.]

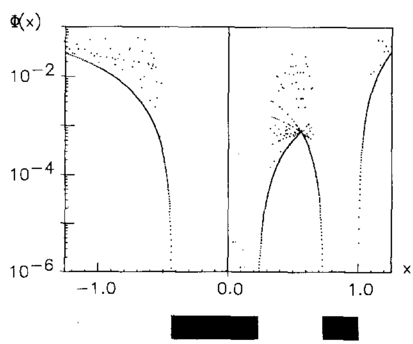

Fig. 2. Quasipotential $\Phi(x)$ for F_μ with $\mu = 1.4320...$ and $r = 2$. As in Fig. 1, $\Phi(x)$ is the lower envelope of all plotted points. The attractor is shown below the abscissa and consists of two intervals in this case.

Instead of discussing the general situation, we give the simplest example here[31] (see also ref. 32), which shows all arguments necessary to handle more complicated forms of unstable Cantor sets. We consider a map with a stable period-three orbit $\{x_1, x_2, x_3\}$ (in increasing order) and a nearby unstable period-three orbit $\{z_1, z_2, z_3\}$ (see Fig. 3a). We denote, for $i = 1, 2, 3$, by $z'_i \neq z_i$ the point nearest to z_i for which $F^3(z'_i) = z_i$ (e.g., $z'_2 = -z_2$). Consider the following decomposition of the interval $[z'_1, z'_3]$ into five subintervals:

$$J_1 = [z'_1, z_1), \qquad I_1 = [z_1, -z_2]$$
$$J_2 = (-z_2, z_2), \qquad I_2 = [z_2, z_3]$$
$$J_3 = (z_3, z'_3]$$

The unstable Cantor set is generated by repeatedly removing the preimages of J_2 from the intervals I_1 and I_2. The transition matrix of this examples is $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 1 \end{smallmatrix}\right)$.

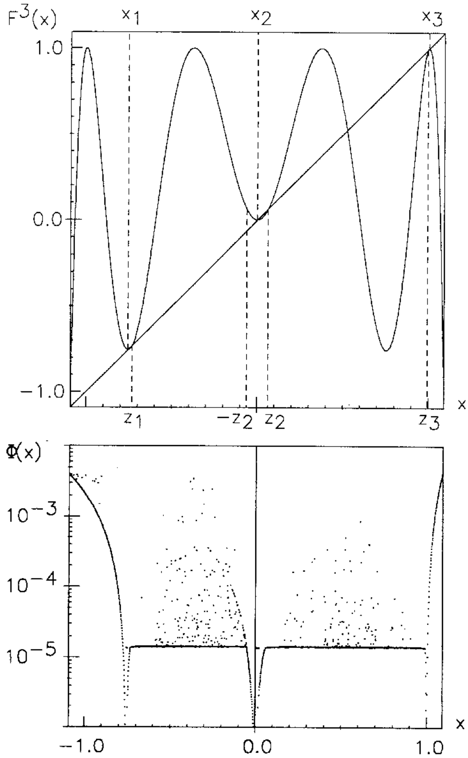

Fig. 3. (a) Graph of the threefold iteration of F_μ with $\mu = 1.7548...$. Nearby the superstable orbit (x_1, x_2, x_3) there is an unstable period-three orbit (z_1, z_2, z_3). (b) Quasipotential $\Phi(x)$ for F_μ, $r = 2$.

Note that $x_i \in J_i$ and

$$F\left(\bigcup_{i=1}^{3} J_i\right) \subset \bigcup_{i=1}^{3} J_i$$

Since $\Phi(z'_1)$ and $\Phi(z'_3)$ both are not smaller than $\Phi(z_i)$, we conclude by the first part of **O3** from Section 3 that

$$\Phi(y) \geqslant \Phi(z_i) \qquad \text{for all} \quad y \in I_1 \cup I_2$$

On the other hand, we know by the Cantor set construction that each iterated preimage of J_2 has an iterated preimage arbitrarily close to z_i. By **O1** and **O2** from Section 3 we infer that

$$\Phi(y) \leqslant \Phi(z_i) \qquad \text{for all} \quad y \in I_1 \cup J_2 \cup I_2$$

Thus we obtain

$$\Phi(y) = \Phi(z_i) \qquad \text{for all} \quad y \in I_1 \cup I_2 \tag{4.5}$$

(see Fig. 3b).

The increase of the quasipotential from x_i toward z_i can, in the vicinity of the unstable periodic orbit, be described as already discussed.

4.2. The Case of Period 2^∞

In the previous subsection we considered in $\mathscr{C}$ for which there is a finite number of basic classes. Now we study the remaining case with an infinite number of basic classes in the presence of ρ_r-noise. The only stable basic class then is a Cantor set all gaps of which contain unstable basic classes. We confine ourselves to the most prominent example, where the unstable basic classes are periodic orbits with period 2^n for arbitrary n: the map occurring as the limit of the Feigenbaum period-doubling sequence (see, e.g., refs. 26 and 33).

In the first step we describe the structure of this example and show that the unstable basic classes are of the third category of Assumption B.

To be specific, we consider a function F with quadratic maximum which satisfies the Feigenbaum–Cvitanovic equation

$$F(x) = -\alpha F(F(\alpha^{-1}x)) \tag{4.6}$$

($\alpha \approx 2.503$) and with critical value $F(0) = 1$.

We introduce the following intervals ($n \geqslant 1$):

$$\Delta_0^{(n)} := [-\alpha^{-n}, \alpha^{-n}]$$
$$\Delta_l^{(n)} := F^l(\Delta_0^{(n)}) \qquad \text{for} \quad l \leqslant 2^n$$

The length of $\Delta_0^{(n)}$, $2\alpha^{-n}$, is larger than the lengths of all intervals $\Delta_l^{(n)}$. The interval $\Delta_{2^n}^{(n)}$ is contained in $\Delta_0^{(n)}$.

