Journal of Dynamics and Differential Equations, Vol. 15, Nos. 2/3, July 2003 (© 2004)

On the Gap between Random Dynamical Systems and Continuous Skew Products*

Arno Berger[1] **and Stefan Siegmund**[2]

Received December 19, 2002

We review the recent notion of a nonautonomous dynamical system (NDS), which has been introduced as an abstraction of both random dynamical systems and continuous skew product flows. Our focus is on fundamental analogies and discrepancies brought about by these two principal classes of NDS. We discuss base dynamics mainly through almost periodicity and almost automorphy, and we emphasize the importance of these concepts for NDS which are generated by differential and difference equations. Nonautonomous dynamics is presented by means of selected examples. We also mention several natural yet unresolved questions.

KEY WORDS: nonautonomous dynamical systems; random dynamical systems; skew product flows; almost periodic equations; almost automorphic equations.

1. INTRODUCTION

Random dynamical systems and continuous skew product flows both describe the dynamical behavior of systems under an external influence—random and deterministic yet nonautonomous, respectively—which is inherent to the concept. In both cases, the external influence is modeled by a dynamical system (the so-called driving system) which is ergodic or continuous, respectively. Both subjects have evolved rather independently up to now, and their formal and practical analogies have been exploited only occasionally.

The theory of random dynamical systems was developed mainly by Ludwig Arnold and his "Bremen Group." As Arnold points out in [3], one of the historical gates in the development of the theory of stochastic

* Dedicated to Victor A. Pliss on the occasion of his 70th birthday.

[1] Institute of Mechanics, Vienna University of Technology, A-1040 Vienna, Austria. E-mail: arno.berger@tuwien.ac.at

[2] Department of Mathematics and Statistics, Boston University, Boston, Massachusetts 02215, USA. E-mail: siegmund@math.bu.edu

1040-7294/03/0700-0237/0

differential equations was the discovery that their solution yields a cocycle over an ergodic dynamical system which models randomness, i.e., a random dynamical system. Results on multiplicative ergodic theory, invariant manifolds, normal forms and bifurcation theory for random dynamical systems are altogether contained in the monograph by Arnold [4]. Recent results on Lyapunov's second method and monotone systems can be found in Arnold and Schmalfuss [9] and Arnold and Chueshov [6–8]. For an overview of random dynamical systems in economics we refer to Schenk-Hoppé [90]. We merely mention that there is also a vast literature on infinite-dimensional random dynamical systems (see, for instance, Flandoli and Schmalfuss [60]).

The study of continuous skew products originated from the ergodic theory of discrete dynamical systems (Anzai [2] and Furstenberg [61]). It was carried out during the 1960s by Miller [80], Millionscikov [82], Miller and Sell [81], and Sell [93–95] where we have mentioned but a few of these authors' publications. We also refer to the slightly later contributions by Artstein [10], Bronstein [28], and Zhikov and Levitan [113]. Meanwhile, a fairly comprehensive theory of continuous skew product flows has emerged. Among the multifarious results, we mention as examples, ranging from classical to more recent, the spectral theory of Sacker and Sell [89] (for its relation to the multiplicative ergodic theorem see Johnson, Palmer, and Sell [65]), the invariant manifold theory by Yi [111], Chow and Yi [42], Aulbach and Wanner [16], the results on monotone and almost automorphic systems by Shen and Yi [99], Chueshov [43], Sell, Shen, and Yi [97], and on inertial manifolds by Koksch and Siegmund [73], see also the book by Sell and You [98] and references therein.

Quite often, results for random dynamical systems and continuous skew product flows are structurally similar, and these similarities partly extend to the corresponding proofs. It is therefore natural to ask whether the analogies could be put on a formal, common basis. Such a unification is provided by the concept of a *nonautonomous dynamical system (NDS)* which may be regarded as an abstraction of both random dynamical systems and skew product flows. Originating also from Arnold's vicinity (see, e.g., the Festschrift [50] for his sixtieth birthday), the notion of a nonautonomous dynamical system is relatively recent and perhaps not yet known to a wider audience. Nevertheless, work on the general theory of NDS is under way (e.g., Caraballo and Langa [31]), and results are promising not only with respect to a forthcoming theory but also for providing some additional insights about the classical concepts and their interrelations.

The present paper takes part in this ongoing development by reviewing some of the fundamental facets inherent to the new concept. We consider it essential for any further progress in the theory of NDS that unifying as well as distinguishing aspects brought about by most important subclasses are seen clearly. Accordingly, it is our main goal to draw the reader's attention to both analogies and discrepancies, thereby stimulating further discussion and exchange. In the spirit we first carefully discuss the driving system level. When comparing ergodic theory and topological dynamics in this context, several analogies and differences may naturally come to one's mind.

- **Existence and Realization**

  Given a continuous dynamical system on a compact metric space, the Krylov–Bogoljubov theorem ensures that there always exist invariant measures, and by a convexity argument there exist ergodic measures as well. It is therefore easy to conceptually switch from the topological to the statistical point of view, although it may be difficult to single out a relevant ergodic measure. On the other hand, the celebrated Jewett–Krieger Theorem asserts that every ergodic map on a (Lebesgue) probability space can be realized as (i.e., is measurably isomorphic to) a uniquely ergodic homeomorphism on a compact metric space (Petersen [86]). Despite its theoretical importance, this result may be of little practical help, as the isomorphism is likely to destroy many particular features of the system under consideration.

- **Recurrence**

  A classical result highlightening the analogy of topology (via Baire category) and measure, is the Poincaré recurrence theorem. Informally, it asserts that for a measure-preserving homeomorphism of a bounded open set $U \subseteq \mathbb{R}^d$ all points in U except a set of first category and zero measure are recurrent. This perfect analogy, however, does not extend far beyond Poincaré's theorem. For example, with respect to the Birkhoff ergodic theorem, Oxtoby says in [85] that "curiously, though, this refinement of Poincaré's theorem turns out to be generally false in the sense of (Baire) category; the set of points where [the asymptotic relative frequency] is defined may be only of first category."

- **Entropy**

  The concept of entropy yields a quantification of a dynamical system's complexity. It may analogously be introduced both in ergodic theory and topological dynamics. An important link between both constructions is provided by the variational principle (Walters [108]).

This later result, however, also makes visible a fundamental discrepancy between the two approaches: measure-theoretic entropy quantifies complexity on average whereas its topological counterpart quantifies the maximal complexity inherent to the system under consideration. Notice also that both concepts trivially coincide in the case of unique ergodicity.

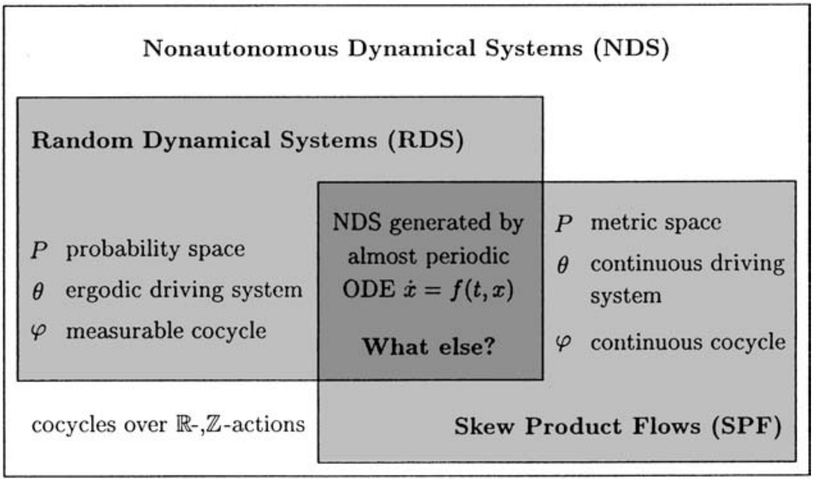
- **Almost Periodicity and Automorphy**

  The classical topic of almost periodic differential equations is a well-known meeting point of the two concepts where one can switch from one interpretation to the other by using the uniquely ergodic Haar measure. The theory of almost periodicity, since established in the 1920s, has had lasting impact on the development of harmonic analysis on groups and also on the development of the theory of both topological and smooth dynamical systems. A vast amount of research has been directed towards the study of almost periodic differential equations in the past (see Fink [59] for a survey). Solutions of almost periodic differential equations, however, are not necessarily almost periodic themselves: Johnson [63] gives an example of an almost periodic differential equation with a bounded solution which is not almost periodic but merely almost automorphic. The more general notion of almost automorphy was introduced by Bochner in 1955. (See Bochner [22] and Veech [106] for classical expositions, Shen and Yi [99] for a definitive recent treatment.)

Among the topics listed above, almost periodicity offers itself as an ideal starting point for our discussion, and we present a fairly complete mathematical treatment under a dynamical systems point of view.

Not least in the light of the above remarks, it is anything but surprising that, as far as the full nonautonomous dynamics is concerned, statistical and topological notions tend to diverge significantly. However, based mainly on examples, our discussion in Section 4 aims at convincing the reader that it is challenging yet highly worthwhile to explore the dark region in the diagram below. (Notions and symbols will be explained in detail in Section 2.)

One should certainly not expect an obvious unification of random dynamical systems and continuous skew products. In the specific situation of [85], Oxtoby observed that "the analogy [...] goes a long way here, but eventually it breaks down." We consider this statement particularly appropriate for the (full) nonautonomous dynamics of NDS. Correspondingly, our presentation is somewhat selective and cursory. Presumably, the time has not yet come to give a clear overall picture of NDS. However, the



power of the concept as well as its inherent subtleties should become apparent through our discussion.

The organization of this paper is as follows. In Section 2 we briefly review random dynamical systems and continuous skew product flows, and we introduce as their common structure the notion of a nonautonomous dynamical system. We explain how these systems may be generated by differential equations, and we recall some notions from measurable and topological dynamics.

Section 3 is devoted to a description of the analogies and discrepancies between random dynamical systems and continuous skew products on the driving system level. We give several illustrating examples. The main result of this section is on the interplay of equicontinuity, recurrence, almost periodicity and almost automorphy of the driving system.

In Section 4 we give a survey on selected topics for random dynamical systems and continuous skew products which illustrate some of the differences between the two concepts but also exhibit properties inherent to their common structure. Concludingly, we summarize in Section 5, and we also suggest starting points for further investigations which could carry forward our discussion.

2. BASIC DEFINITIONS AND PRELIMINARIES

The concept of a *random dynamical system* is an extension of the deterministic concept of a dynamical system, and it reduces to the latter if the noise is absent. It is tailor-mode to treat under a dynamical systems perspective many interesting systems which are under the influence of some "randomness," such as random or stochastic differential equations

$$dx_t = f_0(x_t)\,dt + \sum_{j=1}^{m} f_j(x_t)\,dW_t^j.$$

For a comprehensive study of random dynamical systems we refer to the monograph by Arnold [4].

The concept of *skew product flows* arose from topological dynamics during the 1960s as a description of dynamical systems with nonautonomy, i.e., showing an explicit time dependence. Since then, skew product flows have extensively been studied. They are tailor-made to nonautonomous systems such as nonautonomous differential equations

$$\dot{x} = f(t, x).$$

In both cases alluded to above, we do not obtain a dynamical system directly from solving the respective differential equation. Instead, the solution gives rise to a so-called *cocycle* over a dynamical system which models, respectively, the "randomness" and the "nonautonomy" of the equation. Before giving formal definitions, we have to recall a few basic notions from measurable and topological dynamics.

Let $\mathbb{T}(=\mathbb{R}, \mathbb{Z})$ denote time and let $\theta: \mathbb{T} \times \Omega \to \Omega$ be a measure-preserving dynamical system in the sense of ergodic theory, i.e., $(\Omega, \mathscr{F}, \mathbb{P})$ is a probability space and $(t, \omega) \mapsto \theta(t)\,\omega$ is a measurable[3] flow which leaves $\mathbb{P}$ invariant, i.e., $\theta(t)\,\mathbb{P} = \mathbb{P}$ for all $t \in \mathbb{T}$. (As it may cause confusion, the synonymously used notion of a *metric* dynamical system will be avoided here.) A set $M \subset \Omega$ is *θ-invariant* (or simply *invariant*) if $\theta(t)\,M = M$ for all $t \in \mathbb{T}$. We say that θ is *ergodic* under $\mathbb{P}$ if every θ-invariant set has probability 0 or 1. Under fairly mild and reasonable assumptions on the space $(\Omega, \mathscr{F}, \mathbb{P})$, it is possible to sensibly decompose $\mathbb{P}$ into ergodic components (see Denker, Grillenberger, and Sigmund [55] or Klünger [72]).

Analogously, in the setting of topological dynamics $\theta: \mathbb{T} \times P \to P$ denotes a continuous dynamical system on a metric space (P, d). Let $p \in P$ be a point and consider its orbit $O(p) := \{\theta(t)\,p: t \in \mathbb{T}\}$. Throughout this paper we will use the symbol $H(p) := \operatorname{cl} O(p)$ to denote the orbit closure. The orbit of p inherits, via $\theta(s)\,p \oplus \theta(t)\,p := \theta(s+t)\,p$, from $\mathbb{T}$ the structure of an abelian group with neutral element $0 := \theta_0\,p = p$. Later we shall investigate whether this structure extends to $H(p)$, thereby making this set a topological abelian group itself.

[3] w.r.t. $\mathscr{B} \otimes \mathscr{F}$ and $\mathscr{F}$, respectively, where $\mathscr{B}$ denotes the Borel σ-algebra of $\mathbb{T}$.

Given a closed invariant set M, there may be non-void proper closed subsets of M which are themselves invariant. If there are none, then $\theta|_M$ is called *minimal*, and M is referred to as a *minimal set* (for θ). Equivalently, M is a minimal set if $H(p) = M$ for every $p \in M$. Given any element $p \in P$, its *ω-limit set* $\omega(p)$ denotes the set of all future accumulation points of $O(p)$, more formally

$$\omega(p) := \bigcap_{t \geqslant 0} \overline{\{\theta(s)\, p \colon s \geqslant t\}}.$$

Analogously, the *α-limit set* $\alpha(p)$ is the set of all accumulation points in the past. Both $\omega(p)$ and $\alpha(p)$ are closed; also, $\theta(t)\,\omega(p) \subset \omega(p)$, and similarly for $\alpha(p)$. A point p is called *positively (negatively) recurrent*, if $p \in \omega(p)$ (respectively, $p \in \alpha(p)$); it is *recurrent* if it is both positively and negatively recurrent. The point p thus is recurrent if and only if for every $\varepsilon > 0$ there exists a monotonically increasing sequence $(t_k)_{k \in \mathbb{Z}}$ in $\mathbb{T}$ with $|t_k| \to \infty$ as $|k| \to \infty$ such that $d(\theta_{t_k} p, p) < \varepsilon$ for all k. If in addition $\sup_{k \in \mathbb{Z}} (t_{k+1} - t_k) < \infty$ then p is called *uniformly recurrent*. Note that the notion of recurrence does not depend on the specific metric as long as the latter induces the topology of P.

We emphasize by a formal definition the structure common to random dynamical systems and skew product flows. We will use the symbol $\mathbb{T}$ for either $\mathbb{R}$ or $\mathbb{Z}$, and we will denote by $\mathbb{T}_+$ all non-negative elements of $\mathbb{T}$.

