

Attractors via Random Perturbations*

Yuri Kifer

Institute of Mathematics, Hebrew University, Jerusalem, Israel, and Department of Mathematics, Cornell University, Ithaca, NY 14850, USA

Abstract. We discuss conditions which ensure that weak limits of invariant measures of small random perturbations of dynamical systems have their supports on attractors.

1. Introduction

Let $F: M \rightarrow M$ be a continuous map of a metric space M . Suppose that $\{Q_x^\varepsilon, x \in M, \varepsilon > 0\}$ is a family of probability distributions on M such that for any $\delta > 0$ and each bounded continuous function g ,

$$\limsup_{\varepsilon \rightarrow 0} \sup_{x \in M} |\int g(y) Q_x^\varepsilon(dy) - g(x)| = 0. \quad (1.1)$$

Then the Markov chains X_n^ε with transition probabilities $p^\varepsilon(x, \Gamma) = P\{X_{n+1}^\varepsilon \in \Gamma | X_n^\varepsilon = x\} = Q_{F(x)}^\varepsilon(\Gamma)$ are called random perturbations of iterates of the transformation F . One may be interested in the asymptotic behavior as $\varepsilon \rightarrow 0$ of invariant measures μ^ε of X_n^ε , i.e. the measures satisfying

$$\mu^\varepsilon(\cdot) = \int d\mu^\varepsilon(x) p^\varepsilon(x, \cdot). \quad (1.2)$$

Under (1.1) all weak limits μ of measures satisfying (1.2) are known to be invariant measures of F (see, for instance [K1], Sect. 1.1), i.e. $\mu(F^{-1}\Gamma) = \mu(\Gamma)$ for any Borel $\Gamma \subset M$.

In physical applications one may think on measures μ obtained as limits of μ^ε as more stable to random perturbations and so having more physical sense than other invariant measures. Usually physically relevant systems are considered near stable invariant sets, i.e. attractors and so their statistical behavior is thought to be described by invariant measures which sit on attractors. Thus it is important to specify conditions on random perturbations which ensure that weak limits of their invariant measures will sit on attractors. This will be the main issue of this

* Talk presented at Durham Symposium on Dynamical Systems, July 1988. Partially supported by the U.S. Army Research Office through the Mathematical Sciences Institute of Cornell University

note. At the end we will compare different order relations corresponding to dynamical systems which play important roles here. Remark that there is another important and usually more difficult question addressed in [K1] about distributions of limiting measures inside invariant sets.

2. General Results

A finite sequence of points x_1, \dots, x_n is called a δ -pseudo-orbit if

$$\text{dist}(Fx_i, x_{i+1}) < \delta \quad \text{for all } i = 1, \dots, n - 1. \tag{2.1}$$

One introduces a preorder “ \succ ” writing $y \succ x$ if for any $\delta > 0$ there is a δ -pseudo-orbit starting at x and ending at y . We also have an equivalence relation: $x \sim y$ if $x \succ y$ and $y \succ x$. By the definition $x \sim x$. An equivalence class containing x will be denoted by $[x]$, and if either $Fx = x$ or $[x]$ contains more than one point $[x]$ is called a basic equivalence class. For brevity by an equivalence class we will always mean a basic equivalence class. Next, one obtains the partial order among equivalence classes writing $[y] \succ [x]$ if $y \succ x$. Maximal in this partial order equivalence classes will be called quasi-attractors.

Ruelle [R] considered random perturbations satisfying the localization condition: $\text{supp } Q_x^\varepsilon \subset U_\varepsilon(x)$ and the “core” condition: $\text{supp } Q_{Fx}^\varepsilon \supset FU_\alpha(x)$ for some $\varepsilon > \alpha > 0$, where $U_\delta(y)$ is the δ -ball around y . He showed that for ε small enough all invariant measures μ^ε of corresponding Markov chains X_n^ε have their support in neighborhoods of quasi-attractors. This is a simple consequence of the fact that each time the process X_n^ε comes close to a nonmaximal equivalence class some “mass” will go away along a pseudo-orbit and will never come back to a neighborhood of this equivalence class, and so eventually all “mass” will be lost from there.

If perturbations are not localized the process may return to any neighborhood of any equivalence class infinitely many times and so the above argument does not work. Moreover as an example from Sect. 1.5 of [K1] shows invariant measures μ^ε may be concentrated mainly in neighborhoods of a nonmaximal equivalence class. This is because, in general, the process X_n^ε may be forced to spend much of the time in a neighborhood of any invariant set by making perturbation there very “slow.” So we shall need additional assumptions both on dynamics and on perturbations.

Assumption 2.1. M is a compact space and there exists only finite number of equivalence classes K_1, \dots, K_ν .

In these circumstances equivalence classes K_i are compact F -invariant sets, and if K_i is a quasi-attractor then K_i is a usual attractor in the sense that there exists an open set $V \supset K_i$ such that $FV \subset V$ and $\bigcap_{n \geq 0} F^n V = K_i$ (see Sect. 1.4 of [K1]).

