INSTITUTE OF PHYSICS PUBLISHING NONLINEARITY

Nonlinearity **17** (2004) 1393–1406 PII: S0951-7715(04)62510-2

To persist or not to persist?

Josef Hofbauer[1,3] **and Sebastian J Schreiber**[2]

[1] Institut für Mathematik, Universität Wien, Strudlhofgasse 4, 1090 Wien, Austria
[2] Department of Mathematics, College of William and Mary, Williamsburg, VA 23187-8795, USA

E-mail: jhofb@math.ucl.ac.uk and sjs@math.wm.edu

Received 22 April 2003, in final form 2 April 2004
Published 7 May 2004
Online at stacks.iop.org/Non/17/1393
DOI: 10.1088/0951-7715/17/4/014

Recommended by J A Glazier

Abstract
Ecological vector fields $\dot{x}_i = x_i f_i(x)$ on the non-negative cone $\mathbf{R}^n_+$ on $\mathbf{R}^n$ are often used to describe the dynamics of n interacting species. These vector fields are called permanent (or uniformly persistent) if the boundary $\partial\mathbf{R}^n_+$ of the non-negative cone is repelling. We construct an open set of ecological vector fields containing a dense subset of permanent vector fields and containing a dense subset of vector fields with attractors on $\partial\mathbf{R}^n_+$. In particular, this construction implies that robustly permanent vector fields are not dense in the space of permanent vector fields. Hence, verifying robust permanence is important. We illustrate this result with ecological vector fields involving five species that admit a heteroclinic cycle between two equilibria and the Hastings–Powell teacup attractor.

Mathematics Subject Classification: 37C20, 37C29, 37N25

(Some figures in this article are in colour only in the electronic version)

1. Introduction

The dynamics of n interacting species are often described using ecological vector fields on the non-negative cone $\mathbf{R}^n_+$ of $\mathbf{R}^n$

$$\dot{x}_i = x_i f_i(x) = F_i(x) \tag{1}$$

where $x = (x_1, \ldots, x_n)$ is the vector of population densities and $f = (f_1, \ldots, f_n)$ is the vector of per-capita growth rates. A fundamental question in ecology and conservation biology is what are the 'minimum' conditions for the long-term persistence and adaption of a collection of species in a given place. The long-term persistence for ecological vector

[3] Current address: Department of Mathematics, University College London, London WC1E 6BT, UK.

fields corresponds (at the very least) to an attractor bounded away from extinction. A strong form of persistence occurs when this attractor is a global attractor. Equivalently, the boundary $\partial\mathbf{R}^n_+$ of the non-negative cone is repelling. When this occurs, the system is *permanent* [16, 27]. While populations persisting at a local attractor may only recover from gentle stirrings of the population state, permanence ensures that populations recover from vigorous shake-ups [17].

In this paper, we present an open class of ecological vector fields for which species persistence remains undecided. More specifically, within this class, systems with repelling boundaries are dense and systems with attractors on the boundaries are dense. These vector fields have a heteroclinic cycle in $\partial\mathbf{R}^n_+$ between equilibria and a chaotic set. Moreover, as they contain no robustly permanent vector fields (i.e. ecological vector fields contained within an open set of permanent vector fields), they provide a counterexample to a conjecture of Jansen and Sigmund [17] about the density of robustly permanent vector fields and help clarify the difficulty in attempts to characterize robustly permanent vector fields [7, 26].

The remainder of the paper is structured as follows. In section 2, we state basic definitions and state the main results (theorem 1 and corollary 1). In section 3, we prove theorem 1. In section 4, we introduce ecological vector fields involving five species of prey, predator, and top predator. The top predators engage in intraguild predation [15, 24], which is shown to result in heteroclinic cycles between equilibria and the Hastings–Powell 'teacup' attractor [9]. In section 5, we discuss our results and provide a few closing remarks.

2. Definitions and main results

Let $\mathbf{R}^n_+$ denote the non-negative orthant of $\mathbf{R}^n$, $\partial\mathbf{R}^n_+$ the boundary of the non-negative orthant, and int $\mathbf{R}^n_+$ the positive orthant. Let $(x, t) \mapsto \phi_t x$ denote the flow of (1). Given a subset $S \subset \mathbf{R}^n_+$, $\alpha(S) = \bigcap_{t<0} \overline{\bigcup_{s\leqslant t}\{\phi_s x : x \in S\}}$ is the α-limit set of S and $\omega(S) = \bigcap_{t>0} \overline{\bigcup_{s\geqslant t}\{\phi_s x : x \in S\}}$ is the ω-limit set of S. A compact set $A \subset \mathbf{R}^n_+$ is an *attractor* if there exists an open neighbourhood U of A such that $\omega(U) = A$. A compact set $R \subset \mathbf{R}^n_+$ is a *repellor* if it is an attractor for the backward flow $(x, t) \mapsto \phi_{-t}x$. (1) is *dissipative* if there exists a compact attractor $A \subset \mathbf{R}^n_+$ such that $\emptyset \neq \omega(x) \subset A$ for all $x \in \mathbf{R}^n_+$. A set S is *invariant* if $\phi_t(S) = S$ for all $t \in \mathbf{R}$. A compact invariant set S is *isolated* if there exists a neighbourhood U of S such that the maximal compact invariant set in U is S. Given a point $x \in \mathbf{R}^n_+$, $W^s(x) = \{y \in \mathbf{R}^n_+ : \omega(y) = x\}$ (respectively $W^u(x) = \{y \in \mathbf{R}^n_+ : \alpha(y) = x\}$) is the *stable (respectively unstable) set of* x.

Let $\mathcal{K}^r(n)$ denote the space of C^r ecological vector fields F on $\mathbf{R}^n_+$ that are dissipative. We endow $\mathcal{K}^r(n)$ with the C^r Whitney topology [10].

In this section, we consider ecological vector fields with a heteroclinic cycle between an equilibrium and a boundary attractor as illustrated in figure 1(*a*). While these ideas can be generalized easily to a heteroclinic cycle between several equilibria and an attractor as illustrated in figure 1(*b*), for expositional clarity we focus on the simpler case. More precisely, we consider ecological vector fields such that $n \geqslant 3$ (later $n \geqslant 5$) and

A1. $r \geqslant 1$,

A2. the invariant subsystem $x_1 = x_2 = 0$ contains two attractors, a linearly stable equilibrium p and an attractor A,

A3. $e_p := f_1(p) > 0$ and $c_p := -f_2(p) > 0$,

A4. $c_\mu := -\int_A f_1\, \mathrm{d}\mu > 0$ and $e_\mu := \int_A f_2\, \mathrm{d}\mu > 0$ for all (ergodic) invariant probability measures μ with support in A, and

A5. the one-dimensional outset $\gamma = \{x \in \mathbf{R}^n_+ : x_2 = 0, \alpha(x) = p, x \neq p\}$ of p in the face $\{x : x_2 = 0\}$ converges to A, i.e. $\omega(\gamma) \subseteq A$. The outset of A in the face $\{x : x_1 = 0\}$ converges to p, i.e. every orbit γ' in that face with $\alpha(\gamma') \subseteq A$ has $\omega(\gamma') = p$.