Each interval of the $(n-1)$th generation can be divided into two intervals of the nth generation and, in between, a gap which is subdivided into two half gaps by a periodic point $z_l^{(n)}$:

$$\Delta_l^{(n-1)} = \Delta_l^{(n)} \cup \Delta_l^{(n)} \cup \{z_l^{(n)}\} \cup \Delta_{l+2^{n-1}}^{(n)} \cup \Delta_{l+2^{n-1}}^{(n)}$$

The periodic points in the gaps constitute unstable basic classes

$$K_0^{(n)} = \{z_l^{(n)}: 1 \leqslant l \leqslant 2^{n-1}\}$$

Let

$$K_0^{[n]} := \bigcup_{l=1}^{2^n} \Delta_l^{(n)} \tag{4.7}$$

and

$$K_0 := \bigcap_n K_0^{[n]} \tag{4.8}$$

The Cantor set K_0 is the stable basic class.

Apart from K_0 and the sets $K_0^{(n)}$, there is a further basic class: the unstable fixed point $-a$.

In order to verify Assumption B, we have to check whether the infinite number of $K_0^{(n)}$ is related to K_0 and the $K_0^{[n]}$ as described for the third category of that assumption. The property 3a is fulfilled by construction. For property 3b, we note that to each point of $K_0^{[n]}$ one can find an (iterated) preimage arbitrarily close to any point outside of $K_0^{[n]}$. This implies, by **O2** of Section 3,

$$V(x, y) = 0 \qquad \text{for all} \quad x \notin K_0^{[n]} \quad \text{and} \quad y \in K_0^{[n]}$$

On the other hand, we have $F(K_0^{[n]}) \subset K_0^{[n]}$ and $\partial K_0^{[n]} \subset F(K_0) = K_0$, from which we conclude by **O1** and the second part of **O3** that

$$V(y, x) = V(K_0, x) \qquad \text{for all} \quad x \notin K_0^{[n]} \quad \text{and} \quad y \in K_0^{[n]}$$

Finally we obtain by the definition (2.4) of ρ_r-noise

$$\rho_{0n} \leqslant \frac{2}{r} \alpha^{-(n-1)r}$$

Hence, the property 3c,

$$\lim_{n \to \infty} n \cdot \rho_{0n} = 0$$

follows for our example.

Now that we know Assumption B to hold true, we investigate the scaling behavior of the quasipotential in the gaps

$$\Lambda_l^{(n)} \cup \{z_l^{(n)}\} \cup \Lambda_{l+2^{n-1}}^{(n)}$$

$(0 < l \leqslant 2^{n-1})$ of nth generation. Since the maximal values of the quasipotential in the gaps at the points $z_l^{(n)} \in K_0^{(n)}$ do only depend on n—we introduce the notation $\Phi^{(n)}$ for these values—it is sufficient to study the behavior of the quasipotential in the half gaps $\Lambda^{(n)} := \Lambda_{2^n}^{(n)}$, which are closest in the nth generation to the critical point.

Consider a minimal N-sequence (q_i) from the critical point $q_0 = 0$ to $q_{N-1} = x \in \Lambda^{(n)}$ for which $N = m2^n + 1$ and which runs in m revolutions through all the half gaps $\Lambda_l^{(n)}$, obviously ε-shadowed by the critical orbit with $\varepsilon = 2\alpha^{-(n-1)}$.

By **O7** from Section 3 we infer that

$$V_N(0, x) = [1 + \mathcal{O}(\varepsilon)] \frac{m}{r} |x - F^{2^n}(0)|^r \left[\sum_{j=1}^{2^n} \prod_{k=j}^{2^n - 1} |F'(F^k(0))|^{r/(r-1)} \right]^{1-r} \tag{4.9}$$

The last factor of (4.9) has a well-known scaling behavior, which in ref. 33 was connected to the free energy $\beta\mathcal{F}(\beta)$ of the set K_0. For the use of the thermodynamic formalism in connection with the scaling behavior of Cantor sets we refer to ref. 34.

As can be shown by the same procedure as in the proof of Theorem 4.2 in ref. 33 (note the different sign convention introduced there in the definition of the free energy), the scaling of the factor in (4.9) is

$$\begin{aligned} &\lim_{n \to \infty} \frac{1}{n} \log \left[\sum_{j=1}^{2^n} \prod_{k=j}^{2^n - 1} |F'(F^k(0))|^{r/(r-1)} \right] \\ &\qquad = \frac{r}{1-r} \left[-\mathcal{F}\left(\frac{r}{1-r}\right) + \log \alpha \right] \end{aligned} \tag{4.10}$$

By length scaling near the critical point, on the other hand, we obtain

$$\lim_{n \to \infty} \frac{1}{n} \log |z_{2^{n-1}}^{(n)} - F^{2^n}(0)| = -\log \alpha \tag{4.11}$$

The result of Eqs. (4.9)–(4.11) is the scaling behavior of the quasipotential maxima in the nth gap of the Cantor set:

$$\lim_{n \to \infty} \frac{1}{n} \log \Phi_r^{(n)} = -r\, \mathcal{F}\left(\frac{r}{1-r}\right) \tag{4.12}$$

where we explicitly marked the dependence on r.

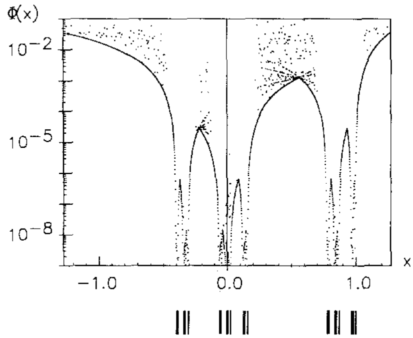

Fig. 4. Same as Fig. 2, for $\mu_\infty = 1.4011...$, the limit of the period-doubling sequence.

Equation (4.12) can be exploited to obtain—for sufficiently high generations—the constant ratio of the quasipotential maxima in gaps of successive generations.

The result for Gaussian noise, i.e., $r=2$ (see Fig. 4), has been announced in ref. 16:

$$\frac{\Phi_2^{(n-1)}}{\Phi_2^{(n)}} \sim \exp[2\mathscr{F}(-2)] \approx (6.619)^2 \tag{4.13}$$

This result involves the noise scaling constant $\kappa \approx 6.619$, which was introduced in refs. 35, 36.

As a further application we discuss a procedure which has been used in refs. 28 and 37 to study noisy maps on the interval by computer experiments. In these experiments one iterates for a given noise strength a noisy map many times to obtain the stationary distribution, i.e., the density of the invariant measure. One then marks the regions of the interval where the stationary distribution exceeds a fixed threshold. This is repeated with varying noise strength. The noise strength is given in the experiments by the standard deviation σ of the noise. This means that in the case of ρ_r-noise we have the following relation to our noise strength η:

$$\sigma \sim \eta^{1/r} \tag{4.14}$$

The boundaries of the marked regions move along lines $\sigma_r(x)$ as σ varies; these lines are the most characteristic features in the pictures obtained in the mentioned computer experiments.