Fix small disjoint neighborhoods V_1, \dots, V_ν of K_1, \dots, K_ν and suppose that K_{i_1}, \dots, K_{i_k} are attractors and $K_{i_{k+1}}, \dots, K_{i_\nu}$ are nonattractors. Put $\delta_0 = \inf \{ \delta > 0 : \text{there exists a finite } \delta\text{-pseudo-orbit starting inside some } V_{i_j}, j = 1, \dots, \text{ and ending outside } V_{i_j} \}$. If neighborhoods V_{i_j} are chosen small enough then, clearly, $\delta_0 > 0$. Remark that (1.1) implies

$$\sup_{x \in M} Q_x^\varepsilon(M \setminus U_\delta(x)) = \beta_\delta(\varepsilon) \rightarrow 0 \quad \text{as } \varepsilon \rightarrow 0 \tag{2.2}$$

for any δ (see Theorem I.1.2. of [K1]).

Let $\tau_V = \min \{n: X_n^\varepsilon \notin V\}$ be the exit time of X_n^ε from $V \subset M$. Then by (2.2) and the definition of δ_0 for $j = i, \dots, k$ one has

$$\begin{aligned} \sup_{x \in V_i} P_x^\varepsilon[\tau_{V_j} < N] &\leq \sup_{x \in V_i} P_x^\varepsilon\{\text{dist}(FX_i^\varepsilon, X_{i+1}) \geq \delta_0 \text{ for some } i < N, i \geq 0\} \\ &\leq N\beta_{\delta_0}(\varepsilon), \end{aligned} \tag{2.3}$$

where $P_x^\varepsilon\{\cdot\}$ is the probability for the process X_n^ε starting at x .

Denote $U_\delta(K_i) = \{y: \text{dist}(y, K_i) < \delta\}$. In the same way as in Lemmas III.2.1 and III.2.2 in [K1] (using moduli of continuity of F in place of the Lipschitz constant there) one concludes that there exist numbers $\delta_1, T_0 > 0$ such that $\overline{U_{\delta_1}(K_i)} \subset V_i$, $i = 1, \dots, v$ and (i) if there exists a δ_1 -pseudo-orbit x_1, \dots, x_n satisfying $x_1 \in U_{\delta_1}(K_i)$, $x_n \in U_{\delta_1}(K_j)$, and $x_k \notin V_i$ for some $k > 1, k \leq n$ then $i \neq j$ and $K_j > K_i$; (ii) any δ_1 -pseudo-orbit x_1, \dots, x_n with $n \geq T_0$ has at least one point in $\bigcup_{i \leq v} U_{\delta_1}(K_i)$.

Remark that by (2.3) except for small probability the exist times from neighborhoods of attractors have at least the order of $\beta_{\delta_0}^{-1}(\varepsilon)$. To conclude that limiting measures will sit on attractors one needs exit times from neighborhood of nonattractors to be of lesser order and for time of this order we would like paths of X_n^ε to go along δ_1 -pseudo-orbits except for small probability which will enable us to use dynamics of pseudo-orbits described by (i) and (ii).

In general, we do not have a priori upper bounds on exit times from neighborhoods of nonattractors which is connected with estimating the time needed for ε -pseudo-orbits to exit from neighborhoods of invariant sets which are nonattractors. So we will need another assumption which we will show later to be satisfied in the main interesting cases.

Assumption 2.2. Neighborhoods V_{k+1}, \dots, V_{i_v} of nonattractors K_{k+1}, \dots, K_{i_v} can be chosen so that

$$\sup_{x \in V_{i_r}} P_x^\varepsilon\{\tau_{V_{i_r}} > \ell(\varepsilon)\} \leq \zeta(\varepsilon) \rightarrow 0 \quad \text{as } \varepsilon \rightarrow 0 \tag{2.4}$$

for some function $\ell(\varepsilon)$ satisfying

$$\ell(\varepsilon)\beta_{\delta_2}(\varepsilon) \rightarrow 0 \quad \text{as } \varepsilon \rightarrow 0, \tag{2.5}$$

where $\delta_2 = \min(\delta_0, \delta_1)$.

Theorem 2.1. *Under Assumptions 2.1 and 2.2 all limit points as $\varepsilon \rightarrow 0$ of invariant measures μ^ε of X_n^ε have support in the union of attractors.*

Proof. Choose a positive integer valued function $n(\varepsilon)$ so that

$$\ell(\varepsilon)/n(\varepsilon) \rightarrow 0 \quad \text{and} \quad n(\varepsilon)\beta_{\delta_2}(\varepsilon) \rightarrow 0 \quad \text{as } \varepsilon \rightarrow 0. \tag{2.6}$$

We will show that for any $x \in M$,

$$P^\varepsilon\left(n(\varepsilon), x, \bigcup_{1 \leq j \leq k} V_{i_j}\right) \rightarrow 1 \quad \text{as } \varepsilon \rightarrow 0, \tag{2.7}$$

where $P^\varepsilon(n, x, \Gamma) = P_x^\varepsilon\{X_n^\varepsilon \in \Gamma\}$ is the n -step transition probability of X_n^ε starting at x . Hence iterating (1.2),

$$\mu^\varepsilon\left(\bigcup_{1 \leq j \leq k} V_{i_j}\right) = \int d\mu^\varepsilon(x) P^\varepsilon\left(n(\varepsilon), x, \bigcup_{1 \leq j \leq k} V_{i_j}\right) \rightarrow 1 \tag{2.8}$$

as $\varepsilon \rightarrow 0$.