Definition 2.1 (Nonautonomous Dynamical System (NDS)). A *non-autonomous dynamical system* (NDS) with (one-sided) time $\mathbb{T}_+$ on a metric space X with base set P consists of two ingredients:

(i) A model of the nonautonomy, namely an action $\theta \colon \mathbb{T} \times P \to P$ of the group $\mathbb{T}$ on P, i.e., the family $\theta(t, \cdot) = \theta(t) \colon P \to P$ of self-mappings of the set P satisfies the *group property*

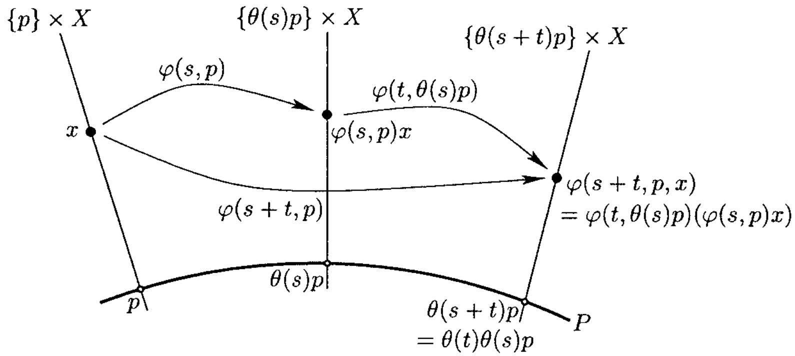
$$\theta(0) = \mathrm{id}_P, \qquad \theta(t+s) = \theta(t) \circ \theta(s),$$

for all $t, s \in \mathbb{T}$.

(ii) A model of the system perturbed or forced by nonautonomy, namely a *cocycle φ over θ*, i.e., a mapping $\varphi \colon \mathbb{T}_+ \times P \times X \to X$, $(t, p, x) \mapsto \varphi(t, p, x)$, such that $(t, x) \mapsto \varphi(t, p, x)$ is continuous for all $p \in P$ and the family $\varphi(t, p, \cdot) = \varphi(t, p) \colon X \to X$ of self-mappings of X satisfies the *cocycle property*

$$\varphi(0, p) = \mathrm{id}_X, \qquad \varphi(t+s, p) = \varphi(t, \theta(s)\, p) \circ \varphi(s, p), \quad (2.1)$$

for all $t, s \in \mathbb{T}_+$ and $p \in P$.



Remark 2.2.

(i) The pair of mappings

$$(\theta, \varphi): \mathbb{T}_+ \times P \times X \to P \times X, \qquad (t, p, x) \mapsto (\theta(t, p)\ \varphi(t, p, x)),$$

is called the corresponding *skew product*. If $P = \{p\}$ consists of a single point, then the cocycle φ is a (semi-)dynamical system.

(ii) If $\mathbb{T} = \mathbb{Z}$ then $\theta(n) = \theta^n$, $n \in \mathbb{T}$, where $\theta := \theta(1)$ is the time one mapping. If $\mathbb{T} = \mathbb{R}$ we often use the less clumsy notation θ_t instead of $\theta(t)$. We also say that φ is an NDS to subsume the situation of Definition 2.1.

(iii) If X is a linear space then an NDS Φ is called *linear* if for any scalar α and $x_1, x_2 \in X$

$$\Phi(t, p)[\alpha(x_1 + x_2)] = \alpha\Phi(t, p)\, x_1 + \alpha\Phi(t, p)\, x_2,$$

for all $t \in \mathbb{T}_+$ and $p \in P$.

We now introduce the notion of a random dynamical systems. Since we want to compare random dynamical systems to continuous skew products, we do not introduce the most general notion but require that the state space be a metric space and assume w.l.o.g. ergodicity of the driving system (see Klünger [72] for the decomposition into ergodic components).

Definition 2.3 (Random Dynamical System (RDS)). A *(continuous) random dynamical system* is an NDS (θ, φ) which in addition has the following properties:

(i) The driving system θ is an ergodic dynamical system $(\Omega, \mathscr{F}, \mathbb{P}, (\theta(t))_{t \in \mathbb{T}})$, i.e., the base $(\Omega, \mathscr{F}, \mathbb{P})$ is a probability space and $(t, \omega) \mapsto \theta(t)\, \omega$ is a measurable flow which is ergodic under $\mathbb{P}$.

(ii) The cocycle $(t, \omega, x) \mapsto \varphi(t, \omega)\, x$ is measurable.[4]

Remark 2.4. In denoting the base space by Ω in Definition 2.3 (and also in Theorem 2.5 below) we pay deference to a strict tradition in random dynamical systems theory. Nevertheless, $(\Omega, \mathscr{F}, \mathbb{P})$ should be viewed as an additional structure on P which by Definition 2.1 is a mere set.

As an important class of examples, we will later consider RDS which are generated by random differential equations. This is the (easy) *real noise case* in which the generator indeed is a certain family of ordinary differential equations with parameter ω, i.e., it can be solved "path-wise" for each fixed ω as a deterministic nonautonomous ordinary differential equation (see Arnold [4]).

Theorem 2.5 (RDS from Random Differential Equation). *Let $(\Omega, \mathscr{F}, \mathbb{P}, (\theta(t))_{t \in \mathbb{R}})$ be a measure-preserving dynamical system, let $f: \Omega \times \mathbb{R}^d \to \mathbb{R}^d$ be measurable, and consider the path-wise random differential equation*

$$\dot{x} = f(\theta_t \omega, x). \tag{2.2}$$

If $(t, x) \mapsto f(\theta_t \omega, x)$ is continuous in (t, x), locally Lipschitz in x for all ω and

$$\|f(\omega, x)\| \leqslant \alpha(\omega)\, \|x\| + \beta(\omega),$$

where $t \mapsto \alpha(\theta_t \omega)$ and $t \mapsto \beta(\theta_t \omega)$ are locally integrable, then (2.2) uniquely generates through its solution

$$\varphi(t, \omega)\, x = x + \int_0^t f(\theta_s \omega, \varphi(s, \omega)\, x)\, ds$$

an RDS φ over θ.

Next we define skew product flows. Again we do not introduce the most general notion but instead require that the base space be a *complete metric* space. For a reduction to this situation from more general base spaces see Johnson, Palmer, and Sell [65].

[4] w.r.t. $\mathscr{B} \otimes \mathscr{F} \otimes \mathscr{B}(X)$ and $\mathscr{B}(X)$, respectively, where $\mathscr{B}(X)$ is the Borel σ-algebra of X.

Definition 2.6 (Skew Product Flow (SPF)). A *(continuous) skew product flow* is an NDS (θ, φ) which in additon has the following properties:

(i) The base is a metric space (P, d), and the driving system $(t, p) \mapsto \theta(t)\, p$ is continuous.

(ii) The cocycle $(t, p, x) \mapsto \varphi(t, p)\, x$ is continuous.

The well-known trick of making a nonautonomous differential equation

$$\dot{x} = f(t, x) \tag{2.3}$$

autonomous by introducing a new variable for the time suggests to investigate a corresponding skew product flow with base $P := \mathbb{R}$ and driving system $(t, s) \mapsto \theta_t s := t + s$. However, as P does not depend on f, we should not expect a specific kind of nonautonomy (e.g., periodicity in t) to be captured by this base dynamics. Moreover, P is not compact which may cause additional difficulties. For a fairly general class of right hand sides f the Bebutov flow $(t, p) \mapsto \theta_t p := p(\cdot + t, \cdot)$ on the *hull* $P := H(f) = \mathrm{cl}\{f(\cdot + t, \cdot) : t \in \mathbb{R}\}$ of f can serve as a model for the nonautonomy (Sell [95]). Here the closure is taken with respect to an adequate topology. The evaluation mapping

$$\bar{f} : P \times \mathbb{R}^d \to \mathbb{R}^d, \qquad (p, x) \mapsto p(0, x)$$

satisfies $\bar{f}(\theta_t p, x) = p(t, x)$ and, since $f \in H(f)$ and therefore $\bar{f}(\theta_t f, x) = f(t, x)$, it is a natural "extension" of f to $P \times \mathbb{R}^d$. As a slight abuse of notation we will sometimes omit the bar. Instead of looking at the single Eq. (2.3) we consider the associated family of equations

$$\dot{x} = \bar{f}(\theta_t p, x), \qquad p \in P = H(f). \tag{2.4}$$

By using standard results about linearly bounded equations as in Amann [1] and Arzela–Ascoli's theorem as in Sell [95] one can prove the following

Theorem 2.7 (SPF from Nonautonomous Differential Equation). *Let $f : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^d$ be a continuous function, and consider the nonautonomous differential equation (2.3). If $(t, x) \mapsto f(t, x)$ is locally Lipschitz in x and*

$$\|f(t, x)\| \leqslant \alpha(t)\, \|x\| + \beta(t),$$

where $t \mapsto \alpha(t)$ and $t \mapsto \beta(t)$ are locally integrable, then the hull $P := H(f)$ is a metric space (where the closure is taken in $C(\mathbb{R} \times \mathbb{R}^d, \mathbb{R}^d)$ with the *compact-open topology), the Bebutov flow* $(t, p) \mapsto \theta_t p = p(\cdot + t, \cdot)$ *is continuous, and (2.3) uniquely generates through the solution*

$$\varphi(t, p)\, x = x + \int_0^t \bar{f}(\theta_s p, \varphi(s, p)\, x)\, ds \qquad (2.5)$$

of the associated family of equations (2.4) an SPF φ *over* θ*. Moreover,* $H(f)$ *is compact if and only if* $(t, x) \mapsto f(t, x)$ *is bounded and uniformly continuous on every set of the form* $\mathbb{R} \times K$ *where* $K \subset \mathbb{R}^d$ *is compact.*

Remark 2.8. (i) Note that the hull $H(f)$ depends on the topology chosen. Therefore it is necessary to determine an adequate topology for a given f in order to make the Bebutov flow continuous. For example, in the case of a linear ordinary differential equation $\dot{x} = A(t)\, x$ in $\mathbb{R}^d$ with a locally integrable matrix function $A \in L^1_{\text{loc}} := L^1_{\text{loc}}(\mathbb{R}, \mathbb{R}^{d \times d})$ it is easy to see (Siegmund [100]) that φ in (2.5) with $f(t, x) = A(t)\, x$ is an SPF over the Bebutov flow on $P := H(A)$ where the closure is taken in L^1_{loc}. Moreover, P is compact if and only if (Sell [95])

(a) there exists $\alpha > 0$ such that $\int_0^1 |A(s+t)|\, ds \leqslant \alpha$ for all $t \in \mathbb{R}$, and

(b) for every $\varepsilon > 0$, there is a $\delta > 0$ such that $\int_0^1 |A(s+t+h) - A(s+t)|\, ds \leqslant \varepsilon$ whenever $|h| \leqslant \delta$ and $t \in \mathbb{R}$.

One could replace L^1_{loc} by the space of essentially bounded A and still obtain an SPF. The hull is then compact in the weak*-topology (Colonius and Kliemann [45]).

(ii) An analogue of Theorem 2.7 holds for nonautonomous difference equation $x_{n+1} = f(n, x_n)$ with the Bebutov flow $(n, p) \mapsto \theta^n p = p(\cdot + n, \cdot)$ on the hull $P := H(f) = \text{cl}\{f(\cdot + n, \cdot) \colon n \in \mathbb{Z}\}$.

3. BASE DYNAMICS

In this section we focus on the following question: Under which conditions can the driving system (P, θ) be interpreted as both an ergodic and a continuous dynamical system on a metric space? Our starting point to tackle this question is provided by almost periodic differential equations. As will soon become clear, almost periodicity gives rise to a rotational flow on a compact metrizable abelian group. By virtue of character theory the dynamics of group rotations is easy to analyze. For our purpose the following well-known theorem is crucial (Walters [108]).

Theorem 3.1 (Haar Measure). *Let G be a compact abelian group. There exists a unique probability measure $\mathbb{P}$ on the Borel σ-algebra $\mathcal{B}(G)$ which is invariant under rotation, i.e., $\mathbb{P}(A) = \mathbb{P}(gA)$ for all $g \in G$ and $A \in \mathcal{B}(G)$.*

Later we will describe what happens if we slightly weaken the assumptions on the dynamics which lead to an application of Theorem 3.1. Instead of elaborating the most general case we consider it more instructive to understand the difficulties by means of explicit examples. Therefore we investigate the base dynamics of nonautonomous dynamical systems generated by nonautonomous difference or differential equations of the form

$$x_{n+1} = f(n, x_n) \qquad \text{or} \qquad \dot{x} = f(t, x),$$

where $f: \mathbb{T} \times \mathbb{R}^d \to \mathbb{R}^d$, by considering the corresponding families

$$x_{n+1} = \bar{f}(\theta^n p, x_n) \qquad \text{or} \qquad \dot{x} = \bar{f}(\theta_t p, x), \quad p \in P,$$

of equations over the Bebutov flow $\theta: \mathbb{T} \times P \to P$ on the hull $P := H(f)$. In the sequel we shall use the symbol $H(f)$ without explicitly referring to a topology; whenever we simultaneously deal with several different topologies (via non-equivalent metrics), clarification will be ensured by means of subscripts.

First we want to review almost periodic differential and difference equations and therefore have to define what it means for a right-hand side f to be almost periodic. For this purpose we consider two different complete metrics on the set $C(\mathbb{T} \times \mathbb{R}^d, \mathbb{R}^d)$ of continuous functions $f: \mathbb{T} \times \mathbb{R}^d \to \mathbb{R}^d$; both metrics emerge from the family of semi-metrics

$$d_{k,l}(f, g) := \sup_{|s| \leqslant k,\, \|x\| \leqslant l} \frac{\|f(s, x) - g(s, x)\|}{1 + \|f(s, x) - g(s, x)\|} \qquad (k, l \in \mathbb{N}).$$

The compact-open topology, i.e., the topology of uniform convergence on compact sets $K \subset \mathbb{T} \times \mathbb{R}^d$, is induced by

$$d_{co}(f, g) := \sum_{k,l=1}^{\infty} 2^{-(k+l)} d_{k,l}(f, g),$$

whereas the somewhat hybrid metric

$$d_{\infty}(f, g) := \lim_{k \to \infty} \sum_{l=1}^{\infty} 2^{-l} d_{k,l}(f, g)$$

gives rise to a mixture of uniform (in the first argument) and locally uniform (in the second argument) topology. Two elementary observations about these metrics and the topologies they induce are contained in

Proposition 3.2. *For all $f, g \in C(\mathbb{T} \times \mathbb{R}^d, \mathbb{R}^d)$*

$$d_{co}(f, g) \leqslant d_\infty(f, g),$$

and the topology induced by d_∞ is strictly finer than the topology induced by d_{co}. The metric d_∞ is invariant under a shift in the first argument, i.e.,

$$d_\infty(f(\cdot+t, \cdot), g(\cdot+t, \cdot)) = d_\infty(f, g) \qquad \text{for all } f, g \text{ and } t \in \mathbb{T}, \tag{3.1}$$

whereas d_{co} is not.

