Jahresber Dtsch Math-Ver (2015) 117:171
DOI 10.1365/s13291-015-0118-x

PREFACE

Preface Issue 3-2015

Hans-Christoph Grunau[1]

Autonomous dynamical systems are systems of differential equations where the right hand side does not explicitly depend on the independent variable, usually interpreted as time t. In order to find out the state $u(t)$ of the system at time t with initial condition u_0 at time t_0, one may start with the same initial condition at time 0 and calculate the state $u(t-t_0)$ at time $(t-t_0)$. To investigate and to understand the dynamical behaviour of systems of this type, there exist well established methods and successful concepts, e.g. invariant sets, limit sets and attractors. In the case of *non*autonomous systems all of these methods, concepts and related issues become more complicated and less research has been carried out. Here, in order to calculate the state $u(t)$ at the present time t one needs indeed to know the starting time t_0, and not only the elapsed time interval $(t-t_0)$. Hans Crauel and Peter E. Kloeden explain in their survey article "Nonautonomous and random attractors" how the notion and theory of invariant sets, limit sets and attractors has to be extended to nonautonomous dynamical systems. The authors explain at the same time how they have developed a closely related theory of *random* dynamical systems, which are intrinsically nonautonomous.

Moritz Kassmann gives a brief account of recent developments in the area of Dirichlet forms and includes some remarks on the book "Semi-Dirichlet forms and Markov processes" by Yoichi Oshima. Heiko von der Mosel gives an informative summary of Xavier Tolsa's recent book on "Analytic capacity, the Cauchy transform, and non-homogeneous Calderón-Zygmund theory". The reviewer has a quite positive opinion of this book. Jean-Claude Yakoubsohn emphasises the dominant role of "Condition" in numerical analysis and reviews the book of the same title written by Peter Bürgisser and Felipe Cucker. Their subtitle is "The geometry of numerical algorithms."

✉ H.-Ch. Grunau
hans-christoph.grunau@ovgu.de

[1] Institut für Analysis und Numerik, Fakultät für Mathematik, Otto-von-Guericke-Universität, Postfach 4120, 39016 Magdeburg, Germany

Jahresber Dtsch Math-Ver (2015) 117:173–206
DOI 10.1365/s13291-015-0115-0

SURVEY ARTICLE

Nonautonomous and Random Attractors

Hans Crauel[1] · **Peter E. Kloeden**[1,2]

Published online: 9 June 2015

Abstract The theories of nonautonomous and random dynamical systems have undergone extensive, often parallel, developments in the past two decades. In particular, new concepts of nonautonomous and random attractors have been introduced. These consist of families of sets that are mapped onto each other as time evolves and have two forms: a forward attractor based on information about the system in the future and a pullback attractor that uses information about the past of the system. Both reduce to the usual attractor consisting of a single set in the autonomous case.

Keywords Skew product flow · 2-parameter semi-group · Pullback attractor · Forward attractor · Random dynamical system · Weak attractor · Mean-square random dynamical system · Mean-field stochastic differential equations

Mathematics Subject Classification (2010) 34D45 · 35B41 · 37-02 · 37B55 · 37C70 · 37H99 · 37L30 · 37L55 · 60H10 · 60H15

1 Introduction

Autonomous dynamical systems are now a very well established area of mathematics. Although nonautonomous systems have been investigated in an ad hoc way for many years, a mathematical theory of nonautonomous dynamical systems has only been developed systematically in recent decades. A closely related theory of random

✉ H. Crauel
crauel@math.uni-frankfurt.de

P.E. Kloeden
kloeden@math.uni-frankfurt.de

1 Institut für Mathematik, Goethe-Universität, Frankfurt am Main, Germany

2 School of Mathematics and Statistics, Huazhong University of Science & Technology, Wuhan, China

dynamical systems, which are intrinsically nonautonomous, arose during the same period with developments in both the deterministic and random cases benefiting interactively from each other.

A characteristic feature of autonomous dynamical systems is their dependence of the elapsed time $t - t_0$ only and not separately on the current time t and initial time t_0, which means that limiting objects exist all the time and not just in the distant future. In contrast, nonautonomous systems depend explicitly on both t and t_0. Hence they form 2-parameter semi-groups, sometimes called processes, rather than by one-parameter semi-groups as in the autonomous case. There is also a more abstract formulation of nonautonomous dynamical systems as skew product flows with a cocycle state space mapping being driven by an autonomous dynamical system modelling the nonautonomicity. For example, this driving system could be the shift operator on the hull of an almost periodic function. The base space on which the driving system acts is often compact, which has many technical advantages and was described by George Sell as "compactifying time". Random dynamical systems are a form of skew product flows, but with the driving system representing the noise acting on the sample space of a probability space and forming a measure theoretical rather than topological group of transformations.

A new feature of nonautonomous and random dynamical systems is that invariant sets are in fact families of sets that are invariant in the sense that they mapped onto each other as time evolves. A single set as an invariant set as in the autonomous case would be too restrictive since it would exclude simple but important behaviour. For example, in the process formulation, a periodic curve can be represented as a family of singleton sets, each consisting of a point on the curve.

Another important new feature of nonautonomous and random dynamical systems is that limiting temporal behaviour must now be characterised in two ways. For an autonomous dynamical system, the elapsed time $t - t_0 \to \infty$ if either $t \to \infty$ with t_0 fixed or if $t_0 \to -\infty$ with t fixed. In the nonautonomous case the limits obtained can be different, if they exist. The former, called forward convergence, involves information about the future of the system, whereas the latter, called pullback convergence, uses information from the past. Note that pullback convergence does not involve the system's running backwards in time, rather it runs forwards from an ever earlier starting time. Two types of nonautonomous attractors arise from these convergences, a forward and a pullback attractor. These consist of families of nonempty compact subsets that are invariant in the above generalised sense and attract other sets (or even families of sets) in the corresponding convergence.

We shall review these developments in this paper after first recalling some basic results about autonomous dynamical systems. Readers are referred to [6] and the monographs [9, 11, 46] for more information. Many of the results presented here were developed by members and visitors, including the authors, of the random dynamical systems group led by Ludwig Arnold at the University of Bremen.

Random attractors were introduced by Crauel, Debussche and Flandoli [13, 16, 19] and also by Schmalfuß [58], who introduced the notion of "backward cocycles" in order to obtain a criterion for the existence of a sort of forward attractors. Based on the ideas in [19], deterministic nonautonomous attractors were proposed in [16] and by Kloeden and Schmalfuß in [49], although it turned out that a related concept had been used earlier by Chepyzhov and Vishik [11] in a more restrictive context. The name "cocycle attractor" used in [49] was later changed to pullback attractor in [33] to distinguish it from a forward attractor.

It is interesting to note that physicists have a heuristic version of a random attractor, introduced in 1990, which they call a *snapshot attractor*, see Romeiras et al. [56].

The concept of "pullback" convergence mentioned above has a long history in one form or another. It has been used in probability theory to construct invariant measures, where it is a form of martingale convergence, see e.g. Ledrappier [52, Lemma 1, pages 64–65]. Khasminskii also used a form of pathwise pullback convergence to construct stationary solutions of stochastic differential equations in his monograph [32] that first appeared in Russian in 1969. A similar idea was also used in the 1960s by Krasnosel'skii [51] to establish the existence of bounded bi-infinite sequences in deterministic systems.

2 Deterministic Dynamical Systems

Deterministic dynamical systems are formulated abstractly in terms of a family of mappings of the state space X into itself, which is parameterised by a time set. The state space is typically assumed to be a complete metric space (X, d_X) and the time set is denoted by $\mathbb{T}$ with $\mathbb{T} = \mathbb{Z}$ for *discrete time* dynamical systems and $\mathbb{T} = \mathbb{R}$ for *continuous time* dynamical systems. For autonomous dynamical systems, the family of mappings form a group under composition, while for autonomous semi-dynamical systems, which are defined only forwards in time with the time set $\mathbb{T}^+ = \mathbb{Z}^+$ or $\mathbb{R}^+$, it forms a semi-group. Nonautonomous and random dynamical systems are somewhat more complicated.

The proximity of sets and corresponding convergence concepts play an important role in the theory of dynamical systems. The *Hausdorff separation* or semi-distance $\mathrm{dist}_X(A, B)$ of nonempty compact subsets A, B of X is defined as

$$\mathrm{dist}_X(A, B) = \max_{a\in A} \mathrm{dist}(a, B) = \max_{a\in A} \min_{b\in B} d(a, b), \tag{1}$$

while

$$H_X(A, B) = \max\big\{\mathrm{dist}_X(A, B), \mathrm{dist}_X(B, A)\big\}$$

defines a metric called the *Hausdorff metric* on the space $\mathscr{K}(X)$ of nonempty compact subsets of X. The subscript X above will often be omitted when there is no confusion about the space X.

These definitions can be extended to the space $\mathscr{C}(X)$ of nonempty closed subsets of X, provided the max and min in (1) are replaced by sup and inf, respectively.

2.1 Autonomous Dynamical Systems

An autonomous difference equation on the state space $\mathbb{R}^d$ has the form

$$x_{n+1} = f(x_n), \tag{2}$$

where $f : \mathbb{R}^d \to \mathbb{R}^d$. Successive iteration of (2) generates the solution mapping $\pi : \mathbb{Z}^+ \times \mathbb{R}^d \to \mathbb{R}^d$ defined by

$$x_n = \pi(n, x_0) = f^n(x_0) := \underbrace{f \circ f \circ \cdots \circ f}_{n \text{ times}}(x_0),$$

which satisfies the *initial condition* $\pi(0, x_0) = f^0(x_0) = x_0$, where f^0 denotes the identity mapping, and the *semi-group property*

$$\pi\big(n, \pi(m, x_0)\big) = f^n\big(\pi(m, x_0)\big) = f^n \circ f^m(x_0) = f^{n+m}(x_0) \ = \pi(n + m, x_0) \quad (3)$$

for n, $m \in \mathbb{Z}^+$ and $x_0 \in \mathbb{R}^d$. Property (3) says that the solution mapping π forms a semi-group under composition; it is typically only a semi-group rather than a group since the mapping f need not be invertible. If the mapping f in the difference equation (2) is at least continuous, then the mappings $\pi(n, \cdot)$ are continuous for every $n \in \mathbb{Z}^+$ and the solution mapping π generates a discrete-time *semi-dynamical system* on $\mathbb{R}^d$.

If the mapping $f : \mathbb{R}^d \to \mathbb{R}^d$ in the autonomous difference equation (2) is *invertible* with inverse f^{-1}, then the solution mapping π can be extended to the time set $\mathbb{Z}^-$ by

$$\pi(n, x_0) = \big(f^{-1}\big)^{|n|}(x_0) := \underbrace{f^{-1} \circ f^{-1} \circ \cdots \circ f^{-1}}_{|n| \text{ times}}(x_0), \quad n \in \mathbb{Z}^-,$$

and hence to the entire time set $\mathbb{Z}$. The semi-group property (3) becomes the *group property*

$$\pi(m + n, x_0) = \pi\big(m, \pi(n, x_0)\big) \quad \text{for all } m, n \in \mathbb{Z}, x_0 \in X.$$

If the inverse mapping $f^{-1} : \mathbb{R}^d \to \mathbb{R}^d$ is also continuous, then $\pi : \mathbb{Z} \times \mathbb{R}^d \to \mathbb{R}^d$ is also continuous in its second variable.

Similarly, under assumptions that ensure the existence and uniqueness of solutions forwards in time, the solution mapping π of an autonomous ordinary differential equation (ODE)

$$\dot{x} = \frac{dx}{dt} = f(x), \quad x(0) = x_0, \tag{4}$$

where $f : \mathbb{R}^d \to \mathbb{R}^d$, defines a continuous-time semi-dynamical system on the state space $\mathbb{R}^d$. In particular, $\pi : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}^d$ satisfies continuous-time analogues of the properties to those for the difference equation (2) with the semi-group property

$$\pi(s + t, x_0) = \pi\big(s, \pi(t, x_0)\big) \quad \text{for all } s, t \in \mathbb{R}^+, x_0 \in \mathbb{R}^d,$$

which is a direct consequence of the uniqueness of solutions of an initial value problem, while the continuity of the mapping $(t, x_0) \mapsto \pi(t, x_0)$ follows from continuity in initial conditions and differentiability, hence continuity, in time of the solutions of the ODE (4).

These two examples motivate the following abstract definition of an autonomous (semi-) dynamical system on a state space X with time set $\mathbb{T}$.

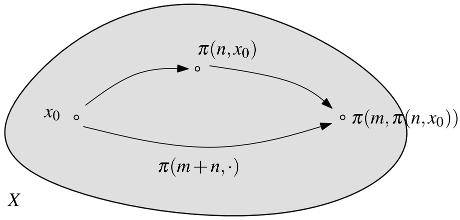


Fig. 1 Semigroup property (ii) Definition 1 of a discrete-time semidynamical system $\pi : \mathbb{Z}^+ \times X \to X$

Definition 1 A mapping $\pi : \mathbb{T}^+ \times X \to X$ satisfying

(i) $\pi(0, x_0) = x_0$ for all $x_0 \in X$,
(ii) $\pi(s+t, x_0) = \pi(s, \pi(t, x_0))$ for all $s, t \in \mathbb{T}^+$ and $x_0 \in X$,
(iii) the mapping $(t, x_0) \mapsto \pi(t, x_0)$ is continuous,

is called an *autonomous semi-dynamical system* on the state space X with the time set $\mathbb{T}$. If these properties hold for the time set $\mathbb{T}$ instead of just $\mathbb{T}^+$, then π is called an *autonomous dynamical system* on the state space X with the time set $\mathbb{T}$.

The continuity in (t, x_0) in property (iii) for $\mathbb{T} = \mathbb{Z}$ is interpreted in terms of the discrete topology on $\mathbb{T}$, in which case (iii) reduces to continuity just in x_0. Actually, the continuity in t is not needed in many proofs and is often omitted from the definition.

By property (ii), the family of mappings $\pi_t(\cdot) := \pi(t, \cdot) : X \to X$, $t \in \mathbb{T}$ (resp. $t \in \mathbb{T}^+$) is a group (resp., semi-group) under composition.

Example 1 The solution mapping $\pi(t, x_0) = x_0 e^{-t}$ of the scalar ODE $\dot{x} = -x$ is defined for all $t \in \mathbb{R}$ and thus forms an autonomous dynamical system on the state space $X = \mathbb{R}$. In contrast, except for those starting at $x_0 \in [-1, 1]$, the solutions

$$\pi(t, x_0) = \frac{x_0 e^t}{\sqrt{1 - x_0^2 + x_0^2 e^{2t}}}$$

of the ODE $\dot{x} = x(1 - x^2)$ do not exist for all negative times. Thus the solution mapping π is defined for all $t \in \mathbb{R}^+$, but only for negative times in small intervals that depend on the initial value x_0 when $x_0^2 > 1$. Here π generates an autonomous semi-dynamical system on the state space $X = \mathbb{R}$ that is not an autonomous dynamical system.

An autonomous (semi-) dynamical system need not be generated explicitly by an autonomous difference or differential equation.

Example 2 Consider the space $\mathbf{X} = \{1, \ldots, r\}^{\mathbb{Z}}$ of bi-infinite sequences $\mathbf{x} = \{k_n\}_{n \in \mathbb{Z}}$ with $k_n \in \{1, \ldots, r\}$ with respect to the group of shift operators $\theta_n := \theta^n$ for $n \in \mathbb{Z}$, where the mapping $\theta : X \to X$ is the left shift operator that is defined by $\theta(\{k_n\}_{n \in \mathbb{Z}}) = \{k_{n+1}\}_{n \in \mathbb{Z}}$. This forms a discrete time autonomous dynamical system on X, which is a compact metric space with the metric

$$d(\mathbf{x}, \mathbf{x}') = \sum_{n \in \mathbb{Z}} 2^{-|n|} \left|k_n - k_n'\right|.$$

2.2 Invariant Sets and Attractors

The dynamical behaviour of an autonomous semi-dynamical system π on a state space X with time set $\mathbb{T}$ is characterised by its invariant sets and what happens in neighbourhoods of such sets.

For a nonempty subset B of X define $\pi(t, B) = \bigcup_{x_0 \in B} \{\pi(t, x_0)\}$ for $t \in \mathbb{T}^+$.

Definition 2 A nonempty subset A of X is called (strictly) *invariant* under π, or π-invariant, if

$$\pi(t, A) = A \quad \text{for all } t \in \mathbb{T}^+.$$

The simplest example of an invariant set A is a steady state solution, in which case A consists of a single point. Note that the union of invariant sets is also an invariant set. See Fig. 1.

Many invariant sets are the ω-limit set of a point $x_0 \in X$. It is defined as

$$\omega^+(x_0) = \left\{y \in X \ : \ \exists t_j \to \infty,\ \pi(t_j, x_0) \to y\right\}.$$

It is nonempty, compact and π-invariant when the forwards trajectory $Tr^+[x_0] := \{\pi(t, x_0); t \in \mathbb{T}^+\}$ is a precompact subset of X and the metric space (X, d) is complete. In addition, the ω-limit sets $\omega^+(x_0)$ of single points are connected for continuous-time systems which are, in addition, continuous in time, but they need not be connected for a discrete-time system. Similarly, the ω-limit set of a subset B of X is defined by

$$\omega^+(B) = \left\{y \in X \ : \ \exists t_j \to \infty, b_j \in B,\ \pi(t_j, b_j) \to y\right\}.$$

Note that, in general, $\bigcup_{x_0 \in B} \omega^+(x_0) \neq \omega^+(B)$.

The asymptotic behaviour of a semi-dynamical system is characterised by its ω-limit sets in general and by its attractors in particular. In defining an attractor it is necessary to use the limit sets of whole sets of initial points rather than individual points in order to capture heteroclinic trajectories inside the attractor.

Definition 3 Let π be an autonomous semi-dynamical system on the state space X with time set $\mathbb{T}$. An *attractor* of π is a nonempty compact π-invariant set A which attracts bounded subsets B of X in the sense that

$$\lim_{t \to \infty} \mathrm{dist}_X\big(\pi(t, B), A\big) = 0. \tag{5}$$

It is called a *local* attractor if the bounded sets are restricted to some bounded neighbourhood of A or otherwise the maximal or *global* attractor when they are unrestricted. Note that a global attractor, if it exists, must be unique.

For later comparison note that, in view of the invariance of A, the attraction (5) can be written equivalently as the *forward convergence*

$$\mathrm{dist}_X\big(\pi(t,B),\pi(t,A)\big)\to 0 \quad \text{as } t\to\infty.$$

Invariant sets can also have a much more complicated geometry; for example, they could be fractal sets. The existence and approximate location of a global attractor follows from that of more easily found absorbing sets, which typically have a convenient simpler shape such as a ball or ellipsoid.

Definition 4 A nonempty subset B of X is called an *absorbing set* of an autonomous semi-dynamical system π on X if for every bounded subset D of X there exists a $T_D\in\mathbb{T}^+$ such that $\pi(t,D)\subset B$ for all $t\geq T_D$ in $\mathbb{T}^+$.

Absorbing sets are often called attracting sets when they are also *forward* or *positively invariant* in the sense that $\pi(t,B)\subseteq B$ holds for all $t\in\mathbb{T}^+$. Attractors differ from attracting sets in that they consist entirely of ω-limit sets of the system and are thus strictly invariant in the sense of Definition 2.

Theorem 1 *Suppose that an autonomous semi-dynamical system π on X has a compact absorbing set B. Then π has a unique global* (*set*) *attractor $A\subset B$ given by*

$$A=\bigcap_{s\geq 0}\overline{\bigcup_{t\geq s}\pi(t,B)}, \quad s,t\in\mathbb{T}^+,$$

or more simply by

$$A=\bigcap_{t\geq 0}\pi(t,B), \quad t\in\mathbb{T}^+,$$

when B is positively invariant.