Thus all weak limits of μ^ε as $\varepsilon \rightarrow 0$ have support in $\overline{\bigcup_{1 \leq j \leq k} V_{i_j}}$, and since they are F -invariant they must have support in $\bigcup_{1 \leq j \leq k} K_{i_j}$, implying the theorem.

To derive (2.7) introduce the event $\mathcal{E}_n = \{\text{dist}(FX_i^\varepsilon, X_{i+1}^\varepsilon) < \delta_2 \text{ for all } i = 0, \dots, n-1\}$. For any Borel set Γ we can write

$$P^\varepsilon(n(\varepsilon), x, \Gamma) = P_x^\varepsilon\{\mathcal{E}_{n(\varepsilon)} \text{ and } X_{n(\varepsilon)}^\varepsilon \in \Gamma\} + r(\varepsilon, x, \Gamma), \tag{2.9}$$

where by (2.2) and (2.6),

$$r(\varepsilon, x, \Gamma) \leq n(\varepsilon)\beta_{\delta_2}(\varepsilon) \rightarrow 0. \tag{2.10}$$

By (i), (ii), and the definition of δ_2 the event $\left\{\mathcal{E}_{n(\varepsilon)} \text{ and } X_{n(\varepsilon)}^\varepsilon \notin \bigcup_{1 \leq j \leq k} V_{i_j}\right\}$ may only occur if for $0 \leq m \leq n(\varepsilon)$ the process X_m^ε stays away from $\bigcup_{0 \leq j \leq k} V_{i_j}$, it never revisits $U_{\delta_1}(K_i)$ after exiting V_i , and the times between subsequent visits to $U_{\delta_1}(K_i)$, $i = 1, \dots, v$ do not exceed T_0 . Thus in this case the process X_m^ε must spend time at least $(n(\varepsilon) - (v + 1)T_0)v^{-1}$ in one of neighborhoods V_j with $j > k$. By (2.4) and (2.6) the probability of such event does not exceed $v\zeta(\varepsilon)$ provided ε is small enough. This yields (2.7). The formal argument uses the strong Markov property of the process

X_m^ε in the following way. Define inductively $\sigma_0 = 0$, $\eta_m = \min\left\{\ell \geq \sigma_{m-1} : X_\ell^\varepsilon \in \bigcup_{1 \leq i \leq v} U_{\delta_1}(K_i)\right\}$, and $\sigma_m = \min\left\{\ell \geq \eta_m : x_\ell^\varepsilon \notin \bigcup_{1 \leq i \leq v} V_{i_j}\right\}$, $m = 1, 2, \dots$. Then according to the explanation above

$$\left\{\mathcal{E}_{n(\varepsilon)} \text{ and } X_{n(\varepsilon)}^\varepsilon \notin \bigcup_{1 \leq j \leq k} V_{i_j}\right\} \subset \left\{\sigma_m - \eta_m \geq (n(\varepsilon) - (v + 1)T_0)v^{-1} \text{ for some } m \leq v \text{ with } X_{\eta_m}^\varepsilon \in \bigcup_{k < j \leq v} U_{\delta_1}(K_{i_j})\right\}.$$

By (2.4) and the strong Markov property of X_n^ε the probability of the last event is bounded for small ε by

$$\sum_{m=1}^v E_x \tilde{\chi} P_{X_{\eta_m}^\varepsilon}^\varepsilon\{\sigma_m - \eta_m \geq (n(\varepsilon) - (v + 1)T_0)v^{-1}\} \leq v\zeta(\varepsilon),$$

where $\tilde{\chi} = 1$ if $X_{\eta_m}^\varepsilon \in \bigcup_{k < j \leq v} U_{\delta_1}(K_{i_j})$ and $= 0$ for otherwise, and E_x denotes the expectation for the process X_n^ε starting at x .

Thus by (2.9) and (2.10),

$$P^\varepsilon\left(n(\varepsilon), x, \bigcup_{1 \leq j \leq k} V_{i_j}\right) \geq 1 - n(\varepsilon)\beta_{\delta_2}(\varepsilon) - v\zeta(\varepsilon),$$

yielding (2.7) and completing the proof of Theorem 2.1. \square

3. Verifying Assumptions

Assumption 2.2 is a kind of global condition on both an invariant set and random perturbations at the same time. It is difficult to check such an assumption and in this section we shall derive sufficient conditions for Assumption 2.2 to be valid assuming only some properties of dynamics of F and the one-step perturbations $Q_x^\varepsilon(\cdot)$ which are much more transparent.

