Math. Proc. Camb. Phil. Soc. (2011), **150**, 353 © *Cambridge Philosophical Society* 2011
doi:10.1017/S0305004110000642
First published online 12 January 2011

On statistical attractors and the convergence of time averages

BY ÖZKAN KARABACAK AND PETER ASHWIN

Mathematics Research Institute, College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter, EX4 4QF.
e-mail: O.karabacak@exeter.ac.uk and P.ashwin@exeter.ac.uk

(*Received* 30 *March* 2010; *revised* 4 *July* 2010)

Abstract

There are various notions of attractor in the literature, including measure (Milnor) attractors and statistical (Ilyashenko) attractors. In this paper we relate the notion of *statistical attractor* to that of the *essential ω-limit set* and prove some elementary results about these. In addition, we consider the convergence of time averages along trajectories. Ergodicity implies the convergence of time averages along almost all trajectories for all continuous observables. For non-ergodic systems, time averages may not exist even for almost all trajectories. However, averages of some observables may converge; we characterize conditions on observables that ensure convergence of time averages even in non-ergodic systems.

1. *Introduction*

Chaotic attractors are subsets of the phase space for dynamical systems to which, in some sense, typical trajectories converge. Exactly what is understood by "typical" and "converge" may give rise to subtly different concepts of attractor; see [**8**, **17**] for a discussion of various such concepts. On the other hand, in many cases one is not interested in the fine details of dynamics, just on the statistical properties of the convergence. This gives rise to the concept of *statistical attractor* or *statistical limit set* introduced by Ilyashenko [**1**].

In this paper, we show, in Theorem 2·3, that the statistical attractor can be defined using the notion of *essential ω-limit set* of trajectories previously defined in [**2**]. We also examine the convergence (or otherwise) of (Birkhoff) time averages of observables along trajectories. This clearly depends on two properties – firstly, the nature of the attractors – secondly, the nature of the particular observable considered. For example, systems with heteroclinic attractors admit non-convergent time averages [**9**] (for more about non-convergence of time averages see [**13**, **18**, **20–22**]). Trajectories for which time averages of "some" observables do not exist are called *historical* by Ruelle [**19**]; typical time-averaged observables do not converge to a constant, and therefore the system retains information about its past. In other words, the system does not admit a pointwise SRB measure in the terminology of [**6**]; we call such systems non-ergodic. For instance, consider the so-called Bowen's example, a dynamical system with an attracting heteroclinic cycle as shown in Figure 1. For a typical continuous observable, the time average of the observable along a typical trajectory oscillates without converging to a constant [**9**]. Although it is known that time averages of observables may not converge for non-ergodic systems, the question as to precisely which observables

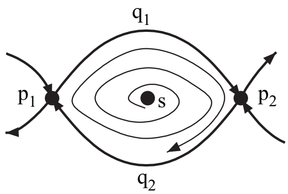


Fig. 1. Bowen's example: An attracting heteroclinic cycle that consists of two hyperbolic equilibria p_1, p_2 and two heteroclinic trajectories q_1 and q_2. A typical trajectory in the region bounded by the heteroclinic cycle converges to the heteroclinic cycle.

have convergent time averages was not previously studied. We give an answer to this, in Theorem 2·5 by proving that time averages of a typical trajectory depend only on the time averages of trajectories in the statistical attractor.

In Section 2, we state the main results of the paper, namely Theorem 2·3 and Theorem 2·5. In Section 3, the statistical attractor is defined in analogy to the Milnor attractor and the relation between Milnor attractors and statistical attractors is discussed. In Section 4, we consider the following question: Given a dynamical system, which observables have convergent time averages for almost all trajectories or for a positive measure subset of initial conditions? We prove that this is determined by the behavior of the observable on the statistical attractors of the system. Finally, in Section 5, we speculate as to how the results in this paper may be sharpened and discuss a particular application of our results to a system of coupled oscillator dynamics where one can find robust non-ergodic behavior (an attracting heteroclinic network) but an important observable (the average frequency difference) still exists.

Notation. We will use ℓ to denote Lebesgue (resp. Riemannian) measure on $\mathbb{R}^n$ (resp. on a compact manifold X). All subsets will be assumed to be measurable (Borel) unless otherwise stated. For two subsets A and B, we will write as in [**4**] $A =_\circ B$ to mean $\ell((A \setminus B) \cup (B \setminus A)) = 0$, and $A \supset_\circ B$ to mean $\ell(B \setminus A) = 0$.

2. *Main results*

We consider throughout a continuous flow (or semiflow) γ_t on a compact manifold X. For a trajectory passing through the point $x \in X$, the *ω-limit set* of x is defined as follows:

$$\omega(x) = \bigcap_{T=0}^{\infty} \left(\overline{\bigcup_{t>T} \gamma_t x} \right).$$

It follows that $\omega(x)$ is closed, connected, non-empty, flow-invariant and it consists of the accumulation points of sequences $\{\gamma_{T_k} x\}$ where $\{T_k\} \to \infty$.

Another type of limit set for a trajectory, that is called *essential ω-limit set*, is defined in [**2**] for maps and in [**3**] for flows. The essential ω-limit set of a trajectory can be thought as the set of points in the ω-limit set whose arbitrary small neighborhoods are visited with a non-vanishing frequency. For an open set $U \subset X$ and a finite orbit $\{\gamma_t x\}_{0 \leqslant t \leqslant T}$, the frequency of the orbit being in U can be represented by the function

$$\rho(x, U, T) = \frac{\ell(\{t : 0 \leqslant t \leqslant T, \gamma_t x \in U\})}{T}. \tag{2·1}$$

We will give a slightly different definition of ω_{ess} to that in [**3**] and show in Theorem 4·3 that these definitions are equivalent.

Definition 2·1 (essential ω-limit set). Let γ_t be a continuous flow on a compact manifold X. For $z \in X$, let $\mathcal{U}_z$ be the set of open neighbourhoods of z. The *essential ω-limit set* is defined as

$$\omega_{ess}(x) = \left\{ z \in X \colon \limsup_t \rho(x, U, t) > 0, \ \forall U \in \mathcal{U}_z \right\}. \tag{2·2}$$

If $z \in \omega_{ess}(x)$ then, for all $U \in \mathcal{U}_z$, there exist arbitrary large values of T for which $\gamma_T x \in U$. Hence,

$$\omega_{ess}(x) \subset \omega(x). \tag{2·3}$$

Milnor defines the likely limit set Λ_{likely} as the smallest closed subset that contains the ω-limit sets of almost all trajectories [**17**]. Similarly, Ilyashenko defines the statistical limit set Λ_{stat} (also called statistical attractor in the literature) as the smallest closed subset for which almost all trajectories spend almost all time near Λ_{stat} [**1**].