The nonemptiness of A follows from the fact that it is the intersection of nonempty nested compact sets. This is easily seen when B is positively invariant since $\pi(t,B)\subset B$, so $\pi(s+t,B)=\pi(s,\pi(t,B))\subset\pi(s,B)$ for all $s,t\in\mathbb{T}^+$.

Similar results hold in infinite dimensional state spaces if the absorbing set is only closed and bounded provided the mapping π is also assumed to be compact or asymptotically compact in some sense.

Global attractors are characterised by the bounded entire solutions of an autonomous semi-dynamical system. A mapping $\phi:\mathbb{T}\to X$ is called an *entire* or *complete solution* of an autonomous semi-dynamical system π if

$$\phi(t)=\pi\big(t-s,\phi(s)\big) \quad \text{for all } t\geq s,\ s,t\in\mathbb{T}.$$

Let A be the global attractor of an autonomous semi-dynamical system π: then $x_0\in A$ if and only if there exists a bounded entire solution ϕ of π with $\phi(t_0)=x_0$ for some $t_0\in\mathbb{T}$ and $\phi(t)\in A$ for all $t\in\mathbb{T}$.

A global attractor is, in fact, uniformly Lyapunov asymptotically stable. The asymptotic stability of attractors and that of attracting sets, in general, can be characterised by Lyapunov functions. Such Lyapunov functions can be used to establish the existence of an absorbing set and hence that of a nearby global attractor in a perturbed system, see e.g., [28, 29, 40].

3 Nonautonomous Dynamical Systems

A fundamental difference between nonautonomous and autonomous dynamical systems is that a nonautonomous system depends on both the actual time t and the starting time t_0, and not just on their difference $t - t_0$, the time elapsed since starting, as in an autonomous dynamical system. This has profound consequences.

There are two different abstract formulations of nonautonomous dynamical systems. The first is a more direct generalisation of the definition of an abstract autonomous semi-dynamical system. It involves a 2-parameter semi-group and was called a *process* by Dafermos [22]. Its properties are based directly on those of the solution mappings of nonautonomous differential and difference equations. The other formulation is somewhat more complicated and involves a built-in "driving system", which provides a much more specific description of the nonautonomity in the system, and is called a *skew product flow*. Later such a driving system will model the noise in random dynamical systems.

3.1 Processes

Define

$$\mathbb{T}^+_{\geq} = \left\{(t, t_0) \in \mathbb{T} \times \mathbb{T} : t \geq t_0\right\}.$$

The definition of a nonautonomous dynamical system on a state space X for a general time set $\mathbb{T}$ as a 2-parameter semi-group or process is given by:

Definition 5 A *process* is a mapping $\phi : \mathbb{T}^+_{\geq} \times X \to X$ with the following properties:

(i) *initial condition:* $\phi(t_0, t_0, x_0) = x_0$ for all $x_0 \in X$ and $t_0 \in \mathbb{T}$,
(ii) 2*-parameter semi-group property:* $\phi(t_2, t_0, x_0) = \phi(t_2, t_1, \phi(t_1, t_0, x_0))$ for all $(t_1, t_0), (t_2, t_1) \in \mathbb{T}^+_{\geq}$ and $x_0 \in X$,
(iii) *continuity:* the mapping $(t, t_0, x_0) \mapsto \phi(t, t_0, x_0)$ is continuous.

See Fig. 2. Assuming existence and uniqueness of the solutions for all nonnegative times, the solution $x(t) = \phi(t; t_0, x_0)$ of a nonautonomous ordinary differential equation (ODE)

$$\frac{dx}{dt} = f(t, x), \quad t \in \mathbb{R},\ x \in \mathbb{R}^d,$$

with initial value $x(t_0) = x_0$ is defined for all $x_0 \in \mathbb{R}^d$ and $t \geq t_0$ in $\mathbb{R}$. The 2-parameter semi-group property is an immediate consequence of the existence and

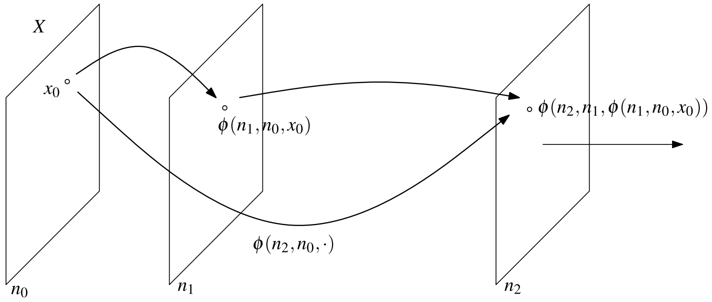


Fig. 2 Property (ii) Definition 5 of a discrete-time process $\phi : \mathbb{Z}^2_{\geq} \times X \to X$

uniqueness of solutions: the solution starting at (t_1, x_1), where $x_1 = \phi(t_1, t_0, x_0)$ is unique, so

$$\phi(t_2, t_0, x_0) = \phi(t_2, t_1, x_1) = \phi\big(t_2, t_1, \phi(t_1, t_0, x_0)\big).$$

An nonautonomous first-order difference equation on $\mathbb{R}^d$ has the form

$$x_{n+1} = f_n(x_n),$$

where the $f_n : \mathbb{R}^d \to \mathbb{R}^d$, $n \in \mathbb{Z}$, are continuous mappings, but, in general, they need not be invertible. Successive iteration from an initial value $x_{n_0} = x_0$ yields the solution

$$\phi(n, n_0, x_0) := f_{n-1} \circ \cdots \circ f_{n_0}(x_{n_0})$$

for all $n > n_0$ in $\mathbb{Z}$ and $x_0 \in \mathbb{R}^d$, with $\phi(n_0, n_0, x_0) := x_0$.

3.2 Skew Product Flows

A skew product flow consist of an *autonomous dynamical system* (full group) on a base space P, which is the source of the nonautonomity in a *cocycle mapping* acting on a state space X. The autonomous dynamical system is often called the *driving system*. Suppose that (P, d_P) and (X, d_X) are metric spaces and consider the time set $\mathbb{T}$. The skew product flow formalism is particularly advantageous when the base space P is compact.

Definition 6 A *skew product flow* (θ, φ) on $P \times X$ consists of an autonomous dynamical system $\theta = \{\theta_t\}_{t \in \mathbb{T}}$ acting on a metric space (P, d_P), which is called the *base space*, i.e.,

$$\text{(i)} \quad \theta_0(p) = p, \qquad \text{(ii)} \quad \theta_{s+t}(p) = \theta_s \circ \theta_t(p),$$
$$\text{(iii)} \quad (t, p) \mapsto \theta_t(p) \text{ continuous}$$

for all $p \in P$ and $s, t \in \mathbb{T}$, and a *cocycle mapping* $\varphi : \mathbb{T}^+ \times P \times X \to X$ acting on a metric space (X, d_X), which is called the *state space*, i.e.,

(1) *initial condition*: $\varphi(0, p, x) = x$ for all $p \in P$ and $x \in X$,
(2) *cocycle property*: $\varphi(s+t, p, x) = \varphi(s, \theta_t(p), \varphi(t, p, x))$ for all $s, t \in \mathbb{T}^+$, $p \in P$ and $x \in X$,
(3) *continuity:* $(t, p, x) \mapsto \varphi(t, p, x)$ is continuous.

The cocycle property is a generalisation of both the semi-group property and the 2-parameter semi-group property. It essentially keeps track of the current state of the driving system.

3.2.1 Examples of Skew Product Flows

A process $\phi : \mathbb{T}^+_{\geq} \times X \to X$ on a state space X with time set $\mathbb{T}$ can be formulated as a skew product flow. Define $\theta_t : \mathbb{T} \to \mathbb{T}$ for each $t \in \mathbb{T}$ by $\theta_t(t_0) = t + t_0$. The $\{\theta_t\}_{t\in\mathbb{T}}$ form a group under addition on $\mathbb{T}$ and an autonomous dynamical system on the base space $P = \mathbb{T}$. In addition, define

$$\varphi(t, t_0, x_0) = \phi(t + t_0, t_0, x_0),$$

where $t \in \mathbb{T}^+$ is the time that has elapsed since starting at $t_0 \in \mathbb{T}$ in φ, while $t + t_0$ and t_0 are absolute times in ϕ. (Note that ϕ denotes a process, whereas φ denotes the cocycle mapping of a skew product flow.)

An autonomous triangular system of scalar ODEs

$$\frac{dx}{dt} = -x + p, \qquad \frac{dp}{dt} = p(1-p),$$

with the initial values $x(0) = x_0$, $p(0) = p_0$, can be formulated as a skew product flow. Note that the p-ODE is decoupled from the x-ODE and that its solutions form an autonomous dynamical system on $P = [0, 1]$. In particular, the solution mapping defined by $\theta_t(p_0) = p(t, p_0)$ for $t \in \mathbb{R}$ forms a group under composition on $P = [0, 1]$. In addition, define $\varphi(t, p_0, x_0) = x(t, p_0, x_0)$. Then φ is a cocycle mapping on the state space $X = \mathbb{R}$. The base space $P = [0, 1]$ is compact here.

The driving system of a skew product flow need not be given in terms of an autonomous ODE. Consider the scalar ODE

$$\frac{dx}{dt} = -x + \cos t. \tag{6}$$

This is like the previous example, the main difference is that $p(t) = \cos t$ here is not given as the solution of an autonomous ODE. Nevertheless it generates a driving system as follows. Define $\theta_t(\cos(\cdot)) := \cos(t + \cdot)$ for $t \in \mathbb{R}$, i.e., θ_t maps the function $\cos(\cdot)$ to the function $\cos(t + \cdot)$ through a *phase shift*. Then define

$$P = \bigcup_{0\leq\tau\leq2\pi} \left\{\cos(\tau + \cdot)\right\} \tag{7}$$

as a subset of the function space $C_b(\mathbb{R}, \mathbb{R})$ of bounded continuous functions $f : \mathbb{R} \to \mathbb{R}$ with the uniform norm $\|f - g\|_\infty := \max_{t\in\mathbb{R}} |f(t) - g(t)|$, for $f, g \in C_b(\mathbb{R}, \mathbb{R})$.

The set P is not a linear subspace of $C_b(\mathbb{R},\mathbb{R})$, but still $d_P(f,g) := \|f-g\|_\infty$, $f, g \in P$, defines a metric on P. In particular, (P, d_P) is a compact metric space and $\theta_t : P \to P$ is continuous in (P, d_P) for each $t \in \mathbb{R}$. This forms the driving system $\{\theta_t\}_{t\in\mathbb{R}}$ on the base space P. Here $p_0(\cdot) = \cos(\tau_0 + \cdot)$ corresponds to the phase τ_0 of $\in [0, 2\pi]$ and

$$\big(\theta_t(p_0)\big)(\cdot) = \theta_t\big(\cos(\tau_0 + \cdot)\big) = \cos(\tau_0 + t + \cdot).$$

Then the solution mapping of the ODE

$$\frac{dx}{dt} = -x + \theta_t(p_0)$$

with the initial value $x(0) = x_0$ defines a cocycle mapping on the state space $X = \mathbb{R}$. The set P defined by (7) is called the *hull* of the function $\cos(\cdot)$. (Alternatively, one could use $\mathbb{S}^1$ as the base space P in this example.)

Finally, consider a nonautonomous difference equation on $X = \mathbb{R}$ given by

$$x_{n+1} = f_{j_n}(x_n), \quad n \in \mathbb{Z},$$

where the functions $f_1, \ldots, f_N$ and the j_n are the components of a bi-infinite sequence $\mathbf{s} = (\ldots, j_{-1}, j_0, j_1, j_2, \ldots)$. Let $P = \{1, \ldots, N\}^{\mathbb{Z}}$ be the totality of all such bi-infinite sequences. Then

$$d_P\big(\mathbf{s}, \mathbf{s}'\big) := \sum_{n=-\infty}^{\infty} 2^{-|n|}\big|j_n - j_n'\big|$$

defines a metric on P.

Define $\theta_n = \theta^n$, the n-fold composition of θ when $n > 0$ and of its inverse θ^{-1} when $n < 0$, where θ is the *left shift operator* on P, i.e., $(\theta\mathbf{s})_n = j_{n+1}$ for $n \in \mathbb{Z}$. Then $\{\theta_n\}_{n\in\mathbb{Z}}$ is a group under composition on P. As in Example 2, (P, d_P) is a compact metric space and the $\theta_n : P \to P$ are continuous in (P, d_P). The $\varphi : \mathbb{Z}^+ \times P \to P$ defined by

$$\varphi(n, \mathbf{s}, x_0) = f_{j_{n-1}} \circ \cdots \circ f_{j_0}(x_0)$$

for $n \geq 1$ is a discrete-time cocycle mapping. It is clear that the mapping $x_0 \mapsto \varphi(n, \mathbf{s}, x_0)$ is continuous due to the continuity of the composition of continuous functions. One can also show that the mappings $\mathbf{s} \mapsto \varphi(n, \mathbf{s}, x_0)$ are continuous on P, see [46].

3.2.2 Skew Product Flows as Semi-dynamical Autonomous Systems

A skew product flow (θ, φ) on $P \times X$ is an autonomous semi-dynamical system Π on the product state space $\mathfrak{X} = P \times X$, where $\Pi : \mathbb{T}^+ \times \mathfrak{X} \to \mathfrak{X}$ is defined by

$$\Pi\big(t, (p_0, x_0)\big) = \big(\theta_t(p_0), \varphi(t, p_0, x_0)\big).$$

The initial condition and continuity properties of Π are straightforward. The semigroup property follows from that of θ and the cocycle property of φ:

$$\begin{aligned}\Pi\big(s+t,(p_0,x_0)\big) &= \big(\theta_{s+t}(p_0),\varphi(s+t,p_0,x_0)\big)\\ &= \big(\theta_s\circ\theta_t(p_0),\varphi\big(s,\theta_t(p_0),\varphi(t,p_0,x_0)\big)\big)\\ &= \Pi\big(s,\big(\theta_t(p_0),\varphi(t,p_0,x_0)\big)\big) = \Pi\big(s,\Pi\big(t,(p_0,x_0)\big)\big).\end{aligned}$$

This representation as an autonomous semi-dynamical system is useful since it provides insights into how one could define invariant sets and attractors for nonautonomous systems.

4 Invariant Sets and Attractors of Nonautonomous Systems

Let ϕ be a process on a metric state space (X, d_X) with time set $\mathbb{T}$. The analogue of an invariant set for an autonomous semi-dynamical system is too restrictive for a process, i.e., a subset A of X such that

$$\phi(t,t_0,A) = A \quad \text{for all } (t,t_0)\in\mathbb{T}^2_{\geq}.$$

Such a subset contains steady state solutions if there are any, but excludes almost everything else such as the periodic solution

$$\bar{x}(t) = \frac{1}{2}\cos t + \frac{1}{2}\sin t \tag{8}$$

of the scalar ODE (6), which has no invariant sets in the above sense.

The next observation gives a hint how to define an invariant set for a nonautonomous process. Consider the autonomous semi-dynamical system Π on $\mathfrak{X} = \mathbb{T}\times X$, defined in terms of a process ϕ on X with time set $\mathbb{T}$ by

$$\Pi\big(\tau,(t_0,x_0)\big) = \big(\tau+t_0,\phi(\tau+t_0,t_0,x_0)\big).$$

Let $\mathfrak{A}\subset\mathfrak{X} = \mathbb{T}\times X$ be an invariant set for Π on $\mathfrak{X}$, i.e.,

$$\Pi(\tau,\mathfrak{A}) = \mathfrak{A}, \quad \text{for all } \tau\in\mathbb{T}^+.$$

Then $\mathfrak{A} = \bigcup_{t\in\mathbb{T}}\{t\}\times A_t$, where $A_t := \{x : (t;x)\in\mathfrak{A}\}$ is a nonempty subset of X for each $t\in\mathbb{T}$ and

$$\phi(t,t_0,A_{t_0}) = A_t \quad \text{for all } (t,t_0)\in\mathbb{T}^2_{\geq}.$$

This suggests the following definition of invariant set in the nonautonomous case.

Definition 7 A family $\mathscr{A} = \{A_t, t\in\mathbb{T}\}$ of nonempty subsets A_t of X is called *invariant* with respect to a process ϕ, or *ϕ-invariant* if

$$\phi(t,t_0,A_{t_0}) = A_t \quad \text{for all } (t,t_0)\in\mathbb{T}^2_{\geq}.$$

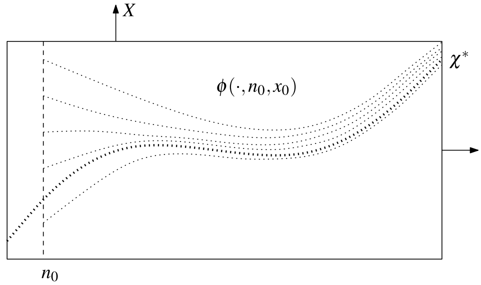


Fig. 3 Forward convergence $n \to \infty$ (discrete time case)

Example 3 The process generated by the ODE (6) has an invariant set $\mathscr{A} = \{A_t, t \in \mathbb{R}\}$ with component sets consisting of individual points on the periodic solution $\bar{x}(t)$, i.e., $A_t = \{\bar{x}(t)\}$ for $t \in \mathbb{R}$.

The periodic solution here is an *entire solution* of the process, i.e., a mapping $e : \mathbb{T} \to X$ such that $e(t) = \phi(t, t_0, e(t_0))$ for all $(t, t_0) \in \mathbb{T}^2_{\geq}$.

It is clear that an attractor of a process ϕ on a state space X with the time set $\mathbb{T}$ should be a ϕ-invariant family $\mathscr{A} = \{A_t, t \in \mathbb{T}\}$ of nonempty compact subsets A_t of X. There is a problem with convergence. This is apparent from the explicit solutions of the ODE (6)

$$x(t) = \left[x_0 - \frac{1}{2}(\cos t_0 + \sin t_0)\right] \mathrm{e}^{-(t-t_0)} + \frac{1}{2}(\cos t + \sin t).$$

In particular there is no limit as $t \to \infty$ (with t_0 fixed), because the non-exponential term is oscillating in t. On the other hand there is a limit $t_0 \to -\infty$ with t held fixed, namely

$$\bar{x}(t) = \frac{1}{2}(\cos t + \sin t),$$

which is called the *pullback limit*. In fact, $\bar{x}(t)$ is the periodic solution (8) of the ODE (6) and the other solutions also converge to it in the usual forwards sense

$$\left|x(t) - \bar{x}(t)\right| = \left|x_0 - \frac{1}{2}(\cos t_0 + \sin t_0)\right| \mathrm{e}^{-(t-t_0)} \to 0 \quad \text{as } t \to \infty \text{ with fixed } t_0.$$

These considerations have introduced two types of convergence:

(i) *Forward convergence* with the initial time t_0 held fixed (see Fig. 3):

$$\lim_{t \to \infty} \left|x(t, t_0, x_0) - \bar{x}(t)\right| = 0$$

(ii) *Pullback convergence* with the final time t held fixed (see Fig. 4):

$$\lim_{t_0 \to -\infty} \left|x(t, t_0, x_0) - \bar{x}(t)\right| = 0.$$

These two convergence concepts are independent.

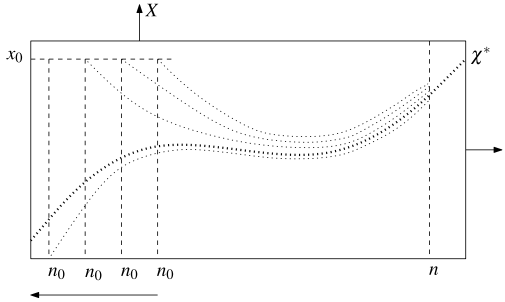


Fig. 4 Pullback convergence $n_0 \to -\infty$ (discrete time case)

Example 4 The solutions of the scalar nonautonomous ODE $\dot{x} = 2tx$ converge in the pullback sense, but not in the forward sense to the zero solution $x^*(t) = 0$, whereas for the scalar nonautonomous ODE $\dot{x} = -2tx$ they converge in the forward sense, but not in the pullback sense.