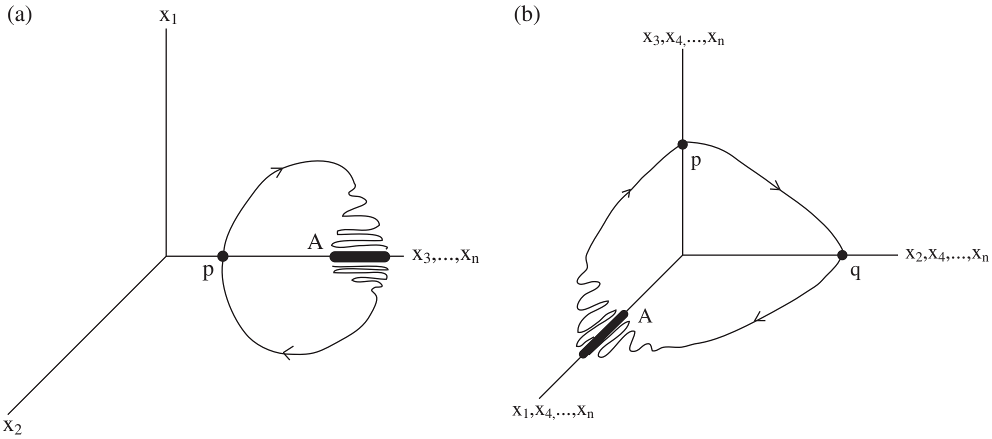

Figure 1. (*a*) Schematic of the heteroclinic cycle studied in lemma 1 and theorems 1 and 2; (*b*) schematic of another heteroclinic cycle discussed in section 4.

Let Γ denote the union of p, A, and all the connecting orbits described in A5. Let $U \subset \{x : x_1x_2 = 0\}$ be a forward invariant, isolating neighbourhood of Γ in the faces $\{x : x_1x_2 = 0\}$. We say that Γ is a C^r *robust attractor* if there is a neighbourhood of (1) in $\mathcal{K}^r(n)$ such that for all vector fields in this neighbourhood, the maximal compact invariant set in U is an attractor in $\mathbf{R}^n_+$. We say that Γ is *unsaturated* if Γ is isolated and $W^s(\Gamma) \subset \partial\mathbf{R}^n_+$. We say that Γ is C^r *robustly unsaturated* if there is a neighbourhood of (1) in $\mathcal{K}^r(n)$ such that for all vector fields in this neighbourhood, the maximal compact invariant set in U is unsaturated. The following lemma provides sufficient conditions that ensure Γ is robustly unsaturated or robustly attracting in $\mathbf{R}^n_+$.

Lemma 1. *Assume A1–A5 hold. If*

$$e_\mu e_p > c_\mu c_p \tag{2}$$

for all (ergodic) invariant probability measures μ with support in A, then Γ is C^r robustly unsaturated in $\mathbf{R}^n_+$. On the other hand, if

$$e_\mu e_p < c_\mu c_p \tag{3}$$

for all (ergodic) invariant probability measures μ with support in A, then Γ is a C^r robust attractor in $\mathbf{R}^n_+$.

Proof. We give two proofs of this result. The first proof is based on work of the second author [26]. Theorem 4.3 in [26] proves that a compact isolated invariant set Γ for the flow restricted to the boundary is robustly unsaturated if

$$\max_{1\leqslant i\leqslant n} \int_\Gamma f_i \,\mathrm{d}\upsilon > 0$$

for all invariant probability measures υ with support in Γ. Since an invariant probability measure υ supported on Γ is a convex combination of the Dirac measure, δ_p, supported by p and an invariant probability measure, μ, supported on A, Γ is robustly unsaturated if

$$\max_{1\leqslant i\leqslant n} \alpha f_i(p) + (1-\alpha) \int_A f_i \,\mathrm{d}\mu > 0$$

for all $\alpha \in [0, 1]$ and for all invariant probability measures μ with support in A. Lemma 5.1 of [26] implies that $\int_A f_i \, d\mu = 0$ and $f_i(p) = 0$ for all $i \geqslant 3$. Thus, Γ is robustly unsaturated if

$$\max_{i=1,2} \left[\alpha f_i(p) + (1-\alpha) \int_A f_i \, d\mu \right] > 0$$

for all $\alpha \in [0, 1]$ and for all invariant probability measures μ with support in A. This condition for being robustly unsaturated is easily seen to be equivalent to (2). Now, assume that (3) holds. The arguments in the previous paragraph applied to the backward flow of (1) imply that Γ is robustly unsaturated for the backward flow of (1). Since Γ is a repellor for the backward flow restricted to $\partial \mathbf{R}^n_+$, Γ is *isolated* (i.e. there exists a neighbourhood V of Γ such that Γ is the maximal compact invariant set in V) and $W^s(\Gamma) = \Gamma$ for the backward flow. These observations imply that Γ is a repellor for the backward flow. Therefore, Γ is an attractor for the forward flow of (1). Since these arguments are robust to structural perturbations within $\mathcal{K}^r$, Γ is a robust attractor for (1).

An alternative proof is based on [7]. By theorem 5.2 in [7], Γ is robustly unsaturated if for some positive constants $\beta_i > 0$ $(i = 1, \ldots, n)$ the function $\prod_i x_i^{\beta_i}$ is an average Lyapunov function for (1) near Γ. By theorem 5.3 in [7] this is equivalent to

$$\sum_{i=1}^{n} \beta_i \int_\Gamma f_i \, d\upsilon > 0$$

for all (ergodic) invariant probability measures υ with support in Γ. Using as above $\int_\Gamma f_i \, d\upsilon = 0$ for $i \geqslant 3$, this reduces to

$$\beta_1 f_1(p) + \beta_2 f_2(p) > 0 \qquad \beta_1 \int_A f_1 \, d\mu + \beta_2 \int_A f_2 \, d\mu > 0$$

for all (ergodic) invariant probability measures μ with support in A. From the sign structure required in A3–A4, this is seen easily to be equivalent to (2). The second part can be shown in a similar way. ■

If the attractor A carries a unique invariant probability measure μ, as in the case of an equilibrium, a periodic orbit, or a quasi-periodic torus, then lemma 1 settles the behaviour near Γ generically. Whenever A is not uniquely ergodic, we prove that there remains an open set of systems for which permanence is undecided. This open set is characterized by the following additional assumptions:

A6. There are two ergodic probability measures $\mu^\pm$ with support in A such that

$$e_p e_{\mu^+} > c_p c_{\mu^+} \qquad e_p e_{\mu^-} < c_p c_{\mu^-}$$

A7. For all ergodic probability measures μ with support in A,

$$-c_\mu < \lambda^-_\mu < 0 < \lambda^+_\mu < e_\mu$$

where λ^+_μ (respectively λ^-_μ) denotes the largest (respectively smallest) Lyapunov exponent for (1) restricted to $\{x : x_1 = x_2 = 0\}$.