Proof. With $a_k := \sum_{l=1}^{\infty} 2^{-l} d_{k,l}(f, g)$ obviously $a_k \leqslant a_{k+1} \leqslant 1$ for all k, and $d_{co}(f, g) = \sum_{k=1}^{\infty} 2^{-k} a_k \leqslant \lim_{k\to\infty} a_k = d_\infty(f, g)$. Fix any $x_0 \in \mathbb{R}^d$ with $\|x_0\| = 1$, define $f_0 \in C(\mathbb{T} \times \mathbb{R}^d, \mathbb{R}^d)$ by $f_0(t, x) := \max\{0, 1-2\,|t|\}\, x_0$, and let $f_n := f_0(\cdot+n, \cdot)$, $n \in \mathbb{N}_0$. Denoting by 0 the null-function, it is easy to see that $d_{co}(f_n, 0) \to 0$ whereas $d_\infty(f_n, 0) = \frac{1}{2}$ for all n, which implies that d_∞ induces a finer topology than d_{co} does.

To prove the second assertion, notice first that for all $k > \lceil |t| \rceil$, for all $l \in \mathbb{N}$ and $f, g \in C(\mathbb{T} \times \mathbb{R}^d, \mathbb{R}^d)$

$$d_{k-\lceil |t| \rceil, l}(f, g) \leqslant d_{k,l}(f(\cdot+t, \cdot), g(\cdot+t, \cdot)) \leqslant d_{k+\lceil |t| \rceil, l}(f, g), \tag{3.2}$$

where, for any $x \in \mathbb{R}$, $\lceil x \rceil$ denotes the smallest integer not smaller than x. Relation (3.1) obviously follows from (3.2). On the other hand, with the functions f_n defined earlier

$$d_{co}(f_0(\cdot+n, \cdot), f_1(\cdot+n, \cdot)) = d_{co}(f_n, f_{n+1}) \to 0 \neq \tfrac{1}{2} = d_{co}(f_0, f_1),$$

which shows that (3.1) cannot hold with d_∞ replaced by d_{co}. □

In the sequel, $\theta(t)$ applied to a function f always denotes the Bebutov flow $(t, f) \mapsto \theta(t) f = f(\cdot+t, \cdot)$. According to Proposition 3.2, d_∞ is invariant under θ. For the following definition, recall that $L \subset \mathbb{T}$ is *relatively dense* if with some number $T \in \mathbb{T}^+$, $[t, t+T] \cap L \neq \varnothing$ for all $t \in \mathbb{T}$.

Definition 3.3 (Almost Periodic Functions). Let $f: \mathbb{T} \times \mathbb{R}^d \to \mathbb{R}^d$ be a continuous function.

(i) f is *Bohr almost periodic* if, for every $\varepsilon > 0$, there exists a relatively dense set $L_\varepsilon \subset \mathbb{T}$ such that

$$d_\infty(\theta(t)\, f, f) < \varepsilon \qquad \text{for all} \quad t \in L_\varepsilon.$$

(ii) f is *Bochner almost periodic* if any sequence in $\mathbb{T}$ contains a subsequence $(t_n)_{n \in \mathbb{N}}$ such that $\theta(t_n)\, f$ converges in the d_∞-topology, i.e., $\lim_{n \to \infty} f(t + t_n, x)$ exists uniformly in $(t, x) \in \mathbb{T} \times K$ for every compact subset $K \subset \mathbb{R}^d$.

Remark 3.4. Bohr and Bochner introduced almost periodicity for functions $f: \mathbb{T} \to \mathbb{C}$. Bohr's definition first appeared in his original paper in 1923 which is most easily found in the collection [26], see also [25]. In 1927, Bochner [20] gave his definition of almost periodicity and showed its equivalence to the Bohr definition.

Example 3.5. Consider the function $h: \mathbb{R} \to \mathbb{C}$ defined as

$$h(t) := 1 + \tfrac{1}{2}\,(e^{2\pi i t} + e^{2\pi i \sqrt{2}\, t})$$

and take $f: \mathbb{R} \to \mathbb{R}^2$ as $f(t) := (\Re h(t), \Im h(t))$. An elementary calculation yields

$$d_\infty(\theta_t f, f) \leqslant |\sin \pi t| + |\sin \sqrt{2}\, \pi t| \qquad (t \in \mathbb{R}),$$

the right-hand side of which becomes smaller than any given ε on an appropriate relatively dense set. Hence f is Bohr almost periodic; by virtue of the invariance of d_∞ and a diagonalization argument f may easily be seen to be Bochner almost periodic, too.

To neatly visualize $H(f)$ let $(\vartheta_t)_{t \in \mathbb{R}}$ denote the minimal Kronecker flow on the two torus $T^2 := \mathbb{R}^2/\mathbb{Z}^2$ according to $\vartheta_t: (x, y) \mapsto (x+t, y + \sqrt{2}\, t)$. Also, define a continuous function $h_0: T^2 \to \mathbb{C}$ as $h_0: (x, y) \mapsto 1 + \frac{1}{2}\,(e^{2\pi i x} + e^{2\pi i y})$. Evidently, $h(t) = h_0(\vartheta_t(0, 0))$ for all $t \in \mathbb{R}$. The assignment $\theta_t f \mapsto \vartheta_t(0, 0)$ may be extended to a homeomorphism $\Psi: H(f) \to T^2$ which satisfies $\Psi \circ \theta_t = \vartheta_t \circ \Psi$ for all t. The Bebutov flow θ on $H(f)$ is therefore flow equivalent to the Kronecker flow ϑ on T^2.

According to Definition 3.3(i), a function f is Bohr almost periodic precisely if it is uniformly recurrent with respect to the d_∞-metric; by (ii), Bochner almost periodicity is equivalent to $H(f)$ being compact in that topology. The following proposition allows one to simply speak of *almost periodic* functions.

Proposition 3.6. *Let $f\colon \mathbb{T}\times\mathbb{R}^d\to\mathbb{R}^d$ be continuous. The following three statements are equivalent:*

(i) *f is Bohr almost periodic.*

(ii) *f is Bochner almost periodic.*

(iii) *Any sequence in $\mathbb{T}$ contains a subsequence $(t_n)_{n\in\mathbb{N}}$ such that for some continuous function g*

$$\theta(t_n)\, f\to g \qquad \text{and} \qquad \theta(-t_n)\, g\to f$$

in $C(\mathbb{T}\times\mathbb{R}^d, \mathbb{R}^d)$ with respect to the d_∞-metric, i.e., uniformly on sets $\mathbb{T}\times K$ where $K\subset\mathbb{R}^d$ is a compact set.

Proof.

(i) $\Leftrightarrow$ (ii). follows similarly as for functions $f\colon\mathbb{R}\to\mathbb{C}$ with the standard arguments of Fink [59] which we do not repeat here.

(ii) $\Leftrightarrow$ (iii). By compactness $d_\infty(\theta(t_n)\, f, g)\to 0$ for some subsequence $(t_n)_{n\in\mathbb{N}}$ and appropriate g. Since d_∞ is invariant under θ, property (iii) follows.

(iii) $\Leftrightarrow$ (ii). This is obvious. □

The classical Definition 3.3 rests on the hybrid metric d_∞. This metric is not easy to operate with. For example, general compactness results are much more sensibly formulated in the d_{co}-topology (e.g., the Arzela–Ascoli theorem). Furthermore, θ may not be continuous, even when restricted to individual orbits. The following lemma allows to circumvent these unpleasant facts; its content will also motivate the formal definition of an almost periodic *point* given later.

Lemma 3.7. *Let $f\colon\mathbb{T}\times\mathbb{R}^d\to\mathbb{R}^d$ be continuous. Then f is almost periodic if and only if, for every $\varepsilon>0$, there exists a relatively dense set $L_\varepsilon\subset\mathbb{T}$ such that*

$$d_{co}(\theta(t)\, g, g)<\varepsilon \qquad \textit{for all} \quad t\in L_\varepsilon, \quad g\in H(f).$$

Proof.

($\Rightarrow$) Let f be almost periodic. It is easy to see that f is bounded and uniformly continuous on every set $\mathbb{T}\times K$ where $K\subset\mathbb{R}^d$ is compact. Hence $H_\infty(f)=H_{co}(f)$, and this set is compact in both topologies. Assume that $d_{co}(\theta(t_n)\, f, h)\to 0$ for some sequence (t_n) in $\mathbb{T}$ and $h\in C(\mathbb{T}\times\mathbb{R}^d, \mathbb{R}^d)$. By compactness every subsequence of $d_\infty(\theta(t_n)\,f, h)$ has a subsequence that converges to 0. Therefore $d_\infty(\theta(t_n)\,f, h) \to 0$ itself, and the two metrics are equivalent on $H(f)$. Since obviously $d_\infty(\theta(t)\,g, g) < \varepsilon$ for all $g \in H(f)$ and t taken from an appropriate relatively dense set L_ε, the claim follows.

($\Leftarrow$) Conversely, if $d_{co}(\theta(t)\,g, g) < \varepsilon$ for all $g \in H(f)$ and $t \in L_\varepsilon$, then $d_{co}(\theta(t+\tau)\,f, \theta(\tau)\,f) < \varepsilon$ for all $\tau \in \mathbb{T}$, $t \in L_\varepsilon$. From this it is easy to see that $d_\infty(\theta(t)\,f, f) < \varepsilon$ for all $t \in L_\varepsilon$, hence f is (Bohr) almost periodic. □

Applications (most prominently from differential equations) may require a notion slightly weaker than almost periodicity. The concept of almost automorphy provides such a weakening of almost periodicity. It was introduced by Bochner [21] in 1955 in an article on differential geometry, and it has subsequently been studied by many others, notably by Veech [106, 107].

Definition 3.8 (Almost Automorphic Functions). Let $f\colon \mathbb{T} \times \mathbb{R}^d \to \mathbb{R}^d$ be a continuous function. This function is *almost automorphic* if every sequence in $\mathbb{T}$ has a subsequence $(t_n)_{n \in \mathbb{N}}$ such that for some function g

$$\theta(t_n)\,f \to g \qquad \text{and} \qquad \theta(-t_n)\,g \to f$$

holds *pointwise*, i.e., $\lim_{n\to\infty} \theta(t_n)\,f(t, x) = g(t, x)$ and $\lim_{n\to\infty} \theta(-t_n)\,g(t, x) = f(t, x)$ for all $(t, x) \in \mathbb{T} \times \mathbb{R}^d$.

Example 3.9. Let $\mathbb{T} = \mathbb{Z}$, fix an irrational real number ν and consider the real-valued function (sequence) $f = (f_k)_{k \in \mathbb{Z}} = (\operatorname{sign}(\cos 2\pi\nu k))_{k \in \mathbb{Z}}$, where, as usual, $\operatorname{sign}(0) = 0$ and $\operatorname{sign}(x) = x/|x|$ for $x \neq 0$. (In essence, this example stems from Furstenberg's seminal paper [62].) Since for any $s \neq t$ the sequences $\theta^s f$ and $\theta^t f$ differ at least at one position, f is certainly not almost periodic. Given any sequence in $\mathbb{T}$, there is a subsequence $(t_n)_{n \in \mathbb{N}}$ such that

$$(\theta^{t_n} f)_k = \operatorname{sign}(\cos(2\pi\nu k + 2\pi\nu t_n)) =: g_{k,n}$$

converges for all k as $n \to \infty$. Ideally, one would like to write

$$g_k := \lim_{n\to\infty} g_{k,n} = \operatorname{sign}(\cos(2\pi\nu k + 2\pi\rho))$$

for this limit, with some $\rho \in [0, 1[$. However, such a representation is not correct if $4\rho \equiv 4\nu l \pmod 1$ for some l, because in this case there are two possible limits for $(\theta^{t_n} f)_{n\in\mathbb{N}}$, see also Example 3.18 below. With $g = (g_k)_{k\in\mathbb{Z}}$, an elementary calculation confirms that nevertheless $\lim_{n\to\infty}(\theta^{-t_n} g)_k = f_k$ holds for all k, i.e., f is an almost automorphic function.

Example 3.10. The pointwise convergence in Definition 3.8 may be complicated to work with in practice. For example, the function g does not need to be continuous. In addition one can easily end up with non-compact hulls, as the following example illustrates. Recall the complex valued funtion h in Example 3.5 but define now $f\colon \mathbb{R}\to\mathbb{R}^2$ as $f(t) := (\Re h(t), \Im h(t))/|h(t)|$. (This function is quite popular in publications on almost automorphy [99, 106].)

We claim that f is almost automorphic in the sense of Definition 3.8 (yet not almost periodic). An easy way to see this is as follows. Let $(\vartheta_t)_{t\in\mathbb{R}}$ denote the same minimal Kronecker flow on the two-torus T^2 as before. With $f_0\colon T^2\setminus\{(\frac{1}{2},\frac{1}{2})\}\to\mathbb{R}^2$ defined as $f_0(x,y) := (\Re h_0(x,y), \Im h_0(x,y))/|h_0(x,y)|$, we clearly have $f(t) = f_0(\vartheta_t(0,0))$ for all $t\in\mathbb{R}$. From this representation we deduce (as in Example 3.9) that f is not almost periodic. Since the function f_0 has directional limits at $(\frac{1}{2},\frac{1}{2})$ it follows that from any sequence in $\mathbb{R}$ we may extract a subsequence $(t_n)_{n\in\mathbb{N}}$ such that $\lim_{n\to\infty}\theta_{t_n} f(t) := g(t)$ exists for all t, and also $\lim_{n\to\infty}\theta_{-t_n} g(t) = f(t)$. Therefore f is almost automorphic. It is, however, easy to choose $(t_n)_{n\in\mathbb{N}}$ in such a way that g is not continuous. Moreover, as f is not uniformly continuous, the hull $H(f)$ is not compact (in the d_{co}-topology), a fact immediately ruling out many techniques from topological dynamics.

For all situations of practical relevance, the following lemma provides a satisfactory class of almost automorphic functions.

Lemma 3.11. *Let $f\colon \mathbb{T}\times\mathbb{R}^d\to\mathbb{R}^d$ be uniformly continuous and bounded on every set $\mathbb{T}\times K$ with $K\subset\mathbb{R}^d$ denoting a compact set. Then f is almost automorphic if and only if any sequence in $\mathbb{T}$ contains a subsequence $(t_n)_{n\in\mathbb{N}}$ such that for some continuous function g*

$$\theta(t_n)\, f \to g \qquad \textit{and} \qquad \theta(-t_n)\, g \to f$$

with respect to the d_{co}-metric, i.e., uniformly on compact subsets of $\mathbb{T}\times\mathbb{R}^d$.

Proof.

($\Rightarrow$) If f is bounded and uniformly continuous on every set $\mathbb{T}\times K$ then every sequence in $\mathbb{T}$ contains a subsequence $(t_n)_{n\in\mathbb{N}}$ such that $d_{co}(\theta(t_n)\, f, g)\to 0$ and $d_{co}(\theta(-t_n)\, g, h)\to 0$ for appropriate continuous functions g, h. If f is almost automorphic then $h = f$, and thus the convergence in Definition 3.8 is in fact locally uniform.

($\Leftarrow$) This is obvious. □

Remark 3.12. The class of functions showing up in Lemma 3.11 coincides with the class of *admissible* functions defined and dealt with exclusively in Shen and Yi [99]; see also Theorem 2.7.