4.1 Nonautonomous Attractors: Processes

There are two different types of nonautonomous attractors for processes, one with pullback convergence and one with forward convergence. Let ϕ be a process on a metric state space (X, d_X) with time set $\mathbb{T}$.

Definition 8 A ϕ-invariant family $\mathscr{A} = \{A_t, t \in \mathbb{T}\}$ of nonempty compact subsets of X is called a

(i) *forward attractor* if it forward attracts all bounded subsets D of X, i.e.,

$$\lim_{t\to\infty} \mathrm{dist}_X\big(\phi(t, t_0, D), A_t\big) = 0 \quad \text{for every (fixed) } t_0$$

(ii) *pullback attractor* if it pullback attracts all bounded subsets D of X, i.e.,

$$\lim_{t_0\to-\infty} \mathrm{dist}_X\big(\phi(t, t_0, D), A_t\big) = 0 \quad \text{for every (fixed) } t.$$

In general, a forward attractor need not be a pullback attractor, or vice versa, cf. Example 4. The family $\mathscr{A} = \{A_t, t \in \mathbb{R}\}$ of singleton sets $A_t = \{\bar{x}(t)\}, t \in \mathbb{R}$, where $\bar{x}(t)$ is given by (8), is both a forward and a pullback attractor for the process generated by the ODE (6), cf. Example 3, but this is a special case. Essentially, pullback attraction uses information about the system from the past,[1] whereas forward attraction uses information about the future of the system.

The existence of a pullback attractor is ensured analogously to that of an autonomous attractor by the existence of a pullback absorbing set, cf. Theorem 1. For greater generality, non-uniformities in the dynamics can be handled by using a pullback absorbing family of sets instead of a single set. In addition, the bounded set D in the definition of pullback attraction by a family $\mathscr{D} = \{D_t, t \in \mathbb{T}\}$ of nonempty

[1] Perhaps *Urzeitattraktor* would be an appropriate name for a pullback attractor in German.

bounded subsets of X and pullback attraction is now written as

$$\lim_{t_0 \to -\infty} \mathrm{dist}_X\big(\phi(t, t_0, D_{t_0}), A_t\big) = 0 \quad \text{for every (fixed) } t.$$

Definition 9 A family $\mathscr{B} = \{B_t, t \in \mathbb{T}\}$ of nonempty compact subsets of X is called a *pullback absorbing family* for a process ϕ on X if for each $t \in \mathbb{T}$ and every family $\mathscr{D} = \{D_t, t \in \mathbb{T}\}$ of nonempty bounded subsets of X there exists a $T_{t,\mathscr{D}} \in \mathbb{T}^+$ such that

$$\phi(t, t_0, D_{t_0}) \subseteq B_t \quad \text{for all } t_0 \le t - T_{t,\mathscr{D}}.$$

Theorem 2 *Suppose that a process ϕ on a complete metric space (X, d) with time set $\mathbb{T}$ has a pullback absorbing family $\mathscr{B} = \{B_t, t \in \mathbb{T}\}$. Then ϕ has a global pullback attractor $\mathscr{A} = \{A_t, t \in \mathbb{T}\}$ with component subsets determined by*

$$A_t = \bigcap_{\tau \le t} \overline{\bigcup_{t_0 \le \tau} \phi(t, t_0, B_{t_0})} \quad \textit{for each } t \in \mathbb{T}. \tag{9}$$

Moreover, if $\mathscr{A}$ is uniformly bounded, i.e., if $\bigcup_{t\in\mathbb{T}} A_t$ is bounded, then $\mathscr{A}$ is the unique pullback attractor with this property.

There are many proofs of Theorem 2 and similar theorems in the literature, starting with [13, 16, 19, 58], see e.g. the monographs [9, 11, 46]. In infinite dimensional spaces it is often more convenient to establish the existence of an absorbing family $\mathscr{B}$ of closed and bounded subsets rather than compact subsets. Compactness of the subsets in (9) is then obtained as a consequence of compactness or asymptotic compactness of the process ϕ, see [8, 16, 39].

4.2 Existence of Forward Attractors for Processes

It is frequently repeated in the literature that there is no expression like (9) for a nonautonomous forward attractor. In fact, the component subsets of a forward attractor, when it exists, are also given by (9). (Strictly speaking, a forward attractor need not be unique, the one constructed here is maximal with respect to the positively invariant family of sets used in the construction.) The important observation is that the expression (9) holds inside any positively invariant family of sets [47], regardless of what is happening outside it. Moreover, a forward attractor is always contained in such a positively invariant family [43, 44].

The situation is somewhat more complicated due to some peculiarities of forward attractors compared to pullback attractors, see e.g. [54]. For example, a forward attractor need not be unique: for the process generated by the ODE $\dot{x} = 0$ if $t \le 0$ and $\dot{x} = -x$ if $t > 0$, every solution is a forward attractor (consisting of singleton component sets). The family $\mathscr{A}_r$ of sets $A_r(t) = r[-1, 1]$ for $t \le 0$ and $A_r(t) = re^{-t}[-1, 1]$ for $t \ge 0$ is also a forward attractor for each $r \in \mathbb{R}^+$. In both of these examples, the component subsets can be determined by the expression (9) for an appropriate positively invariant family.

The expression (9) is, however, only a necessary condition for the family of subsets so defined to be a forward attractor. Consider the process formed by the piecewise autonomous ODE $\dot{x} = 0$ if $t \le 0$ and $\dot{x} = x(1 - x^2)$ if $t > 0$. The expression (9) gives subsets $A(t) = \{0\}$, which are in fact the component subsets of a pullback attractor, but not a forward attractor as the forward ω-limit set is $[-1, 1]$, see [45]. Conditions excluding this case are discussed in [43, 44].

4.3 Nonautonomous Attractors: Skew Product Flows

Since a process is a special case of a skew product flow with the left shift operator defined by $\theta_t(t_0) = t - t_0$ on $P = \mathbb{T}$ as its driving system the definitions of the previous subsection can be translated to a skew product flow (θ, φ) in an obvious way. The essential difference is that the state of the driving system is used instead of the initial time.

Definition 10 A family $\mathscr{A} = \{A_p, p \in P\}$ of nonempty subsets A_p of X is called *φ-invariant* for a skew product flow (θ, φ) on $P \times X$ with time set $\mathbb{T}$ if

$$\varphi(t, p, A_p) = A_{\theta_t(p)} \quad \text{for all } p \in P \text{ and } t \in \mathbb{T}^+.$$

There are counterparts of pullback and forward attractors for skew product flows.

Definition 11 A φ-invariant family $\mathscr{A} = \{A_p, p \in P\}$ of nonempty compact subsets of X is called a

(i) *forward attractor* if it forward attracts all families $\mathscr{D} = \{D_t, t \in \mathbb{T}\}$ of nonempty bounded subsets of X, i.e.,

$$\lim_{t\to\infty} \operatorname{dist}_X\big(\varphi(t, p, D_p), A_{\theta_t(p)}\big) = 0 \quad \text{for each } p \in P.$$

(ii) *pullback attractor* if it pullback attracts all bounded subsets D of X, i.e.,

$$\lim_{t\to\infty} \operatorname{dist}_X\big(\varphi\big(t, \theta_-t(p), D_{\theta_-t(p)}\big), A_p\big) = 0 \quad \text{for each } p \in P.$$

Note the reformulation of the definition of *pullback attraction* here. For a process one starts earlier at time $t - \tau$ and finish at the fixed time t, whereas for a skew product flow the driving system starts at the earlier state $\theta_{-t}(p)$ and ends at time t later at the fixed state $\theta_0(p) = p$.

As for a process, the existence of a pullback attractor for skew product flow is ensured by that of a pullback absorbing family.

Definition 12 A family $\mathscr{B} = \{B_p, p \in P\}$ of nonempty compact subsets of X is called a *pullback absorbing family* for a skew product flow (θ, φ) on $P \times X$ if for each $p \in P$ and every family $\mathscr{D} = \{D_t, t \in \mathbb{T}\}$ of nonempty bounded subsets of X there exists a $T_{p,\mathscr{D}} \in \mathbb{T}^+$ such that

$$\varphi\big(t, \theta_{-t}(p), D_{\theta_{-t}(p)}\big) \subseteq B_p \quad \text{for all } t \ge T_{p,\mathscr{D}}.$$

Theorem 3 *Let (P, d_P) and (X, d_X) be complete metric spaces and suppose that a skew product flow (θ, φ) on $P \times X$ with the time set $\mathbb{T}$ has a pullback absorbing family $\mathscr{B} = \{B_p, p \in P\}$. Then (θ, φ) has a pullback attractor $\mathscr{A} = \{A_p, p \in P\}$ with component subsets determined by*

$$A_p = \bigcap_{t \leq 0} \overline{\bigcup_{s \geq t} \varphi\big(t, \theta_{-t}(p), B_{\theta_{-t}(p)}\big)} \quad \textit{for each } p \in P.$$

Moreover, if the components sets of $\mathscr{A}$ are uniformly bounded then is the unique pullback attractor with this property.

Remark 1 Nonautonomous semi-dynamical systems or skew product semi-flows with a *semi*-dynamical system as the driving system were investigated in [35], where the driving system is extended backwards in time to a set-valued dynamical system. This allows an "accumulative" pullback attraction to be defined corresponding to different "histories".

4.3.1 Another Kind of Nonautonomous Attractor

A skew product flow (θ, φ) on $P \times X$ also has a third type of attractor, namely the autonomous attractor $\mathfrak{A} \subset \mathfrak{X}$ of the corresponding autonomous semi-dynamical system Π on the state space $\mathfrak{X} = P \times X$, defined by

$$\Pi\big(t, (p_0, x_0)\big) = \big(\theta_t(p_0), \varphi(t, p_0, x_0)\big).$$

As seen above (for processes) $\mathfrak{A} = \bigcup_{p \in P} \{p\} \times A_p$, where A_p is a nonempty compact subset of X for each $p \in P$ and

$$\varphi(t, p, A_p) = A_{\theta_t p} \quad \text{for all } t \in \mathbb{T}^+, p \in P.$$

The family of sets $\mathscr{A} = \{A_p, p \in P\}$ is a pullback attractor for the skew product flow. In the converse direction, however, if $\mathscr{A} = \{A_p, p \in P\}$ is a pullback attractor for the skew product flow, then $\mathfrak{A} = \bigcup_{p \in P} \{p\} \times A_p$ need not be an attractor for the corresponding autonomous semi-dynamical system Π, but it is the maximal Π-invariant subset in $\mathfrak{X}$, see [10, 46].

4.3.2 Pullback Attractor of a Nonautonomous Scalar ODE

It was seen above that the scalar ODE (6), i.e., $\dot{x} = -x + \cos t$, generates a skew product flow on the state space $X = \mathbb{R}$ with the *hull* of the functions $\cos(\cdot)$ as the base space P and the left shift operator $\theta_t \cos(\cdot) = \cos(t + \cdot)$, $t \in \mathbb{T}$.

For $p_0 \in P$ with $p_0(\cdot) = \cos(\cdot)$, the ODE has a unique solution $x(0) = x_0$ given by

$$\varphi(t, p_0, x_0) = x_0 \mathrm{e}^{-t} + \mathrm{e}^{-t} \int_0^t \mathrm{e}^s p_0(s)\, ds.$$

On replacing $p_0(\cdot)$ by $\theta_{-\tau}(p_0(\cdot))$ for some $\tau > 0$ this can be rewritten as

$$x(\tau, \theta_{-\tau} p_0, x_0) = x_0 \mathrm{e}^{-t} + \mathrm{e}^{-t} \int_0^t \mathrm{e}^s \theta_{-\tau}\big(p_0(s)\big)\, ds = x_0 \mathrm{e}^{-t} + \int_{-\tau}^0 \mathrm{e}^{\eta} p_0(\eta)\, d\eta$$

with the change of variable $\eta = s - \tau$. Then pullback convergence has a single limit point

$$\lim_{\tau \to -\infty} x(t, p_0, x_0) = \int_{-\infty}^0 \mathrm{e}^{\eta} p_0(\eta)\, d\eta.$$

Moreover, any two solutions approach each other forward in time, which means that the pullback attractor consists of singleton subsets

$$A_p = \left\{ \int_{-\infty}^0 \mathrm{e}^{\eta}\, p(\eta)\, d\eta \right\}, \quad p \in P.$$

These are, in fact equivalent, to the pullback attractor component sets

$$A_t = \left\{ \frac{1}{2} \cos t + \frac{1}{2} \sin t \right\}, \quad t \in \mathbb{R},$$

of the process representation of the ODE. The pullback attractor is also a forward attractor here since any two solutions convergence together forward in time.

4.3.3 Pullback Attractor of a Nonautonomous Scalar Difference Equation

The nonautonomous scalar difference equation

$$x_{n+1} = a x_n + b_n,$$

where $a \in (0, 1)$ and $b_n \in [-B, B]$, i.e., the b_n are uniformly bounded with $|b_n| \leq B$, generates discrete-time skew product flow with the base space $P := [-B, B]^{\mathbb{Z}}$ of bi-infinite sequences $\mathbf{b} = (\ldots, b_{-2}, b_{-1}, b_0, b_1, b_2, \ldots)$ and the left shift operator $\mathbf{b}' := \theta \mathbf{b}$ for all $n \in \mathbb{Z}$. As in Example 2 the space (P, ρ) is a compact metric space with the metric

$$\rho\big(\mathbf{b}, \mathbf{b}'\big) := \sum_{n \in \mathbb{Z}} 2^{-|n|} \big|b_n - b_n'\big|.$$

Since the difference of any two solutions also converge together in the forward sense the pullback attractor consists of singleton component sets

$$A_{\mathbf{b}} = \big\{a(\mathbf{b})\big\} = \left\{ \sum_{k=-\infty}^{0} a^{-1-k} b_k \right\}.$$

Moreover, it is also forward convergent, hence it is a forward attractor too.

4.4 Another Kind of Forward Convergence for Pullback Attractors

The skew product formalism often provides more information about the dynamics, especially when the base space P is compact. Consider a pullback attractor $\mathscr{A} = \{A_p, p \in P\}$ of a skew product flow with a compact base space P and define $A(P) = \bigcup_{p\in P} A_p$. Then $A(P)$ forwards attracts the cocycle trajectories. See [46]. This a bit like the orbital stability of a limit cycle. Chepyzhov and Vishik [11] call $A(P)$ the *uniform* attractor, though it need not be invariant.

Note that $A(P) = [-1, 1]$ for the pullback attractor for the skew product flow generated by the ODE (6), which is quite coarse in view of the fact that the pullback attractor consists of singleton sets and is also a forward attractor.

5 Random Dynamical Systems

A random dynamical system (RDS) is essentially a nonautonomous dynamical system formulated as a skew product flow with the sample space Ω of a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ rather than a topological space as its base space. The driving system θ on Ω is now a *measurable*[2] dynamical system, i.e., with measurability rather than continuity in the base space variables. As with deterministic nonautonomous systems both pullback and forward random attractors are considered, but there are some surprising differences and technical complications since randomness always allows the possibility of exceptional nullsets. However, attractors of RDS have much stronger uniqueness properties than those of general nonautonomous systems. Different types of convergence can also be considered. Ludwig Arnold's monograph [2] is a definitive reference on RDS. See also his review article [1] in this journal.

Definition 13 Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space and let X be a topological space. A *random dynamical system* (θ, φ) on X consists of an autonomous measurable and measure-preserving dynamical system $\theta = \{\theta_t\}_{t\in\mathbb{T}}$ acting on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$, i.e.

$$\text{(i)}\quad \theta_0(\omega) = \omega, \qquad \text{(ii)}\quad \theta_{s+t}(\omega) = \theta_s \circ \theta_t(\omega),$$
$$\text{(iii)}\quad (t, \omega) \mapsto \theta_t(\omega) \text{ measurable}$$

for all $\omega \in \Omega$ and $s, t \in \mathbb{T}$, such that $\theta_t\mathbb{P} = \mathbb{P}$ for every $t \in \mathbb{T}$, where $\theta_t\mathbb{P}$ denotes the image measure of $\mathbb{P}$ under θ_t, and a *cocycle mapping* $\varphi : \mathbb{T}^+ \times \Omega \times X \to X$, i.e.,

(1) *initial condition:* $\varphi(0, \omega, x) = x$ for all $\omega \in \Omega$ and $x \in X$,
(2) *cocycle property:* $\varphi(s+t, \omega, x) = \varphi(s, \theta_t(\omega), \varphi(t, \omega, x))$ for all $s, t \in \mathbb{T}^+$, $\omega \in \Omega$ and $x \in X$,
(3) *measurability:* $(t, \omega, x) \mapsto \varphi(t, \omega, x)$ is measurable,
(4) *continuity:* $x \mapsto \varphi(t, \omega, x)$ is continuous for all $(t, \omega) \in \mathbb{T} \times \Omega$.

[2] The term *metric* or *metrical* is often used in the literature for historical reasons.

Measurability refers to joint measurability with respect to the Borel σ-algebras on $\mathbb{T}$ or $\mathbb{T}^+$ and X, and $\mathscr{F}$. The cocycle property is assumed to hold for all $\omega \in \Omega$, or at least in a subset of probability one. It is usually not too difficult to show that a "crude" cocycle property holds in specific examples, i.e., where the nullsets depend on the times s and t, but the verification of the "perfect" cocycle property as in (2) is a much more delicate issue.

Remark 2 Often *joint* continuity of $(t,x) \mapsto \varphi(t,\omega,x)$ for every ω is assumed instead of (4). Since the theory of attractors only makes use of continuity in the state space, not in the time, continuity in the time is not assumed here. Nevertheless many examples of RDS, in particular those being induced by random or stochastic differential equations, come with joint continuity in time and space.

For RDS the state space X is usually assumed to be a Polish space, i.e. X is a separable topological space such that there exists a complete metric on X inducing the topology.

Remark 3 Most state spaces which are of relevance in applications, in particular Euclidean spaces as well as separable Hilbert and Banach spaces, come equipped with their canonical metric, which is complete. However, this is not the case, for instance, for open nonempty subsets of all of these spaces, while nevertheless they are Polish. Also the set of Borel probability measures on (bounded or even compact subsets of) Euclidean, Hilbert or Banach spaces do not have a canonical metric for the topology of weak convergence. For an investigation of random attractors induced by RDS on the space of probability measures on the state space, equipped with the topology of weak convergence, see [15]. Furthermore, attractors for continuous dynamical systems are a concept which is not related to the properties of the metric of the space, but only to its topology.

In the following the state space X is assumed to be a Polish space. Several assertions are formulated in terms of a metric d on X which is referred to without further mentioning. This metric will always going to be assumed to be complete, even if some of the assertions hold also if this is not the case.

The driving system θ on Ω can be considered to be a canonical representation of the noise. For example, if the RDS is generated by an Itô stochastic differential equation with an m-dimensional two-sided Wiener process W_t, i.e., defined for $t \in \mathbb{R}$, then the probability space can be taken to be the canonical $\Omega = C_0(\mathbb{R}, \mathbb{R}^m)$ of continuous functions $\omega : \mathbb{R} \to \mathbb{R}^m$ with $\omega(0) = 0$ and θ defined by the increment shift operators $\theta_t(\omega)(\cdot) = \omega(t + \cdot) - \omega(t)$. No topological properties of Ω are used here. See [2] for more details and examples on RDS.

5.1 Random Attractors

Attractors of RDS have characteristic measurability properties. These are described using the notion of random sets.

Definition 14 Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space and X a Polish space. A *random set* C is a measurable subset of $X \times \Omega$ with respect to the product σ-algebra of the Borel σ-algebra of X and $\mathscr{F}$.

The ω-section of a random set C is defined by

$$C(\omega) = \{x : (x, \omega) \in C\}, \quad \omega \in \Omega.$$

In the case that a set $C \subset X \times \Omega$ has closed or compact ω-sections it is a random set as soon as the mapping $\omega \mapsto d(x, C(\omega))$ is measurable (from Ω to $[0, \infty)$) for every $x \in X$, see [14, Chap. 2]. Then C will be said to be a *closed* or a *compact*, respectively, random set. An *open random set* is a set $U \subset X \times \Omega$ such that U^c is a closed random set. It will be assumed that closed random sets satisfy $C(\omega) \neq \emptyset$ for all or at least for $\mathbb{P}$-almost all $\omega \in \Omega$.