Definition 2·2 (statistical limit set [**1**, **12**]). Let ρ be defined as in (2·1). The statistical limit set Λ_{stat} is the smallest closed subset of X for which any open neighbourhood U of Λ_{stat} satisfies $\lim_t \rho(x, U, t) = 1$ for almost all $x \in X$.

Using the concept of essential ω-limit set, one can characterize the statistical limit set as follows:

THEOREM 2·3. *The statistical limit set is the smallest closed subset that contains the essential ω-limit sets of almost all trajectories.*

This implies that one can define statistical attractors in an analogy to Milnor attractors [**17**] by replacing the ω-limit set with the essential ω-limit set (see Section 3).

For example, consider Bowen's example in Figure 1. The phase space X is assumed to be the union of the heteroclinic cycle and the region bounded by the heteroclinic cycle. For almost all points in X, except the points on the heteroclinic cycle and the unstable equilibrium s, the ω-limit set is the whole cycle, whereas the essential ω-limit set is the union of two equilibria p_1 and p_2. Therefore, the likely limit set of the system is the heteroclinic cycle while the statistical limit set is $\{p_1, p_2\}$.

Remark 2·4. As pointed out by a referee, for a point $z \in \omega_{ess}(x)$, there can exist $U \in \mathcal{U}_z$ such that $\liminf_t \rho(x, U, t) = 0$. Clearly, the set $\omega_{min}(x) := \{z \in X \colon \liminf_t \rho(x, U, t) > 0, \ \forall U \in \mathcal{U}_z\}$ may be of interest for an alternative definition of attractor. For example, consider the modified Bowen's example studied by Kleptsyn [**16**], namely the heteroclinic cycle shown in Figure 1 where p_1 is non-hyperbolic with exponential contraction and p_2 is hyperbolic. For this example, as shown below, $\omega_{ess}(x) = \{p_1, p_2\}$ and $\omega_{min}(x) = \{p_1\}$ for a typical initial point $x \in X$, which suggests that ω_{min} may be related to the *minimal attractor*[1] introduced by Ilyashenko (the point p_1 for the modified Bowen's example). However, there is no reason to assume that $\omega_{min}(x)$ is non-empty in general, and in addition,

[1] In [**16**], the minimal attractor is defined as the complement of the union of open sets U that satisfy $\frac{1}{T}\int_0^T \ell(\gamma_{-t}(U))\, dt \to 0$ as $T \to \infty$

the minimal attractor is a set-wise definition. These suggest that the relationship between minimal attractor and ω_{min} may be non-trivial. In order to see that $\omega_{ess}(x) = \{p_1, p_2\}$ and $\omega_{min}(x) = \{p_1\}$, let $\tau_{n,1}(x)$ and $\tau_{n,2}(x)$ denote the period of time spent by the orbit $\gamma_t x$ in some neighbourhoods of p_1 and p_2, respectively, on the nth turn (the first turn starting from the first entrance of the trajectory to the neighbourhood of p_1). By [**16**, Proposition 1] a typical trajectory asymptotically spends comparable periods of time near p_1 and p_2, namely $\tau_{n,2}(x)/\tau_{n,1}(x) \to c \neq 0$ as $n \to \infty$. Therefore, $\omega_{ess}(x) = \{p_1, p_2\}$. However, [**16**, equations 1 and 2] imply that $\tau_{n+1,1}(x)/\tau_{n,1}(x) \to \infty$ as $n \to \infty$. Hence, for a sufficiently small open neighbourhood U_2 of p_2, $\liminf_t \rho(x, U_2, t) = 0$, but for all open neighbourhoods U_1 of p_1, $\liminf_t \rho(x, U_1, t) > 0$. Namely, $\omega_{min}(x) = \{p_1\}$.

Our second main result is on the time averages of continuous observables along typical trajectories. We will prove the following in Section 4:

THEOREM 2·5. *If time averages of an observable f have the same limit for all trajectories in the statistical limit set, then time averages of f exist for almost all trajectories.*

Remark 2·6. Batchourin gives a similar result to Theorem 2·5 for the trajectories of the induced dynamics on the space of Borel probability measures: "Weak time averages of continuous functions depend only on their restrictions to the *minimal attractor*[1]". [**11**, Theorem 1]

3. *Statistical attractors*

In this section, we define statistical attractors analogously to Milnor attractors [**17**]; namely, we say a closed set is a statistical attractor if it is the smallest closed subset that contains the essential ω-limit set of almost all points in a given positive measure subset of X. If one replaces the term "essential ω-limit set" with "ω-limit set", one obtains the definition for Milnor attractors. Both of these attractors can be defined using the following set valued set functions:

Definition 3·1. For a given subset Y of X, the *likely limit set of Y* and the *statistical limit set of Y* are defined, respectively, as follows:

$$\Lambda_M(Y) := \bigcap_{V =_\circ Y} \left(\overline{\bigcup_{x \in V} \omega(x)} \right) \tag{3·1}$$

$$\Lambda_S(Y) := \bigcap_{V =_\circ Y} \left(\overline{\bigcup_{x \in V} \omega_{ess}(x)} \right). \tag{3·2}$$

LEMMA 3·2. *Let ζ be any map from X to the set of closed subsets of X. Let $\{U_i\}$ be a countable base for X. Consider two subsets $A, Y \subset X$; then the following statements are equivalent*:

(*a*) $A^c = \bigcup\{U_i : U_i \cap \zeta(x) = \varnothing$ *for ℓ-a.e.* $x \in Y\}$;

(*b*) $A = \bigcap_{V =_\circ Y} \left(\overline{\bigcup_{x \in V} \zeta(x)} \right)$;

(*c*) *A is the smallest closed subset of X that contains $\zeta(x)$ for almost every point in Y. In other words, there exists a full measure subset W of Y such that $A = \overline{\bigcup_{x \in W} \zeta(x)}$, and for any other $W' \subset Y$ with $W' =_0 Y$, $A \subset \overline{\bigcup_{x \in W'} \zeta(x)}$.*