We shall start with the following smooth model. Let $F: M \rightarrow M$ be a C^1 map of a C^1 compact Riemannian manifold. For each equivalence class K_i defined in the previous section put

$$\mathcal{P}_\delta(K_i) = \limsup_{N \rightarrow \infty} \frac{1}{N} \log \text{vol} \left\{ x: \max_{0 \leq n \leq N} \text{dist}(F^n x, K_i) \leq \delta \right\}, \quad (3.1)$$

and $\mathcal{P}(K_i) = \lim_{\delta \rightarrow 0} \mathcal{P}_\delta(K_i)$. Suppose

Assumption 3.1. For all nonattractors K_{i_j} , $j = k + 1, \dots, v$ one has $\mathcal{P}(K_{i_j}) < 0$.

We will need also a kind of a “core” condition

Assumption 3.2. There exist constants $c, \gamma > 0$, $\gamma < \frac{1}{2}(\|DF\| + 1)^{-1}$ such that $Q_x^\varepsilon(U_{\gamma\varepsilon}(y)) \geq c$ for all $x, y \in M$ with $\text{dist}(x, y) < \varepsilon/2$, where $\|DF\|$ is the norm of the differential of F .

In fact, we shall need the last assumption for x and y in neighborhoods of nonattractors. It will be needed to ensure bounds on exit times from these neighborhoods. Remark that Assumption 3.1 holds true for Axiom A diffeomorphism and Assumption 3.2 is satisfied for all random perturbations considered in Sect. 2.7 of [K1], i.e. when measures $Q_x^\varepsilon(\cdot)$ have densities $q_x^\varepsilon(y)$ with respect to the Riemannian volume which behave locally as $\varepsilon^{-d} r_x(1/\varepsilon \exp_x^{-1} y)$, where $d = \dim M$ and $\{r_x, x \in M\}$ is a family of nonnegative functions on tangent spaces $T_x M$, $x \in M$ such that $\int r_x(\xi) d\xi = 1$ and r_x is positive and continuous at $0 \in T_x M$.

Theorem 3.1. *Under Assumptions 3.1 and 3.2 one can choose a number $\kappa > 0$ and neighborhoods V_i of nonattractors K_{i_j} , $j = k + 1, \dots, v$ so that (2.4) holds true for any $\ell(\varepsilon) \geq \varepsilon^{-\kappa}$.*

Proof. Let $j > k$ then there exists $\delta_3 > 0$ such that $\mathcal{P}_{\delta_3}(K_{i_j}) < -b$ for some $b > 0$. Then for any $x \in U_{\delta_3}(K_{i_j})$ there exists $y_x \in U_{\gamma\varepsilon}(x)$ such that $F^n y_x \notin U_{\delta_3}(K_{i_j})$ for some $n \leq 2d/b \log 1/\gamma\varepsilon$ provided ε is small enough. Indeed, if $F^n y_x \in U_{\delta_3}(K_{i_j})$ for all $y \in U_{\gamma\varepsilon}(x)$ and $n = 0, 1, \dots, N$ then $U_{\gamma\varepsilon} \subset \left\{ z: \max_{0 \leq n \leq N} \text{dist}(F^n z, K_{i_j}) \leq \delta_3 \right\}$, and so the volume of the last set must have at least order of $(\gamma\varepsilon)^d$, where $d = \dim M$. But this volume does not exceed $e^{-(b/2)n}$ provided N is large enough.

Put $V_{i_j} = U_{(1/2)\delta_3}(K_{i_j})$ and $m(\varepsilon) = 1 + \text{integral part of } (2d/b \log(1/\gamma\varepsilon))$.

Then by Assumption 3.2 for any $x \in V_{i_j}$ and ε small enough

$$P_x^{\varepsilon} \{ \tau_{V_{i_j}} \leq m(\varepsilon) \} \geq P_x^{\varepsilon} \{ X_i^t \in U_{\gamma\varepsilon}(F^t y_x) \text{ for all } i = 0, \dots, m(\varepsilon) \} \geq c^{m(\varepsilon)}. \quad (3.2)$$

Then by the Markov property for ε small enough

$$P_x^\varepsilon \{ \tau_{V_{i,j}} > nm(\varepsilon) \} = E_x^\varepsilon \chi_{\tau_{V_{i,j}} > m(\varepsilon)} E_{X_{m(\varepsilon)}^\varepsilon} \chi_{\tau_{V_{i,j}} > m(\varepsilon)} \dots$$

$$E_{X_{(n-1)m(\varepsilon)}^\varepsilon} \chi_{\tau_{V_{i,j}} > m(\varepsilon)} \leq (1 - c^{m(\varepsilon)})^n \leq (1 - \varepsilon^R)^n, \tag{3.3}$$

where $R = (3d/b) \log(1/c)$ and χ_A is the indicator of an event A . Thus if $\ell(\varepsilon)$ in (2.6) is bigger than $\varepsilon^{-\kappa}$ with $\kappa > R$ then Assumption 2.2 will be satisfied. \square

Corollary 3.1. *If Assumptions 2.1, 3.1 and 3.2 are satisfied and $\varepsilon^{-\kappa} \beta_\delta(\varepsilon) \rightarrow 0$ as $\varepsilon \rightarrow 0$ for each $\delta > 0$ then all limit points as $\varepsilon \rightarrow 0$ of invariant measures μ^ε of X_n^ε have support in the union of attractors.*