A8. A is *transitive* (i.e. has a dense orbit), is the closure of its periodic orbits, and is *hyperbolic*: there exists a continuous $D\phi_t$-invariant splitting of $TA = E^s \oplus E^c \oplus E^u$ and constants $\epsilon > 0$, $\lambda > 0$ such that E^c is a one-dimensional bundle tangent to the flow of (1),

$$\|D\phi_t v\| \leqslant \exp(-\lambda t + \epsilon)\|v\|$$

for all $v \in E^s$ and $t \geqslant 0$, and

$$\|D\phi_t v\| \geqslant \exp(\lambda t - \epsilon)\|v\|$$

for all $v \in E^u$ and $t \geqslant 0$.

A9. $r_p > c_p$, where $-r_p$ is the largest real part of an eigenvalue of $DF(p)$ for (1) restricted to $\{x : x_1 = x_2 = 0\}$.

Theorem 1. *The set of ecological vector fields satisfying A1–A9 is open. Furthermore, in this open set, the vector fields for which* Γ *is attracting are dense and the vector fields for which* Γ *is unsaturated are dense.*

A vector field in $\mathcal{K}^r(n)$ is *permanent* (equivalently *uniformly persistent*) [2, 27] if there exists a compact global attractor in the positive orthant. A vector field in $\mathcal{K}^r(n)$ is C^r *robustly permanent* if this vector field is contained in an open neighbourhood of permanent vector fields. As a corollary of theorem 1, we can prove the existence of an open set of ecological vector fields in which permanent and non-permanent vector fields are dense. In particular, this class of vector fields provides a counterexample to Jansen and Sigmund's conjecture that all permanent vector fields can be approximated by robustly permanent vector fields [17].

Recall that a collection of sets $\{M_1, \ldots, M_k\}$ is a *Morse decomposition* for a compact invariant set K if $M_1, \ldots, M_k$ are pairwise disjoint, compact isolated invariant sets for $\phi|K$ with the property that for each $x \in K$ there are integers $l = l(x) \leqslant m = m(x)$ such that $\alpha(x) \subseteq M_m$ and $\omega(x) \subseteq M_l$ and if $l = m$ then $x \in M_l = M_m$.

Corollary 1. *The set of ecological vector fields satisfying A1–A9 and*

A10. there is a Morse decomposition $\{M_1, M_2, \ldots, M_l = \Gamma\}$ *of the global attractor of* $\partial\mathbf{R}^n_+$ *such that* $M_1, \ldots, M_{l-1}$ *are robustly unsaturated*

is open. Furthermore, in this open set, permanent vector fields are dense, and non-permanent vector fields are dense.

Proof. Since Morse decompositions are robust to structural perturbations, theorem 1 implies that ecological vector fields satisfying A1–A10 are open in $\mathcal{K}^r(n)$. Theorem 1 implies that there is a dense set of ecological vector fields satisfying A1–A10 such that Γ is an attractor. Hence, vector fields in this dense set are not permanent. Alternatively, theorem 1 implies there exists a dense set of vector fields satisfying A1–A10 such that Γ is unsaturated. For these vector fields, there is a Morse decomposition of $\partial\mathbf{R}^n_+$ for which each of the components of the decomposition is unsaturated. The work of Garay [6] and Hofbauer [13] implies that these vector fields are permanent. ∎

Theorem 1 and corollary 1 can be extended easily to heteroclinic cycles between multiple equilibria and an attractor as illustrated in figure 1(*b*). This type of situation arises quite naturally in ecologically plausible scenarios as we illustrate in section 4.

3. Proof of theorem 1

Let (1) satisfy A1–A9. For notational convenience, we let $(x, t) \mapsto x \cdot t$ and $\phi_t x$ denote the flow of (1). We begin by showing why A1–A9 are open conditions in $\mathcal{K}^r(n)$. Assumptions A1–A3 together with A9 are clearly open conditions. A standard result in ergodic theory (see, e.g. [21]) implies that A4 is equivalent to the existence of $T > 0$ such that

$$\sup_{x\in A} \int_0^T f_1(x \cdot t)\,\mathrm{d}t < 0 \qquad \inf_{x\in A} \int_0^T f_2(x \cdot t)\,\mathrm{d}t > 0.$$

This reformulation of A4 is an open condition as A is compact. A5 is an open condition as A2 and A4 imply that A is an attractor in the $x_2 = 0$ subspace. A result of Sigmund [28, theorem 1] and A8 imply that invariant measures supported by periodic orbits of A are dense in the set of invariant probability measures supported by A. This fact combined with A6 implies that there exist periodic points $q^{\pm}$ in A such that

$$e_p \int_0^{T^-} f_2(q^- \cdot t)\,\mathrm{d}t < -c_p \int_0^{T^-} f_1(q^- \cdot t)\,\mathrm{d}t \tag{4}$$

and

$$e_p \int_0^{T^+} f_2(q^+ \cdot t)\,\mathrm{d}t > -c_p \int_0^{T^+} f_1(q^+ \cdot t)\,\mathrm{d}t \tag{5}$$

where $T^{\pm}$ are the periods of $q^{\pm}$. Since the periodic orbits $q^{\pm}$ are hyperbolic, these inequalities persist under perturbations. To see that A7 is an open condition, define $h : \mathbf{R} \times A \to \mathbf{R}$ by

$$h(t, x) = \ln \| D\phi_t|_{\{x:x_1=x_2=0\}} \| - \int_0^t f_2(\phi_s x)\,\mathrm{d}s$$

where $\phi_t x = x \cdot t$. By the subadditive and multiplicative ergodic theorems [18, 22], for any ergodic probability measure μ,

$$\lambda_\mu^+ - e_\mu = \inf_{t>0} \frac{1}{t} \int_A h(t, x)\,\mathrm{d}\mu(x).$$

A result of the second co-author [25, theorem 1] implies that the right-hand inequality of A7 is equivalent to the existence of $T > 0$ such that $\sup_{x \in A} h(T, x) < 0$. This reformulation of the right-hand inequality is an open condition as A is compact. Similarly, the remaining inequalities of A7 can be shown to form an open condition. The stability of hyperbolic sets [11, 29] and assumption A8 imply that for vector fields G sufficiently C^1 close to F, the maximal invariant set $A(G)$ of G near A is hyperbolic and the flow of G restricted to $A(G)$ is topologically conjugate to the flow of F restricted to A. In particular, $A(G)$ is transitive and is the closure of the periodic orbits for G restricted to $A(G)$. Hence, vector fields satisfying A1–A9 are open in $\mathcal{K}^r(n)$.