Proof. $(a) \Leftrightarrow (b)$: We need to show that $\mathcal{U}^c := \overline{\bigcup\{U_i : U_i \cap \zeta(x) = \emptyset \text{ for almost every } x \text{ in } Y.\}^c} = \bigcap_{V=_\circ Y}(\overline{\bigcup_{x\in V}\zeta(x)})$. We show the contrapositives for both inclusions: "$\subset$" Assume that $x \notin \bigcap_{V=_\circ Y}(\overline{\bigcup_{x\in V}\zeta(x)})$, then there exists $V =_\circ Y$ such that $x \notin \overline{\bigcup_{x\in V}\zeta(x)}$. Then, there exists an open neighbourhood $U \in \{U_i\}$ of x such that $U \cap \overline{\bigcup_{x\in V}\zeta(x)} = \emptyset$. Hence, $U \cap \zeta(x) = \emptyset$ for almost all x in Y, namely $x \notin \mathcal{U}^c$. "$\supset$" Assume $x \notin \mathcal{U}^c$, that is, there exists an open neighbourhood $U \in \{U_i\}$ of x such that $U \cap \zeta(x) = \emptyset$ for almost all x in Y. Hence, $U \cap \bigcup_{x\in V}\zeta(x) = \emptyset$ for some $V =_\circ Y$. Since U is open, we have $U \cap \overline{\bigcup_{x\in V}\zeta(x)} = \emptyset$. Thus, $x \notin \bigcap_{V=_\circ Y}(\overline{\bigcup_{x\in V}\zeta(x)})$. $(b) \Rightarrow (c)$: (b) implies that A is closed and contained in any closed subset that contains $\zeta(x)$ for almost every x in Y. Therefore, we only need to show that $A \supset \overline{\bigcup_{x\in V}\zeta(x)}$ for some $V \subset Y$ with $V =_\circ Y$. From (a), for each $U_i \in \{U_i\}$ with $U_i \subset \mathcal{U}$, there exists $V_i \subset Y$ with $V_i =_\circ Y$ and $U_i \cap \overline{\bigcup_{x\in V_i}\zeta(x)} = \emptyset$. Let W be the intersection of such (at most countably many) V_i's. Hence, $W \subset Y$, $W =_\circ Y$ and $\zeta(W) \subset \mathcal{U}^c$. $(c) \Rightarrow (b)$: this follows from the statement of (c).

By Lemma 3·2, $\Lambda_M(Y)$ (resp. $\Lambda_S(Y)$) is the smallest closed subset of X that contains the ω-limit set (resp. essential ω-limit set) of almost every point in Y. Now, we can define statistical attractor in analogy to the definition of Milnor attractor as follows:

Definition 3·3 (Milnor attractor and statistical attractor). Let Λ_M and Λ_S be defined as in Definition 3·1. A subset A of X is called a *Milnor attractor* if there exists a subset $V \subset X$, $V \neq_0 \emptyset$ such that $A = \Lambda_M(V)$. A is called *statistical attractor* if there exists a subset $V \subset X$, $V \neq_0 \emptyset$ such that $A = \Lambda_S(V)$.

Note that the maximal Milnor attractor $\Lambda_M(X)$, that is, the smallest closed subset that contains the ω-limit sets of almost all points in X is the *likely limit set* Λ_{likely}. Similarly, by Theorem 2·3, the maximal statistical attractor $\Lambda_S(X)$ is equal to the statistical limit set Λ_{stat}. In other words, Definition 3·3 covers the previous definition of a statistical limit set, introduced by Ilyashenko [**1**] as a special case (the latter also called the "statistical attractor"; see [**11**, **16**]).

Proof of Theorem 2·3. By definition, $z \notin \omega_{ess}(x)$ if and only if there exists an open neighborhood of U of z such that $\lim_t \rho(x, U, t) = 0$. In other words, for any open U, $\omega_{ess}(x) \cap U = \emptyset$ if and only if $\lim_t \rho(x, U, t) = 0$. In addition, Definition 2·2 implies that $\Lambda^{\text{c}}_{\text{stat}} = \bigcup\{U \subset X : U \text{ is open and } \lim_t \rho(x, U, t) = 0 \text{ for } \ell\text{-a.e. } x \in X\}$. Therefore, $\Lambda^{\text{c}}_{\text{stat}} = \bigcup\{U \subset X : U \text{ is open and } \omega_{ess}(x) \cap U = \emptyset \text{ for } \ell\text{-a.e. } x \in X\}$. Thus, the statement follows from Lemma 3·2.

We say a Milnor (statistical) attractor is *minimal* if it does not strictly contain any other Milnor (statistical) attractor. Note that both Milnor attractors and statistical attractors are closed and invariant under the flow. We can define the Milnor basin and statistical basin for a subset A as follows:

$$\mathcal{B}_M(A) = \{x \in X \mid \omega(x) \subset A\} \tag{3·3}$$

$$\mathcal{B}_S(A) = \{x \in X \mid \omega_{ess}(x) \subset A\}. \tag{3·4}$$

LEMMA 3·4. *Let Λ_M, Λ_S, $\mathcal{B}_M$ and $\mathcal{B}_S$ be defined as in* (3·1), (3·2), (3·3) *and* (3·4), *respectively. Then the following statements hold, where A, Y_i, Y are Borel subsets of X*:

(i) $\mathcal{B}_M(A) \subset \mathcal{B}_S(A)$ *for any subset* A *of* X;

(ii) $\Lambda_S(Y) \subset \Lambda_M(Y)$ *for any subset* Y *of* X;

(iii) $\Lambda_M(Y_2) \subset \Lambda_M(Y_1)$ *and* $\Lambda_S(Y_2) \subset \Lambda_S(Y_1)$ *if* $Y_1 \supset_\circ Y_2$;

(iv) $\mathcal{B}_M(\Lambda_M(Y)) \supset_\circ Y$ *and* $\mathcal{B}_S(\Lambda_S(Y)) \supset_\circ Y$ *for any subset* Y;

(v) $\Lambda_M(\mathcal{B}_M(A)) \subset \overline{A}$ *and* $\Lambda_S(\mathcal{B}_S(A)) \subset \overline{A}$ *for any subset* A *of* X;

(vi) $\Lambda_M(\mathcal{B}_M(A)) = A$ *if* A *is a Milnor attractor, while* $\Lambda_S(\mathcal{B}_S(A)) = A$ *if* A *is a statistical attractor.*