If the quasipotential $\Phi_r(x)$ is known, one can obtain $\sigma_r(x)$ as follows: For low noise, Theorem 1 states that the invariant density exceeds a given threshold if $\Phi_r(x)$ is smaller than a certain constant times η. The relation between the boundary σ values and the quasipotential is, due to (4.14),

$$\sigma_r(x) \sim [\Phi_r(x)]^{1/r} \tag{4.15}$$

A first consequence is that, for all r, $\sigma_r(x)$ increases linearly near stable basic classes, since the quasipotential increases with power r.

For the period-2^∞ stable class we obtain from (4.12) and (4.15):

$$\frac{\sigma_r^{(n-1)}}{\sigma_r^{(n)}} \sim \exp\left[\mathscr{F}\left(\frac{r}{1-r}\right)\right] \tag{4.16}$$

In ref. 37, a computer experiment is reported where the lines $\sigma(x)$ are shown for equidistributed noise, localized on an interval of length proportional to σ (Fig. 2 of ref. 37). Though the theory of quasipotentials was originally not designed to treat noise not satisfying Assumption A, we can draw a conclusion on the scaling behavior in this case, too, by taking the limit $r \to \infty$. Equation (4.16) becomes

$$\frac{\sigma_\infty^{(n-1)}}{\sigma_\infty^{(n)}} \sim \exp[\mathscr{F}(-1)] \approx 8.490 \tag{4.17}$$

Here we benefitted from a recent calculation of $\mathscr{F}(-1)$ by Kovács.[38]

This is in excellent agreement with the value 8.477 resulting from the numerical data in ref. 37.

5. CONCLUSIONS

Theorems 1 and 2, formulated in Section 2, give a rigorous justification not only for various existing explicit and implicit applications of nonequilibrium-potential or quasipotential methods to noisy maps. The only condition to be fulfilled by the underlying deterministic system, Assumption A, is weak enough to admit systems with such complicated dynamical features as fractal basin boundaries, fractal repellers, strange attractors, or Cantor attractors with nonopen basins of attraction. Our one-dimensional applications in Section 4 revealed some typical consequences of such complicated structures for the quasipotential:

Fractal repellers bring about regions of constant quasipotential. In these regions, the stationary distribution does not depend exponentially on the noise strength.

Fractal attractors imply a certain scaling behavior of the quasipotential in the attractor gaps. The scaling exponents characterizing this behavior describe by which factor noise has to be decreased to be able to resolve the next finer gap. An explicit example for such an application was our treatment of the period-2^∞ attractor in Section 4.2. The quasipotential approach makes direct contact with the observable stationary distribution. This is advantageous as compared to the traditional approach[35,36] and has led to a new insight: Though the universal scaling factor $\kappa \approx 6.619$ describes how to renormalize the noise strength independent of the form of the noise distribution, κ cannot be detected independent of the noise distribution by the observation of stationary distributions. Rather, as we have shown, the scaling of the stationary distribution is described by κ only for (truncated) Gaussian noise, whereas, e.g., localized equidistributed noise leads to a universal scaling factor 8.490.

The general theory exposed in Sections 2 and 3 and in the Appendix should be applicable to higher-dimensional systems, too. Among the aspects which cannot be studied in our example of S-unimodal maps but can occur in higher dimensions are strange attractors with fractal structure or systems with several coexisting attractors. The treatment of the latter problem can profit by the graph-theoretic formulation sketched in Appendix A.

It makes sense to divide the observations from Section 3 into two groups: **O1–O4** with topological tenor and **O5–O7** with analytical content. Applications of the former group to higher-dimensional systems can be carried out as obvious generalizations of our one-dimensional applications. We briefly sketch an example: Consider the map $z \mapsto F_a(z) = z - P_a(z)/P'_a(z)$ on the complex plane, where $P_a(z) = z^3 + (a-1)\,z - a$ with complex parameter a is a cubic polynomial.[39,40] F_a describes Newton's algorithm for finding zeros of P_a. In the presence of ρ_r-noise, the zeros of P_a are stable basic classes. For $a = 1$, the three third roots of unity are the only stable classes. A further basic class, the Julia set, which forms the boundary of the basins of attraction to the zeros, is unstable. This situation is analogous to the example concerning the second type of unstable basic class ("fractal repeller") in Section 4.1. In that example the basin of attraction consisted of the gap intervals cut out of a Cantor set; now the basins of attraction consist of disjoint connected components of more complicated shape. By exactly the same combination of arguments as in the one-dimensional example, we can infer from **O1**, **O2**, and the first part of **O3** that the quasipotential must be constant not only on the Julia set, but even on all connected components of the attractor basins which do not contain the zeros. Only in the three immediate basins of attraction do quasipotential wells with minima in the zeros exist. For $a = 1$, the three wells have equal depth. For other values of a, the depth of quasipotential wells may be different (and there may be more than three stable basic classes, namely stable periodic orbits), supplying a criterion of stability of the stable classes against ρ_r-noise.

The analytical methods inspired by the second group of observations are much harder to transfer to higher dimensions. For two-dimensional hyperbolic attractors, however, an adaptation of **O7** holds true (see ref. 16 for the consequences). A treatment of the quasipotential for more general types of two-dimensional systems remains for future work.

APPENDIX. ON THE PROOF OF THEOREMS 1 AND 2

Theorems 1 and 2 are formed after the model of Theorems 4.3 and 5.3 of Chapter 6 in ref. 4. But there are two new aspects here: We deal with maps instead of flows and we loosen the assumption about the structure of the basic classes by allowing the category 3 in Assumption B.

The fundamental modifications that are necessary for maps in place of flows were presented in ref. 8. A proof of Theorem 1, restricted to dynamical systems with a finite number of basic classes, can be found there.

Rather than giving the full proof of the theorems, we outline the main steps and mention technical details only in those places where alterations in the arguments of refs. 3, 4, and 8 are essential.