Lemmas 3.7 and 3.11, respectively, characterize almost periodicity and almost automorphy of a function in terms of the corresponding Bebutov flow. It is natural to imitate these results on a formal level so as to provide a definition of almost periodic and almost automorphic *points* under arbitrary flows.

Definition 3.13 (Almost Periodic and Almost Automorphic Point). Let $\theta\colon \mathbb{T} \times P \to P$ be a continuous flow on a complete metric space (P, d).

(i) A point $p \in P$ is called *almost periodic* (abbreviated henceforth as *a.p.*), if, for every $\varepsilon > 0$, there exists a relatively dense set $L_\varepsilon \subset \mathbb{T}$ such that

$$d(\theta(t)\, q, q) < \varepsilon \qquad \text{for all} \quad t \in L_\varepsilon, \quad q \in H(p).$$

(ii) A point $p \in P$ is called *almost automorphic* (a.a.), if any sequence in $\mathbb{T}$ contains a subsequence $(t_n)_{n \in \mathbb{N}}$ such that for some $q \in P$

$$\theta(t_n)\, p \to q \qquad \text{and} \qquad \theta(-t_n)\, q \to p.$$

The orbit of an a.p. or a.a. point p naturally inherits from $\mathbb{T}$ the structure of an abelian group. For Theorem 3.1 to be useful in this dynamical context (so as to yield for instance a uniquely ergodic system), we have to check whether $H(p)$ is a compact group. A simple condition which ensures that the group structure of the orbit $O(p)$ can be extended to $H(p)$ reads as follows.

Proposition 3.14. *If the family $(\theta(t))_{t \in \mathbb{T}}$ is equicontinuous, then $H(p)$ is a group. Whenever $H(p)$ is a group, then it is also a minimal set.*

Proof. Let $(\theta(s_n)\, p)_{n \in \mathbb{N}}$, $(\theta(t_n)\, p)_{n \in \mathbb{N}}$ denote two sequences converging to q, r in $H(p)$, respectively. We are going to show that $\lim_{n \to \infty} \theta_{s_n - t_n} p$ yields a well-defined element $q \ominus r \in H(p)$. Indeed, given $\varepsilon > 0$, we find by equicontinuity that $(\theta(-t_n)\,p)_{n\in\mathbb{N}}$ is a Cauchy sequence and hence converges. But then also

$$\begin{aligned} &d(\theta(s_n-t_n)\,p, \theta(s_m-t_m)\,p) \\ &\qquad \leqslant d(\theta(s_n-t_n)\,p, \theta(s_n-t_m)\,p)+d(\theta(s_n-t_m)\,p, \theta(s_m-t_m)\,p)<\varepsilon \end{aligned}$$

for m, n sufficiently large. Therefore $(\theta(s_n-t_n)\,p)_{n\in\mathbb{N}}$ is a Cauchy sequence itself. A completely similar reasoning reveals that the limit thereof does not depend on the choice of the approximating sequences. Consequently, $H(p)$ is a topological abelian group.

In order to prove the second statement, we notice that the orbit of every $q\in H(p)$ is just a rotated version of $O(p)$, more precisely $O(q)=q\oplus O(p)$. Since rotations are homeomorphisms, every orbit is dense in $H(p)$. □

Remark 3.15.

(i) It should be pointed out that the proof of Proposition 3.14 (and also of Theorem 3.21 below) only requires the family $(\theta_t)_{t\in\mathbb{T}}$ to be *locally* equicontinuous. Since, however, even simple one-dimensional examples show that (local) equicontinuity is not necessary for $H(p)$ to be a group, we have refrained from relaxing the assumptions in Proposition 3.14 by using the rather unusual notion of *local* equicontinuity.

(ii) Notice that minimality is not sufficient for the closure $H(p)$ to be a group, even if it happens to be compact. Take for example $\mathbb{T}=\mathbb{Z}$ and θ a minimal diffeomorphism of the two-torus T^2 which admits more than one ergodic measure (see Katok and Hasselblatt [67] for an explicit construction). It is easy to see that each of these measures would be invaraint under rotations of the compact set $H(p)=T^2$, which therefore cannot be a topological group. (Evidently, the algebraic structure induced by θ does *not* coincide with the usual group structure of T^2.)

Corollary 3.16. *If $H(p)$ is a group, then it is either homeomorphic and isomorphic to $\langle\mathbb{T},+\rangle$, or else every point $q\in H(p)$ is recurrent.*

Proof. Suppose first that $\alpha(p)\cup\omega(p)=\varnothing$. Then $H(p)=O(p)$, and the map $\psi\colon O(p)\to\mathbb{T}$ with $\psi(\theta(t)\,p):=t$ is easily seen to be both a homeomorphism and a group isomorphism. If $\alpha(p)\cup\omega(p)$ is not empty, then minimality yields $\alpha(q)=\omega(q)=H(p)$ for every $q\in H(p)$. □

Example 3.17. Consider once more the almost periodic function f introduced in Example 3.5. An inspection of the first half in the proof of Lemma 3.7 shows that $(\theta_t)_{t\in\mathbb{R}}$ is equicontinuous; hence $H(f)$ is a compact group. The map Ψ in Example 3.5 constitutes a group isomorphism and homeomorphism between $\langle H(f), \oplus\rangle$ and $\langle T^2, +\rangle$.

Example 3.18. In general, $H(p)$ is not a group, not even if p is a.a. To see this, let f be again the sequence $(\operatorname{sign}(\cos 2\pi\nu k))_{k\in\mathbb{Z}}$ of Example 3.9. By taking appropriate sequences in $\mathbb{T}=\mathbb{Z}$ it is easy to see that both $g^+=(g_k^+)_{k\in\mathbb{Z}}$ and $g^-=(g_k^-)_{k\in\mathbb{Z}}$ are elements of $H(f)$, where

$$g_k^\pm := \begin{cases} \operatorname{sign}(\sin 2\pi\nu k) & \text{if} \quad k\neq 0,\\ \pm 1 & \text{if} \quad k=0. \end{cases}$$

Let (t_n) denote a sequence in $\mathbb{T}$ for which $2\nu t_n$ decreases to 0 (mod 1). Then

$$\theta^{t_n} f\to f,\quad \theta^{t_n} g^+\to g^+ \qquad \text{and} \qquad \theta^{t_n} g^-\to g^+ \qquad \text{as} \quad n\to\infty.$$

If there were a continuous group structure on $H(f)$, then

$$g^+=\lim_{n\to\infty}\theta^{t_n} g^- = \lim_{n\to\infty}\theta^{t_n} f\oplus g^- = f\oplus g^- = g^-,$$

an obvious contradiction. □

This example shows that the orbit closure of an a.a. point need not be a group and therefore Theorem 3.1 may be out of reach. However, if $H(p)$ is compact then there always exist θ-invariant probability measures on $H(p)$ by virtue of the Krylov–Bogoljubov theorem. According to the next proposition the orbit closure of an a.a. point is compact. Moreover, almost automorphy apparently is not far from almost periodicity.

Proposition 3.19. *If p is almost automorphic then $H(p)$ is compact and p is uniformly recurrent. If p is uniformly recurrent then $H(p)$ is a minimal set. (For compact $H(p)$ the converse of the latter statement is also true.)*

Proof. By its very definition, every a.a. point has a precompact orbit. Furthermore, if $q\in H(p)$ then $\theta(t_n)\,p\to q$ for an appropriate sequence, and (suppressing subscripts) also $\theta(-t_n)\,q\to p$ by automorphy. Hence $p\in H(q)$, and $H(p)$ is minimal. Consider any open set U containing p. It is easily seen that $O(p)\subset H(p)\subset\theta(t_1)\,U\cup\cdots\cup\theta(t_n)\,U$ for appropriate $t_1,\ldots,t_n\in\mathbb{T}$. This shows that p is uniformly recurrent. Assume in turn that p is uniformly recurrent and $q\in H(p)$. Given $\varepsilon>0$ there exists $\delta_1>0$ such that $d(\theta(t)\,p,\theta(t)\,r)<\varepsilon/3$ and $d(\theta(t)\,p,p)<\varepsilon/3$ whenever $d(p,r)<\delta_1$

and $|t| < \delta_1$. By uniform recurrence we also have $d(\theta(t_k)\,p, p) < \delta_1$ where $T := \sup_{k \in \mathbb{Z}} (t_{k+1} - t_k) < \infty$. Continuity at q implies $d(\theta(\delta_1 l)\,q, \theta(\delta_1 l)\,s) < \varepsilon/3$ whenever $d(q, s) < \delta_2$, provided that $|l| \leqslant T/\delta_1 + 1$. Putting all estimates together we see that $d(p, \theta(\delta_1 l)\,q) < \varepsilon$ for an appropriate l. Hence $p \in H(q)$, and $H(p)$ is minimal. Finally, if $H(p)$ is compact and minimal, then the same reasoning as above shows that p is uniformly recurrent. □

A general result from topological dynamics due to Ellis and Gottschalk [58] asserts that one can assign to the restriction $\theta|_M$ of the flow $\theta\colon \mathbb{T} \times P \to P$ to any compact minimal set $M \subset P$ a dynamical system $\theta'\colon \mathbb{T} \times M' \to M'$ which is a *maximal equicontinuous factor* of $(M, \theta|_M)$ in the following sense:

(i) (M', θ') is a factor of $(M, \theta|_M)$, i.e., $\pi \circ \theta(t)|_M = \theta'(t) \circ \pi$ for some continuous surjective map $\pi\colon M \to M'$ and all $t \in \mathbb{T}$;

(ii) $(\theta'(t))_{t \in \mathbb{T}}$ is equicontinuous;

(iii) every other equicontinuous factor of $(M, \theta|_M)$ is a factor of (M', θ').

It is easy to see that all maximal equicontinuous factors of (M, θ_M) are flow equivalent (and hence justifiably referred to as *the* maximal equicontinuous factor). Moreover, the points at which the factor map is one-to-one are characterized by the following celebrated theorem.

Theorem 3.20 (Veech [106]). *Let $M \subset P$ be a compact minimal set. Then*

$$\{m \in M : m \text{ is an a.a. point}\} = \{m \in M : \#\pi^{-1}\{\pi(m)\} = 1\},$$

i.e., the a.a. points in a compact minimal set are exactly those points with one-point π-fibers over the maximal equicontinuous factor.

We are now in a position to prove the main result of this section. It characterizes the situation when $H(p)$ is a compact abelian group and relates this to equicontinuity, recurrence, almost periodicity, and almost automorphy. (Remember that throughout this section $\theta(t)$ stands for the restriction of $\theta(t)$ to $H(p)$.)

Theorem 3.21. *The following statements are equivalent:*

(i) The family $(\theta(t))_{t \in \mathbb{T}}$ is equicontinuous, and p is recurrent.

(ii) $H(p)$ is a compact abelian group.

(iii) p is almost periodic.

(iv) Every point $q \in H(p)$ is almost automorphic.

Proof.

(i) ⇒ (ii) Let us first show that p if *uniformly* recurrent. Given $\varepsilon > 0$ there exists $\delta > 0$ such that $d(\theta(t)\,a, \theta(t)\,p) < \varepsilon/3$ whenever $d(a, p) < \delta$, and also $d(\theta(t_k)\,p, p) < \delta$ for some sequence $(t_k)_{k \in \mathbb{Z}}$. But then $d(\theta(t_k + t_l)\,p, p) < \varepsilon$ for all $k, l \in \mathbb{Z}$, implying that p is uniformly recurrent. It remains to show that $H(p)$ is compact. As P is complete this boils down to verifying that $H(p)$ is totally bounded. Consider the sequence $(t_k)_{k \in \mathbb{Z}}$ with $d(\theta(t_k)\,p, p) < \delta$ from above. For any t we have $d(\theta(t)\,p, \theta(t - \hat{t})\,p) < \varepsilon/3$ where $\hat{t}$ denotes the element of $(t_k)_{k \in \mathbb{Z}}$ closest to and lower than t. As every finity $\varepsilon/3$-cover of the compact set $\{\theta(t)\,p\colon 0 \leqslant t \leqslant \sup_{k \in \mathbb{Z}} (t_{k+1} - t_k)\}$ naturally yields a finite ε-cover of $H(p)$, the latter set is compact.

(ii) ⇒ (iii) Observe that $(\theta_t)_{t \in \mathbb{T}}$ is a family of translations, due to $\theta(t)\,q = q \oplus \theta(t)\,p$ for all $q \in H(p)$. Consequently, $H(p)$ is a minimal set. It is well known that there exists a metric d' on $H(p)$, uniformly equivalent to d, with respect to which θ_t is an isometry for all $t \in \mathbb{T}$ (see Walters [108]). More formally, given $\varepsilon > 0$ there exist $\delta_1(\varepsilon), \delta_2(\varepsilon) > 0$ such that $d'(a, b) < \varepsilon$ whenever $d(a, b) < \delta_1$, and $d(a, b) < \varepsilon$ whenever $\delta'(a, b) < \delta_2$. Fix now $\varepsilon > 0$ and argue as in the proof of Proposition 3.19 to find a uniformly recurrent point $r \in H(p)$, i.e., $d(\theta(t_k)\,r, r) < \delta_1(\delta_2(\varepsilon))$ for all k, and also $\sup_{k \in \mathbb{Z}} (t_{k+1} - t_k) < \infty$. Then

$$d'(\theta(t_k)\,q, q) = d'(\theta(t_k)\,r, r) < \delta_2(\varepsilon)$$

for all k and all $q \in H(p)$, implying that p is almost periodic.

(iii) ⇒ (iv) We observe that by virtue of a diagonalization argument every sequence in $H(p)$ has a Cauchy subsequence. By completeness $H(p)$ is compact. Given $q \in H(p)$ and any sequence in $O(q)$ we therefore may find a subsequence such that $\theta(t_n)\,q \to r$ as well as $\theta(-t_n)\,r \to s$ for appropriate $r, s \in H(p)$. It remains to show that $s = q$. Given $\varepsilon > 0$ and $d(\theta(t'_k)\,a, a) < \varepsilon$ for all $a \in H(p)$ and $k \in \mathbb{Z}$ where $T := \sup_{k \in \mathbb{Z}} (t'_{k+1} - t'_k) < \infty$ we see that for some t with $|t| \leqslant T$ both $d(r, \theta(t)\,q) < \varepsilon$ and $d(r, \theta(t)\,s) < \varepsilon$ hold. Since ε was arbitrary, $q = s$ follows; hence q is a.a.

(iv) $\Rightarrow$ (i) We just have to recall that $H(p)$ is compact and then apply Veech's Theorem 3.20 to see that π is in fact a homeomorphism. Therefore $(\theta(t))_{t\in\mathbb{T}}$ is equicontinuous. By Proposition 3.19, p is (uniformly) recurrent. □

Remark 3.22. The conditions in Theorem 3.21(i) are independent. On the one hand, $\theta(t)$: $x\mapsto x+t$ is equicontinuous ($\mathbb{T}=\mathbb{R}=P$), and $H(p)$ is a group which is not compact because no point is recurrent. On the other hand, consider the space $\{0, 1\}^{\mathbb{Z}}$ endowed with the product topology, and let $\theta^1=\sigma$ be the left-shift ($\mathbb{T}=\mathbb{Z}$). If p is a recurrent point with dense orbit, then $H(p)$ is compact but certainly not a group (with its algebraic structure being induced by σ).