Proof. (i) and (ii) follow from (2·3). For the remainder of the results, we give the proofs for Milnor attractors; proofs for the statistical versions will be similar. (iii) From Lemma 3·2, there exists $W_1 \subset Y_1$ with $\ell(W_1) = \ell(Y_1)$ such that $\Lambda_M(Y_1) = \overline{\bigcup_{x\in W_1} \omega(x)}$. Let $W_2 := W_1 \bigcap Y_2$, then $W_2 \subset Y_2$ and $\ell(W_2) = \ell(Y_2)$. Hence, $\Lambda_M(Y_2) \subset \overline{\bigcup_{x\in W_2} \omega(x)} \subset \Lambda_M(Y_1)$. (iv) From Lemma 3·2, there exists $W \subset Y$ with $W =_\circ Y$ such that $\Lambda_M(Y) = \overline{\bigcup_{x\in W} \omega(x)}$. Therefore, $\mathcal{B}_M(\Lambda_M(Y)) \supset W$. This implies $\ell(Y \setminus \mathcal{B}_M(\Lambda_M(Y))) \leqslant \ell(Y \setminus W) = 0$. Hence $\mathcal{B}_M(\Lambda_M(Y)) \supset_\circ Y$. (v) Consider $x \in \Lambda_M(\mathcal{B}_M(A))$. If $\mathcal{B}_M(A)$ has zero measure then (v) is trivial. Hence, we assume $\mathcal{B}_M(A)$ has positive measure. Choose a subset $V \subset \mathcal{B}_M(A)$ with $V =_\circ \mathcal{B}_M(A)$. Then $x \in \overline{\bigcup_{y\in V} \omega(y)}$. Since for all $y \in \mathcal{B}_M(A)$, $\omega(y) \subset A$, we have $\overline{\bigcup_{y\in V} \omega(y)} \subset \overline{A}$. Hence, $x \in \overline{A}$. Finally, for (vi), assume A is a Milnor attractor. Then A is closed and therefore (v) implies $\Lambda_M(\mathcal{B}_M(A)) \subset A$. We will show that $\Lambda_M(\mathcal{B}_M(A)) \supset A$. Since A is a Milnor attractor, there exists a positive measure subset V such that $A = \Lambda_M(V)$. By Lemma 3·2, there exists $W \subset V$ such that $\Lambda_M(V) = \overline{\bigcup_{x\in W} \omega(x)}$. Hence, for all $x \in W$, $\omega(x) \subset A$, therefore, $W \subset \mathcal{B}_M(A)$. From (iii), this implies $A = \Lambda_M(V) = \Lambda_M(W) \subset \Lambda_M(\mathcal{B}_M(A))$.

Milnor attractors and statistical attractors can be related to each other as shown in Lemma 3·4. However, they are not in one-to-one correspondence as we will see in Example 3·7 below. Nevertheless, we show that for each Milnor attractor there is a smaller statistical attractor, and for each statistical attractor there is a larger Milnor attractor:

THEOREM 3·5. *Suppose that* γ_t *is a continuous flow on a compact manifold* X.

(*a*) *If* A *is a Milnor attractor for the flow, then* $A_S = \Lambda_S(\mathcal{B}_M(A))$ *is a statistical attractor contained in* A *with* $\mathcal{B}_S(A_S) \supset_\circ \mathcal{B}_M(A)$.

(*b*) *If* A *is a statistical attractor for the flow, then* $A_M = \Lambda_M(\mathcal{B}_S(A))$ *is a Milnor attractor that contains* A *with* $\mathcal{B}_M(A_M)) \supset_\circ \mathcal{B}_S(A)$.

Proof. If A is a Milnor attractor, then $\mathcal{B}_M(A)$ has positive measure and therefore $\Lambda_S(\mathcal{B}_M(A))$ is a statistical attractor. From Lemma 3·4(iv), $\mathcal{B}_S(A_S) = \mathcal{B}_S(\Lambda_S(\mathcal{B}_M(A))) \supset_\circ \mathcal{B}_M(A)$. It remains to show that $A_S \subset A$. From Lemma 3·4(ii) and (vi), $A_S = \Lambda_S(\mathcal{B}_M(A)) \subset \Lambda_M(\mathcal{B}_M(A)) = A$. The proof for (*b*) is similar.

The simplest examples where statistical attractors are different than Milnor attractors are systems such as Bowen's example mentioned in Section 1 that admits a heteroclinic cycle. We discuss two other illustrative examples for the remainder of this section.

Example 3·6. The heteroclinic cycle illustrated in Figure 2 arises as a minimal Milnor attractor of the flow given by

$$\begin{aligned} \dot{x} &= kx + (ax^2 + by^2 + cz^2)x \\ \dot{y} &= ky + (ay^2 + bz^2 + cx^2)y \\ \dot{z} &= kz + (az^2 + bx^2 + cy^2)z \end{aligned} \tag{3·5}$$

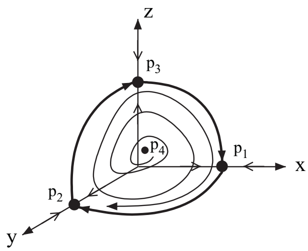


Fig. 2. Example 3·6 showing a trajectory converging towards an attracting heteroclinic cycle on the boundary of the positive orthant of $\mathbb{R}^3$. The ω-limit set of each point except p_4 in the interior of the orthant $\{x, y, z > 0\}$ is the whole cycle. Due to the symmetry, there exist seven symmetric copies of this cycle in the other regions that are attracting almost all points in their interiors and therefore are Milnor attractors. Similarly, there are eight statistical attractors each of which consists of three fixed points contained in a heteroclinic cycle. There is a one-to-one correspondence between Milnor attractors and statistical attractors in this example.

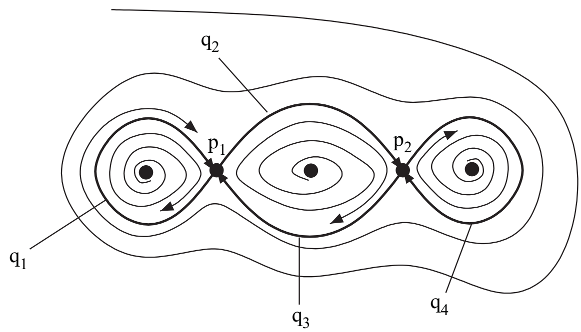


Fig. 3. An invariant set $\{p_1, p_2, q_1, q_2, q_3, q_4\}$ containing seven Milnor attractors and three statistical attractors. The Milnor attractors are two homoclinic cycles $\{p_1, q_1\}$ and $\{p_2, q_4\}$, the heteroclinic cycle $\{p_1, q_2, p_2, q_3\}$ and various combinations of these, namely $\{p_1, p_2, q_1, q_2, q_3\}$, $\{p_1, p_2, q_2, q_3, q_4\}$, $\{p_1, p_2, q_1, q_4\}$ and $\{p_1, p_2, q_1, q_2, q_3, q_4\}$. The statistical attractors are $\{p_1\}$, $\{p_2\}$, and $\{p_1, p_2\}$. Hence, in this example there is no one-to-one correspondence between Milnor attractors and statistical attractors.

where the parameters are chosen such that $k > 0$, $a < 0$ and $b < -c < 0$ [**7**, **10**]. The system admits the rotation symmetries $(x, y, z) \to (z, x, y)$ and the reflection symmetries $(x, y, z) \to (\pm x, \pm y, \pm z)$. The reflection symmetries mean that each orthant is invariant and there exist symmetric copies of this cycle that attract almost all initial conditions in each orthant. These are clearly minimal Milnor attractors. On the other hand, the statistical attractor in the positive orthant consist of the equilibria p_1, p_2, and p_3 only and in each region there is a different minimal statistical attractor that consist of three equilibria each of which is a symmetric copy of p_1, p_2, and p_3.