In Sect. 2.7 of [K1] I have shown that for basic hyperbolic nonattractors (2.4) holds true for any $\ell(\varepsilon)$ which tends to infinity faster than $\log(1/\varepsilon)$ as $\varepsilon \rightarrow 0$, and since we need here only the upper bound (II.7.14) from [K1] this is true in the full generality of the set up of Chapter II in [K1] without additional assumptions of Sect. 2.7 there which implies the conclusion of Theorem 2.1 for Chapter II of [K1] type of random perturbations of Axiom $-A$ dynamical systems. Moreover (2.4) can be established for some nonmaximal equivalence classes K_i even when $\mathcal{P}(K_i) = 0$. For instance, I can do this for K_i being a degenerate fixed point or closed orbit of F , i.e. when there are no eigenvalues of DF^ℓ , $\ell > 0$ whose absolute value are bigger than one but, on the other hand, not all eigenvalues have absolute value less than one. In this case one can get (2.4) with $\ell(\varepsilon)$ of order of negative power of ε .

Next, we will consider another model of random perturbations which was treated in Sect. 1.5 of [K1] and in [K2]. Let again M be a metric space with a continuous map $F: M \rightarrow M$. Suppose that for any open set $U \subset M$ uniformly in $x \in M$ transition probabilities $P^\varepsilon(x, \cdot)$ of Markov chains X_n^ε satisfy

$$\lim_{\varepsilon \rightarrow 0} \varepsilon \log P^\varepsilon(x, U) = - \inf_{y \in U} \rho(x, y), \tag{3.4}$$

where $\rho(x, y) \geq 0$ is a continuous function on $M \times M$. Everything works without any map F at all (see [K2]) and then one can consider this model as a random perturbation of the multi-valued map $x \rightarrow \{y: \rho(x, y) = 0\}$ but we will consider here only random perturbations of F in which case $P^\varepsilon(x, \cdot) = Q_{Fx}^\varepsilon(\cdot)$ and $\rho(x, y) = 0$ if and only if $y = Fx$. This model emerges when, for instance, the measures $Q_x^\varepsilon(\cdot)$ have densities $q_x^\varepsilon(y)$ whose main term behave like $e^{-\rho(x,y)/\varepsilon}$.

For any finite sequence of points $\xi = (\xi_1, \dots, \xi_N)$ define

$$A_N^{(\rho)}(\xi) = \sum_{1 \leq i \leq N-1} \rho(\xi_i, \xi_{i+1}) \tag{3.5}$$

for $N > 1$ and $A_1 \equiv 0$. For any pair of points $x, y \in M$ denote

$$B_\rho(x, y) = \inf \{ A_n^{(\rho)}(\xi) : n \geq 1, \xi = (\xi_1, \dots, \xi_n), \xi_1 = x, \xi_n = y \}. \tag{3.6}$$

The function B_ρ induces a preorder writing $y \overset{\rho}{\succ} x$ provided $B_\rho(x, y) = 0$. This yields an equivalence relation saying $x \overset{\rho}{\sim} y$ if $x \overset{\rho}{\succ} y$ and $y \overset{\rho}{\succ} x$. Equivalence classes corresponding to $\overset{\rho}{\sim}$ are closed sets which are called ρ -equivalence classes and again we will speak only on basic ones, i.e. which either a fixed point of F or

contain more than one point. A ρ -equivalence class containing x will be denoted by $[x]_\rho$. We obtain a partial order among ρ -equivalence classes saying $[y]_\rho \succ [x]_\rho$ if $y \succ x$. Suppose

Assumption 3.3.

M is compact and there exists only finite number of ρ -equivalence classes K_1, \dots, K_v .

In general, ρ -equivalence classes are smaller than the equivalence classes we obtained via pseudo-orbits and we will compare corresponding relations in the next section. It is not difficult to conclude (see [K1]) that under Assumption 3.3 the number of equivalence classes defined in the previous section will not exceed v and all maximal ρ -equivalence classes are attractors (see [K1]).

Lemma 3.1. *Under Assumption 3.3 $[x] = [x]_\rho$ for any $x \in M$, i.e. ρ -equivalence classes and equivalence classes from the previous section coincide.*

Proof. Let $[x]_\rho = K_i$ and $y \sim x$. If $y \notin \bigcup_{1 \leq i \leq v} K_i$, then by Lemma I.5.3. from [K1] only bounded pieces of the orbit $\{F^n y, n \in \mathbb{Z}\}$ may stay away from fixed neighborhoods of compacts K_i $i = 1, \dots, v$. One easily concludes from here that $\text{dist}(F^n y, K_i) \rightarrow 0$ as $n \rightarrow \pm \infty$ which implies that y belongs to the ρ -equivalence class K_i . If $y \in K_j$ and $y \sim x$ then in the same way as in Corollary III.2.1 if [K1] we derive that there exist points z_+, z_- such that $\text{dist}(F^{\pm n} z_\pm, K_i) + \text{dist}(F^{\mp n} z_\pm, K_j) \rightarrow 0$ as $n \rightarrow \infty$. This yields that $K_i = K_j$ concluding the proof. \square