In topological dynamics, the following corollary is sometimes drawn on for another yet equivalent definition of almost periodicity (Brown [29]).

Corollary 3.23. *Let $H(p)$ be compact. The family $(\theta(t))_{t\in\mathbb{T}}$ is equicontinuous if and only if p is almost periodic.*

Although it is an immediate consequence of Theorem 3.21, we wish to emphasize the following observation by singling it out as a statement of its own.

Corollary 3.24. *Let θ: $\mathbb{T}\times P\to P$ be a compact minimal flow. Then every point in the maximal equicontinuous factor of (P, θ) is a.p.*

Example 3.25. Consider again the function $f=(f_k)_{k\in\mathbb{Z}}=(\text{sign}(\cos 2\pi\nu k))_{k\in\mathbb{Z}}$ of Example 3.9. As in Example 3.18 we see that every limit point in $H(f)$ either is of the form $(\text{sign}(\cos(2\pi\nu k+2\pi\rho)))_{k\in\mathbb{Z}}$ where $4\rho\notin\mathbb{Z}+4\nu\mathbb{Z}$, or otherwise equals $\theta^l g^{\pm}$ for some $l\in\mathbb{Z}$. The assignment $\theta^t f\mapsto e^{2\pi i\nu t}$ $(t\in\mathbb{Z})$ may be extended to yield a continuous map π from $H(f)$ onto S^1 for which $\pi\circ\theta^1=R_\nu\circ\pi$ holds, with R_ν denoting the rotation of S^1 by an angle $2\pi\nu$. Moreover, except for the countable set $M_e:=O(g^+)\cup O(g^-)\subseteq H(f)$ the factor map π is one-to-one. It is easy to check that every point in $H(f)\setminus M_e$ is a.a. whereas no point in M_e is. This identifies (S^1, R_ν) as the maximal equicontinuous factor of $(H(f), \theta)$.

We summarize by arranging in a diagram below some of the dynamical properties we have been dealing with in this section. Implications which have been proved are indicated by double-line arrows. (The other ones are obvious.) None of these implications can be reversed in general, as we have had occasion to observe above.

$$
\begin{array}{ccccccc}
p \text{ periodic} \longrightarrow & p \text{ a.p.} & \Longrightarrow & p \text{ a.a.} & \longrightarrow & H(p) \text{ compact} & \\
& \Downarrow & & \Downarrow & & & \\
& H(p) \text{ group} & & p \text{ unif. recurrent} & \longrightarrow & p \text{ recurrent} & \\
& & \searrow & \swarrow & & & \\
& & H(p) \text{ minimal} & & & &
\end{array}
$$

With all these results at hand, we can now put into perspective the well-known fact that almost periodicity gives rise to a uniquely ergodic flow.

***Corollary* 3.26.** *Under any of the equivalent conditions in Theorem 3.21, the family $(\theta(t))_{t \in \mathbb{T}}$ is uniquely ergodic on $H(p)$.*

Proof. The orbit of the neutral element is dense in $H(p)$. The claim thus follows directly from Theorem 3.1. □

At this point it is worth recalling that the dynamical systems studied in their own right here will serve as driving systems for general NDS in the next section. As has been pointed out earlier, assumptions on the base dynamics are typically imposed in order to make the driven NDS fall into the scope of general techniques. Dynamical regularity, ensured, e.g., by unique ergodicity, plays a crucial role in this context. For the rest of the present section we therefore discuss the topic of unique ergodicity and also the somewhat related property of vanishing entropy.

Example 3.27. Recall that Example 3.25 yielded (S^1, R_ν) as the maximal equicontinuous factor of the discrete-time system $(H(f), \theta)$ from Example 3.9. Under the factor map π the exceptional set M_e projects onto

$$\pi(M_e) = \{\pm i e^{2\pi i \nu k} \colon k \in \mathbb{Z}\} = \{z \in S^1 \colon \#\pi^{-1}(\{z\}) > 1\},$$

a set of vanishing Haar measure, $\lambda_{S^1}(\pi(M_e)) = 0$. Therefore $(H(f), \theta)$ is uniquely ergodic with respect to a probability measure that projects onto λ_{S^1} under π. For the topological entropy of $(H(f), \theta)$ we have $h_{\text{top}}(\theta) = 0$ by virtue of the variational principle (Walters [108]). These facts consistently indicate that even though the system $(H(f), \theta)$ fails to be a.p., it is nevertheless rather regular and dynamically well-behaved.

We mention in passing that everything that has been said about the dynamics of $(f_k)_{k \in \mathbb{Z}} = (\text{sign}(\cos(2\pi\nu k)))_{k \in \mathbb{Z}}$ under $\theta = \theta^1$ may be given a continuous-time analogue by means of a straightforward suspension construction. Indeed, the function

$$f(t) := \sum_{k \in \mathbb{Z}} f_k \max\{0, 1-2\,|t-k|\}$$

is a.a. under $(\theta_t)_{t \in \mathbb{R}}$, and $(H(f), \theta)$ has as its maximal equicontinuous factor the suspension flow of (S^1, R_ν), a minimal Kronecker flow on T^2. Essentially the same argument as before show that the system $(H(f), \theta)$ is uniquely ergodic and has zero entropy. (See Berger, Siegmund, and Yi [18] for details and more examples.)

The observations in the last example concerning unique ergodicity and zero entropy are fairly general. Let M be a minimal set under θ and denote by (M', θ') the maximal equicontinuous factor. According to Corollary 3.26 there exists a unique θ'-invariant probability measure μ' on M'. Let M_e denote the exceptional set in M, that is

$$M_e := \{m \in M \colon \#\pi^{-1}\{\pi(m)\} > 1\}.$$

By invariance of M_e and (unique) ergodicity of the almost periodic (maximal equicontinuous) factor we get $\mu'(\pi(M_e)) \in \{0, 1\}$. If $\mu'(\pi(M_e)) = 1$ then (M, θ) is also uniquely ergodic. (This situation occurs in Example 3.27.) If on the other hand $\mu'(\pi(M_0)) = 0$ then the question of unique ergodicity is more delicate and has to be tackled by other means. Examples show that both positive and negative results may be found ([18, 78, 99]). As far as entropy is concerned, we have the following simple fact.

Proposition 3.28. *Let P, Q denote compact metric spaces, and let (Q, ϑ) be a factor of (P, θ) via $\pi\colon P \to Q$. If $\sup_{q \in Q} \#\pi^{-1}(\{q\}) < \infty$, then $h_{\mathrm{top}}(\vartheta) = h_{\mathrm{top}}(\theta)$.*

Proof. We have to show that $h_{\mathrm{top}}(\theta) \leqslant h_{\mathrm{top}}(\vartheta)$ because the reverse inequality is obvious. By a theorem due to Bowen [27] the estimate $h_{\mathrm{top}}(\theta) \leqslant h_{\mathrm{top}}(\vartheta) + \sup_{q \in Q} h_{\mathrm{top}}(\theta, \pi^{-1}(\{q\}))$ holds. Since the last summand vanishes for all q due to the finiteness of $\pi^{-1}(\{q\})$, the result follows. □

As an application of this proposition, we could have proved $h_{\mathrm{top}}(\theta) = 0$ in Example 3.27 without invoking the variational principle.

Example 3.29. We briefly sketch an a.a. system which lacks unique ergodicity and also has positive entropy; all the relevant details may be found in [18]. Let $\Sigma_2 = \{0, 1\}^{\mathbb{Z}}$ denote the space of bi-infinite sequences on two symbols, endowed with any metric inducing the product topology. The (left) shift map on Σ_2 is denoted by σ. Fix $\omega = (\omega_k)_{k \in \mathbb{Z}} \in \Sigma_2$ and consider the sets

$$J_n := (j_n + 2^{N+n}\mathbb{Z}) \cup (-j_n + 2^{N+n}\mathbb{Z}) \subseteq \mathbb{Z} \quad (n \in \mathbb{N}_0),$$

where $N \geqslant 2$, and the integers j_n are determined inductively according to

$$j_0 := 0 \qquad \text{and} \qquad j_{n+1} := \min\left\{k \geqslant 0 \colon k \notin \bigcup_{i=0}^{n} J_i\right\}.$$

Evidently, $\mathbb{Z}$ equals the disjoint union of the sets J_n. Define a point $x(\omega) \in \Sigma_2$ by setting

$$(x(\omega))_k = x_k(\omega) := \sum_{n=0}^{\infty} \omega_n \mathbf{1}_{J_n}(k).$$

In other words, $x_k(\omega) = \omega_n$ whenever $k \in J_n$. It is easy to see that $x(\omega)$ is a.a. with respect to σ. Moreover, $x(\omega)$ is not a.p. unless $(\omega_k)_{k \in \mathbb{Z}}$ is constant eventually. It is shown in [18] that for every $\varepsilon > 0$ there exists an integer $N(\varepsilon)$ and a residual set $\Omega_\varepsilon \subseteq \Sigma_2$ such that for $N \geqslant N(\varepsilon)$ and for every $\omega \in \Omega_\varepsilon$ the symbolic dynamical system $(H(x(\omega)), \sigma)$ lacks unique ergodicity and has entropy $h_{\text{top}}(\sigma|_{H(x(\omega))}) > \log 2 - \varepsilon$, which is close to the largest possible value $\log 2$. The corresponding maximal equicontinuous factor of $(H(x(\omega)), \sigma)$ turns out to be $\langle \mathbb{Z}_2, \oplus_1 \rangle$, i.e., the (totally disconnected) group of dyadic integers with the addition by one. □

Example 3.30. In Example 3.10 we already observed that for the almost automorphic function f given there, the hull $H(f)$ is not compact in the compact-open topology. Nevertheless, a neat description of $H(f)$ may again be given by means of a function resembling the function Ψ in Example 3.5. To this end let P denote the complement of the "forbidden" trajectory of $(\frac{1}{2}, \frac{1}{2})$ under ϑ, i.e., $P := T^2 \setminus \{\vartheta_t(\frac{1}{2}, \frac{1}{2}) \colon t \in \mathbb{R}\}$, metricized as a subspace of T^2. Defining $\Psi \colon P \to C(\mathbb{R}, \mathbb{R}^2)$ by $\Psi(x, y) := f_0(\vartheta_{\bullet}(x, y))$ again yields a continuous one-to-one map. Furthermore $\Psi \circ \vartheta_t = \theta_t \circ \Psi$ for all $t \in \mathbb{R}$, and Ψ maps P onto $H(f)$. Since $\Psi(x, y)(t) = f_0(\vartheta_t(x, y))$ varies rapidly whenever $\vartheta_t(x, y)$ is near $(\frac{1}{2}, \frac{1}{2})$, two points (x, y) and (x', y') of P are near to each other only if $\Psi(x, y)$ is close to $\Psi(x', y')$, i.e., the inverse map Ψ^{-1} is also continuous. Therefore, $(P, \vartheta|_P)$ and $(H(f), \theta)$ are flow equivalent. As a consequence, the latter system is uniquely ergodic, its unique invariant probability measure being $\Psi(\lambda_{T^2}|_P)$.

Example 3.31. As the preceding example shows, a lack of compactness of $H(p)$ does not necessarily rule out (finite) ergodic theory as a tool. However, it may happen that no invariant probability measure at all exists on $H(p)$. Consider for instance the Bebutov flow θ of the function

$$A(t) := 2t \cos t^2.$$

We claim that there does not exist any θ-invariant probability measure on $H(A)$. This may easily be seen as follows.

Let $(\vartheta_t)_{t \in \mathbb{R}}$ denote the flow corresponding to a unit velocity motion to the right on the real line, i.e., $\vartheta_t \colon x \mapsto x+t$. A cumbersome yet elementary analysis confirms that for some $\alpha > 0$ the estimate

$$\sup_{|t| \leqslant 1} |\theta_{s_1} f(t) - \theta_{s_2} f(t)| \geqslant \alpha \min\{1, |s_1 - s_2|\}$$

holds, where α does not depend on s_1, s_2. The assignment $\Psi \colon \theta_t f \mapsto t = \vartheta_t(0)$ therefore is uniformly continuous (with respect to the metric d_{co}), and clearly $\Psi(\theta_t f) = \vartheta_t \circ \Psi(f)$. Hence any θ-invariant probability measure on $H(A)$ would induce via Ψ a ϑ-invariant probability measure on $\mathbb{R}$, an object which evidently does not exist.

As indicated by the above examples, the case of non-compact $H(p)$ typically needs a refined analysis, and invariant probability measures have to be looked for by means of techniques tailored to the particular system under consideration. As merely one general fact we mention the following consequence of Prokhorov's theorem (Stroock [105]). Recall that we have constantly assumed P to be a complete space.

Proposition 3.32. *If the metric space P is separable (and hence Polish), then there exists a θ-invariant probability measure if and only if for some probability measure μ and any $\varepsilon > 0$ one can find a compact set $K_\varepsilon \subset P$ such that $\mu(\theta_t(K_\varepsilon)) > 1 - \varepsilon$ holds for all $t \in \mathbb{T}$.*

As we have seen throughout this section, almost periodicity is easy to grasp from a topological as well as from a statistical viewpoint. Almost automorphy, though seemingly still denominating a regular pattern of recurrence, may already indicate a certain dynamical complexity, noticeable, e.g., through a lack of unique ergodicity, positive entropy etc. Even for an a.a. point p, however, $H(p)$ is a compact set, and notions from topological dynamics *and* ergodic theory both apply.

4. NONAUTONOMOUS DYNAMICS

In this section we give a series of examples which illustrate the current gap between random dynamical systems and continuous skew products. We survey selected topics and compare the results achieved for both concepts.

Example 4.1 (Lyapunov's Second Method and Attractors). In order to recall Lyapunov's second method, let φ be a continuous dynamical system on a locally compact metric space X and let A be a nonvoid compact set which is invariant under φ. The following two statements are equivalent (Bhatia and Szegö [19], see also Sell and You [98] for a definition of stability which allows for the semiflow to have a singularity at $t=0$ and references on the related LaSalle Invariance Principle):

(a) A is asymptotically stable, i.e.,

(i) for all $\varepsilon > 0$ there exists a $\delta > 0$ such that $\varphi(t)\, U_\delta(A) \subset U_\varepsilon(A)$ for all $t \geqslant 0$, where $U_\varepsilon(A) := \{x \in X : d(x, A) < \varepsilon\}$, $d(x, A) := \inf_{y \in A} d(x, y)$, is the open ε-neighborhood of A,

(ii) A is the attractor of φ, i.e., $\lim_{t \to \infty} d(\varphi(t, x), A) = 0$ for all $x \in X$.

(b) There exists a Lyapunov function $V: X \to \mathbb{R}^+$ for A, i.e.,

(i) V is continuous,

(ii) V is uniformly bounded, i.e., for all $C > 0$ there exists a compact set $K \subset X$ such that $V(x) \geqslant C$ for all $x \notin K$,

(iii) V is positive-definite, i.e., $V(x) = 0$ if $x \in A$, and $V(x) > 0$ if $x \notin A$,

(iv) V is strictly decreasing along orbits of φ, i.e., $V(\varphi(t, x)) < V(x)$ for $x \notin A$ and $t > 0$.