Remark 4 It should be noted that in the literature very often it is tried to define random sets by demanding $\omega \mapsto d(x, C(\omega))$ to be measurable for every $x \in X$. Obviously this is satisfied, for instance, by $C(\omega) = N$ for all ω, where N is some non-measurable subset of X, and also by $C = (U \times F) \cup (\bar{U} \times F^c)$ for some open set $U \subset X$ and $F \notin \mathscr{F}$. In both cases $\omega \mapsto d(x, C(\omega))$ is constant, hence measurable, for every $x \in X$. However, both cases give $C \subset X \times \Omega$ which is not an element of the product σ-algebra of the Borel σ-algebra of X and $\mathscr{F}$.

Forward and pullback random attractors are defined similarly to their counterparts for deterministic skew product flows. The probability space $(\Omega, \mathscr{F}, \mathbb{P})$ will not always be stated explicitly.

Definition 15 Let (θ, φ) be an RDS on a Polish space X and let $\mathscr{B}$ be a family of random sets. A compact random set $A \subset X \times \Omega$ that is strictly φ-invariant, i.e.,

$$\varphi\big(t, \omega, A(\omega)\big) = A\big(\theta_t(\omega)\big) \quad \text{for every } t \in \mathbb{T}^+,\ \mathbb{P}\text{-a.s.},$$

is called a *random pullback attractor for* $\mathscr{B}$ if it pullback attracts $\mathscr{B}$, i.e. if

$$\lim_{t\to\infty} \operatorname{dist}\Big(\varphi\big(t, \theta_{-t}(\omega), B\big(\theta_{-t}(\omega)\big)\big), A(\omega)\Big) = 0 \quad \mathbb{P}\text{-a.s., for every } B \in \mathscr{B},$$

and a *random forward attractor for* $\mathscr{B}$ if it forward attracts $\mathscr{B}$, i.e. if

$$\lim_{t\to\infty} \operatorname{dist}\big(\varphi\big(t, \omega, B(\omega)\big), A\big(\theta_t(\omega)\big)\big) = 0 \quad \mathbb{P}\text{-a.s., for every } B \in \mathscr{B},$$

where dist is given by (1) with a complete metric d metrizing the topology of X.

Note that the convergence here is P-almost surely, where the nullset may depend on the set $B \in \mathscr{B}$ which is attracted.

Scheutzow [57] has constructed examples of attractors of RDS which are pullback but not forward, and vice versa, but these are much more difficult to find than in deterministic nonautonomous systems.

Definition 16 Let (θ,φ) be an RDS on a Polish space X and let $\mathscr{B}$ be a family of random sets. An *attracting set for* $\mathscr{B}$ is a random set K such that

$$\lim_{t\to\infty} \operatorname{dist}\big(\varphi\big(t,\theta_{-t}(\omega), B\big(\theta_{-t}(\omega)\big)\big), K(\omega)\big) = 0 \quad \mathbb{P}\text{-a.s.},$$

for every $B \in \mathscr{B}$.

An RDS with a compact attracting set for $\mathscr{B}$ is also said to be *asymptotically compact* with respect to $\mathscr{B}$, and for $\mathscr{B}$ consisting of all compact deterministic sets just asymptotically compact.

Remark 5 Note that the for a compact random set the property of being attracting for a family $\mathscr{B}$ of random sets does not depend on the choice of the metric d metrizing the topology of X.

Remark 6 Attracting sets are defined with reference to pullback convergence. While it would be straightforward to define attracting sets also for forward convergence, it is not evident how to make use of the corresponding concept, or whether it is useful at all. Therefore, when dealing with attracting sets no reference to 'in the pullback sense' will be made.

Remark 7 Clearly an absorbing set, defined for RDS in analogy with Definition 12, is automatically attracting, so existence of an absorbing compact set is a considerably stronger property than existence of an attracting compact set. While for RDS in Euclidean spaces it does not make a difference—whenever a compact set is attracting, its closed δ-neighbourhood is both compact and absorbing—this distinction is of considerable relevance for RDS on infinite-dimensional spaces. In order for an absorbing compact set for bounded sets to exist the system must be eventually compact.

The main result on the existence of random pullback attractors is given by the following result.

Theorem 4 *Suppose that (θ,φ) is an RDS on a Polish space X and let $\mathscr{B}$ be a family of random sets such that there exists a compact random set K which is attracting for $\mathscr{B}$. Then there exists a random pullback attractor A for $\mathscr{B}$.*

Furthermore, there exists a unique minimal *random pullback attractor $A_{\mathscr{B}}$ for $\mathscr{B}$.*

A proof when $\mathscr{B}$ consists of general deterministic sets is given in [13], where versions for $\mathscr{B}$ consisting of deterministic *bounded* sets appear in [16, 19]. The extension of the argument to families of random sets is straightforward. The condition for the existence of an attractor given in Theorem 4 is necessary and sufficient, since an attractor for $\mathscr{B}$ is obviously a $\mathscr{B}$-attracting set.

Remark 8 The assertion of Theorem 4 does not depend on the choice of a metric metrizing (the topology of) X, see Remark 5. Existence of an attractor for a family $\mathscr{B}$ of random sets is a purely topological property. But note that in the case where $\mathscr{B}$ is defined in terms of a metric then $\mathscr{B}$ changes with the metric. Most evidently this is the case for the choice of "all bounded sets" (random or not) for $\mathscr{B}$. For different choices of the metric the condition of Theorem 4 then have to be verified for different families $\mathscr{B}$ of random sets.

There is no analogous result for the existence of a random forward attractor. However, θ_t-invariance of the probability measure $\mathbb{P}$ implies

$$\mathbb{P}\big\{\omega \in \Omega : \mathrm{dist}\big(\varphi\big(t, \theta_{-t}(\omega), B\big(\theta_{-t}(\omega)\big)\big), A(\omega)\big) \geq \varepsilon\big\}$$
$$= \mathbb{P}\big\{\omega \in \Omega : \mathrm{dist}\big(\varphi\big(t, \omega, B(\omega)\big), A\big(\theta_t(\omega)\big)\big) \geq \varepsilon\big\}$$

for each $\varepsilon > 0$. Since P-almost sure convergence implies convergence in probability, a random pullback attractor also converges forwards, but only in the weaker sense of convergence in probability. The same argument gives that a random forward attractor is also a pullback attractor, but only in probability. See Sect. 5.3 below for more details.

5.2 The Global Random Pullback Set Attractor

The following result from [12] exhibits the strong uniqueness properties of random set attractors. It is used thereafter to describe the unique minimal random attractor $A_{\mathscr{B}}$ for a general family $\mathscr{B}$ of random sets in more detail.

Theorem 5 *Suppose that (θ, φ) is an RDS on a Polish space X such that there exists a compact attracting set for the family of all compact deterministic subsets of X. Then there exists a random pullback attractor A, and this attractor is unique in the sense that whenever A' is a random pullback attractor for every compact deterministic set then $A = A'$, $\mathbb{P}$-a.s. Furthermore, every random compact φ-invariant set is a subset of A, $\mathbb{P}$-a.s.*

As an immediate corollary one obtains

Corollary 1 *If $\mathscr{B}$ is an arbitrary collection of random sets with a random pullback attractor $A_{\mathscr{B}}$, then $A_{\mathscr{B}} \subset A$, $\mathbb{P}$-a.s. Furthermore, if $\mathscr{B}$ contains every compact deterministic set, then $A_{\mathscr{B}} = A$, $\mathbb{P}$-a.s.*

The large class of possible random pullback attractors for different families $\mathscr{B}$ of sets which are attracted thus reduces to just one unique attractor. Provided there exists an attractor A for the compact deterministic sets, then for an arbitrary family $\mathscr{B}$ of random sets there are just three possibilities:

1. There is no random pullback attractor for $\mathscr{B}$.
2. There is a random pullback attractor $A_{\mathscr{B}}$ for $\mathscr{B}$, then
   (a) either $A_{\mathscr{B}}$ is unique, then $A_{\mathscr{B}} = A$ (this is the case if $\mathscr{B}$ contains every compact deterministic set), or
   (b) $A_{\mathscr{B}}$ is not unique, then $A_{\mathscr{B}} \subset A$, with inequality for some $\mathscr{B}$-attractor.

Definition 17 Suppose that (φ, θ) is an RDS on a Polish space X such that there exists a compact attracting set for the family of all compact deterministic subsets of X. Then A is called the *global random pullback set attractor* or just *global random set attractor* or *set attractor* for short.

The global set attractor for RDS has several properties, which require some additional notation.

The *omega-limit set* of a random set B is defined by

$$\Omega_B(\omega) = \bigcap_{T \geq 0} \overline{\bigcup_{t \geq T} \varphi\big(t, \theta_{-t}(\omega), B\big(\theta_{-t}(\omega)\big)\big)}. \tag{10}$$

It is always a closed set and is also given by

$$\Omega_B(\omega) = \Big\{x \in X : \text{ there exist } t_n \in \mathbb{T},\ t_n \to \infty, \text{ and } b_n \in B\big(\theta_{-t_n}(\omega)\big),\ n \in \mathbb{N}, \\ \text{such that } x = \lim_{n \to \infty} \varphi\big(t_n, \theta_{-t_n}(\omega), b_n\big)\Big\}.$$

The omega-limit set Ω_B is always forward φ-invariant. In general, it may not be strictly φ-invariant. However, if the RDS φ has two-sided time, or if B is attracted by some compact random set, then the omega-limit set of B is strictly φ-invariant.

The σ-algebra $\mathscr{F}^- = \sigma\{\varphi(s, \theta_{-t}(\omega), x) : x \in X,\ 0 \leq s \leq t\} \subset \mathscr{F}$ is called the *past* of an RDS (θ, φ).

Theorem 6 *Suppose that (θ, φ) is an RDS on a Polish space X which has a global random pullback set attractor $\omega \mapsto A(\omega)$. Then*

(i) *A is measurable with respect to the past of φ, i.e., $\omega \mapsto d_X(x, A(\omega))$ is $\mathscr{F}^-$-measurable for every $x \in X$.*
(ii) *$\Omega_B(\omega) \subset A(\omega)$, $\mathbb{P}$-a.s., for every random set B attracted by A. This holds, in particular, for every compact deterministic set B.*
(iii) *$A(\omega) = \overline{\bigcup \Omega_C(\omega)}$, where the union is taken over all compact deterministic $C \subset X$.*
(iv) *if $(\theta_t)_{t \in \mathbb{T}}$ is ergodic then*

$$A(\omega) = \Omega_K(\omega) \tag{11}$$

for every compact deterministic $K \subset X$ with $\mathbb{P}\{A(\omega) \subset K\} > 0$.
(v) *The global random pullback set attractor is connected $\mathbb{P}$-a.s. provided that X has the additional property that for every compact $K \subset X$ there exists a compact connected $C \subset X$ with $K \subset C$.*

Theorem 6, which is proved in [12, 13], gives an idea of how to characterise a minimal attractor for a general class of random sets $\mathscr{B}$. Indeed, from [13] one has

Theorem 7 *Suppose that (θ, φ) is an RDS on a Polish space X and that $\mathscr{B}$ is a collection of random sets for which there exists a random pullback attractor. Then*

(i) *The omega-limit set Ω_B of every element $B \in \mathscr{B}$ is a subset of the attractor.*
(ii) *The minimal $\mathscr{B}$-attractor is given by*

$$A_{\mathscr{B}}(\omega) = \overline{\bigcup_{B \in \mathscr{B}} \Omega_B(\omega)}. \tag{12}$$

Measurability of the global set attractor with respect to the past follows from (11) together with (10). The omega-limit sets of random sets in general need not be measurable with respect to the past. In fact, unstable random equilibria are in general not measurable with respect to the past. Such an equilibrium induces an invariant (random one point) set, which therefore coincides with its omega-limit set, which is thus not measurable with respect to the past. It is interesting to observe that (12) holds as well for larger collections $\mathscr{B}$ of random sets, which possibly also contain omega-limit sets which are not measurable with respect to the past. Nevertheless, (12) still coincides with the global set attractor, which is measurable with respect to the past.

5.3 Weak Random Attractors

The random pullback and forward attractors considered so far involve pathwise convergence. Random attractors defined with respect to the weaker convergence in probability were introduced by Gunter Ochs [53].

Definition 18 Suppose that (θ, φ) is an RDS on a Polish space X. Let $\mathscr{B}$ be a family of random sets. Then a compact random set $A \subset X \times \Omega$ which is strictly φ-invariant is called a *weak random attractor for* $\mathscr{B}$ if A attracts $\mathscr{B}$ weakly, i.e.,

$$\lim_{t\to\infty} \operatorname{dist}\big(\varphi\big(t, \theta_{-t}(\omega) B\big(\theta_{-t}(\omega)\big), A(\omega)\big)\big) = 0 \quad \text{in probability} \tag{13}$$

for every $B \in \mathscr{B}$.

Since the probability $\mathbb{P}$ is invariant under θ_t for every t, the convergence (13) is equivalent to

$$\lim_{t\to\infty} \operatorname{dist}\big(\varphi\big(t, \omega, B(\omega)\big), A\big(\theta_t(\omega)\big)\big) = 0 \quad \text{in probability.}$$

Hence there is no difference between weak pullback and weak forward attractors.

There is no counterpart of the notion of weak attractor for general deterministic nonautonomous systems. Scheutzow [57] pointed out that a weak attractor need not be either a pullback or forward attractor in the stronger sense of almost sure convergence.

The following three theorems, the proofs of which are given in [17], provide a characterisation of conditions for the existence of pullback and of weak attractors for bounded and for compact deterministic sets. Throughout (θ, φ) is an RDS on a Polish space X and d is a particular choice of a complete metric inducing the topology of X. The closed δ-neighbourhood of a subset $S \subset X$ is denoted by $S^\delta = \{x \in X : d_X(x, S) \le \delta\}$.

First suppose that $\mathscr{B}$ is the collection of all bounded and closed deterministic subsets of X, so $\mathscr{B}$ obviously depends on the choice of the metric d.

Theorem 8 *The following are equivalent*:

1. *φ has a pullback $\mathscr{B}$-attractor.*
2. *For every $\varepsilon > 0$ there exists a compact $C_\varepsilon \subset X$ such that*

$$\mathbb{P}\Big\{\bigcup_{s\geq 0}\bigcap_{t\geq s}\varphi\big(t,\theta_{-t}(\omega),B\big)\subset C_\varepsilon^\delta\Big\}\geq 1-\varepsilon$$

for every $\delta > 0$ and for every $B \in \mathscr{B}$.

3. *φ is asymptotically compact with respect to $\mathscr{B}$.*

Theorem 9 *The following are equivalent*:

1. *φ has a weak $\mathscr{B}$-attractor.*
2. *For every $\varepsilon > 0$ there exists a compact $C_\varepsilon \subset X$ such that for every $\delta > 0$ and for every $B \in \mathscr{B}$ there is a $t_0 > 0$ such that for every $t \geq t_0$*

$$\mathbb{P}\big\{\varphi(t,\omega,B)\subset C_\varepsilon^\delta\big\}\geq 1-\varepsilon.$$

3. *There exists a compact weakly $\mathscr{B}$-attracting set $\omega \mapsto K(\omega)$, i.e., φ is weakly asymptotically compact with respect to $\mathscr{B}$.*

Now let $\mathscr{K}$ denote the collection of all compact deterministic subsets of X. Concerning the existence of a pullback attractor, Theorem 8 carries over verbatim with $\mathscr{K}$ replacing $\mathscr{B}$, although the proofs are more complicated.

Theorem 10 *The following are equivalent*:

1. *φ has a weak $\mathscr{K}$-attractor.*
2. *For every $\varepsilon > 0$ there exists a compact subset C_ε such that for every $\delta > 0$ and every $K \in \mathscr{K}$ there is a $t_0 > 0$ such that for all $t \geq t_0$*

$$\mathbb{P}\big\{\varphi(t,\omega,K)\subset C_\varepsilon^\delta\big\}\geq 1-\varepsilon.$$

3. *There exists a compact weakly attracting set $\omega \mapsto K(\omega)$ for $\mathscr{K}$ (so φ is weakly asymptotically compact).*

5.4 Random Morse Decompositions

For deterministic dynamical systems the concept of a splitting of a state space into attractor-repeller pairs is well known. Consider a two-sided time dynamical system on a compact state space, for simplicity. The state space might, for instance, have been obtained as the attractor of a dynamical system defined on a large space, and one continues investigations by considering the restriction of the dynamical system to the attractor.

So suppose that φ is a two-sided time system on a compact state space X. A *local attractor* A is a compact subset of X, which is strictly invariant in the sense of Definition 2 and for which there exists a neighbourhood U which is attracted by A.

One may then construct an associated repeller, and there are, in fact, several ways to achieve this. One possibility is to take $R = \{x \in X : \omega(x) \cap A = \emptyset\}$, where $\omega(x)$ denotes the (deterministic) omega-limit set of $x \in X$, and then to verify that R is compact, strictly invariant, and that it attracts every $x \notin A$ backward in time.

As a next step one may then perform the same procedure for the restriction of the dynamical system to A and to R, respectively, Continuing in this manner this results in an increasing sequence of local attractors with an associated decreasing sequence of local repellers, from which a *Morse decomposition* can be constructed by taking appropriate intersections.

It should be noted that the sequence of attractors and associated repellers is not unique in general, and also it does not have to be finite.

This problem has recently also been investigated for RDS. The deterministic approach sketched above does not work for RDS, however. Still, using a different characterisation of a repeller associated with a random attractor, corresponding results can be established for RDS, see [4, 18]. It is interesting to note that these results in general work for weak attractors and repellers only.

6 Mean-Square Random Attractors

Mean-square properties are of traditional interest in engineering and physics, in particular the mean-square stability of a zero solution of an Itô stochastic differential equation and the mean-square ultimate boundedness of solutions. In more complicated situations it is often possible to establish ultimate boundedness using Lyapunov function techniques. Ultimate boundedness is, essentially, equivalent to the existence of an absorbing set and it is natural to ask if there then exists a random attractor in the mean-square sense.

Of course, many Itô SDE can also be investigated as pathwise random dynamical systems and have random attractors under pathwise convergence, but this is not possible for a *mean-field* SDE, i.e., an SDE with expectations in coefficients such as

$$dX_t = f\left(X_t, \mathbb{E}X_t^2\right) dt + g\left(X_t, \mathbb{E}X_t^2\right) dW_t$$

or the more general SDE with non-local sample path dependence that were introduced in [41].

6.1 Mean-Square Random Dynamical Systems

As before consider the time set $\mathbb{T} = \mathbb{Z}$ (discrete-time case) or $\mathbb{R}$ (continuous-time case) and recall that $\mathbb{T}^2_{\geq} := \{(t, t_0) \in \mathbb{T}^2 \ : \ t \geq t_0\}$.

Let $(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_{t\in\mathbb{T}}, \mathbb{P})$ be a complete filtered probability space satisfying the usual hypothesis, i.e., $\{\mathscr{F}_t\}_{t\in\mathbb{T}}$ is an increasing and right continuous family of σ-sub-algebras of $\mathscr{F}$ that contains all $\mathbb{P}$-null sets. Essentially, $\mathscr{F}_t$ represents the information about the randomness at time t.

Finally, define $\mathfrak{X} := L^2(\Omega, \mathscr{F}; \mathbb{R}^d)$ and $\mathfrak{X}_t := L^2(\Omega, \mathscr{F}_t; \mathbb{R}^d)$ for each $t \in \mathbb{T}$. Note that $\mathfrak{X}_{t_1} \subset \mathfrak{X}_{t_2} \subset \mathfrak{X}$ for all $(t_2, t_1) \in \mathbb{T}^2$. These are all Banach spaces with the norm $\|X\|_2 := \sqrt{\mathbb{E}\|X\|^2}$, where $\|\cdot\|$ is the Euclidean norm of $\mathbb{R}^d$.