In Example 3·6, there is a one-to-one correspondence between Milnor attractors and statistical attractors. The following example shows that this is not always the case.

Example 3·7. Figure 3 shows the phase portrait of a flow with a Milnor attractor that contains seven smaller Milnor attractors. Three of them are minimal Milnor attractors: one heteroclinic cycle ($\{p_1, q_2, p_2, q_3\}$) and two homoclinic cycles ($\{p_1, q_1\}$ and $\{p_2, q_4\}$). The statistical limit set consists of two points p_1 and p_2 each of which is a minimal statistical attractor. There are two minimal statistical attractors but three minimal Milnor attractors. As a result, unlike Example 3·6, there is no one-to-one correspondence between Milnor attractors and statistical attractors in this case.

4. *Convergence of time averages*

It is known that time averages of trajectories asymptotic to a heteroclinic cycle (or more generally non-ergodic attractors) do not converge in general [**9**, **20**]; they show historical behavior in the terminology of Ruelle [**19**]. However, for a particular system and observable it may be non-trivial to check that an average converges or not; e.g. [**18**]. We will show that, in order to prove the convergence of the time average or otherwise, one only needs to consider the values of the observable on the statistical attractors of the flow.

In this section, we attempt to precisely characterize those observables whose time averages converge for a given statistical attractor. Our main theoretical result will be Theorem 4·2, for which we first need to develop some notation. In terms of applications, we give two corollaries that have more easily checkable assumptions.

Consider the measures $\mu_{x,T} = (1/T)\int_0^T \delta_{\gamma_t x}\,dt$ where $T > 0$ and δ_x is the Dirac measure. Note that

$$\mu_{x,T}(U) = \rho(x, U, T). \tag{4·1}$$

For each $f \in \mathcal{C}(X)$ we have $\int_X f\, d\mu_{x,T} = (1/T)\int_0^T f(\gamma_t x)\,dt$. Define a functional on $\mathcal{C}(X)^*$ by $\varphi_{x,T} : f \to \int_X f d\mu_{x,T}$. Then, $|\varphi_{x,T}| = \sup_{\|f\|=1} |(1/T)\int_0^T f(\gamma_t x)dt| \leqslant \sup_{\|f\|=1}\|f\| \leqslant 1$. Since the unit $(|\cdot|)$-ball in $\mathcal{C}(X)^*$ is weak* compact (Alaoglu's theorem), the set of accumulation points of $\varphi_{x,T}$ as $T \to \infty$ in the weak* topology of $\mathcal{C}(X)^*$, namely,

$$\Theta(x) = \bigcap_{T>0} \overline{\{\varphi_{x,\tilde{T}} : \tilde{T} > T\}} \tag{4·2}$$

is non-empty and bounded, where the closure is in the weak* topology. The Riesz Representation Theorem implies that for each $\tilde{\varphi} \in \Theta(x)$ there exists a unique Borel probability measure $\mu(\tilde{\varphi})$ such that

$$\tilde{\varphi}(f) = \int_X f d\mu(\tilde{\varphi}). \tag{4·3}$$

The set of such measures $\{\mu(\tilde{\varphi}) : \tilde{\varphi} \in \Theta(x)\}$ can also be written as

$$\Omega(x) = \bigcap_{T>0} \overline{\{\mu_{x,\tilde{T}} : \tilde{T} > T\}}, \tag{4·4}$$

where the closure is under the weak topology of measures. We say a sequence of measures $\{\mu_k\}$ converges weakly to the measure μ ($\mu_k \rightharpoonup \mu$) if and only if $\lim_k \int f d\mu_k = \int f d\mu$ for every continuous function f [**5**]. Since $\Theta(x)$ is non-empty, $\Omega(x)$ is also non-empty. We can use $\Theta(x)$ and $\Omega(x)$ to classify the behavior of time averages of continuous observables as follows.

LEMMA 4·1. *For every $f \in \mathcal{C}(X)$,*

$$\bigcap_{T>0} \overline{\left\{ \frac{1}{\tilde{T}} \int_0^{\tilde{T}} f(\gamma_t x)\, dt \ : \ \tilde{T} > T \right\}} = \{\varphi(f) \ : \ \varphi \in \Theta(x)\}, \tag{4·5}$$

where $\Theta(x)$ is defined in (4·2).

Proof. We prove equality in two stages. "$\subset$" Let $\bar{f} \in \mathbb{R}$ be a limit point of $(1/T)\int_0^T f(\gamma_t x)\,dt$ as $T \to \infty$. Then, there exists a sequence $\{t_k\} \to \infty$ such that $(1/t_k)\int_0^{t_k} f(\gamma_t x)\,dt = \int_X f\, d\mu_{x,t_k} = \varphi_{x,t_k}(f) \to \bar{f}$. By Alaoglu's theorem, there exists a subsequence $\{t_{n_k}\}$ such that $\{\varphi_{x,t_{n_k}}\}$ converges to some functional $\tilde{\varphi}$ in weak* topology. That is, for each $f \in \mathcal{C}(X)$, $\varphi_{x,t_{n_k}}(f) \to \tilde{\varphi}(f)$. Therefore, $\bar{f} = \tilde{\varphi}(f)$. It is clear that $\tilde{\varphi} \in \Theta(x)$. "$\supset$" Let $\bar{f} = \tilde{\varphi}(f)$ for some $\tilde{\varphi} \in \Theta(x)$. By the definition of $\Theta(x)$, there exists a sequence $\{t_k\} \to \infty$ such that $\varphi_{x,t_k}(f) \to \tilde{\varphi}(f)$ for any $f \in \mathcal{C}(X)$. Therefore, $(1/t_k)\int_0^{t_k} f(\gamma_t x)\,dt \to \bar{f}$.

Using (4·3), we can rewrite (4·5) as

$$\bigcap_{T>0} \overline{\left\{ \frac{1}{\tilde{T}} \int_0^{\tilde{T}} f(\gamma_t x)\,dt \;:\; \tilde{T} > T \right\}} = \left\{ \int_X f\,d\mu \;:\; \mu \in \Omega(x) \right\}. \tag{4·6}$$

We can now state the main result of this section:

THEOREM 4·2. *Let γ_t be a continuous flow on a compact metric space X and $x(t) \subset X$ be an orbit with $x_0 = x(0)$. Let $f : X \to \mathbb{R}$ be a continuous function. Then the following limit*

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T f(\gamma_t x_0)\,dt = \bar{f} \tag{4·7}$$

exists, if and only if,

$$\int_X f\,d\mu = \bar{f} \text{ for all } \mu \in \Omega(x_0). \tag{4·8}$$

Proof. This follows as a special case of Lemma 4·1 where the sets on both sides of (4·6) reduce to single points.