In [9], Arnold and Schmalfuss generalize Lyapunov's second method for RDS. It is a special feature of their work that it identifies the matching random notions of stability, Lyapunov functions, and attraction which allow for a coherent extension of the deterministic result. The definition of an attractor has been generalized to the random case by Crauel and Flandoli [49], Crauel, Debussche, and Flandoli [48], and Schmalfuss [92]. Meanwhile there are at least two different notions of a random attractor, that is a random compact set A which is invariant, i.e., satisfies $\varphi(t, \omega)\, A(\omega) = A(\theta(t)\, \omega)$ for all $t \in \mathbb{T}$:

(i) A is a *cocycle or pullback attractor* if for any bounded set $D \subset X$

$$\lim_{t \to \infty} d(\varphi(t, \theta(-t)\,\omega)\,D \,|\, A(\omega)) = 0 \qquad \mathbb{P}\text{-almost surely},$$

where $d(A \,|\, B) := \sup_{x \in A} d(x, B)$ is the Hausdorff semi-metric. Note that this implies $\mathbb{P}\text{-}\lim_{t \to \infty} d(\varphi(t, \cdot\,)\,D \,|\, A(\theta(t)\,\cdot\,)) = 0$.

(ii) A is a *weak attractor* or *attractor in probability* if, for any bounded set $D \subset X$,

$$\mathbb{P}\text{-}\lim_{t \to \infty} d(\varphi(t, \cdot\,)\,D, A(\theta(t)\,\cdot\,)) = 0,$$

i.e., if $\lim_{t \to \infty} \mathbb{P}\{\omega\colon d(\varphi(t, \omega)\,D, A(\theta(t)\,\omega)) > \varepsilon\} = 0$ for all $\varepsilon > 0$.

Similarly to (i), a *forward attractor* can be defined by requiring that

$$\lim_{t \to \infty} d(\varphi(t, \omega)\,D \,|\, A(\theta(t)\,\omega)) = 0.$$

Another notion of attractor using a "background measure" is the *Milnor attractor*, see Ashwin [11]. One can show by means of examples that all these concepts lead to different notions of attractors. For a comparison see Ashwin and Ochs [14], Caraballo and Langa [30], Cheban, Kloeden, and Schmalfuss [40], and Scheutzow [91]. An attractor may also depend on the class of attracted sets: Allowing the attracted sets D to depend on ω, one can distinguish between local attractors attracting different families $\{D(\omega)\}$, the so-called attracting universes. Note that in the autonomous case all these concepts conincide.

The above notions of attractors (i) and (ii) describe the attraction of (bounded) sets. As in the autonomous case, there is a difference between set and point attractors. A point attractor version of (ii) which allows for attraction of arbitrary random variables reads as follows.

(iii) A is a *weak (point) attractor* if for any random variable x

$$\mathbb{P}\text{-}\lim_{t \to \infty} d(\varphi(t, \cdot\,)\,x(\cdot\,), A(\theta(t)\,\cdot\,)) = 0.$$

It turns out that (iii) is the adequate notion to prove Lyapunov's second method for RDS, which is a strong indication that (iii) is the natural definition of a *random attractor*, see [9] for details. However, it is the notion (i) of a cocycle attractor which has an obvious correspondence for SPF (by omitting "$\mathbb{P}$-almost surely"), whereas (ii) and (iii) are random notions which intrinsically make use of the probability measure on the base. Moreover, Caraballo and Langa [31] (see also [33]) show that a cocycle attractor A_ε of an abstract NDS φ_ε which is a perturbation of an NDS φ, converges upper semicontinuously to the unperturbed attractor A, provided that A_ε is contained in a compact absorbing set which converges in the Hausdorff semidistance to a compact forward invariant absorbing set of A. They apply this upper semicontinuity result to a nonautonomous perturbation of an autonomous attractor and to a random perturbation of a nonautonomous attractor, thereby showing that the concept of cocycle attractor for abstract NDS is a meaningful notion for both SPF and RDS.

We should point out that cocycle attractors for NDS are not unique in general. Uniqueness follows, e.g., if the attractor is bounded uniformly over the base, or if the base flow preserves a probability measure, i.e., in the situation of RDS. For attractors in the extended state space $P \times X$ with compact P, see Babin and Sell [17] and Sell and You [98], for their relation to pullback and forward attractors, see Cheben, Kloeden, and Schmalfuss [40]. Pullback and forward attractors for a nonautonomous Lotka–Volterra system are studied by Langa, Robinson, and Suarez [75, 76], results on attractors for multivalued RDS are contained in Caraballo, Langa, and Valero [32, 36, 37], and finiteness of the attractor dimension for specific infinite-dimensional systems is shown, e.g., by Debussche [52, 53], Caraballo, Langa, and Robinson [34], and Caraballo, Langa, and Valero [38].

For SPF, Lyapunov's second method has not been established yet. This seems to be an interesting open problem, as any answer is likely to indicate a "correct" generalization of the attractor notion for SPF. However, there are already many results available on nonautonomous Lyapunov functions, and especially on pullback attractors by Kloeden (see, e.g., [69]) and Cheban, Kloeden, and Schmalfuss (see [39] and references therein).

The next example shows two numerical NDS which are viewed neither as RDS nor as SPF, and it suggests that numerics for RDS which are done pathwise, i.e., for fixed ω; have the same structure as numerics for SPF.

Example 4.2 (Numerical NDS). Consider a numerical scheme for an ODE $\dot{x} = f(x)$ as in Kloeden, Keller, and Schmalfuss [70]. An explicit one-step numerical scheme with variable time steps $h_n > 0$ is often written as

$$x_{n+1} = F_{h_n}(x_n) := x_n + hf_{h_n}(x_n),$$

with increment function f_{h_n} (e.g., in the Euler scheme $h_n = h > 0$, $f_h(x) = f(x)$, and in the Heun scheme $h_n = h > 0$, $f_h(x) = \frac{1}{2}[f(x) + f(x + hf(x))]$). Let the base P be the set of positive bi-infinite sequence $h = (h_j)_{j \in \mathbb{Z}}$ which form divergent series in both directions with θ as the shift operator, i.e., with $h' = \theta^n h$ defined by $h'_j := h_{n+j}$. Then the cocycle

$$\varphi(0, h)\, x_0 := x_0, \qquad \varphi(n, h)\, x_0 := F_{h_{n-1}} \circ \cdots \circ F_{h_0}(x_0),$$

for $n \geqslant 1$, $h = (h_j)$ and $x_0 \in \mathbb{R}^d$ defines an NDS which models the numerical scheme. For an application to the discretization of attractors see Kloeden and Schmalfuss [71].

Another class of numerical NDS describing a numerical method for both RDS and SPF is the box algorithm of Dellnitz and Junge (see, e.g., Dellnitz, Froyland, and Junge [54] for an introduction). It was used by Keller and Ochs [68] to compute random attractors, and by Siegmund [103] to approximate nonautonomous invariant manifolds. The abstract formulation for an NDS $\varphi\colon \mathbb{Z}_+ \times P \times \mathbb{R}^d \to \mathbb{R}^d$ is as follows: Choose a compact set $Q \subset \mathbb{R}^d$, a step size $K \in \mathbb{N}$ and a finite collection $\mathscr{B} = (B_i)_{i=1}^m$ of connected, closed subsets B_i of Q with (i) $\bigcup_{i=1}^m B_i = Q$ and (ii) $\operatorname{int} B_i \cap \operatorname{int} B_j = \emptyset$ if $i \neq j$, $1 \leqslant i, j \leqslant m$. Then $\hat{\varphi}\colon \mathbb{Z}_+ \times P \times \mathscr{P}(\mathscr{B}) \to \mathscr{P}(\mathscr{B})$ is the *box NDS* over $\hat{\theta} = \theta^K$, where

$$\mathscr{P}(\mathscr{B}) := \left\{ B_I = \bigcup_{i \in I} B_i : I \subset \{1,..., m\} \right\}$$

and $\hat{\varphi}(n, p) = \hat{\varphi}(1, \theta^{(n-1)K} p) \circ \cdots \circ \hat{\varphi}(1, p)$ is defined by

$$\hat{\varphi}(1, p)\, B_I := B_J = \bigcup_{j \in J} B_j$$

with $J = \{ j \in \{1,..., m\} : \varphi(K, p)\, B_I \cap B_j \neq \emptyset \}$. Here the state space $\mathscr{P}(\mathscr{B})$ is a metric space with the Hausdorff metric.

The next example is on linear theory for RDS and SPF. It explains the relation between the Lyapunov exponents as provided by the multiplicative ergodic theorem of Oseledets and the spectral intervals of the Sacker–Sell spectrum, which is well-understood if the base is compact (Johnson, Palmer, and Sell [65]), for related early work in the russian literature see Millionščikov [83].

Example 4.3 (Oseledets and Sacker–Sell Spectra). Assume that P is a compact metric space. Let $\Phi\colon \mathbb{T} \times P \to Gl(d, \mathbb{R})$ be a linear continuous skew product flow over a driving system $\theta\colon \mathbb{T} \times P \to P$. Let $\Phi_\lambda(t, p)\, x := e^{-\lambda t} \Phi(t, p)\, x$ be the shifted cocycle. Recall that Φ_λ has an *exponential* *dichotomy over* P if there is a (continuous) projector $(p, x) \mapsto (p, Q(p)\,x)$ on $\mathbb{R}^d$ and constants $K \geqslant 1$, $\alpha > 0$, such that

$$\|\Phi_\lambda(t, p)\, Q(p)\, \Phi_\lambda^{-1}(s, p)\| \leqslant Ke^{-\alpha(t-s)}, \quad t \leqslant s,$$
$$\|\Phi_\lambda(t, p)[I - Q(p)]\, \Phi_\lambda^{-1}(s, p)\| \leqslant Ke^{\alpha(t-s)}, \quad t \geqslant s,$$

for all $p \in P$ and $t, s \in \mathbb{T}$. The set of $\lambda \in \mathbb{R}$ for which Φ_λ fails to have an exponential dichotomy over P is defined to be dyn Σ, the *dynamical* (or *dichotomy* or *Sacker–Sell*) *spectrum*. The spectral theorem (Sacker and Sell [89]) assures that dyn $\Sigma = \bigcup_{i=1}^k [a_i, b_i]$ is the union of k nonoverlapping compact intervals, where $1 \leqslant k \leqslant d$. The boundary of dyn Σ is the finite collection of end points $\{a_1,..., a_k, b_1,..., b_k\}$. If μ is an ergodic measure on P then the set $\Sigma(\mu) = \{\lambda_1,..., \lambda_k\}$ of Lyapunov exponents of Φ is the Oseledets spectrum (w.r.t. μ). If, additionally, P is connected then

$$\text{boundary dyn } \Sigma \subset \bigcup_\mu \Sigma(\mu) \subset \text{dyn } \Sigma,$$

where the union is taken over all ergodic measure μ on P. (A slightly more technical inclusion result with the union taken over all invariant probability measures on P can be found in [65]).

What can be said if the base P is not compact? Consider the scalar differential equation $\dot{x} = A(t)\,x$ with $A(t) = 2t\cos(t^2)$, $t \in \mathbb{R}$, from Example 3.31. Let φ be the continuous skew product flow over the Bebutov flow θ on the hull $H(A) = \text{cl}\{A(\cdot + s) \colon s \in \mathbb{R}\}$ in some topology (e.g., the compact-open topology). Then for $\theta_s A = 2(\cdot + s)\cos(\cdot + s)^2 \in H(A)$ we get

$$\varphi(t, \theta_s A)\, x = \exp(\sin(s+t)^2 - \sin s^2)\, x.$$

Now we show that the hull $H(A)$ cannot be compact. Arguing negatively, assume that $H(A)$ is compact. Using the continuity there is an $\varepsilon > 0$ such that

$$|\varphi(t, \sigma(s, A))| < 2 \qquad \text{for} \quad |t| < \varepsilon \qquad \text{and all} \quad s \in \mathbb{R}.$$

Now choose $k \in \mathbb{N}$ with $t := \sqrt{\pi/2 + 2k\pi} - \sqrt{-\pi/2 + 2k\pi} < \varepsilon$ and $s := \sqrt{-\pi/2 + 2k\pi}$ to obtain the contradiction

$$\exp(2) = \exp(\sin(s+t)^2 - \sin s^2) < 2,$$

proving that the hull is not compact with respect to any topology. A simple computation shows that the dynamical spectrum of this equation is dyn $\Sigma = \{0\}$ and that the Lyapunov exponent exists as a limit and equals 0.

On the other hand the multiplicative ergodic theorem is not applicable because there exists no ergodic measure on the hull as we have seen in Example 3.31.

In the following example, one can see that once the linear theories are provided, one can expect analog qualitative theories for RDS and SPF. Also, the proofs are similar if only one knows how to cope with the non-uniformity of RDS.

Example 4.4 (Hartman–Grobman, and Normal Forms). Invariant manifold theory is one of the cornerstones of qualitative theory. It dates back to Hadamard (graph transformation method) as well as Lyapunov and Perron (Lyapunov–Perron method). We certainly cannot survey the vast literature on the subject, but we have to mention some results for SPF with compact base space which use exponential dichotomy and the Sacker–Sell spectral theory, namely Chow and Yi [42] and Yi [111, 112] for classical as well as Sell [96] for generalized center manifolds. Invariant manifold theory for RDS is part of smooth ergodic theory and was initiated by Pesin in 1976. His technique to cope with the non-uniformity of the linear theory provided by the multiplicative ergodic theory can be adapted to RDS. More recently, Wanner [110] used this technique to transfer the deterministic construction of center manifolds and foliations to RDS, and he was thus able to prove the Hartman–Grobman result for RDS by using the same ideas as in the deterministic case [109]. Once an Oseledets splitting $\mathbb{R}^d = E_1(\omega) \oplus \cdots \oplus E_n(\omega)$ for the linearized cocycle Φ is given, a *random norm* $\|x\|_{\kappa,\omega}$ is constructed which crucially improves the uniformity in the behavior of Φ by means of the estimate

$$e^{\lambda_i t - \kappa|t|} \leqslant \|\Phi(t, \omega)_{E_i(\omega)}\|_{\kappa, \omega, \theta(t)\omega} \leqslant e^{\lambda_i t + \kappa|t|} \qquad \text{for all} \quad t \in \mathbb{T},$$

where $\lambda_1,..., \lambda_n$ are the corresponding Lyapunov exponents provided by the multiplicative ergodic theorem of Oseledets.