To prove the density of vector fields with Γ attracting (respectively repelling), we proceed in three steps. First, we make three C^r small perturbations of (1) that make the vector field smoothly linearizable near p and smoothly linearizable near the periodic orbit q^- (respectively q^+) and connect the outset of p to the stable manifold of q^- (respectively q^+). Second, we erect Poincaré sections near p and the orbit of q^- and study the behaviour of the associated Poincaré maps. Third, we use the information from the Poincaré maps to deduce Γ is an attractor (respectively repellor).

Step 1. *Three perturbations of (1).* In each of the following three perturbations of (1), we assume the perturbations are sufficiently small to ensure that A1–A9 hold for the resulting vector field. First, we C^r perturb (1) in a neighbourhood of p such that the eigenvalues $\lambda = (\lambda_1, \ldots, \lambda_n)$ of $DF(p)$ are distinct and non-resonant (e.g. for any i, $\lambda_i \neq \sum \lambda_i m_i$, where m_i are non-negative integers that satisfy $\sum m_i \geqslant 2$). The Sternberg linearization theorem (see, e.g. [8]) and the invariance of the faces of $\mathbf{R}^n_+$ imply there is a smooth change of coordinates (x_1, x_2, z) of $\mathbf{R} \times \mathbf{R} \times \mathbf{R}^{n-2}$ such that p corresponds to $(0, 0, 0)$ in this coordinate system and such that the flow near p is given by

$$(x_1, x_2, z) \cdot t = (\exp(e_p t)x_1, \exp(-c_p t)x_2, \exp(-R_p t)z) \tag{6}$$

where $e_p = f_1(p) > 0$, $-c_p = f_2(p) < 0$, and R_p is an $(n-2) \times (n-2)$ block diagonal matrix whose smallest diagonal entry equals $r_p > 0$, see A9, and whose blocks are either 1×1 blocks or 2×2 blocks of the form $[\begin{smallmatrix} a & b \\ -b & a \end{smallmatrix}]$.

Second, let $q = q^-$ (respectively q^+). We will C^r perturb (1) in a neighbourhood of $\mathrm{o}(q) := \{q \cdot t : t \in \mathbf{R}\}$ such that the Lyapunov exponents of this orbit are distinct and non-resonant. Let $k+1$ equal the dimension of the stable manifold of $o(q)$. The Sternberg linearization theorem and the invariance of the faces of $\mathbf{R}^n_+$ imply there is a smooth change of coordinates $(x_1, x_2, \theta, s, u) \in \mathbf{R} \times \mathbf{R} \times S^1 \times \mathbf{R}^k \times \mathbf{R}^{n-3-k}$ in a neighbourhood of $o(q)$ such that with respect to these coordinates $(0,0) \times S^1 \times (0,0)$ equals $o(q)$ and such that the local flow is given by

$$(x_1, x_2, \theta, s, u) \cdot t = (\exp(-c_q t)x_1, \exp(e_q t)x_2, \theta + t, \exp(-R_q t)s, \exp(Et)u) \tag{7}$$

where $c_q > 0$, $e_q > 0$, R_q is a block diagonal matrix whose smallest diagonal entry equals $r_q > 0$, E is a block diagonal matrix with positive diagonal entries, the largest of which equals $\eta > 0$, and the blocks of R_q and E are either 1×1 blocks or 2×2 blocks of the form $[\begin{smallmatrix} a & b \\ -b & a \end{smallmatrix}]$.

Third, we C^r perturb (1) within the face $\{x : x_2 = 0\}$ to connect the outset of p to the stable manifold of q. Let U in $\{x : x_1 = x_2 = 0\}$ be a forward invariant neighbourhood of A. By taking U sufficiently small, A7 implies that we can construct a vector field F' such that (1) $F' = F$ in a neighbourhood $V \subset \mathbf{R}^n_+$ of A and (2) U is a normally hyperbolic (non-compact) manifold for F'. By shrinking V as necessary, the theory of normally hyperbolic (non-compact) manifolds [12, theorem 6.1] implies that the flow of F' (hence F) restricted to V is topologically conjugate to the normally linearized flow

$$(x_1, x_2, z) \cdot t = \left(\exp\left(\int_0^t f_1((0,0,z)\cdot t)\right)x_1, \exp\left(\int_0^t f_2((0,0,z)\cdot t)\right)x_2, z \cdot t\right).$$

Since the outset of p converges to A, the outset enters the neighbourhood V. In the local coordinate system, let $(x_1^o, 0, z^o)$ with $x_1^o > 0$ denoting a point of the outset of p. The local product structure and transitivity of A imply that the stable manifold of $o(q)$ is dense in A. Hence, there exists a point $(0, 0, z^s)$ in the stable manifold of $o(q)$ arbitrarily close to $(0, 0, z^o)$. Perturb (1) within a flow box in the face $\{x_2 = 0\}$ around the points $(x_1^o, 0, z^o)$ and $(x_1^o, 0, z^s)$ (see, e.g. the construction of the perturbation used in lemma 2.4 in [23]) so that the outset of p is contained in the stable manifold of $o(q)$.

Step 2. *Poincaré sections and maps.* To understand the behaviour of the system near Γ, we define appropriate Poincaré sections and determine the corresponding Poincaré first return maps. With respect to the local coordinates (x_1, x_2, z) around p, define Poincaré sections as follows:

$$H_1^{\text{in}} = \{(x_1, 1, z) : x_1^2 + \|z\|^2 \leqslant 1\}$$

and

$$H_1^{\text{out}} = \{(1, x_2, z) : x_2^2 + \|z\|^2 \leqslant 1\}.$$

Using the linearization (6) about p, we get the following Poincaré map, ϕ_1, from a neighbourhood of $(0, 1, 0)$ in H_1^{in} into H_1^{out}:

$$\phi_1(x_1, 1, z) = (1, x_1^{c_p/e_p}, x_1^{(r_p-\epsilon)/e_p} A_1(x_1)z)$$

where $A_1(x_1)$ is a block diagonal matrix with entries that depend continuously on x_1 for any $0 < \epsilon \leqslant \min\{r_p - c_p, r_q, e_q - \eta\}$. Assumptions A7 and A9 guarantee that $\epsilon > 0$ exists.