There exist $r_0 > 0$ such that if K_1 is a nonattractor then for any $\delta > 0$ and $x \in U_{r_0}(K_i)$ there exists a sequence $\xi = (\xi_1, \dots, \xi_n)$ satisfying $\xi_1 = x$, $\xi_n \notin U_{r_0}(K_i)$, and $A_n^{(\rho)}(\xi) < \delta$. Using the lower bound of Theorem 1.5.2 from [K1] for the probability that X_n^ε stays in a small tube near ξ we derive that for any $\delta > 0$ there is N_δ such that

$$P_x^\varepsilon \{ \tau_{V_i} \leq N_\delta \} > e^{-\delta/\varepsilon} \tag{3.7}$$

for an open $V_i \supset K_i$, $V_i \subset U_{\frac{1}{2}r_0}(K_i)$ provided ε is small enough. Thus the Markov property argument as in (3.3) yields

$$P_x^\varepsilon \{ \tau_{V_i} > e^{2\delta/\varepsilon} \} \rightarrow 0 \text{ as } \varepsilon \rightarrow 0. \tag{3.8}$$

On the other hand, (3.4) implies that $\beta_{\delta_2}(\varepsilon) \leq e^{-\lambda/\varepsilon}$, where $\lambda = \frac{1}{2} \inf \{ \rho(x, y) : \text{dist}(Fx, y) \geq \delta_2 \}$, δ_2 is the same as in Assumption 2.2, and $\beta_\delta(\varepsilon)$ is defined by (2.2). Now if δ in (3.8) is small enough then (2.5) holds true and so the conclusion of Theorem 2.1 remains valid, i.e. limiting measures will sit on attractors.

A more elaborate technique enables one to derive from (3.4) more precise results. For all $x \in K_i$ and $y \in K_j$ the number $B_{ij}^{(\rho)} = B_\rho(x, y)$ is the same. Let $L = \{1, \dots, v\}$. Given $i \in L$, a graph consisting of arrows $m \rightarrow n$ is called i -graph if any vertex $m \neq i$ originates exactly one arrow and the graph has no cycles. The set of all i -graphs will be denoted by $G(i)$. Put

$$B_\rho(i) = \min_{g \in G(i)} \sum_{(m \rightarrow n) \in g} B_{mn}^{(\rho)}$$

and

$$L_{\min}^{(\rho)} = \left\{ i \in L : B_{\rho}(i) = \min_{j \in L} B_{\rho}(j) \right\}.$$

The following result was proved in [K1], Sect. 1.5 and [K2].

Theorem 3.2. *Suppose that (3.4) and Assumption 3.3 hold true. Then for each $i \in L_{\min}^{(\rho)}$, K_i is an attractor and any weak limit of measures μ^ε as $\varepsilon \rightarrow 0$ has support in $\bigcup_{i \in L_{\min}^{(\rho)}} K_i$.*

We note that, in general, the set $L_{\min}^{(\rho)}$ depends on ρ though maximal δ -equivalence classes are attractors for any continuous ρ . Thus the less precise statement saying that limiting measures sit on attractors is more robust than Theorem 3.2. Remark also that this large deviation approach based on (3.4) can be extended to a locally compact case under certain assumptions on the function ρ making it extremely difficult for the Markov chain X_n^ε to go far away from a bounded domain.

4. Comparing Order Relations

By Lemma 3.1 Assumption 3.3 implies Assumption 2.1 but, in general, vice versa is not true. This and other reasons make it interesting to compare preorder relations \succ and \succ^{ρ} . If M is compact and $\rho(x, y) > 0$ unless $y = Fx$ which we assume throughout this section then, clearly, $z \succ^{\rho} v$ implies $z \succ v$. To get a sufficient condition for an inverse assertion define $\delta_{\rho}(\varepsilon) = \frac{1}{3} \sup \{ \delta > 0 : \rho(v, w) < \varepsilon \text{ for all } v, w \in M \text{ satisfying } \text{dist}(Fv, w) < \delta \}$. Let $N_{\delta}(\varepsilon)$ be the minimal number of open $\delta_{\rho}(\varepsilon)$ -balls needed to cover M .

Proposition 4.1. *If*

$$\lim_{\varepsilon \rightarrow 0} \varepsilon N_{\rho}(\varepsilon) = 0,$$

then $y \succ x$ implies $y \succ^{\rho} x$ for any pair of points $x, y \in M$.

Proof. Let $U_{\delta_{\rho}(\varepsilon)}(z_i)$, $i = 1, \dots, N_{\rho}(\varepsilon)$ be a cover of M . Suppose that $y \succ x$, then there exists a $\delta_{\rho}(\varepsilon)$ -pseudo-orbit x_1, \dots, x_n such that $x_1 = x$ and $x_n = y$. Exclude any subsequence $x_{j_1+1}, \dots, x_{j_2}$, $n > j_2 > j_1$ if Fx_{j_1} and Fx_{j_2} belong to the same element of the cover. After this procedure we will obtain a sequence $x, x_{i_1}, \dots, x_{i_k}, y$ with $k < N_{\rho}(\varepsilon)$, such that $\text{dist}(Fx_{i_j}, x_{i_{j+1}}) < 3\delta_{\rho}(\varepsilon)$ for all $j = 0, \dots, k$ where $x_{i_0} = x$ and $x_{i_{k+1}} = y$. Then by the definition of $\delta_{\rho}(\varepsilon)$ we obtain that $\rho(x_{i_j}, x_{i_{j+1}}) \leq \varepsilon$, and so $B_{\rho}(x, y) \leq \varepsilon N_{\rho}(\varepsilon)$ for any $\varepsilon > 0$. This together with (4.1) yields $y \succ^{\rho} x$. \square