Theorem 4·2 means that the time average of a given observable exists if and only if the observable has a constant integral with respect to all measures in $\Omega(x)$. We now show the relation between $\Omega(x)$ and the essential ω-limit set $\omega_{ess}(x)$.

THEOREM 4·3. *Let γ_t be a continuous flow on a compact manifold X and $\Omega(x)$ be defined as in* (4·4). *Then, for all $x \in X$,*

$$\omega_{ess}(x) = \overline{\bigcup_{\mu\in\Omega(x)} \operatorname{supp} \mu}. \tag{4·9}$$

This theorem implies that our definition of essential ω-limit set is equivalent to the original definition in [**3**]. Note that Definition 2·1 is in some sense simpler in that it does not depend on measure theoretical notions such as weak convergence. It also makes the relation between the concepts *essential ω-limit set* and *statistical attractor* clearer.

To prove Theorem 4·3 we use the following lemmas:

LEMMA 4·4. *Let Ω be any set of measures on X. Then $z \in \overline{\bigcup_{\mu\in\Omega} \operatorname{supp} \mu}$ if and only if for every open neighbourhood U of z, there exists a $\mu \in \Omega$ such that $\mu(U) > 0$.*

Proof. ('only if') Assume that there exists a sequence $\{z_k\} \to z$ such that for all k, $z_k \in \operatorname{supp} \mu_k$ where $\mu_k \in \Omega$. Then, for each open neighbourhood U of z there exists $K > 0$ such that $z_K \in U$. Choose an open neighbourhood V of z_K such that $V \subset U$. Since $\mu_K(V) > 0$, then $\mu_K(U) > 0$. ('if') Assume that given any open neighbourhood U of z there exists a

$\mu \in \Omega$ such that $\mu(U) > 0$. Hence, $\text{supp}\ \mu \cap U \neq \emptyset$, that is, there exists a $\bar{z} \in U$ such that $\bar{z} \in \text{supp}\ \mu$. This implies that $z \in \overline{\bigcup_{\mu\in\Omega} \text{supp}\ \mu}$.

The following result we quote from [**5**, Theorem 2·1] without proof.

LEMMA 4·5. *Let μ and μ_k, $k = 1, 2, \ldots$ be Borel probability measures. The following statements are equivalent*:

(i) $\mu_k \rightharpoonup \mu$, *(i.e. μ_k converges weakly to μ);*
(ii) $\liminf \mu_k(U) \geqslant \mu(U)$ *for every open set U;*
(iii) $\limsup \mu_k(F) \leqslant \mu(F)$ *for every closed set F.*

Proof of Theorem 4·3 ('$\supset$') Assume $z \in \overline{\bigcup_{\mu\in\Omega(x)} \text{supp}\ \mu}$. From Lemma 4·4, for each open neighbourhood U of z, there exists a $\mu \in \Omega(x)$ such that $\mu(U) > 0$. Note that $\mu \in \Omega(x)$ implies there exists a sequence $\{T_k\} \to \infty$ such that $\mu_{x,T_k} \rightharpoonup \mu$. Hence, from Lemma 4·5 and equation (4·1), $\liminf_k \rho(x, U, T_k) = \liminf_k \mu_{x,T_k}(U) \geqslant \mu(U) > 0$. Therefore, $\limsup_t \rho(x, U, t) > 0$, and from Definition 2·1, $z \in \omega_{ess}(x)$. ('$\subset$') Assume that $z \in \omega_{ess}(x)$ and U is an arbitrary open neighbourhood of z. Let U' and F be open and closed neighbourhoods of z, respectively, that satisfy $U' \subset F \subset U$. From Definition 2·1, $\limsup_t \rho(x, U', t) > 0$ since $z \in U'$ and U' is open. Then, $\limsup_t \rho(x, F, t) \geqslant \limsup_t \rho(x, U', t) > 0$. Therefore, there exists a sequence $\{t_k\} \to \infty$ such that $\lim_k \rho(x, F, t_k) = \lim_k \mu_{x,t_k}(F) > 0$. By compactness, there exists a subsequence $\{t_{k_m}\} \to \infty$ such that $\mu_{x,t_{k_m}}$ converges weakly to a measure $\mu \in \Omega(x)$. From Lemma 4·5, $\limsup_m \mu_{x,t_{k_m}}(F) \leqslant \mu(F)$. Hence, $\mu(U) \geqslant \mu(F) \geqslant \limsup_m \mu_{x,t_{k_m}}(F) = \lim_k \mu_{x,t_k}(F) > 0$. Finally, Lemma 4·4 implies $z \in \overline{\bigcup_{\mu\in\Omega} \text{supp}\ \mu}$.

Using the Ergodic Decomposition Theorem [**15**, Theorem 4·1·12] one can restrict the condition on measures in Theorem 4·2 to ergodic measures supported on the essential ω-limit set. Let $\mathcal{E}(X)$ denote the set of ergodic γ_t-invariant probability measures supported in X. Namely, $\mathcal{E}(X) = \{\mu \in \mathcal{M}(X)\colon \text{supp}(\mu) \subset X\}$, where $\mathcal{M}(X)$ is the set of invariant ergodic measures of the flow (X, γ_t). Since the supports of the measures in $\Omega(x)$ are contained in $\omega_{ess}(x)$, the Ergodic Decomposition Theorem implies

$$\Omega(x) \subset \text{conv}(\mathcal{E}(\omega_{ess}(x))). \tag{4·10}$$

Using this, we can conclude in the following result that the time average of an observable along a trajectory depends only on the time averages of trajectories in the essential ω-limit set.

THEOREM 4·6. *Suppose that γ_t is a continuous flow on a compact manifold X. Assume that, for a given continuous function $f : X \to \mathbb{R}$, there exists a constant $\bar{f} \in \mathbb{R}$ such that, for all $y \in \omega_{ess}(x_0)$, $\lim_{T\to\infty}(1/T)\int_0^T f(\gamma_t y)\,dt = \bar{f}$. Then the limit $\lim_{T\to\infty}(1/T)\int_0^T f(\gamma_t x_0)\,dt$ exists and equal to $\bar{f}$.*

Proof. Let μ be an ergodic measure supported on $\omega_{ess}(x_0)$, namely $\mu \in \mathcal{E}(\omega_{ess}(x_0))$. From a corollary of Birkhoff Ergodic Theorem [**15**, Corollary 4·1·4], there exists a point $y \in \omega_{ess}(x_0)$ such that $\lim_{T\to\infty}(1/T)\int_0^T f(\gamma_t y)\,dt = \int f d\mu$. Then, by assumption, $\int f d\mu = \bar{f}$. Let $\mathcal{F}$ denote the set of accumulation points of $(1/T)\int_0^T f(\gamma_t x_0)\,dt$ as $T \to \infty$. From (4·6) $\mathcal{F} = \{\int_X f\,d\mu \;:\; \mu \in \Omega(x_0)\}$. Using (4·10), we conclude

$$\mathcal{F} \subset conv\left(\left\{\int_X f\,d\mu \;:\; \mu \in \mathcal{E}(\omega_{ess}(x_0))\right\}\right) = conv(\{\bar{f}\}) = \{\bar{f}\}.$$

Finally, these results are related to statistical attractors in the following way: If an observable f has a constant time average along all trajectories in a statistical attractor, then there is a positive measure subset of initial states for which time averages of f exists. Both the following corollary and the second main result, Theorem 2·5, in Section 2 follow directly from Theorem 4·6.