The importance of the action (2.5) derives from the following estimate (which is Theorem I.5.2 of ref. 8) for the probability of nearly realizing a given sequence: Given an arbitrary $\beta > 0$ and $\tilde{N} > 0$, one has for sufficiently small δ and η

$$\begin{aligned} \exp\left[-\frac{(S_N[(q_i)]+\beta)}{\eta}\right] \leqslant P_x^\eta\{\max_{0\leqslant j\leqslant N-1} d(X_j^\eta, q_j) < \delta\} \\ \leqslant \exp\left[-\frac{(S_N[(q_i)]-\beta)}{\eta}\right] \end{aligned} \tag{A.1}$$

for each sequence $(q_i)_{0\leqslant i\leqslant N-1}$, $N \leqslant \tilde{N}$, starting in x. Thus, paths with small action are especially probable.

On the other hand, one can show that it is improbable that the random sequence deviates much from the most probable paths (Corollary I.5.2. of ref. 8): We define with $D \subset M$, $\bar{D}$ compact,

$$\mathscr{A}_x(s) = \{(q_0,\ldots, q_{N-1}) \in D^{N-1} \times F(D): q_0 = x; S_N[(q_i)] \leqslant s\}$$

Then for any $\tilde{N}, \delta, \beta > 0$ there is $\tilde{\eta} > 0$ such that for all $\eta \leqslant \tilde{\eta}$, $N \leqslant \tilde{N}$, and $s \geqslant 0$,

$$P_x^\eta\{\inf_{(q_i)\in\mathscr{A}_x(s)} \max_{0\leqslant j\leqslant N-1} d(X_j^\eta, q_j) \geqslant \delta\} \leqslant \exp\left(-\frac{s-\beta}{\eta}\right) \tag{A.2}$$

The estimates (A.1) and (A.2) suggest for small η the replacement of the probability of a transition from x to a neighborhood of y in $N-1$ steps by the probability of the most probable path: $e^{-V_N(x,y)/\eta}$. This replacement is in the spirit of Laplacian or saddle point approximations of integrals, and in the physical literature it is usually justified by such an approximation in the Chapman–Kolmogorov equation for the transition probability.

However, the heuristic arguments of the physical literature that lead to an asymptotic expression for the stationary probability in terms of least actions involve a risky limit $N \to \infty$ after the above replacement. This is the point where rigor is lost in those arguments. Precise conditions under which such an asymptotic expression is valid require a more subtle reasoning.

The method of Freidlin and Wentzell takes advantage of the lucid results for invariant measures and mean exit times of finite Markov chains. The original Markov sequence on M related to the given perturbed dynamical system is boiled down to a Markov sequence on a finite union of well-chosen subsets of M. The choice of subsets will be described later, but we mention that the basic classes play an important role in this connection.

We now cite the lemmas of ref. 4, Chapter 6, Paragraph 3, which exploit the results for finite Markov chains. We start with some notational definitions.

Let L be an alphabet with letters α, β, etc. Let l be the number of letters of the alphabet. A graph (i.e., a set of arrows between letters) consisting of $l-1$ arrows is called a β-graph if from every letter in $L\backslash\{\beta\}$ there is exactly one path of arrows leading to β. The set of all β-graphs is denoted by $G(\beta)$. The β-graphs do not contain cycles of arrows. A graph without cycles and consisting of $l-2$ arrows starting in $L\backslash\{\beta\}$ which does not contain any path leading from $\alpha \in L\backslash\{\beta\}$ to β is called an $(\alpha \not\to \beta)$-graph [the set of all such graphs is $G(\alpha \not\to \beta)$].

We now consider a Markov sequence on a space U which is a finite disjoint union of l sets U_β ($\beta \in L$). If there are numbers $p_{\alpha\beta}$ ($\alpha, \beta \in L$, $\alpha \neq \beta$) and a constant $a > 1$ such that the transition probabilities $\tilde{P}(x, U_\beta)$ ($x \in U_\alpha$, $\alpha \neq \beta$) of the sequence satisfy the inequalities

$$a^{-1} p_{\alpha\beta} \leqslant \tilde{P}(x, U_\beta) \leqslant a p_{\alpha\beta} \tag{A.3}$$

then, according to ref. 4, Lemmas 3.2 and 3.4, the following estimates hold true, where, for a graph g, $\pi(g)$ denotes the product $\prod_{(\gamma \to \delta) \in g} p_{\gamma\delta}$:

For an invariant probability measure $\tilde{\mu}$ of the Markov sequence one obtains

$$a^{-2(l-1)} \frac{\sum_{g \in G(\alpha)} \pi(g)}{\sum_{\beta \in L} \sum_{g \in G(\beta)} \pi(g)} \leqslant \tilde{\mu}(U_\alpha) \leqslant a^{2(l-1)} \frac{\sum_{g \in G(\alpha)} \pi(g)}{\sum_{\beta \in L} \sum_{g \in G(\beta)} \pi(g)} \tag{A.4}$$

For the expectation value $\langle\tilde{\theta}_{U_\beta}\rangle_x$ of the number of steps till the first entrance into U_β for sequences starting in $x \in U_\alpha$ ($\alpha \neq \beta$), one has

$$a^{-4^{l-1}} \frac{\sum_{g \in G(\alpha \not\to \beta)} \pi(g)}{\sum_{g \in G(\beta)} \pi(g)} \leqslant \langle\tilde{\theta}_{U_\beta}\rangle_x \leqslant a^{4^{l-1}} \frac{\sum_{g \in G(\alpha \not\to \beta)} \pi(g)}{\sum_{g \in G(\beta)} \pi(g)} \tag{A.5}$$

In the case $l = 2$ the numerator has to be taken equal to 1.

In the next step, we define, related to the sequence (X_n^η) in M, to a set $U \subseteq M$ with $M \backslash D \subseteq U$, and to an integer k, a sequence $(\tilde{X}_i^{(\eta, D, U, k)})$ in U for which we shall later prove an estimate of the type (A.3).