With respect to the local coordinates (x_1, x_2, θ, s, u) around $o(q)$, define Poincaré sections by

$$H_2^{\text{in}} = \{1\} \times [0, 1] \times S^1 \times [-1, 1]^{n-3}$$

and

$$H_2^{\text{out}} = [0, 1] \times \{1\} \times S^1 \times [-1, 1]^{n-3}.$$

Using the linearization (7) about $o(q)$, the Poincaré map ϕ_2 from a neighbourhood of $(1, 0) \times S^1 \times (0, 0)$ in H_2^{in} into H_2^{out} is given by

$$\phi_2(1, x_2, \theta, s, u) = (x_2^{c_q/e_q}, 1, \theta + \tau(x_2), x_2^{(r_q-\epsilon)/e_q} A_2(x_2)s, x_2^{-(\eta+\epsilon)/e_q} A_3(x_2)u)$$

where $\tau(x_2) = -(1/e_q)\ln x_2$ and $A_i(x_2)$ $(i = 2, 3)$ are block diagonal matrices whose entries depend continuously on x_2. Note: we used the fact that r_q is the smallest diagonal entry of R_q and η is the largest diagonal entry of E, and the choice of ϵ.

Next, we need the connecting maps. With respect to the local coordinates near γ (i.e. the one-dimensional outset of p in the face $\{x : x_2 = 0\}$), let $(1, 0, \theta^*, s^*, 0)$ denote the point $\gamma \cap H_2^{\text{in}}$. Then the map ψ_1 from a neighbourhood of $(1, 0, 0)$ in H_1^{out} to a neighbourhood of $(1, 0, \theta^*, s^*, 0)$ in H_2^{in} is given by

$$\psi_1(1, x_2, z) = (1, a_1(x_2, z)x_2, \theta^* + \Theta_1(x_2, z), s^* + S_1(x_2, z), U_1(x_2, z))$$

where $a_1 > 0$, Θ_1, S_1, and U_1 are continuous functions. The last three are of order 1, i.e. they satisfy an estimate $|g(x_2, z)| \leqslant C(|x_2| + |z|)$ for some positive constant C. For each $\theta \in S^1$, let $z^*(\theta)$ be such that $(0, 1, z^*(\theta)) \in H_1^{\text{in}}$ equals the intersection of the forward orbit of $(0, 1, \theta, 0, 0) \in H_2^{\text{out}}$ with H_1^{in}. The connecting map ψ_2 from a neighbourhood $\{0\} \times \{1\} \times S^1 \times \{0\} \times \{0\}$ in H_2^{out} to a neighbourhood of $(0, 1, z^*(S^1))$ in H_1^{in} is given by

$$\psi_2(x_1, 1, \theta, s, u) = (a_2(x_1, \theta, s, u)x_1, 1, z^*(\theta) + Z_1(x_1, \theta, s, u))$$

where $a_2 > 0$ and Z_1 are continuous functions. Z_1 is of order 1 in (x_1, s, u).

To get the Poincaré return map for the Poincaré section H_1^{in}, we need to compose these four maps. Composing the first two, $\psi_1 \circ \phi_1$, and using the relation that $r_p - \epsilon > c_p$ (cf our choice of ϵ) yields

$$(x_1, 1, z) \mapsto (1, B_2(x_1, z)x_1^{c_p/e_p}, \theta^* + x_1^{c_p/e_p}\Theta_2(x_1, z), s^* + x^{c_p/e_p}S_2(x_1, z), x_1^{c_p/e_p}U_2(x_1, z))$$

where $B_2 > 0$ and Θ_2, S_2, and U_2 are continuous. Composing the first three, $\phi_2 \circ \psi_1 \circ \phi_1$, gives

$$(x_1, 1, z) \mapsto (B_3(x_1, z)x_1^{c_pc_q/e_pe_q}, 1, \Theta_3(x_1, z), x_1^{c_p(r_q-\epsilon)/e_pe_q}S_3(x_1, z), x_1^{c_p/e_p(1-(\eta+\epsilon)/e_q)}U_3(x_1, z))$$

where $B_3 > 0$, S_3 and U_3 are continuous for all (x_1, z), and Θ_3 is continuous for $x_1 > 0$. Composing one last time with ψ_2 gives us the desired return map for H_1^{in}:

$$(x_1, 1, z) \mapsto (B(x_1, z)x_1^{c_pc_q/e_pe_q}, 1, z^*(\Theta(x_1, z)) + x_1^{\alpha}Z(\Theta(x_1, z), x_1, z)) \tag{8}$$

where $B > 0$ is continuous, $\alpha = \min\{c_pc_q/e_pe_q, (c_p/e_p)(1-(\eta+\epsilon)/e_q), c_p(r_q - \epsilon)/e_pe_q\} > 0$ by our choice of ϵ, Θ is continuous for $x_1 > 0$, and Z is continuous.

Step 3. *Putting it all together.* Let us begin with the case of approximating (1) by a vector field such that Γ is an attractor. In this case, we use $q = q^-$ in steps 1 and 2. Equation (4) implies that $c_pc_q > e_pe_q$, in which case $(0, 1, z^*(S^1))$ is an attractor for the Poincaré map defined by (8). Since any positive orbit starting sufficiently close to Γ passes through the Poincaré section H_1^{in}, it follows that Γ is an attractor. Alternatively, to approximate (1) by a vector field such that Γ is unsaturated, we apply steps 1 and 2 with $q = q^+$. Equation (5) implies that $c_pc_q < e_pe_q$, in which case, $(0, 1, z^*(S^1))$ is unsaturated for the Poincaré map defined by (8). Since any positive orbit starting sufficiently close to Γ passes through the Poincaré section H_1^{in}, it follows that Γ is unsaturated.

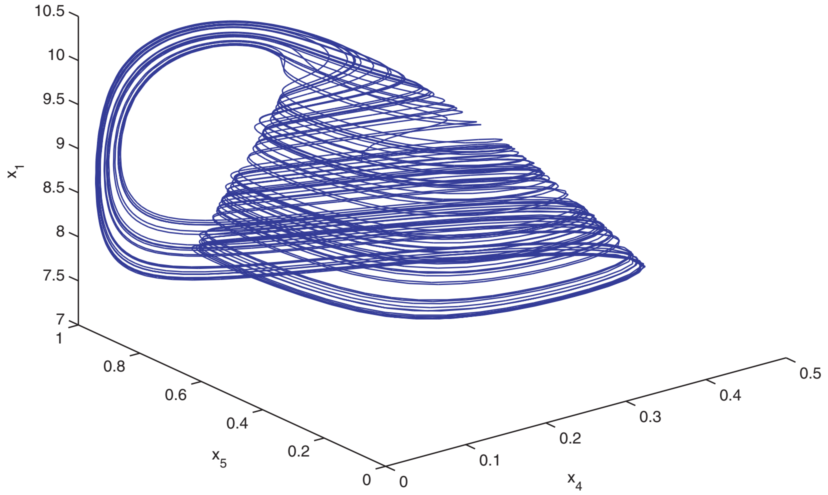

Figure 2. The teacup attractor of Hastings and Powell for the x_1–x_4–x_5 food chain with parameters $a_5 = 5.0$, $a_4 = 0.1$, $b_5 = 3.0$, $b_4 = 2.0$, $d_4 = 0.4$, $d_1 = 0.01$, and $\theta_4 = \theta_5 = 1$.