In contrast to the topological linearization of hyperbolic systems accomplished by the Hartman–Grobman theorem, the aim of normal form theory is to simplify (ultimately linearize) a system by means of a smooth coordinate transformation. However, some "resonant" terms may defy elimination, and so the "simplest possible" form in general is nonlinear. It was Poincaré who founded the normal form theory in his thesis in 1879. For a system $\dot{x} = f(x)$ with equilibrium 0 and eigenvalues $\lambda_1,..., \lambda_n$ of the linearization $Df(0)$, he formulated a nonresonance condition

$$\ell_1\lambda_1 + \cdots + \ell_n\lambda_n \neq \lambda_j,$$

which ensures the existence of a smooth transformation eliminating the jth component of the Taylor coefficient $\frac{1}{\ell_1! \cdots \ell_n!} D_{x_1}^{\ell_1} \cdots D_{x_n}^{\ell_n} f(0)$ of the nonlinearity. Normal form theory for RDS generated by random differential/difference equations or stochastic differential equations is elegantly developed and described in Arnold [4]. The linear theory of course again builds on the multiplicative ergodic theorem, and the eigenvalues in Poincaré's nonresonance condition are replaced by the Lyapunov exponents. Later, Siegmund [101, 102] and Colonius and Siegmund [46] extended normal form theory to SPF generated by nonautonomous differential/difference equations and control systems (see Colonius and Kliemann [45]), respectively. For these systems, the linear theory relies on the *Sacker–Sell* or *dichotomy spectrum* which consists of disjoint compact intervals $\lambda_i = [a_i, b_i]$, where $b_i < a_{i+1}$, $i = 1, \ldots, n-1$. (We also write $\lambda_1 < \cdots < \lambda_n$ in this situation). The corresponding nonresonance condition which generalizes Poincaré's condition takes the form

$$\ell_1 \lambda_1 + \cdots + \ell_n \lambda_n < \lambda_j \qquad \text{or} \qquad \ell_1 \lambda_1 + \cdots + \ell_n \lambda_n > \lambda_j,$$

where multiples and sums of sets are taken pointwise. □

An extension to RDS and SPF of bifurcation results for dynamical systems in anything but obvious, and it seems that new concepts have to be found.

Example 4.5 (Bifurcation Theory). Bifurcation theory for NDS is a relatively new branch which has been developed almost independently for RDS and SPF so far. Remarkable succes has been achieved for RDS by Ludwig Arnold and his "Bremen Group." The basic concepts of their theory are laid down in Chapter 9 (Bifurcation Theory) of the monograph [4]. One approach is *dynamical* or *D-bifurcation* which is related to sign changes of Lyapunov exponents $\lambda_i(\mu_\alpha)$ of φ_α-invariant measures μ_α, with α denoting a bifurcation parameter, as opposed to qualitative changes of stationary densities of the corresponding Markov process (denoted *P-bifurcations*), the latter being not related to the stability measured by Lyapunov exponents of the RDS. Examples show that the concept of D-bifurcation is more adequate to generalize deterministic bifurcation scenarios. Moreover, as it is summarized by Arnold in [5], the complete analysis of the case of stochastic differential equations (SDE) with state space $\mathbb{R}$, done by Crauel, Imkeller, and Steinkamp [51], shows that in dimension one, invariant measures bifurcate from invariant measures at parameter values where the Lyapunov exponent is equal to zero, and this is all that can happen! In higher dimensions, less is know. A promising approach of Ashwin [11] and Ashwin and Ochs [14] is to use a notion of Milnor attractor to relate D-bifurcations to blowout bifurcations of chaotic attractors, see also [12, 13]. For a result on a stochastic pitchfork bifurcation in the infinite-dimensional setting, see Caraballo, Langa, and Robinson [35]. Bifurcation theory for SPF seems to be intricate, too, as the few papers on the subject suggest. Langa, Robinson, and Suarez [74] give, by means of relatively simple examples of nonautonomous pitchfork and saddle-node bifurcations, an illustration of the idea of a bifurcation as a change in the structure and stability of a cocycle attractor. In more technical papers, Johnson and Yi investigate a Hopf bifurcation from non-periodic solutions of differential equations, thereby continuing earlier investigations on the bifurcation of invariant tori (see [64, 66] and references therein). In analogy to RDS, they use the linearization and its spectrum to analyze the nonlinear bifurcating system. More precisely, they assume that a one-parameter family of C^3 vector fields admits a one-parameter family $\mu \mapsto Y_\mu$ of compact invariant sets such that (i) Y_μ is an asymptotically stable attractor for $\mu < 0$, (ii) Y_μ is no longer an attractor for $\mu > 0$ and (iii) all Y_μ are homeomorphic (but not diffeomorphic) to a 2-torus. Using a variant of center manifold theory and rotation numbers, they provide assumptions under which a parameter interval $(0, \delta)$ contains an open and dense subset such that the system admits a stable attracting 2-torus Z_μ which depends in a strongly discontinuous way on $\mu \in (0, \delta)$.

Center manifolds and the reduction principle of Pliss [87] are indispensable tools in the bifurcation theory of dynamical systems [77]. The counterparts for RDS and SPF are developed for various situations (see, e.g., [4, 15, 42, 79, 84, 110] and the references therein) and they are expected to be similarly useful in the development of a nonautonomous bifurcation theory for reduced lower-dimensional NDS.

Today, monotone methods and comparison arguments are increasingly intertwined with dynamical systems theory (see, e.g., Smith [104] and the literature quoted therein). Monotonicity simplifies the investigation of the long-time behavior and of the invariant objects of a dynamical system.

Example 4.6 (Monotone Systems). In a series of papers, Arnold and Chueshov [6–8] and Chueshov [43] (see also the book [44]) developed a systematic study of order-preserving (or monotone) RDS and SPF culminating in a limit set trichotomy for order preserving systems. Their definition of an order preserving RDS, resp. SPF, naturally extends to NDS. Thereto, let $X \neq \emptyset$ be a subset of a real Banach space V and let $V_+ \subset V$ be a closed convex cone such that $V_+ \cap (-V_+) = \{0\}$. This cone defines a partial order on X as follows: We have $x \geqslant y$ if and only if $x - y \in V_+$, and we write $x > y$ when $x \geqslant y$ but $x \neq y$. If V_+ has nonempty interior $\mathrm{int}(V_+)$ we say that V is strongly ordered, and we write $x \gg y$ if $x - y \in \mathrm{int}(V_+)$. Moreover, assume that every bounded set B in X is contained in an order interval. Then an NDS (θ, φ) is said to be *strongly order-preserving* if

$$x > y \quad \text{implies} \quad \varphi(t, p)\, x \gg \varphi(t, p)\, y \quad \text{for all} \quad t \geqslant 0 \quad \text{and} \quad p \in P.$$

Shen and Yi show in [99] that a scalar parabolic PDE generates a strongly order-preserving SPF. Exploiting zero number properties which hold for this special class of PDEs they get, e.g., the result that if the driving system is compact and has a unique ergodic probability measure $\mathbb{P}$ then unique ergodicity of a minimal set $E \subset P \times X$ of the SPF (θ, φ) is equivalent to

$$\mathbb{P}\{p \in P : \#E(p) = 1\} = 1,$$

where $\#E(p)$ denotes the cardinality of the fiber of E over p. The unique ergodic probability measure on E then is a lifting of $\mathbb{P}$. The lifting problem for RDS is investigated by Crauel [47], and Eckmann and Hairer [56, 57]. It would be interesting to formulate these results for NDS in order to extract the common core of the theory for RDS and SPF and to obtain, for instance, a limit set trichotomy theorem for abstract order-preserving NDS.

5. CONCLUSIONS

What is the gap between random dynamical systems (RDS) and continuous skew products (SPF)? In Section 2, we observed as their obvious common structure that they both constitute a cocycle over a group action of time (the driving-system). Following the recent coining of the term in the Festschrift [50], we called this common structure a *nonautonomous dynamical system* (NDS). From the point of view adopted here, it is thus mainly the gap between ergodicity and continuity of the driving system which causes the gap between RDS and SPF. Moreover, a typical feature of RDS is the lack of compactness in the base, and therefore, in general, also a lack of uniformity. As we pointed out by quoting from Oxtoby [85] earlier, one should not expect too much from analogies alone, even though the concepts do look similar. However, in the compact and uniquely ergodic case for example the gap disappears, and the concepts coincide to a large extent. Almost periodic dynamics (which are always conjugate to group rotations) provide the best-known examples for such a good match. The concept of almost automorphy, on the other hand, "is essential and fundamental in the qualitative study of almost periodic differential equations" (Shen and

Yi [99]). Example 3.18 shows that almost automorphic dynamics do not yield a group in general, although an almost automorphic flow always is compact and minimal by Proposition 3.19. Nevertheless, Example 3.27 clearly points out that in order to better understand almost automorphic dynamics, it is indispensable to have a closer look at the associated maximal equicontinuous (hence almost periodic) factor. If the set of points supporting one-point fibers over this factor has full Haar measure, then the almost automorphic flow has zero entropy. In other words, an almost automorphic flow can have positive (topological) entropy only if the set of one-point fibers has zero Haar-measure, as is the case in Example 3.29.

Quantitative concepts like entropy (or, more generally, topological pressure, see Walters [108]) or spectra may equally stimulate a broader view on NDS in the spirit of the present article (notwithstanding the fact that the technicalities inherent to these topics largely precluded a detailed discussion here). For example, in its measure-theoretic form, entropy quantifies *on average* the complexity inherent to the system whereas its topological counterpart measures the *maximal* complexity. The well-known variational principle asserts that the topological entropy actually equals the supremum over the measure-theoretic entropy taken at all ergodic measures. A similar relation exists between the Oseledets and the Sacker–Sell spectrum, as outlined in Example 4.3: The union of the Lyapunov exponents over all ergodic measures is an inner approximation for the Sacker–Sell spectrum which measures the exponential growth rates uniformly over the base. For RDS, entropy has been developed in Bogenschütz [23]. It turns out that most of the classical notions and results (e.g., the Shannon–McMillan–Breiman theorem, the variational principle, etc.) carry over to the nonautonomous context in a very natural way, see also Bogenschütz and Crauel [24].

It seems as if some issues of uniformity or maximality of concepts for SPF might be handled successfully by considering the corresponding concepts for the family of RDS in view of all ergodic measures of the driving system. As has been pointed out in Section 4, so far only ad-hoc techniques exist in this context. A careful, systematic analysis also assessing the potential of the approach therefore constitutes one of the directions along which—in our opinion—further investigations will be worthwhile. Another pressing problem suggests itself through Example 4.1 on Lyapunov's second method: Has the natural definition of an attractor for SPF really been found yet? Does it allow a coherent generalization of Lyapunov's second method to SPF, just as the notion of random attractor does for RDS? It is the similar structure of results for RDS and SPF, as apparent, e.g., in Example 4.2 and 4.4, which encourages this simultaneous study. It is an interesting problem to find out to which extent such a structure also exists for the monotone systems in Example 4.6. As a possible strategy, one could compare the limit set theorem for monotone RDS [8] with the result for SPF [43] in order to understand up to which extent the new findings reflect a property of the abstract NDS alone, not depending on the particular driving system.

It is beyond doubt that any refined knowledge about the similarities and differences between RDS and SPF will be an invaluable help when trying to profit from the work in *both* areas. As one field of future research which certainly requires such a two-eyed perspective, we finally have to mention the development of a comprehensive bifurcation theory for NDS. To the best of our knowledge, this highly interesting though challenging subject has not entered the phase of a substantial breakthrough yet.

ACKNOWLEDGMENT

The first author was supported by a Max Kade fellowship, the second author by an Emmy Noether fellowship.

REFERENCES

1. Amann, H. (1990). *Ordinary Differential Equations*, de Gruyter, Berlin.
2. Anzai, H. (1951). Ergodic skew product transformations on the torus. *Osaka Math. J.* **3**, 83–99.
3. Arnold, L. (1997). The unfolding of dynamics in stochastic analysis. *Mat. Apl. Comput.* **16**, 3–25.
4. Arnold, L. (1998). *Random Dynamical Systems*, Springer, Berlin, Heidelberg, New York.
5. Arnold, L. (2001). Recent progress in stochastic bifurcation theory. In Narayanan, S., and Iyengar, R. N. (eds.), *IUTAM Symposium on Nonlinearity and Stochastic Structural Dynamics (Madras, Chennai, India, 1999)*, Kluwer, Dordrecht, pp. 15–27.
6. Arnold, L., and Chueshov, I. D. (1998). Order-preserving random dynamical systems: Equilibria, attractors, applications. *Dynam. Stability Systems* **13**, 265–280.
7. Arnold, L., and Chueshov, I. D. (2001). Cooperative random and stochastic differential equations. *Discrete Contin. Dynam. Systems* **7**, 1–33.
8. Arnold, L., and Chueshov, I. D. (2001). A limit set trichotomy for order-preserving random systems. *Positivity* **5**, 95–114.
9. Arnold, L., and Schmalfuss, B. (2001). Lyapunov's second method for random dynamical systems. *J. Differential Equations* **177**, 235–265.
10. Artstein, Z. (1998). Topological dynamics of an ordinary differential equation. *J. Differential Equations* **23**, 216–223.
11. Ashwin, P. (1999). Minimal attractors and bifurcations of random dynamical systems. *R. Soc. Lond. Proc. Ser. A Math. Phys. Eng. Sci.* **455**, 2615–2634.
12. Ashwin, P. (1999). Attractors of a randomly forced electronic oscillator. *Phys. D* **125**(3–4), 302–310.
13. Ashwin, P., Buescu, J., and Stewart, I. (1996). From attractor to chaotic saddle: A tale of transverse instability. *Nonlinearity* **9**(3), 703–737.
14. Ashwin, P., and Ochs, G. (2003). Convergence and local random attractors. *Dyn. Syst.* **18**(2), 139–158.