The new sequence $(\tilde{X}_i)$ results from the old one by just keeping each kth member that hits $U \cap D$, as long as the sequence does not leave D:

$$\begin{aligned} \tilde{X}_i &= X^\eta_{\theta_{U \cap D, i \cdot k}} \qquad \text{if} \quad \theta_{U \cap D, i \cdot k} < \theta_{M \backslash D, 1} \\ \tilde{X}_i &= X^\eta_{\theta_{M \backslash D, 1}} \qquad \text{if} \quad \theta_{U \cap D, i \cdot k} \geqslant \theta_{M \backslash D, 1} \end{aligned} \tag{A.6}$$

where we have introduced the notation (for $W \subset M$)

$$\begin{aligned} &\theta_{W, 0} = 0 \\ &\theta_{W, j} = \min\{i > \theta_{W, j-1} : X_i^\eta \in W\}, \qquad j > 0 \end{aligned}$$

The following lemma guarantees the relation (A.3) for the transition probabilities $\tilde{P}(x, U_\beta)$ of $(\tilde{X}_i)$, provided a proper choice of U is made.

Lemma. Consider a dynamical system on M perturbed by ρ-noise. Suppose that the basic classes K_ρ, $0 \leqslant \rho < \lambda + \kappa$, and $K_\nu^{(j)}$, $0 \leqslant \nu < \lambda_1$, $j = 1, 2, \ldots$, in a domain $D \subset M$ ($\bar{D}$ compact) satisfy Assumption B.

For each $\chi > 0$ there exist $s > 0$ and integers J and k such that the following is true:

For each letter α of the alphabet

$$L = \{(\nu j) : 0 \leqslant \nu < \lambda_1,\ 0 \leqslant j \leqslant J\} \cup \{\nu : \lambda_1 \leqslant \nu < \lambda + \kappa\}$$

define a compact set K_α according to the following rules:

$$\begin{aligned} &K_{(\nu 0)} = K_\nu^{[J]} \qquad 0 \leqslant \nu < \lambda_1 \\ &K_{(\nu j)} = K_\nu^{(j)} \qquad 1 \leqslant j \leqslant J, \quad 0 \leqslant \nu < \lambda_1 \\ &\qquad K_\nu \quad \text{as already defined for} \quad \lambda_1 \leqslant \nu < \lambda + \kappa \end{aligned}$$

Choose for all $\alpha \in L$ pairwise disjoint neighborhoods $U_\alpha \subset B_s(K_\alpha)(\subset D)$ of these compact sets. [Here $B_s(K_\alpha)$ denotes the set $\{x \in M : d(x, K_\alpha) < s\}$.]

Let $U = \bigcup_{\alpha \in L} U_\alpha \cup (M \backslash D)$. Consider the sequence $(\tilde{X}_i^{(\eta, D, U, k)})$ defined by (A.6). For some integer k and sufficiently small $\eta > 0$, the transition probabilities satisfy

$$\exp\left(-\frac{V^D_{\alpha\beta}+\chi}{\eta}\right) \leqslant \tilde{P}(x, U_\beta) \leqslant \exp\left(-\frac{V^D_{\alpha\beta}-\chi}{\eta}\right) \tag{A.7}$$

and

$$\exp\left(-\frac{V^D_{\alpha C}+\chi}{\eta}\right) \leqslant \tilde{P}(x, M\backslash D) \leqslant \exp\left(-\frac{V^D_{\alpha C}-\chi}{\eta}\right) \tag{A.8}$$

for $x \in U_\alpha$. Here we introduced the following abbreviations:

$$\begin{aligned} V^D_{\alpha\beta} &= V^D(K_\alpha, K_\beta) \\ V^D_{\alpha C} &= \inf_{y \in M\backslash D} V^D(K_\alpha, y) \end{aligned} \tag{A.9}$$

Note that by **O3** and **O4** from Section 3, for monotone deviation rates, the last quantity can be written as

$$V^D_{\alpha C} = \min_{y \in \partial D \cup \overline{F(D)}\backslash D} V^D(K_\alpha, y)$$

The statement is similar to Lemma I.5.4 of ref. 8. We give here the proof of the estimate (A.7) [(A.8) can be proved analogously] in order to exemplify the treatment of the new features: the restriction to a domain D and the new basic classes of the third category in Assumption B.

Proof. (a) Upper estimate in (A.7): This estimate does not need a specific choice of J and k. Consider the realizations of (X^η_n) for which $X^\eta_0 = z \in U_{\alpha'}$ and $\theta_{U,1} = \tilde{\imath}$, $X^\eta_{\tilde{\imath}} \in U_{\beta'}$, α', $\beta' \in L$, $\alpha' \neq \beta'$. We have $\overline{d(X^\eta_{\tilde{\imath}}, M\backslash B_{s+\delta}(K_{\beta'}))} \geqslant \delta$ for each $\delta > 0$. Now choose δ and s such that $\overline{B_\delta(K_{\beta'})} \subset U_{\beta'}$ and

$$\sup\{|\rho(u, v) - \rho(u', v')| : d(u, u'), d(v, v') \leqslant s + \delta\} < \frac{\chi}{3k}$$

u, v, u', $v' \in D$. By the latter condition, we ensure that for each sequence $(q_i)_{0 \leqslant i < N}$ starting in $q_0 = z \in U_{\alpha'}$, ending in $q_{N-1} \in B_{s+\delta}(K_{\beta'})$, and not leaving D, there is the following lower bound for the action: $S_N[(q_i)] > V^D_{\alpha'\beta'} - 2\chi/3k$. Here we have used the definitions (2.5)–(2.8) and the condition 3b of Assumption B. We conclude from the above inequalities and (A.2) that

$$\begin{aligned} &P^\eta_z\{\theta_{U,1} = \tilde{\imath} \leqslant N; X^\eta_{\tilde{\imath}} \in U_{\beta'}\} \\ &\quad\leqslant P^\eta_z\{\inf_{(q_i) \in \mathscr{A}_z(V^D_{\alpha'\beta'} - 2\chi/3k)} \max_{0 \leqslant j < N} d(X^\eta_j, q_j) \geqslant \delta\} \leqslant \exp\left(-\frac{V^D_{\alpha'\beta'} - 5\chi/6k}{\eta}\right) \end{aligned}$$

for sufficiently small η.

On the other hand, one can show [Lemma I.5.3(b) of ref. 8] that

$$P_z^\eta\{\theta_{U,1} > N\} \leqslant \exp(-V^D_{\alpha'\beta'}/\eta)$$

for sufficiently large N and sufficiently small η, since $M\backslash U$ does not contain any complete orbit of the unperturbed system starting in D.