COROLLARY 4·7. *If A is a statistical attractor and* $\lim_{T\to\infty}(1/T)\int_0^T f(\gamma_t x)\,dt = \bar{f}$ *for all $x \in A$. Then, for all $x \in \mathcal{B}_S(A)$,* $\lim_{T\to\infty}(1/T)\int_0^T f(\gamma_t x)\,dt$ *exists and equal to $\bar{f}$.*

5. *Discussion*

This paper has introduced the idea of a "non-maximal" statistical attractor associated with a basin of attraction and has related it to Milnor attractors. There are clearly many open questions as to the decomposition of statistical attractors and the subtleties of their relation to Milnor attractors or other types of attractor, though we should clearly indicate that in many cases (in particular those of ergodic attractors) the notions coincide. Nonetheless, for systems with symmetries and/or invariant subspaces non-ergodic attractors may be robust, giving a large and relevant class of systems where the distinction is non-trivial.

Moving on to the question of which observables will give convergence of time averages even to non-ergodic attractors, in Theorem 4·2 we give a necessary and sufficient condition for convergence of an averaged observable on a specific trajectory. Theorem 4·6 relates the convergence of averages to statistical attractors by giving a sufficient condition for convergence. An open question is how much one can weaken the assumptions on this theorem to give necessary and sufficient conditions. Clearly, for the special case of heteroclinic attractors (where the statistical attractors are just finite sets of points) this is much easier than for more general non-ergodic attractors where there may be quite wild behavior of a small number of trajectories.

To finish with, we present an application of Theorem 4·6 to an example system with an attracting non-ergodic heteroclinic network. We show that an observable of relevance can be demonstrated to converge.

5·1. *Example of a non-trivial convergent average for a non-ergodic attractor*

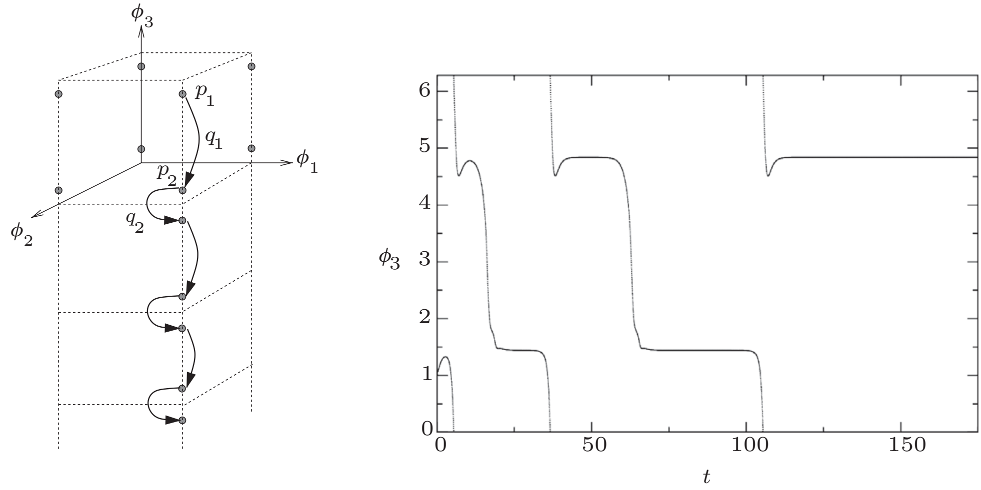
Consider the dynamics on $\mathbb{T}^3 = \{\phi_1, \phi_2, \phi_3\}$ generated by the ODE

$$\begin{aligned} \dot{\phi}_1 &= g(\phi_1, \phi_2 - \phi_3, 0) - g(0, \phi_1, \phi_2 - \phi_3) \\ \dot{\phi}_2 &= g(\phi_2, \phi_1 + \phi_3, 0) - g(0, \phi_1 + \phi_3, \phi_2) \\ \dot{\phi}_3 &= g(\phi_3, \phi_1 + \phi_3, \phi_2) - g(0, \phi_1 + \phi_3, \phi_2) + 2\sin\phi_1, \end{aligned} \tag{5·1}$$

where we set

$$\begin{aligned} g(x, y, z) &+ h(x - y) + h(x - z), \\ h(x) &= -\sin(x + 1.4) + 0.3\sin(2x) - 0.1\sin(3x). \end{aligned}$$

Note that the subspaces $\phi_1 = 0$ and $\phi_2 = 0$ are invariant under the dynamics. One can verify, using numerical simulation and examination of the flows in invariant subspaces, that there is an attracting heteroclinic network (a type of heteroclinic ratchet [**14**]) contained within these subspaces (see Figure 4). This heteroclinic network consists of a heteroclinic cycle between the equilibria p_1 and p_2 that winds in $-\phi_3$ direction. Therefore, when lifted to $\mathbb{R}^3$, a typical trajectory converging to the attractor has $\phi_3^L \to -\infty$ (ϕ^L denotes the lifted trajectory).



(a) (b)

Fig. 4. A heteroclinic attractor ($\{p_1, p_2, q_1, q_2\}$) for the system (5·1). (a) Schematic illustration of the attractor in $\mathbb{T}^3 = \{\phi_1, \phi_2, \phi_3\}$ lifted to $\mathbb{R}^3$. (b) The ϕ_3 component of a trajectory converging to the attractor; note that a lift of this component will be unbounded, but it grows so slowly that its average converges to zero.