4. A motivating example

In this section, we introduce an explicit system of ecological equations for which lemma 1 does apply and for which an appropriate generalization of theorem 1 might apply. We consider a system consisting of five species: a prey (species 5), a predator (species 4), and three top predators (species 1, 2 and 3). One of the top predators (species 1) is an intraguild predator [15,24]—a top predator that also predates on another top predator (species 2). In order to generate a chaotic attractor, we assume that species 4 and 1 have saturating functional responses with respect to their primary prey. For the sake of simplicity, we assume that all the other functional responses are linear and consider a model of the following form,

$$\begin{aligned}
\frac{dx_1}{dt} &= \frac{\theta_4 a_4 x_4 x_1}{1 + b_4 x_4} - d_1 x_1 + \theta_2 a_2 x_1 x_2 \\
\frac{dx_2}{dt} &= \theta_4 a_4 x_2 x_4 - d_2 x_2 - a_2 x_1 x_2 \\
\frac{dx_3}{dt} &= \theta_4 a_4 x_3 x_4 - d_3 x_3 \\
\frac{dx_4}{dt} &= \frac{\theta_5 a_5 x_5 x_4}{1 + b_5 x_5} - d_4 x_4 - \frac{a_4 x_1 x_4}{1 + b_4 x_4} - a_4 (x_2 + x_3) x_4 \\
\frac{dx_5}{dt} &= x_5 (1 - x_5) - \frac{a_5 x_5 x_4}{1 + b_5 x_5}
\end{aligned} \tag{9}$$

where x_i is the density of species i, the a_i are proportional to the rate at which the predator encounters the prey species i, the b_i are proportional to the handling times of the predators, the d_i are proportional to the *per capita* mortality rates of the predators, and the θ_i convert prey eaten to predators produced.

To generate a chaotic attractor for the food chain consisting of species 1, 4, and 5, we choose the parameters found in Hastings and Powell (1991): $a_5 = 5.0$, $a_4 = 0.1$, $b_5 = 3.0$, $b_4 = 2.0$, $d_4 = 0.4$, $d_1 = 0.01$, and $\theta_4 = \theta_5 = 1$. This yields the 'teacup' attractor illustrated in figure 2. While there exists no proof of this attractor being hyperbolic or singular hyperbolic, Deng and Hines [3–5] using singular perturbation theory have proven for different parameter values that food chains can exhibit attractors containing a shift on two symbols. For the other food chains, choose $d_2 = 0.001$ and $d_3 = 0.01$. For these parameter values, the subsystems $\{x : x_1 = x_2 = 0\}$ and $\{x : x_1 = x_3 = 0\}$ have a unique positive asymptotically stable equilibrium. Call these equilibria p and q, respectively. Since $d_1 = d_3$, we have

$$f_3(x) = \theta_4 a_4 x_4 - d_3 > \frac{\theta_4 a_4 x_4}{1 + b_4 x_4} - d_1 = f_1(x)$$

whenever $x_4 > 0$. Hence,

$$\lim_{t\to\infty} (x \cdot t)_1 = 0$$

for any $x = (x_1, x_2, x_3, x_4, x_5)$ with $x_2 = 0$ and $x_3 x_4 x_5 > 0$. In other words, species 3 displaces species 1 (figure 3(*a*)). Similarly, since $d_2 < d_3$, species 2 displaces species 3, i.e.

$$\lim_{t\to\infty} (x \cdot t)_3 = 0$$

for any $x = (x_1, x_2, x_3, x_4, x_5)$ with $x_1 = 0$ and $x_2 x_4 x_5 > 0$. To complete a heteroclinic cycle between the equilibria and the attractor as illustrated in figure 1(*b*), we need to choose a_2 and $\theta_2 a_2$ sufficiently large to ensure that species 1 displaces species 2. For instance, simulations (figure 3(*b*)) suggest that $a_2 \geqslant 0.01$ and $\theta_2 a_2 \geqslant 0.0025$ suffice.

To understand the attracting/repelling properties of the heteroclinic cycle Γ formed by p, q, and A, define the expansion and contraction rates at q and p:

$$\begin{aligned}
e_q &= f_1(q) = \frac{d_2}{1 + (b_4 d_2/a_4\theta_4)} - d_1 + \theta_2 a_2 q_2 \\
c_q &= -f_3(q) = d_3 - d_2 \\
e_p &= f_2(p) = d_3 - d_2 \\
c_p &= -f_1(p) = d_1 - \frac{d_3}{1 + (b_4 d_3/a_4\theta_4)}.
\end{aligned}$$

Notice that $c_q = e_p$. Given any invariant probability measure supported on A, define

$$\begin{aligned}
e_\mu &= \int_A f_3(x)\,\mathrm{d}\mu(x) = \theta_4 a_4 \int_A x_4\,\mathrm{d}\mu(x) - d_3 \\
c_\mu &= -\int_A f_2(x)\,\mathrm{d}\mu(x) = d_2 + a_2 \int_A x_1\,\mathrm{d}\mu(x) - \theta_4 a_4 \int_A x_4\,\mathrm{d}\mu(x).
\end{aligned}$$

These parameters can be arranged conveniently in the characteristic matrix of the heteroclinic cycle Γ (see [14]):

	x_1	x_2	x_3	x_4	x_5
A	0	$-c_\mu$	e_μ	0	0
q	e_q	0	$-c_q$	0	0
p	$-c_p$	e_p	0	0	0

A modification of lemma 1 (in the spirit of theorem 1a and corollary 2 in [14]) implies that Γ is a robust attractor if

$$e_q e_p e_\mu < c_q c_p c_\mu$$

for all invariant probability measures μ supported on A. This can be achieved if θ_2 is sufficiently small (and $\theta_2 a_2$ is sufficiently large). For instance, numerical simulations (figure 4(*a*)) suggest that $\theta_2 = 0.25$ and $a_2 = 0.01$ will do. Alternatively, a modification of lemma 1 implies that Γ is robustly unsaturated if (2)