Summing up both contributions, we obtain

$$\begin{aligned} P_z^\eta\{X^\eta_{\theta_{U,1}} \in U_{\beta'}\} &\leqslant \left[\exp\left(-\frac{V^D_{\alpha'\beta'}}{\eta}\right)\right]\left(1+\exp\frac{5\chi/6k}{\eta}\right) \\ &\leqslant \exp\left(-\frac{V^D_{\alpha'\beta'}-\chi/\kappa}{\eta}\right) \end{aligned}$$

for all $z \in U_{\alpha'}$, provided that η is sufficiently small.

This directly proves (with $\alpha' = \alpha$ and $\beta' = \beta$) the right inequality of (A.7) for $k = 1$. For arbitrary k, one obtains the upper estimate of (A.7) by inserting $k-1$ intermediate steps, i.e., by $(k-1)$ times applying the Chapman–Kolmogorov formula. The probability $\tilde{P}(x, U_\beta)$ is then expressed as $k-1$ integrals over products of probabilities of the type $P_z^\eta\{X^\eta_{\theta_{U,1}} \in U_{\beta'}\}$ for which an upper bound can trivially be obtained using the above estimate.

(b) Lower estimate in (A.7): Here we have to choose J and k depending on χ. The choice of k will be described below. Take J large such as to fulfill $\rho_{vJ} < \chi/6l$ for $0 \leqslant v < \lambda_1$, where $l = \kappa + \lambda + J\lambda_1$ is the number of letters in the alphabet L; this is possible according to the condition 3c of Assumption B.

Choose s and δ so small that $\delta < \chi/6l$, $B_{2\delta}(K_\alpha) \subset U_\alpha$ for all $\alpha \in L$, and

$$\sup\{|\rho(u,v)-\rho(u',v')| : d(u,u'), d(v,v') \leqslant s+\delta\} < \frac{\chi}{6l}$$

We are going to construct a sequence $(Q_i)_{0 \leqslant i < N}$ with the following properties:

1. $Q_0 = x \in U_\alpha$, $Q_{N-1} \in K_\beta$
2. Exactly $k+1$ members of (Q_i) are contained in $\bigcup_{\alpha \in L} B_\delta(K_\alpha)$
3. All the other members of (Q_i) are not in $\bigcup_{\alpha \in L} B_{s+\delta}(K_\alpha)$, nor in $M\backslash D$
4. $S_N[(Q_i)] < V^D_{\alpha\beta} + \chi$

If we succeed in this construction, the estimate is proved, since the above properties imply that from $X_0^\eta = x$ and $\max_{0 \leqslant j < N} d(X_j^\eta, Q_j) < \delta$ it follows that $\tilde{X}_1 \in U_\beta$ for the random sequence defined by (A.6). Therefore we have, according to (A.1),

$$\tilde{P}(x, U_\beta) \geqslant P_x^\eta\{\max_{0 \leqslant j < N} d(X_j^\eta, Q_j) < \delta\} > \exp\left(-\frac{V_{\alpha\beta}^D + \chi}{\eta}\right)$$

The construction of the sequence (Q_i) proceeds as follows: Let $\tilde{x}$ be the point in K_α which is nearest to $x \in U_\alpha$. By the choice of $s+\delta$, we have $\rho(x, F(\tilde{x})) < \chi/6l$. By definition and by condition 3b of Assumption B, there is a sequence $(q_i)_{0 \leqslant i < n}$ in D such that $q_0 = F(\tilde{x}) \in K_\alpha$, $q_{n-1} \in K_\beta$, and $S_n[(q_i)] < V_{\alpha\beta}^D + \chi/6$. Let q_{k_1} be the last point of this sequence which is in $B_{s+\delta}(K_\alpha)$. Let $q_{j_2} \in B_{s+\delta}(K_{\gamma_2})$ be the first point in $\bigcup_{\gamma \in L} B_{s+\delta}(K_\gamma)$ of the subsequence $(q_i)_{k_1 < i < n}$, and q_{k_2} be the last point of this subsequence in $B_{s+\delta}(K_{\gamma_2})$. Unless $\gamma_2 = \beta$, iterate this procedure [now with $(q_i)_{k_2 < i < n}$], defining γ_3, j_3, and k_3, and further until $B_{s+\delta}(K_\beta)$ is reached for $\gamma_m = \beta$, $m \leqslant l$.

We know that

$$\sum_{\iota=1}^{m-1} S_{j_{\iota+1}-k_\iota+1}[(q_{k_\iota},..., q_{j_{\iota+1}})] < V_{\alpha\beta}^D + \frac{\chi}{6}$$

Because of the choice of $s+\delta$, we know even after replacing q_{k_ι} by the nearest point in K_{γ_ι}, $\tilde{q}_{k_\iota}$, and $q_{j_{\iota+1}}$ by the nearest point in $K_{\gamma_{\iota+1}}$, $\tilde{q}_{j_{\iota+1}}$, for all $0 < \iota < m$, that

$$\sum_{\iota=1}^{m-1} S_{j_{\iota+1}-k_\iota+1}[(\tilde{q}_{k_\iota}, q_{k_\iota+1},..., \tilde{q}_{j_{\iota+1}})] < V_{\alpha\beta}^D + \frac{\chi}{2}$$

Now take for each ι, $0 < \iota < m$, a sequence $(p_j^{(\iota)})_{0 \leqslant j < n^{(\iota)}}$ in $B_\delta(K_{\gamma_\iota})$ which connects $\tilde{q}_{j_\iota}$ and $\tilde{q}_{k_\iota}$ ($\tilde{q}_{j_1} := q_0$) such that $S_{n^{(\iota)}}[(p_j^{(\iota)})] < \chi/2l$, where the length $n^{(\iota)}$ is bounded by some integer $\tilde{n}$ that depends on χ. This is possible if K_{γ_ι} is a basic class because of Lemma I.5.2 of ref. 8 (which is a simple consequence of the definitions and the compactness of basic classes) and remains valid if K_{γ_ι} is a set $K_\nu^{[J]}$ due to the choice of J.