However, one can use Theorem 4·6 to show that $(1/T)\int_0^T \phi_3 dt$ converges to zero, as follows. Let $\mathcal{L}_\gamma f : X \to \mathbb{R}$ be the Lie derivative of the function f along the flow ϕ, that is $\mathcal{L}_\gamma f(x) = (d/dt)f(\gamma_t x)|_{t=0}$. Consider the function $\tilde{f}(\phi) = \phi_3$. Let $\tilde{\phi}(t)$ be a trajectory converging to the heteroclinic attractor in Figure 4. That is, $\omega(\tilde{\phi}(t)) = \{p_1, p_2, q_1, q_2\}$ and $\omega_{ess}(\tilde{\phi}(t)) = \{p_1, p_2\}$. Then, from Theorem 4·6 we have

$$\lim_{T\to\infty} \frac{\phi_3^L(T)}{T} = \lim_{T\to\infty} \frac{1}{T}\int_0^T \mathcal{L}_\gamma \tilde{f}(\phi(t))\, dt. \qquad (5\cdot2)$$

exists and equal to zero since $\mathcal{L}_\gamma \tilde{f}$ is zero at fixed points p_1 and p_2.

We remark that system (5·1) can be seen as the phase difference equations for a system of four coupled phase oscillators, $\theta_i \in \mathbb{T}^1 = [0, 2\pi)$,

$$\begin{aligned}
\dot{\theta}_1 &= \omega + g(\theta_1, \theta_2, \theta_3) + 2\sin(\theta_1 - \theta_3)\\
\dot{\theta}_2 &= \omega + g(\theta_2, \theta_1, \theta_4)\\
\dot{\theta}_3 &= \omega + g(\theta_3, \theta_1, \theta_2) + 2\sin(\theta_1 - \theta_3)\\
\dot{\theta}_4 &= \omega + g(\theta_4, \theta_1, \theta_2).
\end{aligned} \qquad (5\cdot3)$$

with a heteroclinic ratchet in the terminology of [**14**]. The reduction to (5·1) follows on setting $\phi_1 := \theta_1 - \theta_3$, $\phi_2 := \theta_2 - \theta_4$ and $\phi_3 := \theta_3 - \theta_4$.

Heteroclinic ratchets are attractors on an N-torus that include heteroclinic cycles winding in one direction but no other heteroclinic cycles winding in the opposite direction. Even though there is no winding for the trajectories in the heteroclinic ratchet in the specified direction, trajectories near the heteroclinic ratchet may wind in this direction repeatedly. Therefore, one may expect that average frequency of windings may not converge, but Theorem 4·6 implies that it converges to zero. Assuming that the system consists of coupled oscillators, this convergence to zero implies a frequency synchronization between oscillators, although the phase differences grow unboundedly. The example here contrasts to those in [**14**] where arbitrary small noise or detuning between the natural frequencies of oscillators is needed to observe ratcheting of a nearby trajectory, and shows that ratcheting of oscillators introduced in [**14**] can occur even without noise or detuning.

Acknowledgments. We would like to thank Mike Field for his comments, and especially for motivating us to find the example in Section 5·1. We also thank Andrei Török and Mark Holland for discussions related to this work. Finally, we thank a referee for suggestions to improve the argument connecting 2·2 to 2·3, and for pointing out the lim inf issue on which we comment in Remark 2·4.

REFERENCES

[**1**] V. I. ARNOLD, V. S. AFRAJMOVICH, YU. S. IL′YASHENKO and L. P. SHIL′NIKOV. *Bifurcation Theory and Catastrophe Theory* (Springer-Verlag, 1999).
[**2**] P. ASHWIN, P. J. ASTON and M. NICOL. On the unfolding of a blowout bifurcation. *Phys. D* **111**(1–4) (1998), 81–95.
[**3**] P. ASHWIN and M. FIELD. Heteroclinic networks in coupled cell systems. *Arch. Ration. Mech. Anal.* **148**(2) (1999), 107–143.
[**4**] P. ASHWIN and J.R. TERRY. On riddling and weak attractors. *Phys. D* **142** (2000), 87–100.
[**5**] P BILLINGSLEY. *Convergence of Probability Measures* (John Wiley & Sons Inc., 1968).
[**6**] M. BLANK and L. BUNIMOVICH. Multicomponent dynamical systems: SRB measures and phase transitions. *Nonlinearity* **16**(1) (2003), 387–401.
[**7**] F. H. BUSSE and R. M. CLEVER. Nonstationary convection in a rotating system. In U. Müller, K. G. Roesner and B. Schmidt, eds., *Recent Developments in Theoretical and Experimental Fluid Dynamics*, pages 376–385 (Springer Verlag, 1979).
[**8**] I. D. CHUESHOV. *Introduction to the theory of infinite-dimensional dissipative systems* (http://www.emis.de/monographs/Chueshov/). (ACTA, Kharkov, 2002).
[**9**] A. GAUNERSDORFER. Time averages for heteroclinic attractors. *SIAM J. Appl. Math.* **52**(5) (1992), 1476–1489.
[**10**] J. GUCKENHEIMER and P. HOLMES. Structurally stable heteroclinic cycles. *Math. Proc. Camb. Phil. Soc.* **103** (1988), 189–192.
[**11**] YU. ILYASHENKO. Minimal attractors. In *EQUADIFF* 2003, pages 421–428 (World Scientific Publishing, 2005).
[**12**] YU. ILYASHENKO and A. NEGUT. Invisible parts of attractors. *Nonlinearity* **23**(5) (2010), 1199.
[**13**] T. JORDAN, V. NAUDOT and T. YOUNG. Higher order Birkhoff averages. *Dyn. Syst.* **24**(3) (2009), 299–313.
[**14**] Ö. KARABACAK and P. ASHWIN. Heteroclinic ratchets in networks of coupled oscillators. *J. Nonlinear Sci.* **20** (2010), 105–129.
[**15**] A. KATOK and B. HASSELBLATT. *Introduction to the Modern Theory of Dynamical Systems* (Cambridge University Press, 1995).
[**16**] V. KLEPTSYN. An example of non-coincidence of minimal and statistical attractors. *Ergodic Theory Dynam. Systems* **26**(3) (2006), 759–768.
[**17**] J. MILNOR. On the concept of attractor. *Commun. Math. Phys.* **99** (1985), 177–195.
[**18**] W. OTT and J. A. YORKE. When Lyapunov exponents fail to exist. *Phys. Rev. E* **78**(5, Part 2) (2008).
[**19**] D. RUELLE. Historical behaviour in smooth dynamical systems. In *Global analysis of dynamical systems*, pages 63–66 (Inst. Phys., Bristol, 2001).
[**20**] K. SIGMUND. Time averages for unpredictable orbits of deterministic systems. *Ann. Oper. Res.* **37**(1–4) (1992), 217–228.
[**21**] F. TAKENS. Heteroclinic attractors: time averages and moduli of topological conjugacy. *Bol. Soc. Brasil. Mat. (N.S.)* **25**(1) (1994), 107–120.
[**22**] F. TAKENS. Orbits with historic behaviour, or non-existence of averages. *Nonlinearity* **21**(3) (2008), 33–36.