Definition 19 A *mean-square random dynamical system* (MS-RDS) ϕ on the underlying space $\mathbb{R}^d$ with a probability set-up $(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_{t\in\mathbb{R}}, \mathbb{P})$ is a family of mappings

$$\phi(t, t_0, \cdot) : \mathfrak{X}_{t_0} \to \mathfrak{X}_t, \quad (t, t_0) \in \mathbb{T}^2_{\geq},$$

satisfying

(i) *initial value property*: $\phi(t_0, t_0, X_0) = X_0$ for every $X_0 \in \mathfrak{X}_{t_0}$ and $t_0 \in \mathbb{T}$;
(ii) 2*-parameter semigroup property*: $\phi(t_2, t_0, X_0) = \phi(t_2, t_1, \phi(t_1, t_0, X_0))$ for every $X_0 \in \mathfrak{X}_{t_0}$ and all (t_2, t_1), $(t_1, t_0) \in \mathbb{T}^2_{\geq}$;
(iii) *continuity property*: $(t, t_0, X_0) \mapsto \phi(t, t_0, X_0)$ is continuous in the space $\mathbb{T}^2 \times \mathfrak{X}$.

Note that the continuity property in which random variables from the different spaces $\mathfrak{X}_t$ are compared is defined in the larger enclosing Banach space $\mathfrak{X}$.

Mean-square random dynamical systems, introduced in [42], are essentially deterministic processes (not to be confused with stochastic processes) or 2-parameter semigroups on the space of mean-square random variables on $\mathbb{R}^d$. The stochasticity is built into or hidden in the time-dependent state spaces $\mathfrak{X}_t$. The definition differs slightly from that in Definition 5 in the use of time-dependent state spaces $\mathfrak{X}_t$.

Many other definitions and results from Sect. 3 carry over to the context of MS-RDS and will not be repeated here except to note that a family $\mathscr{B} = \{B_t, t \in \mathbb{T}\}$ of nonempty subsets of $\mathfrak{X}$ with $B_t \subset \mathfrak{X}_t$ for each $t \in \mathbb{T}$ will be called a *family of subsets of* $\{\mathfrak{X}_t, t \in \mathbb{T}\}$. Such a family $\mathscr{B}$ is said to be *uniformly bounded* if there is an $R := R_{\mathscr{B}} < \infty$ such that for every $t \in \mathbb{T}$, the estimate

$$\mathbb{E}\|X_t\|^2 \leq R^2 \quad \text{for all } X_t \in B_t \in \mathscr{B}.$$

6.2 Mean-Square Attractors

A mean-square random attractor is just a pullback attractor of a mean-square random dynamical system with its components sets taking values in the corresponding time-indexed state spaces. Mean-square forward random attractors can be defined similarly.

The proof of the following theorem establishing the existence of a mean-square random attractor is essentially the same as for Theorem 2 in Sect. 2.1. The only difference is that the spaces $\mathfrak{X}_t$ depends on time, but that is more a notational than technical difference since they are all subspaces of a common space $\mathfrak{X} = L^2(\Omega, \mathscr{F}; \mathbb{R}^d)$.

Theorem 11 *Suppose that a mean-square RDS ϕ on $\mathbb{R}^d$ has a ϕ-positively invariant pullback absorbing uniformly bounded family $\mathscr{B} = \{B_t, t \in \mathbb{T}\}$ of nonempty closed subsets of $\{\mathfrak{X}_t, \in \mathbb{T}\}$ and that the mappings $\phi(t, t_0, \cdot) : \mathfrak{X}_{t_0} \to \mathfrak{X}_t$ are pullback compact (respectively, eventually or asymptotically compact) for all $(t, t_0) \in \mathbb{T}^2_{\geq}$.*

Then ϕ has a unique global pullback attractor $\mathscr{A} = \{A_t, t \in \mathbb{T}\}$ with its component sets determined by

$$A_t = \bigcap_{t_0 \leq t} \phi(t, t_0, B_{t_0}) \quad \textit{for each } t \in \mathbb{T}.$$

If $\mathscr{B}$ is not ϕ-positively invariant, then

$$A_t = \bigcap_{s \geq 0} \overline{\bigcup_{t_0 \leq t-s} \phi(t, t_0, B_{t_0})} \quad \text{for each } t \in \mathbb{T}.$$

An explicit example of a trivial mean-square attractor bifurcating off the zero solution of a two-dimensional mean-field SDE can be found in [25].

Remark 9 The main difficulty in applying this theorem is to show that the process ϕ is compact in some sense. This is related to *the lack of criteria characterising compact subsets of the space $L^2(\Omega, \mathscr{F}; \mathbb{R}^d)$ of mean-square of random variables*. Sometimes the special structure of a system allows one to establish the existence of the attractor in another way.

6.3 Uniformly Strictly Contracting Property

Compactness can be circumvented in establishing the existence of an mean-square random attractor when the mean-square RDS satisfies a uniform strictly contracting property. This allows the construction of a Cauchy sequence and completeness, then ensure the existence of limit points; for a proof see [7, 42].

Let $\mathbb{B}_R$ be the closed and bounded ball in $\mathfrak{X}$ about the origin of radius $R > 0$ in the mean-square norm.

Definition 20 A mean-square RDS ϕ is said to satisfy a *uniform strictly contracting property* if for each $R > 0$, there exist constants $K, \alpha > 0$ such that

$$\mathbb{E}\left\|\phi(t, t_0, X_{t_0}) - \phi(t, t_0, Y_{t_0})\right\|^2 \leq K\mathrm{e}^{-\alpha(t-t_0)} \cdot \mathbb{E}\|X_{t_0} - Y_{t_0}\|^2$$

for all $(t, t_0) \in \mathbb{T}^2_{\geq}$ and $X_{t_0}, Y_{t_0} \in \mathbb{B}_R \cap \mathfrak{X}_{t_0}$.

This property holds, for example, in mean-square RDS generated by an SDE with a drift coefficient that satisfies a one-sided dissipative Lipschitz condition as well as other regularity assumptions.

The uniform strictly contracting property suffices in combination with a pullback absorbing set to ensure the existence of a mean-square attractor in both the forward and pullback sense that consists of singleton sets. Instead of compactness the proof uses completeness of the state space to obtain a limit of the Cauchy sequence of random variables.

Theorem 12 *Suppose that a mean-square RDS ϕ on $\mathbb{R}^d$ is uniformly strictly contracting on a ϕ-positively invariant pullback absorbing family $\mathscr{B} = \{B_t, t \in \mathbb{T}\}$ of nonempty uniformly bounded closed subsets of $\{\mathfrak{X}_t, t \in \mathbb{T}\}$.*

Then the mean-square RDS ϕ has a unique global forward and pullback attractor $\mathscr{A} = \{A_t, t \in \mathbb{T}\}$ with component sets consisting of singleton sets, i.e., $A_t = \{X^*_t\}$ for each $t \in \mathbb{T}$, where $\{X^*_t, t \in \mathbb{T}\}$ is a mean-square two-sided stochastic process in $\mathbb{R}^d$.

Note that the single process constituting the mean-square random attractor here need not be a stationary process.

As an example consider a scalar mean-field SDE

$$dX_t = -\big(X_t + (\mathbb{E}X_t)^3\big)\,dt + dW_t. \tag{14}$$

The "drift" function $f(x,y) = -x - y^3$ satisfies a one-sided dissipative Lipschitz condition with constant $L = 1$ in the mean-square sense, i.e.,

$$\mathbb{E}\big[(X-Y)\big(f(X,\mathbb{E}X) - f(Y,\mathbb{E}Y)\big)\big] \le -\mathbb{E}|X-Y|^2 \quad \text{for all } X, Y \in \mathfrak{X},$$

since

$$\mathbb{E}(X-Y)\big((\mathbb{E}X)^3 - (\mathbb{E}Y)^3\big) = (\mathbb{E}X - \mathbb{E}Y)\big((\mathbb{E}X)^3 - (\mathbb{E}Y)^3\big) \ge 0$$

and $(a-b)(a^3-b^3) \ge 0$ for all $a, b \in \mathbb{R}$.

Subtracting integral versions of the mean-field SDE (14) for two solutions X_t and Y_t, then taking expectations gives

$$\mathbb{E}(X_t - Y_t) = \mathbb{E}(X_0 - Y_0) - \int_0^t \big[\mathbb{E}(X_s - Y_s) + (\mathbb{E}X_s)^3 - (\mathbb{E}Y_s)^3\big]\,ds.$$

Since the expectations here are continuous functions of time, the Fundamental Theorem of (deterministic) Calculus says that $\mathbb{E}(X_t - Y_t)$ is differentiable, so

$$\frac{d}{dt}\mathbb{E}(X_t - Y_t) = -\mathbb{E}(X_t - Y_t) - (\mathbb{E}X_t)^3 - (\mathbb{E}Y_t)^3,$$

from which it follows that

$$\frac{d}{dt}\mathbb{E}|X_t - Y_t|^2 = 2\mathbb{E}(X_t - Y_t)\frac{d}{dt}\mathbb{E}(X_t - Y_t) \le -2\mathbb{E}|X_t - Y_t|^2.$$

Hence

$$\frac{d}{dt}\mathbb{E}|X_t - Y_t|^2 \le -2\mathbb{E}|X_t - Y_t|^2$$

which is integrated to give the mean-square uniform strict contracting property

$$\mathbb{E}|X_t - Y_t|^2 \le -\mathrm{e}^{-2(t-t_0)}\mathbb{E}|X_0 - Y_0|^2.$$

To show that there is a mean-square pullback aborbing family, subtract an Ornstein-Uhlenbeck process O_t, i.e., a solution of the linear Itô SDE with the same Wiener process, from the mean-field SDE (14) in integral form and take expectations to obtain

$$\mathbb{E}(X_t - O_t) = \mathbb{E}(X_0 - O_0) + \int_0^t \big[-\mathbb{E}(X_s - O_s) - (\mathbb{E}X_s)^3\big]\,ds.$$

Again by the continuity of the expected values of functions here, the Fundamental Theorem of Calculus says that $\mathbb{E}(X_t - O_t)$ is differentiable, so pathwise

$$\frac{d}{dt}\mathbb{E}(X_t - O_t) = -\mathbb{E}(X_t - O_t) - (\mathbb{E}X_t)^3,$$

from which it follows that

$$\frac{d}{dt}\mathbb{E}|X_t - O_t|^2 = 2\mathbb{E}(X_t - O_t)\big(-\mathbb{E}(X_t - O_t) - (\mathbb{E}X_t)^3\big)$$
$$\leq -\mathbb{E}|X_t - O_t|^2 + (\mathbb{E}O_t)^6.$$

Integration yields

$$\mathbb{E}|X_t - O_t|^2 \leq \mathrm{e}^{-(t-t_0)}\mathbb{E}|X_0 - O_0|^2 + \mathrm{e}^{-t}\int_{t_0}^{t} \mathrm{e}^s(\mathbb{E}O_s)^6\,ds.$$

Hence for $t_0 \leq T$, depending on suitable bounded sets of initial values and the end time t,

$$\mathbb{E}|X_t - O_t|^2 \leq \bar{R}_t := 1 + \mathrm{e}^{-t}\int_{-\infty}^{t} \mathrm{e}^s(\mathbb{E}O_s)^6\,ds.$$

The balls in $\mathfrak{X}_t$ centered on O_t of radius R_t thus form a pullback absorbing family. Theorem 12 then gives the existence a unique global forward and pullback attractor $\mathscr{A} = \{A_t, t \in \mathbb{R}\}$ with component sets consisting of singleton sets $A_t = \{X_t^*\}$, where $\{X_t^*, t \in \mathbb{R}\}$ is a mean-square stochastic process.

6.4 Mean-Square Weak Compactness

A closed and bounded subset of the space $L^2(\Omega, \mathscr{F}; \mathbb{R}^d)$ of mean-square of random variables is compact in the weak topology. Since closedness and boundedness are easily shown in this space, this suggests using weak compactness instead of norm mean-square compactness to define a weak mean-square random attractor. Moreover, since weak compactness is metrisable the above theory of mean-square RDS carries over to this new topology. The main difficulty in applications is to establish the continuity $X_0 \mapsto \phi(t, t_0, X_0)$ in specific examples. See [42] for such an example.

7 Future Developments and Other Results

The basic theories of nonautonomous and random attractors have now been established and open the door to new developments. The internal structure of such attractors is a very interesting issue and beginning to attract attention. Carvalho et al. [9] have results on a Morse-Smale like structure of nonautonomous attractors, in particular for systems that are not just perturbations of autonomous systems with a Morse-Smale structure. Closely related is the nature of forward attraction in nonautonomous systems. Rasmussen [54] has investigated Morse decompositions in the deterministic case.

Bifurcations of nonautonomous and random dynamical systems and their attractors is still very much undeveloped terrain with specific examples but no general theory, see e.g. [2, 20, 30, 31, 36, 46, 50, 55].

Estimates of the dimension of nonautonomous and random attractors of infinite dimensional systems have been extensively investigated in the literature, see e.g.

[9, 11, 21, 23, 24, 26]. It is interesting to note that in the case of an ergodic base flow $(\Omega, \mathscr{F}, \mathbb{P}, (\theta_t))$ (see Definition 14) dimensions of a random attractor are deterministic constants. One phenomenon observed in certain cases is that random attractors associated with a perturbation of a deterministic system may have a much smaller dimension than the dimension of the attractor of the perturbed system, see [5, 20]. In fact, the random attractor may consist of one random point, given by a random variable, while the deterministic system may have a very high-dimensional attractor. This has connections with the phenomen addressed to as *multistability*. A better understanding, possibly even a classification, of random dynamical systems with a random pullback attractor which is not pathwise forward attracting, and vice versa, is still an open problem.

Lyapunov functions for pullback attractors have been constructed in [27, 33, 34] and for random attractors in [3].

The numerical approximation of nonautonomous and random attractors has been discussed in the literature under strong assumptions, see, e.g., [31, 37, 38, 48, 49]. The efficient computation of pullback convergence remains a challenging open problem.

References

1. Arnold, L.: Random dynamical systems (Zufällige dynamische Systeme). Jahresber. Dtsch. Math.-Ver. **96**, 85–100 (1994)
2. Arnold, L.: Random Dynamical Systems. Springer, Berlin (1998)
3. Arnold, L., Schmalfuß, B.: Lyapunov's second method for random dynamical systems. J. Differ. Equ. **177**, 235–265 (2001)
4. Callaway, M.: On attractors, spectra and bifurcations of random dynamical systems. Ph.D. Thesis, Imperial College London (2014)
5. Caraballo, T., Crauel, H., Langa, J.A., Robinson, J.C.: The effect of noise on the Chafee-Infante equation: a nonlinear case study. Proc. Am. Math. Soc. **135**, 373–382 (2007)
6. Caraballo, T., Kloeden, P.E., Langa, J.A.: Atractores globales para sistemas diferenciales no autonomos. Cubo Mat. Educ. **5**, 305–329 (2003)
7. Caraballo, T., Kloeden, P.E., Schmalfuß, B.: Exponentially stable stationary solutions for stochastic evolution equations and their perturbation. Appl. Math. Optim. **50**, 183–207 (2004)
8. Caraballo, T., Lukaszewicz, G., Real, J.: Pullback attractors for asymptotically compact nonautonomous dynamical systems. Nonlinear Anal. (TMA) **64**, 484–498 (2006)
9. Carvalho, A.N., Langa, J.A., Robinson, J.C.: Attractors of Infinite Dimensional Nonautonomous Dynamical Systems. Springer, Berlin (2012)
10. Cheban, D.N., Kloeden, P.E., Schmalfuß, B.: The relationship between pullback, forward and global attractors of nonautonomous dynamical systems. Nonlinear Dyn. Syst. Theory **2**, 125–144 (2002)
11. Chepyzhov, V.V., Vishik, M.I.: Attractors for Equations of Mathematical Physics. American Mathematical Society, Providence (2002)
12. Crauel, H.: Global random attractors are uniquely determined by attracting deterministic compact sets. Ann. Mat. Pura Appl., IV. Ser. **CLXXVI**, 57–72 (1999)
13. Crauel, H.: Random point attractors versus random set attractors. J. Lond. Math. Soc., II Ser. **63**, 413–427 (2001)
14. Crauel, H.: Random Probability Measures on Polish Spaces. Taylor & Francis, London and New York (2002)
15. Crauel, H.: Measure attractors and Markov attractors. Dyn. Syst. **23**, 75–107 (2008)
16. Crauel, H., Debussche, A., Flandoli, F.: Random attractors. J. Dyn. Differ. Equ. **9**, 307–341 (1997)
17. Crauel, H., Dimitroff, G., Scheutzow, M.: Criteria for strong and weak random attractors. J. Dyn. Differ. Equ. **21**, 233–247 (2009)
18. Crauel, H., Hoang Duc, L., Siegmund, S.: Towards a Morse theory for random dynamical systems. Stoch. Dyn. **4**, 277–296 (2004)

19. Crauel, H., Flandoli, F.: Attractors for random dynamical systems. Probab. Theory Relat. Fields **100**, 365–393 (1994)
20. Crauel, H., Flandoli, F.: Additive noise destroys a pitchfork bifurcation. J. Dyn. Differ. Equ. **10**, 259–274 (1998)
21. Crauel, H., Flandoli, F.: Hausdorff dimension of invariant sets for random dynamical systems. J. Dyn. Differ. Equ. **10**, 449–473 (1998)
22. Dafermos, C.M.: An invariance principle for compact processes. J. Differ. Equ. **9**, 239–252 (1971)
23. Debussche, A.: On the finite dimensionality of random attractors. Stoch. Anal. Appl. **15**, 473–491 (1997)
24. Debussche, A.: Hausdorff dimension of a random invariant set. J. Math. Pures Appl. **77**, 967–988 (1998)
25. Doan, T.S., Rasmussen, M., Kloeden, P.E.: The mean-square dichotomy spectrum and a bifurcation to a mean-square attractor. Discrete Contin. Dyn. Syst., Ser. B **20**, 875–887 (2015)
26. Efendiev, M.: Attractors for Degenerate Parabolic Type Equations. American Math. Soc., Providence (2013)
27. Grüne, L., Kloeden, P.E., Siegmund, S., Wirth, F.R.: Lyapunov's second method for nonautonomous differential equations. Discrete Contin. Dyn. Syst., Ser. A **18**, 375–403 (2007)
28. Hale, J.K.: Asymptotic Behavior of Dissipative Systems. American Mathematical Society, Providence (1988)
29. Henry, D.: Geometric Theory of Semilinear Parabolic Equations. Springer, Berlin (1981)
30. Johnson, R.A., Kloeden, P.E., Pavani, R.: Two-step transition in nonautonomous bifurcations: an explanation. Stoch. Dyn. **2**, 67–92 (2002)
31. Keller, H., Ochs, G.: Numerical approximation of random attractors. In: Crauel, H., Gundlach, V.M. (eds.) Stochastic Dynamics, pp. 93–115. Springer, New York (1999)
32. Khasminskii, R.: Stochastic Stability of Differential Equations, 2nd edn. Springer, Heidelberg (2012)
33. Kloeden, P.E.: Pullback attractors in nonautonomous difference equations. J. Differ. Equ. Appl. **6**, 91–102 (2000)
34. Kloeden, P.E.: A Lyapunov function for pullback attractors of nonautonomous differential equations. Electron. J. Differ. Equ. **05**, 91–102 (2000)
35. Kloeden, P.E.: Pullback attractors of nonautonomous semidynamical systems. Stoch. Dyn. **3**, 101–112 (2003)
36. Kloeden, P.E.: Pitchfork and transcritical bifurcations in systems with homogeneous nonlinearities and an almost periodic time coefficient. Commun. Pure Appl. Anal. **3**, 161–173 (2004)
37. Kloeden, P.E., Keller, H., Schmalfuß, B.: Towards a theory of random numerical dynamics. In: Crauel, H., Gundlach, V.M. (eds.) Stochastic Dynamics, pp. 259–282. Springer, New York (1999)
38. Kloeden, P.E., Kozyakin, V.S.: Uniform nonautonomous attractors under discretization. Discrete Contin. Dyn. Syst. **10** (1 & 2), 423–433 (2004)
39. Kloeden, P.E., Langa, J.A.: Flattening, squeezing and the existence of random attractors. Proc. R. Soc. Lond., Ser. A **463**, 163–181 (2007)
40. Kloeden, P.E., Lorenz, J.: Stable attracting sets in dynamical systems and in their one-step discretizations. SIAM J. Numer. Anal. **23**, 986–995 (1986)
41. Kloeden, P.E., Lorenz, T.: Stochastic differential equations with nonlocal sample dependence. Stoch. Anal. Appl. **28**, 937–945 (2010)
42. Kloeden, P.E., Lorenz, T.: Mean-square random dynamical systems. J. Differ. Equ. **253**, 1422–1438 (2012)
43. Kloeden, P.E., Lorenz, T.: Pullback and forward attractors of nonautonomous difference equations. In: Bohner, M., Ding, Y., Dosly, O. (eds.) Proc. ICDEA Wuhan 2014. Springer, Heidelberg (2015)
44. Kloeden, P.E., Lorenz, T.: Construction of nonautonomous forward attractors. Proc. Am. Math. Soc. Published online on May 28, 2015. doi:10.1090/proc/12735
45. Kloeden, P.E., Pötzsche, C., Rasmussen, M.: Limitations of pullback attractors of processes. J. Differ. Equ. Appl. **18**, 693–701 (2012)
46. Kloeden, P.E., Rasmussen, M.: Nonautonomous Dynamical Systems. Amer. Math. Soc., Providence (2011)
47. Kloeden, P.E., Marín-Rubio, P.: Negatively invariant sets and entire solutions. J. Dyn. Differ. Equ. **23**, 437–450 (2011)
48. Kloeden, P.E., Schmalfuß, B.: Lyapunov functions and attractors under variable time-step discretization. Discrete Contin. Dyn. Syst. **2**, 163–172 (1996)
49. Kloeden, P.E., Schmalfuß, B.: Cocycle attractors of variable time–step discretizations of Lorenzian systems. J. Differ. Equ. Appl. **3**, 125–145 (1997)