Now we specify k to be $k = (l-1)\tilde{n} + 1$. The concatenation of the above segments leads to a sequence

$$(x, \tilde{q}_{j_1}, p_1^{(1)},..., p_{n^{(1)}-2}^{(1)}, \tilde{q}_{k_1}, q_{k_1+1},..., \tilde{q}_{j_2}, p_1^{(2)},..., \tilde{q}_{j_m})$$

which has all the properties required for the sequence (Q_i), except that the number of its points in $\bigcup_{\alpha \in L} B_\delta(K_\alpha)$ is $\tilde{k} + 1 = 1 + \sum_{\iota=1}^{m-1} n^{(\iota)} + 1$, which may be smaller than $k+1$. In that case, we define (Q_i) to be the above sequence, followed by the $k - \tilde{k}$ points $F^s(\tilde{q}_{j_m})$, $s \leqslant k - \tilde{k}$, which are all contained in K_β. ∎

It is possible to modify the proof in such a way that the lemma remains valid if we add a letter z to the alphabet L and define $K_z = \{z\}$, where z is an arbitrary point in $D \backslash \bigcup_{\alpha \in L} K_\alpha$.

Setting $D = M$, we can derive from the lemma and (A.4) estimates for the invariant measure $\tilde{\mu}$, and then for the invariant measure μ^η, since by Proposition I.5.3 of ref. 8,

$$\tilde{\mu}(U_\alpha) = \frac{\mu^\eta(U_\alpha)}{\mu^\eta(U)}$$

Thus, for each $\chi > 0$ and all $x \in M$, one has, for all sufficiently small ε and η,

$$\mu^\eta(B_\varepsilon(x)) \lesseqgtr \exp\left[-\frac{1}{\eta} \left\{ \min_{v \in L_s} \left[\min_{g \in G_s(v)} \sum_{(\sigma \to \rho) \in g} V(K_\sigma, K_\rho) + V(K_v, x) \right] \right.\right.$$
$$\left.\left. - \min_{v \in L_s} \left[\min_{g \in G_s(v)} \sum_{(\sigma \to \rho) \in g} V(K_\sigma, K_\rho) \right] \mp \chi \right\} \right] \tag{A.10}$$

Here we denote by L_s the alphabet $\{0, 1,..., \lambda - 1\}$, which enumerates the stable basic classes, and by G_s the corresponding graphs. Concerning the reasons which allow us to confine ourselves to the stable classes, we refer to Paragraph 4 in Chapter 6 of ref. 4.

The relation (A.10) proves Theorem 1, and we can now give the general formula for the quasipotential:

$$\Phi(x) = \min_{v \in L_s} \left[\min_{g \in G_s(v)} \underset{\langle \sigma \to \rho \rangle \in g}{J} V(K_\sigma, K_\rho) + V(K_v, x) \right]$$
$$- \min_{v \in L_s} \left[\min_{g \in G_s(v)} \sum_{(\sigma \to \rho) \in g} V(K_\sigma, K_\rho) \right] \tag{A.11}$$

Theorem 2 can be proved similarly to Theorem 5.3 in Chapter 6 of ref. 4. The connection between the mean exit time of (X_n^η) and the sequence $(\tilde{X}_n)$ described by the above lemma is given by the following equation:

$$\langle \tau_D^\eta \rangle_x = \sum_{j=0}^{\infty} \int_{\{\omega' \in \Omega : \tilde{X}_j(\omega') \in D\}} \left[\int_\Omega \Theta_{\tilde{X}_j(\omega')}(\omega)\, dP^\eta_{\tilde{X}_j(\omega')}(\omega) \right] dP_x^\eta(\omega') \tag{A.12}$$

Here $\Theta_y(\omega)$, for a sequence realization which is characterized by ω and starts in $y \in U \cap D$, denotes the index of the next member of the sequence kept in the tilded sequence according to the rule (A.6), namely $\min\{\theta_{U \cap D,k}, \theta_{M \backslash D,1}\}$. Its expectation value [which is the inner integral in (A.12)] is not smaller than 1, but [owing to Lemma I.5.3(b) of ref. 8] does not exceed $\exp(\chi'/\eta)$ for any arbitrary $\chi'>0$, provided η is sufficiently small, and therefore does not influence the further estimates. From Eq. (A.12), just $\sum_{j=0}^{\infty} P_x^\eta\{\tilde{X}_j \in D\}$ remains to be estimated. This can be done using (A.5), since the above sum of probabilities is equal to the mean number of steps, $\langle\tilde{\theta}_{M\setminus D}\rangle_x$, till the sequence $(\tilde{X}_j)$ enters $M\setminus D$. We give the result: For every $\chi>0$ and sufficiently small η, $\langle\tau_D^\eta\rangle_x$ is between

$$\exp\left[\frac{1}{\eta}\left\{\min_{g\in G_c(C)}\sum_{(\sigma\to\rho)\in g} V_{\sigma\rho}^D - \min_{v\in L_c}\left[V^D(x,K_v)+\min_{g\in G_c(v\nrightarrow C)}\sum_{(\sigma\to\rho)\in g} V_{\sigma\rho}^D\right]\mp\chi\right\}\right]$$

provided the exponent is positive. Here the alphabet L_s is enlarged by the letter C [see (A.9)]: $L_c=L_s\cup\{C\}$. (Concerning the reasons which allow us to confine ourselves to the stable classes, we refer to Paragraph 5 in Chapter 6 of ref. 4.)

This proves Theorem 2, setting

$$\Delta\Phi_x^D=\min_{g\in G_c(C)}\sum_{(\sigma\to\rho)\in g} V_{\sigma\rho}^D-\min_{v\in L_c}\left[V^D(x,K_v)+\min_{g\in G_c(v\nrightarrow C)}\sum_{(\sigma\to\rho)\in g} V_{\sigma\rho}^D\right] \tag{A.13}$$

or $\Delta\Phi_x^D=0$, if Eq. (A.13) gives a negative value.

Acknowledgments

We would like to thank T. Tél for fruitful discussions and Z. Kovás for imparting to us the exact value of $\mathscr{F}(-1)$ for the Feigenbaum attractor. We gratefully acknowledge support by the Deutsche Forschungsgemeinschaft (Sonderforschungsbereich 237).