The results become more meaningful if

$$\rho(x, y) = \rho_r(x, y) = (\text{dist}(Fx, y))^r \text{ for some } r > 0. \tag{4.2}$$

Clearly,

$$y \succ^{r_1} x \text{ if } r_2 > r_1.$$

Define

$$r(F, M) = \sup_{x, y \in M} \inf \left\{ r : y \succ_r x \text{ implies } y \succ_{\rho_r} x \right\}. \tag{4.3}$$

Let $N(\epsilon)$ denote the minimal number of ϵ -balls needed to cover M . Then limiting capacity $\kappa(M)$ of M defined as

$$\kappa(M) = \limsup_{\epsilon \rightarrow 0} \frac{\ln N(\epsilon)}{|\ln \epsilon|}. \tag{4.4}$$

We derive from Proposition 4.1

Corollary 4.1. *One has*

$$r(F, M) \leq \kappa(M). \tag{4.5}$$

In particular, if M is a compact manifold then

$$r(F, M) \leq \dim M,$$

Proposition 4.2. *Suppose that $F = hGh^{-1}$, where $h: M \rightarrow M$ is a Holder continuous homeomorphism, i.e. for some $\alpha, C > 0$*

$$\max(\text{dist}(hx, hy), \text{dist}(h^{-1}x, h^{-1}y)) \leq C(\text{dist}(x, y))^\alpha.$$

Then

$$\alpha r(G, M) \leq r(F, M) \leq \alpha^{-1} r(G, M). \tag{4.6}$$

In particular, if both h and h^{-1} are Lipschitz continuous then $r(F, M) = r(G, M)$.

Proof. Since both h and h^{-1} transform δ -pseudo-orbits to $C\delta^\alpha$ -pseudo-orbits and

$$C^{-1} \sum_i (\text{dist}(Gx_i, x_{i+1}))^{r/\alpha} \tag{4.7}$$

$$\sum_i (\text{dist}(Fhx_i, hx_{i+1}))^r \leq C \sum_i (\text{dist}(Gx_i, x_{i+1}))^{r\alpha},$$

the assertion follows. \square

A dynamical system $\{F^n, n \in \mathbb{Z}\}$ is called transitive if the orbit $\bigcup_{n \in \mathbb{Z}} F^n y$ of some point y is dense in M . A dynamical system $\{F^n, n \in \mathbb{Z}\}$ is said to have the shadowing property if for any $\epsilon > 0$ there is $\delta > 0$ such that for each δ -pseudo-orbit x_1, \dots, x_n one can choose a point y satisfying $\text{dist}(F^i y, x_i) < \epsilon$ for all $i = 1, \dots, n$. The following is straightforward.

Proposition 4.3. *If a dynamical system $\{F^n, n \in \mathbb{Z}\}$ is either transitive or it has the shadowing property then $r(F, M) = 0$.*

Further properties of the number $r(F, M)$ were established in a conversation with A. Katok. Recall, that a point x is called nonwandering if it does not have an open neighborhood U such that all sets $F^{-n}U, n \geq 0$ are disjoint.

Proposition 4.4. *Suppose that M is a smooth finite dimensional manifold, $F: M \rightarrow M$ is a continuous map, and the set of nonwandering points $\Omega(F)$ is dense in M (in particular, periodic points are dense). Then $r(F, M) \leq 1$.*

Proof. Let $x, y \in M$ be arbitrary points. Connect them by a smooth curve $\gamma(t)$, $\gamma(0) = x$, $\gamma(1) = y$. Fix a small $\varepsilon > 0$. Choose $z \in \Omega(F)$ with $\text{dist}(z, x) < \varepsilon$. There exists minimal n_1 satisfying $\text{dist}(z, F^{n_1}z) < \varepsilon$. We put $z_1 = x$ and $z_{i+1} = F^i z$, $i = 1, \dots, n_1 - 1$. Set $t_1 = \sup \{t < 1 : \text{dist}(x, \gamma(t)) \leq \varepsilon\}$. If $t_1 = 1$ then put $z_{n_1+1} = y$. Otherwise, choose z_{n_1+1} from $\Omega(F)$ satisfying $\text{dist}(z_{n_1+1}, \gamma(t_1)) < \varepsilon$. We continue by induction. Suppose that the numbers n_ℓ , t_ℓ , and the points z_i are already defined for $\ell = 1, \dots, k$ and $i = 0, 1, \dots, n_{k+1}$. There exists minimal n_{k+1} such that $\text{dist}(z_{n_{k+1}}, F^{n_{k+1}-n_k}z_{n_k}) < \varepsilon$. We put $z_{n_k+i} = F^i z_{n_{k+1}}$, $i = 0, 1, \dots, n_{k+1} - n_k - 1$. Set $t_{k+1} = \sup \{t \leq 1 : \text{dist}(\gamma(t_k), \gamma(t)) \leq \varepsilon\}$. If $t_{k+1} = 1$, then put $z_{n_{k+1}+1} = y$ and end the procedure. Otherwise, choose $z_{n_{k+1}+1}$ from $\Omega(F)$ satisfying $\text{dist}(z_{n_{k+1}+1}, \gamma(t_{k+1})) < \varepsilon$. On some step this procedure will end and we will obtain points z_1, \dots, z_n with $z_1 = x$ and $z_n = y$. It is easy to see that $\text{dist}(Fz_k, z_{n_k+1}) \leq 4\varepsilon$ for all $k \geq 1$ and the number of k -th in this construction does not exceed ε^{-1} times length of γ . Thus