50. Kloeden, P.E., Siegmund, S.: Bifurcation and continuous transition of attractors in autonomous and nonautonomous systems. Int. J. Bifurc. Chaos **15**, 743–762 (2005)
51. Krasnosel'skii, M.A.: The Operator of Translation along Trajectories of Differential Equations. Translations of Mathematical Monographs, vol. 19. Amer. Math. Soc., Providence (1968)
52. Ledrappier, F.: Positivity of the exponent for stationary sequences of matrices. In: Arnold, L., Wihstutz, V. (eds.) Lyapunov Exponents, Bremen, 1984. Lecture Notes in Math., vol. 1186, pp. 56–73. Springer, Berlin (1986)
53. Ochs, G.: Weak random attractors. Report Nr. 449, Institut für Dynamische Systeme, Universität Bremen (1999)
54. Rasmussen, M.: Attractivity and Bifurcation for Nonautonomous Dynamical Systems. Lecture Notes in Mathematics, vol. 1907. Springer, Heidelberg (2007)
55. Rasmussen, M.: Morse decompositions of nonautonomous dynamical systems. Trans. Am. Math. Soc. **359**, 5091–5115 (2007)
56. Romeiras, F., Grebogi, C., Ott, E.: Multifractal properties of snapshot attractors of random maps. Phys. Rev. A **41**, 784–799 (1990)
57. Scheutzow, M.: Comparison of various concepts of a random attractor: a case study. Arch. Math. **78**, 233–240 (2002)
58. Schmalfuß, B.: Backward cocycles and attractors of stochastic differential equations. In: Reitmann, V., Riedrich, T., Koksch, N. (eds.) International Seminar on Applied Mathematics—Nonlinear Dynamics: Attractor Approximation and Global Behaviour, pp. 185–192. Technische Universität Dresden, Dresden (1992)

Hans Crauel received the Diploma degree in Mathematics in 1981 and the Doctoral degree from the Universität Bremen in 1987, where he continued as a PostDoc until 1990. From 1990 to 1997 he worked at the Universität des Saarlandes, during which time he had a Vertretungsprofessur at the Universität Trier from 1993 to 1994 and a Visiting Professor's position at the Scuola Normale Superiore in Pisa, Italy. He finished his habilitation at the Universität Bremen in 1994. From 1997 to 2000 he stayed at the Technische Universität Berlin as a research associate. In 2000 and 2001 he was a research associate at the University of Exeter, UK. From 2000 to 2007 he was first an Oberassistent and then a Hochschuldozent at the Technische Universität Ilmenau, during which time he had a Vertretungsprofessur at the Universität des Saarlandes in 2004/2005. In 2007 he moved to the Goethe-Universität Frankfurt, where he has a position as außerplanmäßiger Professor. His research interests are random dynamical systems, in particular stability properties, Lyapunov exponents, hyperbolicity and random attractors.

Peter E. Kloeden retired as Professor of Applied and Instrumental Mathematics at the Goethe University Frankfurt am Main towards the end of 2014. He had previously worked in various universities in Australia. He is currently Research Professor for Stochastic Dynamics at the Huazhong University of Science & Technology in Wuhan China under the "1000 Expert" program and is also a Felix-Klein research fellow in the Felix-Klein-Zentrum für Mathematik at the Technische Universität Kaiserslautern. Peter Kloeden is an inaugural Fellow of SIAM and was awarded SIAM's W.T. & Idalia Reid Prize in 2006.

Jahresber Dtsch Math-Ver (2015) 117:207–215
DOI 10.1365/s13291-015-0110-5

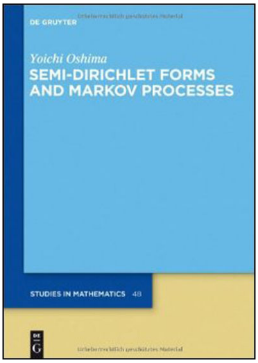
BOOK REVIEW

On Dirichlet Forms and Semi-Dirichlet Forms

Remarks on the Book "Semi-Dirichlet Forms and Markov Processes" by Yoichi Oshima

Moritz Kassmann[1]

Published online: 20 February 2015



Abstract One aim of this note is to give an overview of some developments in the area of Dirichlet forms. A second aim is to review the new book "Semi-Dirichlet forms and Markov processes" by Yoichi Oshima. This book was published by de Gruyter (Berlin) in 2013, but first versions were written as lecture notes 25 years ago. We first give a rather short and light introduction into the field of Dirichlet forms with a special emphasis on the subjects presented in the book under consideration. After a small account on the history of Dirichlet forms we comment on the book by Oshima against the background of related works.

Keywords Dirichlet forms · Potential theory · Markov processes

Mathematics Subject Classification 31-XX · 60-XX

1 Dirichlet Forms and Semi-Dirichlet Forms

In short, the theory of Dirichlet forms is an advancement and an abstraction of potential theory. The theory of Dirichlet forms has created a lot of interesting and interrelated research since 1970. Dirichlet forms are related to probability theory,

The author would like to thank several participants of the Oberwolfach workshop "Dirichlet Form Theory and its Applications" in 2014 for helpful hints and discussions.

✉ M. Kassmann
moritz.kassmann@uni-bielefeld.de

1 Fakultät für Mathematik, Universität Bielefeld, Postfach 100131, 33501 Bielefeld, Germany

Riemannian geometry, pseudo-differential operators and mathematical physics. The strength and the beauty of this theory are that it provides a framework that connects spectral theory, functional analysis and stochastic processes in a natural way. Given some Hilbert space $L^2(X,m)$ of square-integrable functions on some topological space X with measure m, a Dirichlet form is a pair $(\mathcal{E},\mathcal{F})$ of a bilinear form $(u,v)\mapsto\mathcal{E}(u,v)$ for u and v from some domain $\mathcal{F}\subset L^2(X,m)$. The domain $\mathcal{F}$ itself, historically, is called Dirichlet space. Before discussing further requirements and examples, let us explain the main characteristics. A Dirichlet form is called symmetric if $\mathcal{E}(u,v)=\mathcal{E}(v,u)$ for all u,v. Whether X is a locally compact separable metric space like a subset of $\mathbb{R}^d$ or whether X is a more general infinite-dimensional state space has significant consequences for the whole theory. The book [55] focuses on lower-bounded semi-Dirichlet forms which are defined on locally compact separable metric spaces X.

Let us look at examples which are basic and were in the mind of those who founded the theory of Dirichlet forms. Assume $D\subset\mathbb{R}^d$ is open. We denote by $H^1(D)$ the space of all elements $u\in L^2(D)$ whose distributional derivatives ∇u are elements of $L^2(D)$. $H^1(D)$ is a Banach space with respect to the norm $\|u\|^2_{H^1(D)}=\|u\|^2_{L^2(D)}+\||\nabla u|\|^2_{L^2(D)}$. $H^1_0(D)$ denotes the closure of $C_c^\infty(D)$ with respect to $\|u\|_{H^1(D)}$. Set $\mathcal{E}^{(1)}(u,v)=\int_D\nabla u\nabla v\,\mathrm{d}\lambda$, $\mathcal{F}^{(1)}=H^1(D)$, $\mathcal{F}^{(2)}=H^1_0(D)$, where λ denotes the Lebesgue measure. Then $(\mathcal{E}^{(1)},\mathcal{F}^{(1)})$ and $(\mathcal{E}^{(1)},\mathcal{F}^{(2)})$ are symmetric Dirichlet forms on $L^2(D,\lambda)$. For $\alpha\geq 0$ and $u,v\in\mathcal{F}$ let us denote $\mathcal{E}_\alpha(u,v)=\mathcal{E}(u,v)+\alpha_0(u,v)$ and $\mathcal{E}_\alpha(u)=\mathcal{E}_\alpha(u,u)$. The following conditions/properties turn out to be important. Assume $\alpha\geq 0$.

($\mathcal{E}$.1) $\mathcal{E}_\alpha(u)\geq 0$ for all $u\in\mathcal{F}$.
($\mathcal{E}$.2) For some $K\geq 1$ and for all $u,v\in\mathcal{F}$ $|\mathcal{E}(u,v)|\leq K\sqrt{\mathcal{E}_\alpha(u)}\sqrt{\mathcal{E}_\alpha(v)}$.
($\mathcal{E}$.3) For all $\beta>\alpha$ the domain $\mathcal{F}$ is a Hilbert space with respect to the scalar product $\mathcal{E}_\beta(u,v)+\mathcal{E}_\beta(v,u)$.
($\mathcal{E}$.4) For all $u\in\mathcal{F}$ the function $u^+\wedge 1$ belongs to $\mathcal{F}$ and $\mathcal{E}(u^+\wedge 1,u-u^+\wedge 1)\geq 0$.

These conditions define what is called a lower-bounded semi-Dirichlet form in [55, Section 1.1]. It makes sense that the author calls such forms Dirichlet forms in [55] but it might be confusing for the reader of this review. The by now classical definition given in [30, 32] requires $\mathcal{E}$ to be symmetric, conditions ($\mathcal{E}$.1), ($\mathcal{E}$.3) to hold with $\alpha=0$ and for all $u\in\mathcal{F}$ the condition $\mathcal{E}(u^+\wedge 1,u^+\wedge 1)\leq\mathcal{E}(u,u)$ which is stronger than ($\mathcal{E}$.4). The notion of a nonsymmetric Dirichlet form is slightly more tricky. Again, one requires ($\mathcal{E}$.1), ($\mathcal{E}$.3) to hold with $\alpha=0$. In this case one observes that ($\mathcal{E}$.4) is equivalent to $\mathcal{E}(u+u^+\wedge 1,u-u^+\wedge 1)\geq 0$. A nonsymmetric Dirichlet form now additionally requires

($\mathcal{E}$.4′) For all $u\in\mathcal{F}$ the function $u^+\wedge 1$ belongs to $\mathcal{F}$ and $\mathcal{E}(u-u^+\wedge 1,u+u^+\wedge 1)\geq 0$.

The above examples obviously are very special cases with additional features. First, they are symmetric forms. Second, we can choose $\alpha=0$ and $K=1$ by the Cauchy-Bunyakovsky-Schwarz inequality. ($\mathcal{E}$.4) is called Markov property because it relates to the Markov property of the related stochastic process. Note that $\mathcal{E}^{(i)}(u,u^+\wedge 1)\geq\mathcal{E}^{(i)}(u^+\wedge 1)$ trivially for $i=1,2$.

When relating these Dirichlet forms to extensions of $(-\Delta, C_c^\infty(D))$, readers might find [64] quite informative despite the fact that the author does not hide his personal view and his preference for another approach.

Let us provide an example which makes use of the flexibility of the above conditions. Define $\mathcal{E}^{(3)}(u,v) = \mathcal{E}^{(1)}(u,v) + \sum_{i=1}^d \int_D b_i \frac{\partial u}{\partial x_i} v \,\mathrm{d}\lambda$ for some functions $b_1, \ldots, b_d : D \to \mathbb{R}$ which either have bounded absolute values in D or (for $d \geq 3$) have the property that $\|\mathbf{b}\|_{L^d(D)}$ is finite and $\operatorname{div} \mathbf{b}$ is bounded from above. Here and below, we write $\mathbf{b}$ for the vector $(b_1, \ldots, b_d)^T$. The assumptions on $\mathbf{b}$ are tailored for conditions ($\mathcal{E}$.1) and ($\mathcal{E}$.2). This time, $\alpha \geq 0$ and $K \geq 1$ are chosen in dependence of $\mathbf{b}$. Thus the term *lower bounded* makes sense because of ($\mathcal{E}$.1). The tuple $(\mathcal{E}^{(3)}, H_0^1(D))$ is a lower bounded semi-Dirichlet form. In general, it is not a nonsymmetric Dirichlet form in the above sense. Note that our previous examples all were local forms in the sense that $\mathcal{E}(u,v) = 0$ if $u, v \in \mathcal{F}$ have disjoint supports. There is a whole universe of nonlocal symmetric/nonsymmetric Dirichlet/semi-Dirichlet forms.

There is a natural link between closed bilinear forms, semi-groups and resolvent operators. Assume $(\mathcal{E}, \mathcal{F})$ satisfies ($\mathcal{E}$.1)–($\mathcal{E}$.3). Then there exist strongly continuous semigroups (T_t), $(\widehat{T}_t)$ on $L^2(X,m)$ such that $\|T_t\| \leq e^{\alpha t}$, $\|\widehat{T}_t\| \leq e^{\alpha t}$, $(T_t f, g) = (f, \widehat{T}_t g)$ and for the resolvents $G_\beta, \widehat{G}_\beta$ given by $G_\beta f = \int_0^\infty e^{-\beta t} T_t f \,\mathrm{d}t$ and analogously for $\widehat{G}_\beta$: $\mathcal{E}_\beta(G_\beta f, u) = (f, u) = \mathcal{E}_\beta(u, \widehat{G}_\beta f)$ for all $f \in L^2(X,m)$, $u \in \mathcal{F}$. The term *semi-Dirichlet form* relates to the fact that, different from (T_t), the dual semi-group is not Markov in general.

What we have explained so far, holds true in infinite dimensions, too. This changes, when it comes to the important concept of regularity. A lower bounded semi-Dirichlet form $(\mathcal{E}, \mathcal{F})$ on $L^2(X,m)$ is called regular if $C_c(X) \cap \mathcal{F}$ is (a) dense in $\mathcal{F}$ with respect to the norm induced by $\mathcal{E}_1(\cdot)$ and (b) dense in $C_c(X)$ with respect to the supremum norm. The concept of regularity needs to be changed significantly when working with infinite dimensional state spaces, which we comment on below. The major achievement of the theory of Dirichlet forms is that there is a correspondence between Hunt processes (= quasi-left strong Markov processes) and regular Dirichlet forms if X is a locally compact separable metric space. More precisely, there exists a Hunt process whose resolvent $R_\alpha f$ is a quasi-continuous modification $G_\alpha f$ for any $f \in L^\infty(X;m)$ and $\alpha > 0$. Note that $R_\alpha f(x) = \int f(y) R_\alpha(x, \mathrm{d}y)$ where $R_\alpha(x, E) = \int_0^\infty e^{-\alpha t} p_t(x, E)\,\mathrm{d}t$ and p_t is the transition function of the Hunt process. It is possible to give a complete characterization of all symmetric and nonsymmetric Dirichlet forms satisfying the sector condition in terms of right processes, see [50].

Given the relation between a given Dirichlet form and the corresponding Hunt process many properties of the form can be studied by investigating the process and vice versa. A fundamental result in the theory of regular symmetric Dirichlet forms is the formula of Beurling-Deny which provides a unique representation. Together with the results of Le Jan it leads to the following beautiful description which we give in a simple setting. Assume $X = D \subset \mathbb{R}^d$ is a domain and $(\mathcal{E}, C_c^\infty(D))$ is a closable symmetric bilinear form on $L^2(D,m)$ satisfying the Markov property ($\mathcal{E}$.4). Then $\mathcal{E}$

can be expressed uniquely by

$$\mathcal{E}(u,v) = \sum_{i,j=1}^{d} \int_D \frac{\partial u}{\partial_{x_i}} \frac{\partial v}{\partial_{x_j}} \nu_{ij}(\mathrm{d}x)$$
$$+ \int_{D\times D} \big(u(y)-u(x)\big)\big(v(y)-v(x)\big) J(\mathrm{d}x\,\mathrm{d}y) + \int_D u(x)v(x)\kappa(\mathrm{d}x),$$

where ν_{ij}, J and k are nondegenerate positive Radon measures satisfying, among other properties, $\sum_{i,j=1}^{d} \xi_i \xi_j \nu_{ij}(K) \geq 0$ and $\int_{K\times K} |x-y|^2 J(\mathrm{d}x\,\mathrm{d}y) < +\infty$ for any compact $K \subset \mathbb{R}^d$. From the probabilistic point of view, a major result in the theory of Dirichlet forms is the Fukushima decomposition of an additive functional into a martingale additive functional and an additive functional of zero energy. This decomposition is similar to the semi-martingale decomposition for Markov processes and leads to results which, in the simplest cases, can be obtained by the Itô formula. We do not elaborate on this important result here.

The development of Dirichlet forms and corresponding Markov processes on infinite-dimensional state spaces is motivated by questions arising in quantum field theory and interacting particle systems. The main mathematical challenge is to find a substitute for the notion of regular Dirichlet forms. To this end, the notion of quasi-regularity for Dirichlet forms is introduced. Leaving technicalities aside, a Dirichlet form is quasi-regular if and only if it is quasi-homeomorphic to a regular Dirichlet form on a locally compact metric space. It can be shown that a Dirichlet form is associated with a nice Markov process if and only if it is quasi-regular. The quasi-homeomorphism allows the results developed for regular Dirichlet forms to be applied to quasi-regular Dirichlet forms on general infinite-dimensional state spaces. Note that the term "quasi" relates to exceptional sets which are defined with respect to capacity. These exceptional sets appear naturally and are abundant in the theory of Dirichlet forms. One task is to overcome them using regularity theory.

2 A Small and Incomplete Account on the History of Dirichlet Forms

The articles [17, 18, 27] study the domains of Dirichlet forms as Hilbert spaces and show that the concept of Dirichlet space captures a lot of the classical potential theory. Usually, these studies are regarded as the birth of the analytic side of Dirichlet form theory. The correspondence with a strong Markov process is provided in [28]. This review is not the right place to list all articles that contributed to this theory. We restrict ourselves to monographs and to those articles which are related to the focus of [55].