REFERENCES

1. L. D. Landau and E. M. Lifshitz, *Statistical Physics* (Pergamon, Oxford, 1969).
2. R. Graham, Macroscopic potentials, bifurcations and noise in dissipative systems, in *Noise in Nonlinear Dynamical Systems*, Vol. 1, F. Moss and P. V. E. McClintock, eds. (Cambridge University Press, Cambridge, 1989).
3. A. D. Wentzell and M. I. Freidlin, On small random perturbations of dynamical systems, *Usp. Math. Nauk* **25**:1, 3 (1970) [*Russ. Math. Surv.* **25**:1, 1 (1970)].
4. M. I. Freidlin and A. D. Wentzell, *Random Perturbations of Dynamical Systems* (Springer, New York, 1984).
5. R. Graham and T. Tél, Nonequilibrium potential for coexisting attractors, *Phys. Rev. A* **33**:1322 (1986).

6. R. L. Kautz, Thermally induced escape: The principle of minimum available noise energy, *Phys. Rev. A* **38**:2066 (1988).
7. P. Grassberger, Noise-induced escape from attractors, *J. Phys. A* **22**:3283 (1989).
8. Yu. Kifer, *Random Perturbations of Dynamical Systems* (Birkhäuser, Boston, 1988).
9. Yu. Kifer, Attractors via random perturbations, *Commun. Math. Phys.* **121**:445 (1989).
10. M. L. Blank, Deterministic properties of stochastically perturbed dynamic systems, *Theory Prob. Appl.* **33**:612 (1988).
11. P. Talkner and P. Hänggi, Discrete dynamics perturbed by weak noise, in *Noise in Nonlinear Dynamical Systems,* Vol. 2, F. Moss and P. V. E. McClintock, eds. (Cambridge University Press, Cambridge, 1989).
12. P. Reimann, Stationäre Wahrscheinlichkeitsverteilungen für diskrete dynamische Systeme mit schwachem Rauschen, Diplomarbeit, Basel (1989), unpublished.
13. P. Reimann and P. Talkner, Probability densities for discrete dynamical systems with weak noise, *Helv. Phys. Acta* **63**:845 (1990); and to be published.
14. R. L. Kautz, Global stability of the chaotic state near an interior crisis, in *Structure, Coherence and Chaos in Dynamical Systems,* P. L. Christiansen and R. D. Parmentier, eds. (Manchester University Press, Manchester, 1989).
15. P. D. Beale, Noise-induced escape from attractors in one-dimensional maps, *Phys. Rev. A* **40**:3998 (1989).
16. R. Graham, A. Hamm, and T. Tél, Non-equilibrium potentials for dynamical systems with fractal attractors or repellers, *Phys. Rev. Lett.* **66**:3089 (1991).
17. D. Ruelle, *Elements of Differentiable Dynamics and Bifurcation Theory* (Academic Press, San Diego, 1989).
18. J. Guckenheimer and P. J. Holmes, *Nonlinear Oscillations, Dynamical Systems, and Bifurcation of Vector Fields* (Springer, New York, 1983).
19. D. Ruelle, Small random perturbations of dynamical systems and the definition of attractors, *Commun. Math. Phys.* **82**:137 (1981).
20. R. Graham and T. Tél, On the weak-noise limit of Fokker–Planck models, *J. Stat. Phys.* **35**:729 (1984).
21. R. Graham and T. Tél, Weak-noise limit of Fokker–Planck models and nondifferentiable potentials for dissipative dynamical systems, *Phys. Rev. A* **31**:1109 (1985).
22. H. R. Jauslin, Melnikov's criterion for nondifferentiable weak-noise potentials, *J. Stat. Phys.* **42**:573 (1986).
23. R. Kubo, K. Matsuo, and K. Kitahara, Fluctuation and relaxation of macrovariables, *J. Stat. Phys.* **9**:51 (1973).
24. H. Lemarchand and G. Nicolis, Stochastic analysis of symmetry-breaking bifurcations: Master equation approach, *J. Stat. Phys.* **37**:609 (1984).
25. G. Hu and H. Haken, Polynomial expansion of the potential of Fokker–Planck equations with a noninvertible diffusion matrix, *Phys. Rev. A* **40**:5966 (1989).
26. P. Collet and J.-P. Eckmann, *Iterated Maps on the Interval As Dynamical Systems* (Birkhäuser, Boston, 1980).
27. P. Holmes and D. Whitley, Bifurcations of one- and two-dimensional maps, *Phil. Trans. R. Soc. Lond. A* **311**:43 (1984).
28. J. P. Crutchfield, J. D. Farmer, and B. A. Huberman, Fluctuations and simple chaotic dynamics, *Phys. Rep.* **92**:46 (1982).
29. H. Haken and G. Mayer-Kress, Chapman–Kolmogorov equation and path integrals for discrete chaos in presence of noise, *Z. Phys. B* **43**:185 (1981).
30. L. Jonker and D. Rand, Bifurcations in one dimension I, *Invent. Math.* **62**:347 (1981).
31. J. Guckenheimer, G. Oster, and A. Ipaktchi, The dynamics of density dependent population models, *J. Math. Biol.* **4**:101 (1977).

32. R. L. Devaney, *An Introduction to Chaotic Dynamical Systems* (Addison-Wesley, Redwood City, 1987).
33. E. B. Vul, Ya. G. Sinai, and K. M. Khanin, Feigenbaum universality and the thermodynamic formalism, *Usp. Math. Nauk* **39**:3, 3 (1984) [*Russ. Math. Surv.* **39**:3, 1 (1984)].
34. T. Bohr and T. Tél, The thermodynamics of fractals, in *Directions in Chaos*, Vol. 2, B.-L. Hao, ed. (World Scientific, Singapore, 1988).
35. J. Crutchfield, M. Nauenberg, and J. Rudnick, Scaling for external noise at the onset of chaos, *Phys. Rev. Lett.* **46**:933 (1981).
36. B. Shraiman, C. E. Wayne, and P. C. Martin, Scaling theory for noisy period-doubling transitions to chaos, *Phys. Rev. Lett.* **46**:935 (1981).
37. G. Mayer-Kress and H. Haken, The influence of noise on the logistic model, *J. Stat. Phys.* **26**:149 (1981).
38. Z. Kovács, Universal $f(\alpha)$ spectrum as an eigenvalue, *J. Phys. A* **22**:5161 (1989); and private communication.
39. J. H. Curry, L. Garnett, and D. Sullivan, On the iteration of a rational function: Computer experiments with Newton's method, *Commun. Math. Phys.* **91**:267 (1983).
40. H.-O. Peitgen, D. Saupe, and F. v. Haeseler, Cayley's problem and Julia sets, *Math. Intell.* **6**:2, 11 (1984).