$$e_q e_p e_\mu > c_q c_p c_\mu$$

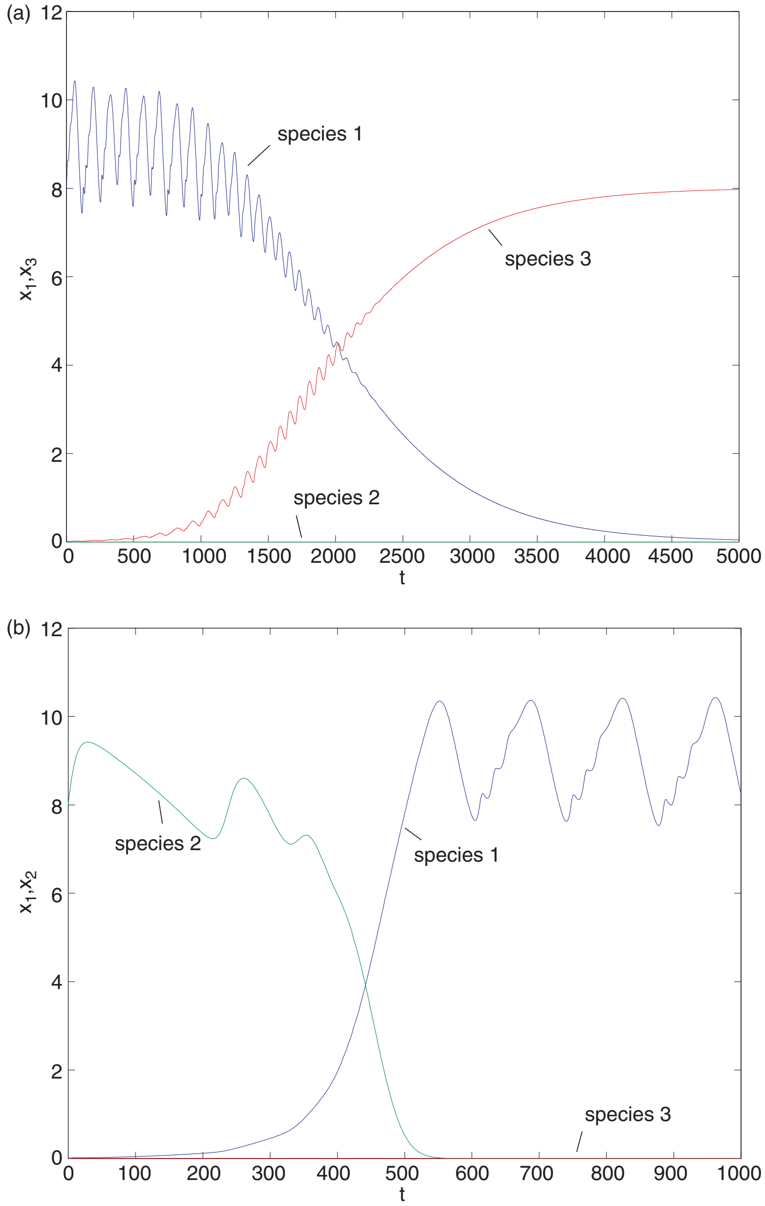

Figure 3. (*a*) Species 1 (in blue) displaced by species 3 (in red) with initial condition $x_1 = 8$, $x_3 = 0.01$, $x_4 = 0.2$, $x_5 = 0.65$. (*b*) Species 2 (in green) displaced by species 1 with $\theta_2 = 0.25$, $a_2 = 0.01$, and initial condition $x_1 = 0.01$, $x_2 = 8$, $x_4 = 0.2$, $x_5 = 0.65$.

for all invariant probability measures μ supported on A. This occurs provided that θ_2 is sufficiently large. For instance, numerical simulations (figure 4(*b*)) suggest that if $\theta_2 = 2.0$ and $a_2 = 0.01$, then Γ is unsaturated.

Now suppose there exist invariant probability measures μ^+ and μ^- supported by A such that

$$\frac{e_{\mu^+}}{c_{\mu^+}} > \frac{e_{\mu^-}}{c_{\mu^-}}.$$

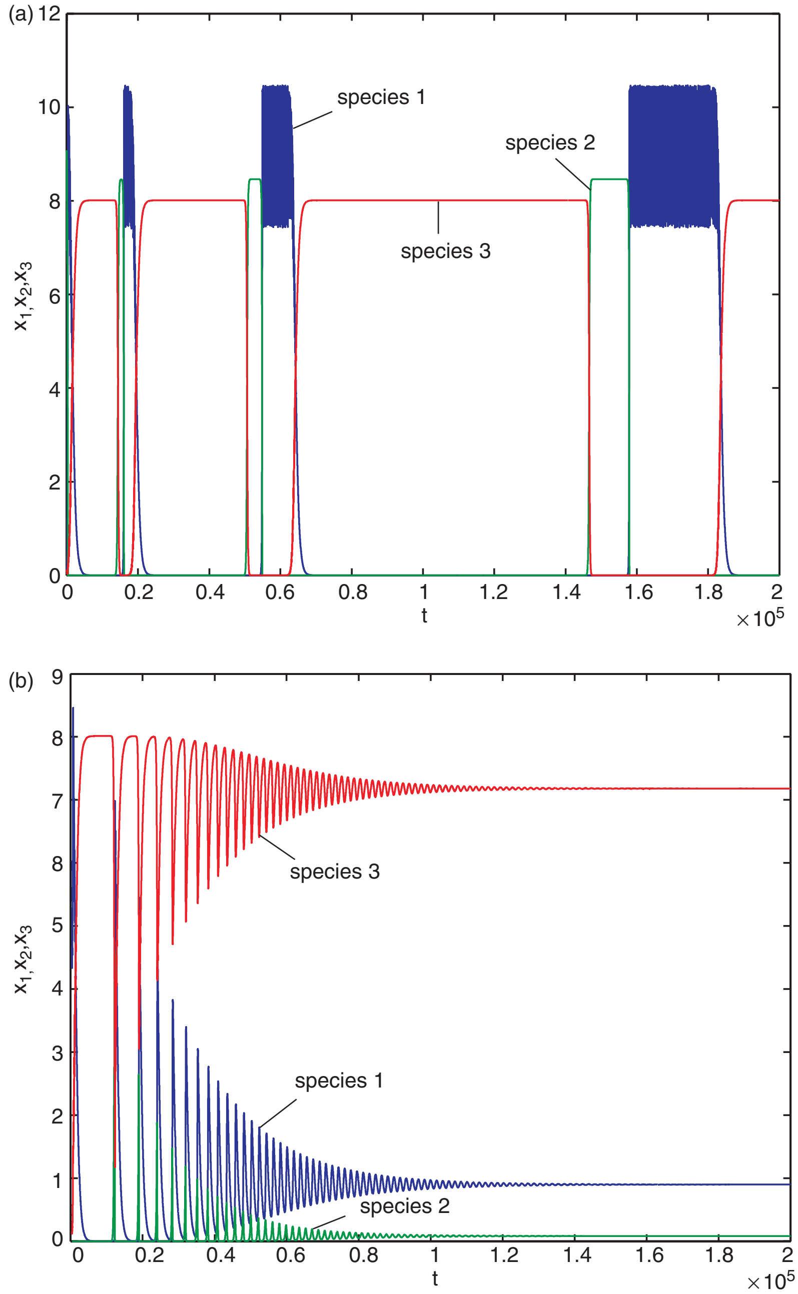

Figure 4. The dynamics of (9) for the initial condition $x_1 = 0.1$, $x_2 = 8.0$, $x_3 = 0.1$, $x_4 = 0.2$, $x_5 = 0.65$. In (*a*), $a_2 = 0.01$, $\theta_2 = 0.25$, and the heteroclinic cycle Γ appears to be attracting. Species 1 is in blue, species 2 in green, and species 3 in red. In (*b*) $a_2 = 0.01$, $\theta_2 = 2$, and the heteroclinic cycle Γ appears to be repelling.