$$\begin{aligned} \sum_{1 \leq i < N} (\text{dist}(Fz_i, z_{i+1}))^r &= (\text{dist}(Fx, Fz))^r + \sum_k (\text{dist}(Fz_{n_k}, z_{n_k+1}))^r \\ &\leq (\text{dist}(Fx, Fz))^r + 4^r \varepsilon^{r-1} \text{length}(\gamma) \end{aligned}$$

which tends to zero as $\varepsilon \rightarrow 0$ for any $r > 1$, and so $r(F, M) \leq 1$. \square

If M is a finite dimensional smooth manifold and $F = \text{id}$ its identity map: $Fx = x$, then, clearly, $r(F, M) = 1$. At the first sight it may seem as the worst case, i.e. that $r(F, M)$ cannot be bigger than 1. The following two dimensional examples shows that this is not true. We can choose a function whose graph has the Hausdorff dimension arbitrarily close to 2. For instance, take the function considered by Besicovitch and Ursell (see [F], pp. 114–115)

$$f(x) = \sum_{i=1}^{\infty} \lambda_i^{s-2} g(\lambda_i x), \quad x \in [0, 1], \quad 1 < s < 2,$$

where g is the “zig-zag” function of period 4 defined by

$$g(4k + x) = \begin{cases} x & \text{for } 0 \leq x < 1, \\ 2 - x & \text{for } 1 \leq x < 3, \\ x - 4 & \text{for } 3 \leq x < 4, \end{cases}$$

for $0 \leq x < 4$ and any integer k . Then for any positive $\{\lambda_i\}$ satisfying

$$\frac{\lambda_{i+1}}{\lambda_i} \rightarrow \infty \quad \text{and} \quad \frac{\log}{\log} \rightarrow 1 \quad \text{as } i \rightarrow \infty$$

the Hausdorff dimension of the graph Γ is s . Place Γ on the two-dimensional sphere S^2 so that the interval $[0, 1]$ becomes the equator. One can easily construct a diffeomorphism F of S^2 so that all points from the set Γ on S^2 are fixed points of F and F repels away from Γ in the direction of meridians attracting all points from $S^2 \setminus \Gamma$ to the North or South poles. It is clear that all points on Γ are

equivalent and Γ is an equivalent class in the sense of Sect. 2 since any pair $x, y \in \Gamma$ can be connected by a δ -pseudo-orbit with arbitrarily small $\delta > 0$. On the other hand, it is easy to see that the sums

$$\sum_{0 \leq i < n} (\text{dist}(Fx_i, x_{i+1}))^r = \sum_{0 \leq i < n} (\text{dist}(x_i, x_{i+1}))^r$$

with $x_0 = x, x_n = y, x_i \in \Gamma$ for all $i = 0, \dots, n$ can be made arbitrarily small for $r > s$ and they are bounded away from zero for $r < s$. Hence, in this case, $r(F, M) = s, M = S^2$. Taking the graph of another function constructed by the Serpinsky method one can get $r(F, M) = 2$. It seems that for any $n \geq 2$ one can produce n -dimensional examples with $r(F, M)$ arbitrarily close to n . The above construction would work if one can produce continuous curves $(x_1, f_1(x_1), \dots, f_{n-1}(x_1)) \in \mathbb{R}^n, x_1 \in [0, 1]$ with f_1, \dots, f_n continuous whose Hausdorff dimension can be made arbitrarily close to n .

It is known that the transitivity is a C^0 -generic property of diffeomorphisms of a compact manifold, and so by Proposition 4.3 C^0 -generically $r(F, M) = 0$.

Conjecture. C^1 -generically $r(F, M) = 0$. By Proposition 4.3 this would follow from another standing conjecture that the shadowing property is C^1 -generic.

References

- [F] Falconer, K. J.: The geometry of fractal sets. Cambridge: Cambridge University Press 1985
- [K1] Kifer, Y.: Random perturbations of dynamical systems. Boston: Birkhäuser 1988
- [K2] Kifer, Y.: A discrete time version of the Wentzell–Freidlin theory. Preprint 1988
- [R] Ruelle, D.: Small random perturbations of dynamical systems and the definition of attractors. *Commun. Math. Phys.* **82**, 137–151 (1981)

Communicated by J.-P. Eckmann
 Received August 5, 1988