15. Aulbach, B. (1982). A reduction principle for nonautonomous differential equations. *Arch. Math. (Basel)* **39**(3), 217–232.
16. Aulbach, B., and Wanner, T. (1996). Integral manifolds for Carathodory type differential equations in Banach spaces. In *Six Lectures on Dynamical Systems (Augsburg, 1994)*, World Scientific Publishing, River Edge, NJ, pp. 45–119.
17. Babin, A. V., and Sell, G. R. (2000). Attractors of non-autonomous parabolic equations and their symmetry properties. *J. Differential Equations* **160**, 1–50.
18. Berger, A., Siegmund, S., and Yi, Y. (2003). On almost automorphic dynamics in symbolic lattices. *Ergodic Theory Dynam. Systems*. In press.
19. Bhatia, N. P., and Szegö, G. H. (1970). *Stability Theory of Dynamical Systems*, Springer, Berlin, Heidelberg, New York.
20. Bochner, S. (1927). Beiträge zur Theorie der fastperiodischen Funktionen, I: Funktionen einer Variablen. *Math. Ann.* **96**, 119–147.
21. Bochner, S. (1955–1956). Curvature and Betti numbers in real and complex vector bundles. *Univ. Politec. Torino, Rend. Sem. Mat.* **15**, 225–253.
22. Bochner, S. (1962). A new approach to almost periodicity. *Proc. Nat. Acad. Sci. U.S.A.* **48**, 2039–2043.
23. Bogenschütz, T. (1992/93). Entropy, pressure, and a variational principle for random dynamical systems. *Random Comput. Dynam.* **1**(1), 99–116.
24. Bogenschütz, T., and Crauel, H. (1992). The Abramov–Rokhlin formula. *Ergodic Theory and Related Topics, III (Güstrow, 1990)*, Lecture Notes in Math., Vol. 1514, Springer, Berlin, pp. 32–35
25. Bohr, H. (1947). *Almost Periodic Functions*. Chelsea, New York. Translation of *Fastperiodische Funktionen*, Springer, Berlin, 1932.
26. Bohr, H. (1952). *Collected Mathematical Works*, Følner, E., and Jessen, B. (eds), Kobenhavn, Dansk Matematisk Forening.
27. Bowen, R. (1971). Entropy for group endomorphisms and homogeneous spaces. *Trans. Amer. Math Soc.* **153**, 401–414.
28. Bronstein, I. U. (1979). *Extensions of Minimal Transformation Groups*. Martinus Nijhoff Publishers, The Hague. (Translated from the Russian).
29. Brown, J. R. (1976). *Ergodic Theory and Topological Dynamics*, Academic Press, New York.
30. Caraballo, T., and Langa, J. A. (2002). On the theory of random attractors and some open problems. *Stochastic Partial Differential Equations and Applications (Trento, 2002)*, Lecture Notes in Pure and Appl. Math., Vol. 227, Dekker, New York, pp. 89–104.
31. Caraballo, T., and Langa, J. A. (2003). On the upper semicontinuity of cocycle attractors for non-autonomous and random dynamical systems. *Dynamics of Continuous, Discrete Math. Anal.* **10**(4), 491–513.
32. Caraballo, T., Langa, J. A., Melnik, V. S., and Valero, J. (2003). Pullback attractors of nonautonomous and stochastic multivalued dynamical systems. *Set Valued Analysis* **11**(2), 153–201.
33. Caraballo, T., Langa, J. A., and Robinson, J. C., (1998). Upper semicontinuity of attractors for small random perturbations of dynamical systems. *Comm. Partial Differential Equations* **23**(9–10), 1557–1581.
34. Caraballo, T., Langa, J. A., and Robinson, J. C. (2000). Stability and random attractors for a reaction-diffusion equation with multiplicative noise. *Discrete Contin. Dynam. Systems* **6**(4), 875–892.
35. Caraballo, T., Langa, J. A., and Robinson, J. C. (2001). A stochastic pitchfork bifurcation in a reaction-diffusion equation. *R. Soc. Lond. Proc. Ser. A Math. Phys. Eng. Sci.* **457**(2013), 2041–2061.

36. Caraballo, T., Langa, J. A., and Valero, J. (2001). Global attractors for multivalued random dynamical systems generated by random differential inclusions with multiplicative noise. *J. Math. Anal. Appl.* **260**(2), 602–622.
37. Caraballo, T., Langa, J. A., and Valero, J. (2002). Global attractors for multivalued random dynamical systems. *Nonlinear Anal. Ser. A: Theory Methods* **48**(6), 805–829.
38. Caraballo, T., Langa, J. A., and Valero, J. Dimension of attractors of nonautonomous reaction-diffusion equations, *ANZIAM*. To appear.
39. Cheban, D. N., Kloeden, P., and Schmalfuss, B. (2001). Pullback attractors in dissipative nonautonomous differential equations under discretization. *J. Dynam. Differential Equations* **13**, 185–213.
40. Cheban, D., Kloeden, P. E., and Schmalfuss, B. (2002). The relationship between pullback, forwards and global attractors of nonautonomous dynamical systems. *Nonlinear Dynam. & Systems Theory* **2**, 9–28.
41. Chepyzhov, V. V., and Vishik, M. I. (2002). Attractors for equations of mathematical physics. *American Mathematical Society Colloquium Publications*, Vol. 49, AMS, Providence, RI.
42. Chow, S.-N., and Yi, Y. (1994). Center manifold and stability for skew-product flows. *J. Dynam. Differential Equations* **6**, 543–582.
43. Chueshov, I. (2001). Order-preserving skew-product flows and nonautonomous parabolic systems. *Acta Appl. Math.* **65**, 185–205.
44. Chueshov, I. (2002). *Monotone Random Systems-Theory and Applications*. Lecture Notes in Mathematics, Springer, Berlin, p. 1779.
45. Colonius, F., and Kliemann, W. (2000). *The Dynamics of Control*, Birkhäuser.
46. Colonius, F., and Siegmund, S. (2003). Normal forms of control systems at singular points. *J. Dynam. Differential Equations*. To appear.
47. Crauel, H. (2002). *Random Probability Measures on Polish Spaces*, Stochastics Monographs, Vol. 11, Taylor & Francis, London.
48. Crauel, H., Debussche, A., and Flandoli, F. (1997). Random attractors. *J. Dynam. Differential Equations* **9**(2), 307–341.
49. Crauel, H., and Flandoli, F. (1994). Attractors for random dynamical systems. *Probab. Theory Related Fields* **100**(3), 365–393.
50. Crauel, H., and Gundlach, M. (eds.) (1999). *Stochastic Dynamics*, Springer, Berlin, Heidelberg, New York.
51. Crauel, H., Imkeller, P., and Steinkamp, M. (1999). Bifurcations of one-dimensional stochastic differential equations. In Crauel, H., and Gundlach, M. (eds.), *Stochastic Dynamics*, Springer-Verlag, New York.
52. Debussche, A. (1997). On the finite dimensionality of random attractors. *Stochastic Anal. Appl.* **15**(4), 473–491.
53. Debussche, A. (1998). Hausdorff dimension of a random invariant set. *J. Math. Pares Appl. (9)* **77**(10), 967–988.
54. Dellnitz, M., and Junge, O. (2001). *The Algorithms Behind GAIO-Set Oriented Numerical Methods for Dynamical Systems*, The DANSE-Book.
55. Denker, M., Grillenberger, C., and Sigmund, K. (1976). *Eryodic Theory on Compact Spaces*, Springer, Berlin, Heidelberg, New York.
56. Eckmann, J.-P., and Hairer, M. (2001). Invariant measures for stochastic partial differential equations in unbounded domains. *Nonlinearity* **14**(1), 133–151.
57. Eckmann, J.-P., and Hairer, M. (2001). Uniqueness of the invariant measure for a stochastic PDE driven by degenerate noise. *Comm. Math. Phys.* **219**(3), 523–565.
58. Ellis, R., and Gottschalk, W. H. (1960). Homomorphisms of transformation groups. *Trans. Amer. Math Soc.* **94**, 258–271.

59. Fink, A. M. (1974). *Almost Periodic Differential Equations*, Springer, Berlin, Heidelberg, New York.
60. Flandoli, F., and Schmalfuss, B. (1999). Weak solutions and attractors for three-dimensional Navier-Stokes equations with nonregular force. *J. Dynam. Differential Equations* **11**, 355–398.
61. Furstenberg, H. (1963). The structure of distal flows. *Amer. J. Math.* **85**, 477–515.
62. Furstenberg, H. (1967). Disjointness in ergodic theory, minimal sets, and a problem in Diophantine approximation. *Math. Systems Theory* **1**, 1–50.
63. Johnson, R. A. (1981). A linear, almost periodic equation with an almost automorphic solution. *Proc. Amer. Math. Soc.* **82**, 206–208.
64. Johnson, R. A. (1989). Hopf bifurcation from nonperiodic solutions of differential equations. I. Linear theory. *J. Dynam. Differential Equations* **1**, 179–198.
65. Johnson, R. A., Palmer, K. J., and Sell, G. R. (1987). Ergodic properties of linear dynamical systems. *SIAM J. Math. Anal.* **18**, 1–33.
66. Johnson, R. A., and Yi, Y. (1994). Hopf bifurcation from nonperiodic solutions of differential equations. II. *J. Differential Equations* **107**, 310–340.
67. Katok, A., and Hasselblatt, B. (1995). *Introduction to the Modern. Theory of Dynamical Systems*, Cambridge University Press.
68. Keller, H., and Ochs, G. (1999). Numerical approximation of random attractors. In Crauel, H., and Gundlach, M. (eds.), *Stochastic Dynamics*, Springer.
69. Kloeden, P. (2000). A Lyapunov function for pullback attractors of nonautonomous differential equations. *Electron. J. Differ. Equ. Conf.* **5**, 91–102.
70. Kloeden, P., Keller, H., and Schmalfuss, B. (1999). Towards a theory of random numerical dynamics. In Crauel, H., and Gundlach, M. (eds.), *Stochastic Dynamics*, Springer.
71. Kloeden, P., and Schmalfuss, B. (1996). Lyapunov functions and attractors under variable time-step discretization. *Discrete Contin. Dynam. Systems* **2**, 163–172.
72. Klünger, M. (1999). Ergodic decomposition of invariant measures, Report Nr. 444, Institut für Dynamische Systeme, Universität Bremen.
73. Koksch, N., and Siegmund, S. (2002). Pullback attracting inertial manifolds for nonautonomous dynamical systems. *J. Dynam. Differential Equations*. **14**(4), 889–941.
74. Langa, J. A., Robinson, J. C., and Suarez, A. (2002). Stability, instability, and bifurcation phenomena in non-autonomous differential equations. *Nonlinearity* **15**(3), 887–903.
75. Langa, J. A., Robinson, J. C., and Suarez, A. (2003). Forwards and pullback behaviour of a non-autonomous Lotka–Volterra system. *Nonlinearity* **16**(4), 1277–1293.
76. Langa, J. A., Robinson, J. C., and Suarez, A. (2003). Pullback permanence in a non-autonomous competitive Lotka–Volterra model. *J. Differential Equations*. **190**(1), 214–238.
77. Marsden, J. E., and McCracken, M. (1976). *The Hopf bifurcation and its applications*, Applied Mathematical Sciences, Vol. 19, Springer-Verlag, New York.
78. Markley, N. G., and Paul, M. E. (1979). Almost automorphic symbolic minimal sets without unique ergodicity. *Israel J. Math.* **34**, 259–272.
79. Mielke, A. (1986). A reduction principle for nonautonomous systems in infinite-dimensional spaces. *J. Differential Equations* **65**, 68–88.
80. Miller, R. K. (1965). Almost periodic differential equations as dynamical systems with applications to the existence of almost periodic solutions. *J. Differential Equations* **1**, 337–345.
81. Miller, R. K., and Sell, G. R. (1970). *Volterra Integral Equations and Topological Dynamics*, Memoirs of the AMS, Vol. 102, Providence.
82. Millionščikov, V. (1968). Proof of the existence of irregular systems of linear differential equations with almost periodic coefficients. *Differencialnye Urav.* **4**, 391–396 (in Russian).

83. Millionščikov, V. (1968). Metric theory of linear systems of differential equations. *Mat. Sb. (N.S.)* **77**(119), 163–173 (in Russian).
84. Van Minh, N. (1993). A reduction principle for topological classification of nonautonomous differential equations. *Proc. Roy. Soc. Edinburgh Sect. A* **123**, 621–632.
85. Oxtoby, J. C. (1980). *Measure and Category*, Springer Graduate Texts in Mathematics 2, Springer, Berlin, Heidelberg, New York.
86. Petersen, K. (1983). *Ergodic Theory*, Cambridge University Press, Cambridge.
87. Pliss, V. A. (1964). A reduction principle in the theory of stability of motion. *Izv. Akad. Nauk SSSR Ser. Mat.* **28**, 1297–1324.
88. Sacker, R. J., and Sell, G. R. (1977). Lifting properties in skew-product flows with applications to differential equations. *Mem. Amer. Math. Soc.* **11**(190), iv–67.
89. Sacker, R. J., and Sell, G. R. (1978). A spectral theory for linear differential systems. *J. Differential Equations* **27**, 320–358.
90. Schenk-Hoppé, K. R. (2001). Random dynamcal systems in economics. *Stoch. Dyn.* **1**, 63–83.
91. Scheutzow, M. (2002). Comparison of various concepts of a random attractor: A case study. *Arch. Math. (Basel)* **78**(3), 233–240.
92. Schmalfuss, B. (1992). Backward cocycles and attractors of stochastic differential equations. In Reitmann, V., Riedrich, T., and Koksch, N. (eds.) *International Seminar on Applied Mathematics-Nonlinear Dynamics: Attractor Approximation and Global Behaviour*, Technische Universität, Dresden, pp. 185–192.
93. Sell, G. R. (1967). Nonautonomous differential equations and topological dynamics. I. The basic theory. *Trans. Amer. Math. Soc.* **127**, 241–262.
94. Sell, G. R. (1967). Nonautonomous differential equations and topological dynamics. II. Limiting equations. *Trans. Amer. Math. Soc.* **127**, 263–283.
95. Sell, G. R. (1971). *Topological Dynamics and Ordinary Differential Equations.* Van Nostrand Reinhold Mathematical Studies, London.
96. Sell, G. R. (1978). The structure of a flow in the vicinity of an almost periodic motion. *J. Differential Equations* **27**, 359–393.
97. Sell, G. R., Shen, W., and Yi, Y. (1995). Topological dynamics and differential equations. *Topological Dynamics and Applications*, Minneapolis, MN, pp. 279–297.
98. Sell, G. R., and You, Y. (2002). *Dynamics of Evolutionary Equations*, Springer, New York.
99. Shen, W., and Yi, Y. (1998). *Almost Automorphic and Almost Periodic Dynamics in Skew-Product Semiflows*, Memoirs of the AXIS 136, No. 647.
100. Siegmund, S. (2002). Dichotomy spectrum for nonautonomous differential Equations. *J. Dynamics and Differential Equations* **14**, 243–258.
101. Siegmund, S. (2002). Normal forms for nonautonomous differential equations. *J. Differential Equations* **178**, 541–573.
102. Siegmund, S. (2002). Normal forms for nonautonomous difference equations. *Advances in Difference Equations IV*, special issue of: Computers and Mathematics with Applications, Vol. 45, pp. 1059–1073.
103. Siegmund, S. (2003). Computation of nonautonomous invariant manifolds. Aulbach, B., Elaydi, S., and Ladas, G. (eds), *Proc. of ICDEA2001,* Taylor & Francis, London.
104. Smith, H. L. (1995). *Monotone Dynamical Systems. An Introduction to the Theory of Competitive and Cooperative Systems*, Amer. Math. Soc., Providence, Rhode Island.
105. Stroock, D. W. (1993). *Probability Theory*, Cambridge University Press, Cambridge.
106. Veech, W. A. (1965). Almost automorphic functions on groups. *Amer. J. Math.* **87**, 719–751.
107. Veech, W. A. (1977). Topological dynamics. *Bull. Amer. Math. Soc.* **83**, 775–830.

108. Walters, P. (1982). *An Introduction to Ergodic Theory*, Springer Graduate Texts in Mathematics 79, Berlin, Heidelberg, New York.
109. Wanner, T. (1991). *Invariante Faserbündel and topologische Äquivalenz bei dynamischen Prozessen.* Diplom-Arbeit, Universitat Augsburg.
110. Wanner, T. (1995). Linearization of random dynamical systems. In Jones, C., Kirchgraber, U., and Walther, H. O. (eds.), *Dynamics Reported,* Springer-Verlag, Berlin, Heidelberg, New York.
111. Yi, Y. (1993). A generalized integral manifold theorem. *J. Differential Equations* **102**, 153–187.
112. Yi, Y. (1993). Stability of integral manifold and orbital attraction of quasi-periodic motion. *J. Differential Equations* **103**, 278–322.
113. Zhikov, V. V., and Levitan, B. M. (1977). Favard theory. *Uspehi Mat. Nauk* **32**, 123–171 (in Russian).