The books [29, 30, 58, 59] lay out the foundations of symmetric Dirichlet forms and corresponding Markov processes. It is important to note that nonsymmetric forms and related stochastic calculus were studied already at the very beginning of the theory, e.g. in [8, 9, 21, 22, 24, 42, 46–49, 56, 60, 61]. Note that these works benefit from the corresponding theory for elliptic differential operators in divergence form worked out in [62]. A standard reference for nonsymmetric forms is the monograph

[50] which appeared roughly at the same time as [23] and the first edition of [32] which is an extension of [30] and nowadays is the main reference for the field. Note that the main emphasis of [50] is on the development of the theory of Dirichlet forms on general state spaces. As mentioned above, the motivation to relax the assumptions (locally compactness) on the state space is closely connected to mathematical physics. First studies in this direction include [2, 3]. Since the book under review does not add results in this direction we do not provide more references on this important development and refer the interested reader to the discussion in [1, 4, 5, 7, 23, 50, 63, 68]. Note that the exposition in [63] goes beyond the scope of other books and covers truly nonsymmetric (without sector condition) and rather general time-dependent Dirichlet forms for infinite-dimensional state spaces.

A second field of current interest which is not touched by the book under consideration is the connection between geometry and Dirichlet forms. Again, we decide not to list many articles but rather give only a few hints where to find more information. The proceedings volume [39] might be a good start because it contains several related articles. On the one hand, Dirichlet forms provide a tool to define a Laplacian on general state spaces. On the other hand, they provide a framework for results which are robust with respect to geometric quantities. The geometric significance of Harnack inequalities and heat kernel bounds for Dirichlet forms are studied in [20, 65–67], see also the references therein. In typical situations, the Gaussian short time off-diagonal behavior of the heat kernel is a function of the intrinsic distance. This holds true for general strongly local Dirichlet forms [11, 36]. Aronson-type bounds have been studied in metric measure spaces and on fractals using the theory of Dirichlet forms. The works [10, 13–16, 33–35, 40, 41, 43, 44] contain several important results. The recent book [12] is a good source for the relation of functional inequalities and local Dirichlet forms in a general context. Note that the aforementioned contributions mainly concentrate on local Dirichlet forms. Similar studies for nonlocal Dirichlet forms and their relation to geometry seem to be much more subtle. Let us also mention that Dirichlet forms can be applied to other areas like discrete groups or random media. Two chapters of [70] address discrete groups and estimates of the decay of convolution powers of probability measures on these groups. For applications to random media see [45].

Last, let us mention the monograph [25]. It provides a 100-pages summary of the theory of symmetric regular and quasi-regular Dirichlet forms which can be used as a first read. Moreover, it treats new developments in more specialized fields, i.e., trace processes, boundary theory and reflected Dirichlet spaces for regular symmetric Dirichlet forms.

As explained above, nonsymmetric Dirichlet forms were studied right on from the beginning of research activities around Dirichlet forms. The lectures of Y. Oshima at Friedrich-Alexander-Universität Erlangen-Nürnberg in 1988 and 1994 contributed significantly to this development. It is not clear to the author of this review when the notion of a semi-Dirichlet form was used for the first time. Regularity resp. quasi-regularity of semi-Dirichlet forms are studied in [6, 52]. Several examples of semi-Dirichlet forms can be found in [31, 57, 69]. Chapter 7 of [19] contains several results on quasi-regular semi-Dirichlet forms. Recently, the Fukushima decomposition and the Beurling-Deny formulae have been studied for semi-Dirichlet forms [26, 37, 38, 51, 53]. See also the very recent survey [54] and which the author is thankful for having been provided with.

3 Remarks on the New Book by Yoichi Oshima

The new book [55] by Y. Oshima extends the theory of Dirichlet forms on locally compact separable metric spaces. After having set up the analytic theory of lower bounded semi-Dirichlet forms in the first two chapters, the relation to Hunt processes is studied in Chapter 3. Chapters 4 and 5 are devoted to additive functionals and decompositions. Thus the book can be viewed as a natural extension of [32]. The book is carefully written and has got a nice layout. It certainly will serve as a standard reference for its respective field. Because of the development of the theory of Dirichlet forms during the past 25 years, it is natural that there are by now other monographs. The standard reference for symmetric Dirichlet forms on locally compact separable metric spaces remains [32]. Symmetric Dirichlet forms in a more general framework are presented nicely also in [25]. If interested in symmetric or nonsymmetric Dirichlet forms satisfying the sector condition on more general state spaces, then [50] is the first choice. A special feature of [55] is the treatment of time-dependent Dirichlet forms in Chapter 6 which is an advancement of parabolic potential theory. These questions are also covered in [63] and even in much greater generality. However, the presentation in [55] is easier accessible and sufficient for many purposes. Altogether, the new book [55] is a welcome addition to the literature. The accuracy of the presentation and the similarity of the style to the one of [32] will be appealing to many researchers. As mentioned above, the main source for the book are the lectures by the author on nonsymmetric Dirichlet forms from 1988 and 1994 in Erlangen. Although the corresponding lecture notes are not available to the public, they have been circulated with the author's permission among interested colleagues. In this way, the book certainly has had an impact on the theory long before its publication.

References

1. Albeverio, S.: Theory of Dirichlet forms and applications. In: Lectures on Probability Theory and Statistics, Saint-Flour, 2000. Lecture Notes in Math., vol. 1816, pp. 1–106. Springer, Berlin (2003)
2. Albeverio, S., Høegh-Krohn, R.: Dirichlet forms and diffusion processes on rigged Hilbert spaces. Z. Wahrscheinlichkeitstheor. Verw. Geb. **40**(1), 1–57 (1977)
3. Albeverio, S., Høegh-Krohn, R.: Dirichlet forms and Markov semigroups on C^*-algebras. Commun. Math. Phys. **56**(2), 173–187 (1977)
4. Albeverio, S., Röckner, M.: Classical Dirichlet forms on topological vector spaces—Closability and a Cameron-Martin formula. J. Funct. Anal. **88**(2), 395–436 (1990)
5. Albeverio, S., Ma, Z.M., Röckner, M.: Quasi-regular Dirichlet forms and Markov processes. J. Funct. Anal. **111**(1), 118–154 (1993)
6. Albeverio, S., Ma, Z.M., Röckner, M.: Characterization of (non-symmetric) semi-Dirichlet forms associated with Hunt processes. Random Oper. Stoch. Equ. **3**(2), 161–179 (1995)
7. Albeverio, S., Ma, Z.-M., Röckner, M.: Quasi-regular Dirichlet forms and the stochastic quantization problem. In: Festschrift in Honour of Prof. M. Fukushima, World Scientific Review Volume. World Scientific, Singapore (2014)
8. Ancona, A.: Sur les espaces de Dirichlet: principes, fonction de Green. J. Math. Pures Appl. (9) **54**, 75–124 (1975)

9. Ancona, A.: Continuité des contractions dans les espaces de Dirichlet. In: Séminaire de Théorie du Potentiel de Paris, No. 2, Univ. Paris, Paris, 1975–1976. Lecture Notes in Math., vol. 563, pp. 1–26. Springer, Berlin (1976)
10. Andres, S., Barlow, M.: Energy inequalities for cutoff functions and some applications. J. Reine Angew. Math. (2015, to appear). doi:10.1515/crelle-2013-0009
11. Ariyoshi, T., Hino, M.: Small-time asymptotic estimates in local Dirichlet spaces. Electron. J. Probab. **10**(37), 1236–1259 (2005) (electronic)
12. Bakry, D., Gentil, I., Ledoux, M.: Analysis and Geometry of Markov Diffusion Operators. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 348. Springer, Cham (2014)
13. Barlow, M.T.: Diffusions on fractals. In: Lectures on Probability Theory and Statistics, Saint-Flour, 1995. Lecture Notes in Math., vol. 1690, pp. 1–121. Springer, Berlin (1998)
14. Barlow, M.T., Bass, R.F., Kumagai, T.: Stability of parabolic Harnack inequalities on metric measure spaces. J. Math. Soc. Jpn. **58**(2), 485–519 (2006)
15. Barlow, M.T., Bass, R.F., Kumagai, T., Teplyaev, A.: Uniqueness of Brownian motion on Sierpiński carpets. J. Eur. Math. Soc. **12**(3), 655–701 (2010)
16. Barlow, M.T., Grigor'yan, A., Kumagai, T.: On the equivalence of parabolic Harnack inequalities and heat kernel estimates. J. Math. Soc. Jpn. **64**(4), 1091–1146 (2012)
17. Beurling, A., Deny, J.: Espaces de Dirichlet. I. Le cas élémentaire. Acta Math. **99**, 203–224 (1958)
18. Beurling, A., Deny, J.: Dirichlet spaces. Proc. Natl. Acad. Sci. USA **45**, 208–215 (1959)
19. Beznea, L., Boboc, N.: Potential Theory and Right Processes. Mathematics and Its Applications, vol. 572. Kluwer Academic, Dordrecht (2004)
20. Biroli, M., Mosco, U.: A Saint-Venant type principle for Dirichlet forms on discontinuous media. Ann. Mat. Pura Appl. (4) **169**, 125–181 (1995)
21. Bliedtner, J.: Dirichlet forms on regular functional spaces. In: Seminar on Potential Theory, II. Lecture Notes in Math., vol. 226, pp. 15–62. Springer, Berlin (1971)
22. Bliedtner, J.: Functional spaces and their exceptional sets. In: Seminar on Potential Theory, II. Lecture Notes in Math., vol. 226, pp. 1–14. Springer, Berlin (1971)
23. Bouleau, N., Hirsch, F.: Dirichlet Forms and Analysis on Wiener Space. De Gruyter Studies in Mathematics, vol. 14. De Gruyter, Berlin (1991)
24. Carrillo-Menendez, S.: Processus de Markov associé à une forme de Dirichlet non symétrique. Z. Wahrscheinlichkeitstheor. Verw. Geb. **33**(2), 139–154 (1975/1976)
25. Chen, Z.-Q., Fukushima, M.: Symmetric Markov Processes, Time Change, and Boundary Theory. London Mathematical Society Monographs Series, vol. 35. Princeton University Press, Princeton (2012)
26. Chen, C., Ma, L., Sun, W.: Stochastic calculus for Markov processes associated with non-symmetric Dirichlet forms. Sci. China Math. **55**(11), 2195–2203 (2012)
27. Deny, J.: Méthodes hilbertiennes en théorie du potentiel. In: Potential Theory, C.I.M.E., I Ciclo, Stresa, 1969, pp. 121–201. Edizioni Cremonese, Rome (1970)
28. Fukushima, M.: Dirichlet spaces and strong Markov processes. Trans. Am. Math. Soc. **162**, 185–224 (1971)
29. Fukushima, M.: Dirichlet Forms and Markov Processes. Kinokuniya, Tokyo (1975) (in Japanese)
30. Fukushima, M.: Dirichlet Forms and Markov Processes. North-Holland Mathematical Library, vol. 23. North-Holland, Amsterdam (1980)
31. Fukushima, M., Uemura, T.: Jump-type Hunt processes generated by lower bounded semi-Dirichlet forms. Ann. Probab. **40**(2), 858–889 (2012)
32. Fukushima, M., Oshima, Y., Takeda, M.: Dirichlet Forms and Symmetric Markov Processes. De Gruyter Studies in Mathematics, vol. 19. De Gruyter, Berlin (2011) (extended edition)
33. Grigor'yan, A.: Heat kernels on metric measure spaces with regular volume growth. In: Handbook of Geometric Analysis, No. 2. Adv. Lect. Math. (ALM), vol. 13, pp. 1–60. International Press, Somerville (2010)
34. Grigor'yan, A., Hu, J.: Heat kernels and Green functions on metric measure spaces. Can. J. Math. **66**(3), 641–699 (2014)
35. Grigor'yan, A., Telcs, A.: Two-sided estimates of heat kernels on metric measure spaces. Ann. Probab. **40**(3), 1212–1284 (2012)
36. Hino, M., Ramírez, J.A.: Small-time Gaussian behavior of symmetric diffusion semigroups. Ann. Probab. **31**(3), 1254–1295 (2003)
37. Hu, Z.-C., Ma, Z.-M.: Beurling-Deny formula of semi-Dirichlet forms. C. R. Math. Acad. Sci. Paris **338**(7), 521–526 (2004)

38. Hu, Z.-C., Ma, Z.-M., Sun, W.: On representations of non-symmetric Dirichlet forms. Potential Anal. **32**(2), 101–131 (2010)
39. Jost, J., Kendall, W., Mosco, U., Röckner, M., Sturm, K.-T.: New Directions in Dirichlet Forms. AMS/IP Studies in Advanced Mathematics, vol. 8. Am. Math. Soc., Providence (1998)
40. Kigami, J.: Resistance forms, quasisymmetric maps and heat kernel estimates. Mem. Am. Math. Soc. **216**(1015), vi+132 (2012)
41. Kigami, J.: Analysis on Fractals. Cambridge Tracts in Mathematics, vol. 143. Cambridge University Press, Cambridge (2001)
42. Kim, J.H.: Stochastic calculus related to nonsymmetric Dirichlet forms. Osaka J. Math. **24**(2), 331–371 (1987)
43. Koskela, P., Zhou, Y.: Geometry and analysis of Dirichlet forms. Adv. Math. **231**(5), 2755–2801 (2012)
44. Koskela, P., Shanmugalingam, N., Zhou, Y.: Geometry and analysis of Dirichlet forms (II). J. Funct. Anal. **267**(7), 2437–2477 (2014)
45. Kumagai, T.: Anomalous random walks and diffusions: From fractals to random media (2014), presentation at ICM 2014, Seoul. See http://www.kurims.kyoto-u.ac.jp/~kumagai/ICM-Proc-TK-new.pdf
46. Kunita, H.: General boundary conditions for multi-dimensional diffusion processes. J. Math. Kyoto Univ. **10**, 273–335 (1970)
47. Le Jan, Y.: Mesures associées à une forme de Dirichlet. Applications. Bull. Soc. Math. Fr. **106**(1), 61–112 (1978)
48. Le Jan, Y.: Dual Markovian semigroups and processes. In: Functional Analysis in Markov Processes, Katata/Kyoto, 1981. Lecture Notes in Math., vol. 923, pp. 47–75. Springer, Berlin (1982)
49. Le Jan, Y.: Balayage et formes de Dirichlet. Z. Wahrscheinlichkeitstheor. Verw. Geb. **37**(4), 297–319 (1976/1977)
50. Ma, Z.M., Röckner, M.: Introduction to the Theory of (nonsymmetric) Dirichlet Forms. Universitext. Springer, Berlin (1992)
51. Ma, Z.-M., Sun, W.: Some topics on Dirichlet forms. In: New Trends in Stochastic Analysis and Related Topics. Interdiscip. Math. Sci., vol. 12, pp. 329–398. World Sci. Publ., Hackensack (2012)
52. Ma, Z.M., Overbeck, L., Röckner, M.: Markov processes associated with semi-Dirichlet forms. Osaka J. Math. **32**(1), 97–119 (1995)
53. Ma, L., Ma, Z.-M., Sun, W.: Fukushima's decomposition for diffusions associated with semi-Dirichlet forms. Stoch. Dyn. **12**(4), 1250003, 31 (2012)
54. Ma, Z.-M., Sun, W., Li-Fei, W.: Quasi-regular semi-Dirichlet forms and beyond. In: Festschrift in Honour of Prof. M. Fukushima, World Scientific Review Volume. World Scientific, Singapore (2014)
55. Oshima, Y.: Semi-Dirichlet Forms and Markov Processes. De Gruyter Studies in Mathematics, vol. 48. De Gruyter, Berlin (2013)
56. Paclet, P.: Formes de Dirichlet non symétriques et contractions normales. Atti Accad. Naz. Lincei, Rend. Cl. Sci. Fis. Mat. Nat. (8) **64**(2), 118–123 (1978)
57. Röckner, M., Schmuland, B.: Quasi-regular Dirichlet forms: examples and counterexamples. Can. J. Math. **47**(1), 165–200 (1995)
58. Silverstein, M.L.: Symmetric Markov Processes. Lecture Notes in Mathematics, vol. 426. Springer, Berlin (1974)
59. Silverstein, M.L.: Boundary Theory for Symmetric Markov Processes. Lecture Notes in Mathematics, vol. 516. Springer, Berlin (1976)
60. Silverstein, M.L.: The sector condition implies that semipolar sets are quasi-polar. Z. Wahrscheinlichkeitstheor. Verw. Geb. **41**(1), 13–33 (1977)
61. Silverstein, M.L.: Application of the sector condition to the classification of sub-Markovian semigroups. Trans. Am. Math. Soc. **244**, 103–146 (1978)
62. Stampacchia, G.: Le problème de Dirichlet pour les équations elliptiques du second ordre à coefficients discontinus. Ann. Inst. Fourier (Grenoble) **15**(fasc. 1), 189–258 (1965)
63. Stannat, W.: The theory of generalized Dirichlet forms and its applications in analysis and stochastics. Mem. Am. Math. Soc. **142**(678) (1999)
64. Stroock, D.W.: Book Review: Dirichlet forms & symmetric Markov processes, Dirichlet forms. Bull. Am. Math. Soc. **3**(1), 87–92 (1996)
65. Sturm, K.-T.: Analysis on local Dirichlet spaces. I. Recurrence, conservativeness and L^p-Liouville properties. J. Reine Angew. Math. **456**, 173–196 (1994)
66. Sturm, K.-T.: Analysis on local Dirichlet spaces. II. Upper Gaussian estimates for the fundamental solutions of parabolic equations. Osaka J. Math. **32**(2), 275–312 (1995)

67. Sturm, K.-T.: Analysis on local Dirichlet spaces. III. The parabolic Harnack inequality. J. Math. Pures Appl. (9) **75**(3), 273–297 (1996)
68. Trutnau, G.: Stochastic calculus of generalized Dirichlet forms and applications to stochastic differential equations in infinite dimensions. Osaka J. Math. **37**(2), 315–343 (2000)
69. Uemura, T.: On multidimensional diffusion processes with jumps. Osaka J. Math. **51**(4), 969–992 (2014)
70. Varopoulos, N.T., Saloff-Coste, L., Coulhon, T.: Analysis and Geometry on Groups. Cambridge Tracts in Mathematics, vol. 100. Cambridge University Press, Cambridge (1992)

Jahresber Dtsch Math-Ver (2015) 117:217–223
DOI 10.1365/s13291-015-0111-4

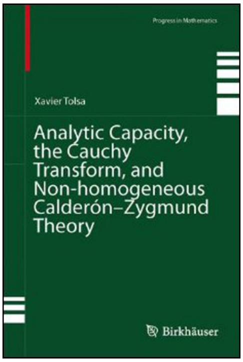
BOOK REVIEW

Xavier Tolsa: "Analytic Capacity, the Cauchy Transform, and Non-homogeneous Calderón-Zygmund Theory"

Birkhäuser, 2014, 396 pp.

Heiko von der Mosel[1]

Published online: 27 January 2015



What is analytic capacity? What does that have to do with the classic question, which subset E of the complex plane $\mathbb{C}$ is removable for a bounded analytic function on $\mathbb{C} \setminus E$? Why is it helpful in view of analytic capacity to develop a Calderón-Zygmund theory for measures that fail to have the doubling condition? And what is the magical relation between the boundedness of the Cauchy transform and the purely geometric concept of Menger curvature? These and many other fascinating questions are treated in the excellent monograph "Analytic capacity, the Cauchy transform, and non-homogeneous Calderón-Zygmund theory" by Xavier Tolsa, who is one of the leading experts in this subarea of harmonic analysis.