The aforementioned numerical observations suggest that by leaving $a_2 = 0.01$, one can find a value of θ_2 between 0.25 and 2.0 such that

$$e_q e_p e_{\mu^+} > c_q c_p c_{\mu^+}$$

and

$$e_q e_p e_{\mu^-} < c_q c_p c_{\mu^-}$$

For such a choice of θ_2, persistence may be undecided for all ecological systems sufficiently close to (9) in the C^1 topology. However, verifying or refuting this assertion will require further advances in our understanding of the teacup attractor and generalizations of theorem 1.

5. Discussion

Permanence in ecological equations has been equated with long-term persistence or co-existence of species. Since any ecological equation is a crude approximation of reality, permanence is only 'observable' if it is robust to structural perturbations of the equations i.e. robustly permanent. Alternatively, as numerical methods for differential equations involve simulating perturbations of the original equation, only robustly permanent systems will appear 'numerically permanent' (see [7] for some results in this direction). Ideally, one might hope that robustly permanent equations together with their counterparts, robustly 'non-permanent' equations, are dense in the space of ecological equations. We have shown that this is not the case. There exist open sets of ecological equations for which persistence remains undecided; within these open sets, permanent ecological equations and ecological equations with a boundary attractor are hopelessly intermingled (i.e. both are dense). Therefore permanence does not persist generically under perturbations, and verifying robust permanence is important.

The basic mechanisms underlying the intermingling of permanence and non-permanent systems appear to be heteroclinic cycles with chaotic sets that generate a tension between attraction towards and repulsion from the boundary. We believe that open sets of equations of this type are likely to intersect many ecologically plausible scenarios. One scenario involving tritrophic food chains and intraguild predation was discussed in section 4. Other forms of linked top predator–predator–prey models of competing species with stage structure, competing species with discrete generations, or higher dimensional 'nonlinear' replicator equations also are likely to generate heteroclinic cycles involving chaotic sets. Given their prevalence and the unanswered questions about heteroclinic cycles involving only equilibria [1, 14, 19, 20], heteroclinic cycles involving chaotic sets are likely to taunt and tease future generations of mathematicians and ecologists.

References

[1] Brannath W 1994 Heteroclinic networks on the tetrahedron *Nonlinearity* **7** 1367–84
[2] Butler G J and Waltman P 1986 Persistence in dynamical systems *J. Diff. Eqns* **63** 255–63
[3] Deng B 2001 Food chain chaos due to junction-fold point *Chaos* **11** 514–25
[4] Deng B and Hines G 2002 Food chain chaos due to Shilnikov's orbit *Chaos* **12** 533–8
[5] Deng B and Hines G 2003 Food chain chaos due to transcritical point *Chaos* **13** 578–85
[6] Garay B M 1989 Uniform persistence and chain recurrence *J. Math. Anal. Appl.* **139** 372–82
[7] Garay B M and Hofbauer J 2003 Robust permanence for ecological differential equations, minimax, and discretizations *SIAM J. Math. Anal.* **34** 1007–39
[8] Hartman P 1982 *Ordinary Differential Equations* 2nd edn (Boston, MA: Birkhäuser)
[9] Hastings A and Powell T 1991 Chaos in a three species food chain *Ecology* **72** 896–903
[10] Hirsch M W 1976 Differential topology *Graduate Texts in Mathematics* vol 33 (New York: Springer)
[11] Hirsch M W and Pugh C C 1970 Stable maifolds and hyperbolic sets *Proc. Symp. Pure. Math.* **14** 133–63
[12] Hirsch M W, Pugh C C and Shub M 1977 *Invariant Manifolds* (*Lecture Notes in Mathematics* vol 583) (Berlin: Springer)
[13] Hofbauer J 1989 A unified approach to persistence *Acta Appl. Math.* **14** 11–22
[14] Hofbauer J 1994 Heteroclinic cycles in ecological differential equations *Tatra Mt. Math. Publ.* **4** 105–16
[15] Holt R D and Polis G A 1997 A theoretical framework for intraguild predation *Am. Nat.* **149** 745–64

[16] Hutson V and Schmitt K 1992 Permanence and the dynamics of biological systems *Math. Biosci.* **111** 1–71
[17] Jansen V A A and Sigmund K 1998 Shaken not stirred: on permanence in ecological communities *Theor. Popul. Biol.* **54** 195–201
[18] Kingman J F C 1973 Subadditive ergodic theory *Ann. Probab.* **1** 883–909
[19] Krupa M 1997 Robust heteroclinic cycles *J. Nonlinear Sci.* **7** 129–76
[20] Krupa M and Melbourne I 1995 Asymptotic stability of heteroclinic cycles in systems with symmetry *Ergod. Theory Dynam. Sys.* **15** 121–47
[21] Mañé R 1983 *Ergodic Theory and Differentiable Dynamics* (New York: Springer)
[22] Oseledec V I 1968 A multiplicative ergodic theorem *Trans. Moscow Math. Soc.* **19** 197–231
[23] Palis J and de Melo W 1982 *Geometric Theory of Dynamical Systems* (New York: Springer)
[24] Polis G A and Holt R D 1992 Intraguild predation: the dynamics of complex trophic interactions *Trends Ecol. Evol.* **7** 151–4
[25] Schreiber S J 1998 On growth rates of subadditive functions for semiflows *J. Diff. Eqns* **148** 334–50
[26] Schreiber S J 2000 Criteria for C^r robust permanence *J. Diff. Eqns* 400–26
[27] Schuster P, Sigmund K and Wolff R 1979 Dynamical systems under constant organization 3: Cooperative and competitive behavior of hypercycles *J. Diff. Eqns* **32** 357–68
[28] Sigmund K 1972 On the space of invariant measures for hyperbolic flows *Am. J. Math.* **94** 31–7
[29] Smale S 1967 Differentiable dynamical systems *Bull. Am. Math. Soc.* **73** 747–817