After an inspiring introduction Tolsa starts out with basic properties of Ahlfors's analytic capacity for compact sets $E \subset \mathbb{C}$,

$$\gamma(E) := \sup\left\{ \left| f'(\infty) \right| : f : \mathbb{C} \setminus E \to \mathbb{C} \text{ analytic}, \|f\|_\infty \le 1 \right\},$$

where $f'(\infty) := \lim_{z\to\infty} z(f(z) - f(\infty))$. Ahlfors showed in 1947 that $\gamma(E) = 0$ if and only if E is removable for bounded analytic functions $f : \mathbb{C} \setminus E \to \mathbb{C}$, that is, f can be extended onto all of $\mathbb{C}$. This is proved here in Chapter 1 with classic tools of complex analysis, and later improved using Vitushkin's localization operator. But Ahlfors's result left open the problem of characterizing removable sets in

✉ H. von der Mosel
heiko@instmath.rwth-aachen.de

1 Aachen, Germany

terms of their metric or geometric properties. After relating analytic capacity to Hausdorff measure Tolsa states at the end of the first chapter David's solution of 1998 to the famous Vitushkin conjecture for compact subsets $E \subset \mathbb{C}$ with one-dimensional Hausdorff measure $\mathscr{H}^1(E) < \infty$: $\gamma(E) = 0$ *if and only if E is purely unrectifiable, that is, if and only if $\mathscr{H}^1(E \cap \Gamma) = 0$ for any rectifiable curve $\Gamma \subset \mathbb{C}$*. In contrast to that, a set $E \subset \mathbb{C}$ is called (countably) 1-rectifiable if it can be covered—up to a set of $\mathscr{H}^1$-measure zero—by a countable union of rectifiable curves. David's proof of the very deep "only if" part uses sophisticated tools from geometric measure theory and harmonic analysis. An alternative proof based on a powerful Tb-theorem obtained by Nazarov, Treil, and Volberg in 2002, is presented in Tolsa's book; the necessary machinery is developed later in Chapters 5–7, probably the most demanding chapters of this book.

The second chapter contains the very useful Calderón-Zygmund theory for non-doubling measures with all the necessary covering lemmas, the various maximal operators, and some standard estimates for singular integral operators. The actual Calderón-Zygmund decomposition is applied to prove the weak $(1,1)$-boundedness of Calderón-Zygmund operators, and one learns about Cotlar's inequality for non-doubling measures proven by Nazarov, Treil, and Volberg in 1998. Here, as in many other places, Tolsa provides elegant alternative arguments taken from his own original papers. Based on a Whitney decomposition of open sets, Tolsa presents a version of the "good-lambda-method" for non-doubling measures, which is a powerful tool to prove the L^p-boundedness of singular integral operators.

The Menger curvature $c(x,y,z)$ defined as the inverse of the circumcircle radius of pairwise disjoint points $x, y, z \in \mathbb{C}$ is discussed in the third chapter, and after a few simple estimates and illuminating explicit calculations for three example sets, Tolsa establishes in an efficient way the magic relation between the Cauchy transform $\mathcal{C}(\mu)(x) := \int (y-x)^{-1}\,d\mu(y)$ and (integrated) Menger curvature

$$c^2(\mu) := \iiint c^2(x,y,z)\,d\mu(x)\,d\mu(y)\,d\mu(z)$$

of a finite Radon measure μ on $\mathbb{C}$ with linear growth:

$$\|\mathcal{C}_\epsilon(\mu)\|^2_{L^2(\mu)} = \frac{1}{6}c^2_\epsilon(\mu) + O\big(\mu(\mathbb{C})\big),$$

where the index ϵ indicates suitable truncations to cut out the singularities on the respective diagonals. This connection, originally discovered by Melnikov and exploited by Melnikov and Verdera in 1995, is used later in the book several times; here, in the third chapter, e.g., to present a new proof of the $T1$-theorem for the Cauchy singular operator giving three equivalent conditions for the mapping $f \mapsto \mathcal{C}_\mu(f) := \mathcal{C}(f\mu)$ to be bounded on L^2, one of them in terms of Menger curvature on squares. By means of additional basic estimates for pointwise Menger curvature Tolsa proceeds to prove the L^2-boundedness of the Cauchy transform $\mathcal{C}_{\mathscr{H}^1 \llcorner \Gamma}$ on Lipschitz graphs $\Gamma \subset \mathbb{C}$, and also on so-called AD-regular curves $\Gamma \subset \mathbb{C}$ characterized by the upper Ahlfors regularity condition

$$\mathscr{H}^1\big(\Gamma \cap B_r(x)\big) \le c_0 r \quad \text{for all } x \in \mathbb{C} \text{ and all } r > 0.$$

Peter Jones’s famous traveling salesman theorem of 1990 is only mentioned, but Jones’s β-numbers, as a scale invariant measure on how well one-dimensional sets can be approximated by straight lines, are discussed in more detail; in particular, how to bound Menger curvature in terms of β-numbers—again with a tricky but quite elementary proof. This can be used to sharpen the L^2-estimates for the Cauchy transform on Lipschitz graphs following ideas of Murai (1986).

The fourth chapter is devoted to the detailed study of an alternative notion of capacity, $\gamma_+(E)$ of compact subsets $E \subset \mathbb{C}$, defined as

$$\gamma_+(E) := \sup\bigl\{\mu(E) : \operatorname{supp}(\mu) \subset E,\ \bigl\|\mathcal{C}(\mu)\bigr\|_{L^\infty} \le 1\bigr\}.$$

Although this concept appears already in Murai’s book [2] in 1988, it is Tolsa’s great achievement to turn this capacity into a central and mighty tool in harmonic analysis. Tolsa could show in 2003 that analytic capacity $\gamma(E)$ and $\gamma_+(E)$ are, in fact, comparable quantities, the proof of which is deferred to Chapter 6, since it uses the deep Tb-theorem of Nazarov, Treil, and Volberg (2002) discussed in detail in Chapter 5. But already the fourth chapter contains a lot of very nice results. To start with, after preliminaries about convolutions, Tolsa uses a kind of representation inequality for Borel sets of Davie and Øksendal (1982) in the rather abstract setting of Radon measures on Hausdorff spaces, to obtain specifically for Radon measures with linear growth on $\mathbb{C}$ a dual form of the weak (1, 1)-inequality for the Cauchy transform. As a now relatively straightforward application Tolsa proves the Denjoy conjecture saying that a compact subset of a rectifiable curve in $\mathbb{C}$ has positive analytic capacity if and only if it has positive one-dimensional Hausdorff measure. It also follows quickly that any set $E \subset \mathbb{C}$ with $\mathscr{H}^1(E) < \infty$ and $\gamma_+(E) = 0$ is purely unrectifiable. Of central importance are several different characterizations of the alternative capacity $\gamma_+(E)$ in terms of suprema of the total variations of Radon measures under different constraints on either their truncated Cauchy transforms or their Menger curvatures, one of which immediately implies the countably semiadditivity of γ_+. By means of Verdera’s potential, the sum of the radial maximal function and pointwise Menger curvature, even a few more useful characterizations of $\gamma_+(E)$ are established, again based on elementary estimates for Menger curvature.

Tolsa’s very elegant style throughout the whole book may be exemplarily described in his proof of the dual characterization of γ_+ as the infimum of total variations of positive Radon measures whose Verdera potential is pointwise above 1. For this, Tolsa considers first elementary length measures on coordinate grids of squares, with internal concentric segments parallel to the coordinate axes, and he proves estimates for their Menger curvature at different points. Then he uses a nice variational argument to find maximizers of the quotient $\|\mu\|^2/(\|\mu\| + c^2(\mu))$, where $\|\mu\|$ denotes the total variation and $c^2(\mu)$ the (integrated) Menger curvature of a Radon measure μ supported on a finite collection of such segments. This maximizing measure satisfies particularly simple estimates on Menger curvature from above, and on Verdera’s potential from below, which is proven by an elementary variational inequality. And it is these estimates that qualify this maximizing measure as a useful comparison measure to prove the dual characterization. Towards the end of Chapter 4 Tolsa reproves Denjoy’s conjecture with a better quantitative lower bound on $\gamma_+(E)$ that is derived by a clever combination of Frostman's lemma with Jones's traveling salesman theorem and the relation between β-numbers and Menger curvature. Computing γ_+ for an explicit Cantor set shows that this lower bound is indeed sharp. The fourth chapter concludes with a brief discussion on how Verdera's potential is related to Riesz capacity which itself turns out to serve as a lower bound for γ_+ on compact subsets of the complex plane.

In Chapter 5 Tolsa gives a fully detailed proof of the deep Tb theorem of Nazarov, Treil, and Volberg, that he needs later to establish the comparability of the two capacities γ and γ_+. Although restricted to the Cauchy transform instead of more general singular integral operators that are treated, e.g., in Volberg's book [3] of 2003, the arguments presented here do not build on the relations between the Cauchy transform and Menger curvature and may be adapted to treat the more general case. It would go way beyond the scope of this review to describe this very deep result and its long proof, but let me point out that Nazarov, Treil, and Volberg used an ingenious decomposition of dyadic lattices to bound the so-called Θ-suppressed singular integral operator (in Tolsa's case the Θ-suppressed Cauchy transform), where the usual kernel is regularized by a Lipschitz function in the denominator. For particular such Lipschitz regularizations bounded from below by the distance to two dyadic lattices, Tolsa shows in every detail, how to obtain L^2-bounds on the Θ-suppressed Cauchy transform on "good" functions, a notion which in turn is defined by a subtle Martingale decomposition. This proof alone takes more than 20 pages, and belongs, as mentioned before, to the most technical parts of the book.

Before proving the comparability of the two capacities γ and γ_+ in Chapter 6, Tolsa points out two immediate consequences. A compact set $E \subset \mathbb{C}$ is not removable for bounded analytic functions if and only if E supports a non-vanishing Radon measure with linear growth and finite Menger curvature. Secondly, since γ_+ is countably subadditive, so is γ. Tolsa also proves first a weaker version of his comparability estimate due to David, saying that a compact subset of the complex plane with finite one-dimensional Hausdorff measure and positive analytic capacity γ must have also positive capacity γ_+. This simpler form (because of finite length of the set E) reveals clearly how to use the deep Theorem of Nazarov, Treil, and Volberg, and in this situation, it is indeed easier to verify all assumptions of that theorem by means of the results of Chapter 4. This model case is accompanied by a very instructive sketch of proof for Tolsa's full comparability result, before actually going into all the technical details needed to verify all assumptions of the Tb theorem. Chapter 6 closes with two nice applications in complex analysis. First he proves a general estimate for the Cauchy integral,

$$\left| \int_{\partial G} f(z)\, dz \right| \le c(G) \|f\|_\infty \gamma(E)$$

for bounded and holomorphic functions on $(G \setminus E) \subset \mathbb{C}$, for a compact subset E of G, under fairly mild conditions on the boundary ∂G, and secondly, he presents an alternative proof of the L^2-boundedness of the Cauchy transform on AD-regular curves.

The central issue of Chapter 7 is the deep rectifiability theorem of David and Léger of 1999: *An* $\mathscr{H}^1$*-measurable subset* $E \subset \mathbb{C}$ *with* $\mathscr{H}^1(E) < \infty$ *and with finite Menger curvature is* 1*-rectifiable*. Before proving this, Tolsa explains why this in connection with his comparability result for γ and γ_+ implies immediately David’s solution of the Vitushkin conjecture, that is, why the analytic capacity $\gamma(E)$ vanishes exactly on those one-dimensional compact subsets $E \subset \mathbb{C}$ of finite length that are purely unrectifiable. In addition, since γ_+ and hence γ is countably semiadditive David’s result immediately extends to compact subsets $E \subset \mathbb{C}$ of only σ-finite length, i.e., to sets that can be written as a countable union of sets of finite length. What follows now in Chapter 7 is probably the best presentation of the very technical proof of this deep theorem, even better in my taste than in Dudziak’s very nice book [1]. The general strategy of Léger’s proof is to decompose E into its rectifiable and purely unrectifiable part, and to show that this unrectifiable part has zero $\mathscr{H}^1$-measure. Assuming to the contrary positive measure of the purely unrectifiable part, one constructs a Lipschitz graph that covers a “good” subset of that part. In order to show that this good subset actually has positive measure (to obtain the contradiction), one needs to carry out very intricate estimates on the “bad” subset of the purely unrectifiable part. And here comes the relation between Menger curvature and β-numbers into effect, since one restricts to a subset with very small Menger curvature, thus controlling also the β-numbers which roughly shows that the set in question is fairly well approximated by lines on different scales. This general strategy sounds appealing, the technical details comprise a lot of precise and subtle estimates. Tolsa adds at the end of this chapter three nice applications: a characterization of 1-rectifiable sets in terms of pointwise Menger curvature, an upper bound on analytic capacity of a Borel subset $E \subset \mathbb{C}$ in terms of the $\mathscr{H}^1$-measure of its rectifiable part, and a new characterization of the analytic capacity for compact subsets $E \subset \mathbb{C}$ as the supremum of $\mu(E)$ over those Radon measures μ whose Menger curvature is bounded by $\mu(E)$ and whose upper density is bounded by one. For one-dimensional sets that are, in addition, AD-regular, Tolsa concludes with a rectifiability proof based on Jones’s traveling salesman theorem.

The L^2-boundedness of singular integral operators does not necessarily imply the existence of the pointwise Cauchy principal values—not even almost everywhere, but for Radon measures μ on $\mathbb{C}$ with linear growth and with $L^2(\mu)$-boundedness of the Cauchy transform $\mathcal{C}_\mu$, Tolsa proves in Chapter 8 that the Cauchy principal value

$$\text{p.v. } \mathcal{C}_\mu f(x) = \lim_{\epsilon\to 0} \mathcal{C}_{\mu,\epsilon} f(x) = \lim_{\epsilon\to 0} \int_{|x-y|>\epsilon} \frac{f(y)}{y-x}\, d\mu(y)$$

exists for every $f \in L^p(\mu)$, $p \in [1,\infty)$ and for μ-a.e. $x \in \mathbb{C}$. As an immediate consequence of this in combination with the $T1$-theorem for the Cauchy transform one can replace μ by some finite Radon measure with finite Verdera potential almost everywhere. Tolsa establishes first the existence of principal values for $\mu = \mathscr{H}^1 \llcorner \Gamma$, where Γ is a Lipschitz graph, and also for general complex finite measures for $\mathscr{H}^1$-a.e. point on a rectifiable subset of the complex plane. Then he studies Radon measures on $\mathbb{R}^d$ with growth of degree n and vanishing n-dimensional density a.e., following the approach of Mattila and Verdera (2009), before proving the main existence result for principal values. The two ingredients, rectifiable sets and measures of zero density, appear naturally through a simple decomposition of the support of $f\mu$, where it suffices to work on the dense set of C^1-functions f. Here, it is worth mentioning that the assumed L^2-boundedness of the Cauchy transform implies finite Menger curvature of the subsets where the upper density is positive, so that the David-Léger theorem of Chapter 7 can be applied to deduce rectifiability. The second part of Chapter 8 discusses the converse: when does the existence of the Cauchy principal value imply the L^2-boundedness of the Cauchy transform—at least on a subset of positive measure? One particular conclusion for one-dimensional subsets $E \subset \mathbb{C}$ with $\mathscr{H}^1(E) < \infty$ is the equivalence of rectifiability, finite Menger curvature, existence of the Cauchy principal value of $\mathscr{H}^1 \llcorner E$, and the boundedness of the maximal Cauchy transform

$$\mathcal{C}_*\big(\mathscr{H}^1 \llcorner E\big)(x) = \sup_{\epsilon>0}\big|\mathcal{C}_\epsilon \mu(x)\big| = \sup_{\epsilon>0}\left|\int_{|x-y|>\epsilon} \frac{1}{y-x}\, d\mu(y)\right|$$

for $\mathscr{H}^1$-a.e. $x \in E$. The main ingredient for this and some related results combining the maximal Cauchy transform with the radial maximal function, or with the upper density of a Radon measure, is again the Tb-theorem of Nazarov, Treil, and Volberg of Chapter 5, albeit not in its full generality this time.

In the final Chapter 9, Tolsa continues the Calderón-Zygmund theory for non-homogeneous spaces by discussing his variant of the space of bounded mean oscillation introduced in 2001 that is adapted to non-doubling measures μ on $\mathbb{R}^d$, so that, e.g., a John-Nirenberg inequality holds, and such that $L^2(\mu)$-bounded singular integral operators are also bounded from $L^\infty(\mu)$ to this new *BMO*-type space, denoted by $RBMO(\mu)$, which stands for *regular bounded mean oscillation*. The additional regularity requirement distinguishing this new space is some explicit control of the difference of two "means" f_Q and f_R for different cubes $Q \subset R$ of non-zero measure, which is, indeed, satisfied for functions $f = T_\mu(g)$, if T_μ is an $L^2(\mu)$-bounded singular integral operator and $g \in L^\infty(\mu)$. It turns out that $RBMO(\mu)$ is a Banach space (modulo additive constants) containing $L^\infty(\mu)$, and that there are various characterizations of this new space. The central boundedness of (e.g. L^2-bounded) singular integral operators as mappings from $L^\infty(\mu)$ to $RBMO(\mu)$ follows from a uniform bound on suitable truncations. Three examples of measures on $\mathbb{C}$ with linear growth are studied to get a better feeling for *RBMO*: if $E \subset \mathbb{C}$ is AD-regular, and $\mu := \mathscr{H}^1 \llcorner E$ then $RBMO(\mu)$ coincides with the usual *BMO*-space; if μ is the planar Lebesgue measure restricted to the unit square then $RBMO(\mu)$ turns out to be $L^\infty(\mu)$ modulo additive constants. And finally, by means of a more complicated measure on $\mathbb{C}$, Tolsa shows that the more traditional weighted *BMO*-norms for non-doubling measures heavily depend on the respective weights in this situation, whereas the *RBMO*-norm does not. The proof of a version of the classic John-Nirenberg inequality adapted to *RBMO* concludes that first part of Chapter 9. In the second part Tolsa introduces an atomic space as a Hardy space that turns out to be the predual of *RBMO* thus completing his *BMO*-theory for non-doubling measures. The central boundedness theorem on singular integral operators specifically asserts that L^2-bounded singular integral operators are also bounded from this new Hardy space into L^1, but there is actually a list of three equivalent conditions of that boundedness. By means of an interpolation result (interpolating between "Hardy $\to L^1$" and "$L^\infty \to RBMO$") in combination with the magic relation between the Cauchy transform and Menger curvature from the third chapter, Tolsa presents an alternative proof of the $T1$-theorem for the Cauchy transform. This is complemented by a general $T1$-theorem for Radon measures μ on $\mathbb{R}^d$ with growth of degree n and n-dimensional singular integral operators T_μ, giving equivalent conditions of the $L^2(\mu)$-boundedness in terms of uniform bounds on the truncated operators in $RBMO(\mu)$, and in weighted BMO-spaces, respectively, together with additional weak bounds.

This is a great book, I studied large portions of it with great benefit and pleasure. It covers a lot of material in this field—much more than I could mention here—with illuminating views from different perspectives. The arguments are presented with just the right amount of details so that the reader can go through the proofs without consulting the original papers. Most chapters could be read by students with a solid background in analysis, and certain parts of the book could serve as the basis for an advanced student seminar on, say, graduate level. Every chapter starts out with a short introduction so that the stage is set from the outset. In addition, at the beginning of almost every section Tolsa reminds the reader what is to be done next; whenever necessary he recalls definitions or central statements from previous chapters. Moreover, every chapter concludes with brief and very informative sections on history and references, containing additional hints towards generalizations, connections to other fields, and to open problems. These sections are treasures for experts and non-experts alike, since it allows you to either dive into some more specific topic, or to veer off to other related directions with the help of the extensive bibliography, which reflects the latest state of the art. To summarize, this outstanding book belongs in every mathematical library.

References

1. Dudziak, J.J.: Vitushkin's Conjecture for Removable Sets. Universitext. Springer, New York (2010)
2. Murai, T.: A Real Variable Method for the Cauchy Transform, and Analytic Capacity. Lecture Notes in Mathematics, vol. 1307. Springer, Berlin (1988)
3. Volberg, A.: Calderon-Zygmund Capacities and Operators on Nonhomogeneous Spaces. CBMS Regional Conference Series in Mathematics, vol. 100 (2003). Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence (2003)

Jahresber Dtsch Math-Ver (2015) 117:225–229
DOI 10.1365/s13291-015-0117-y

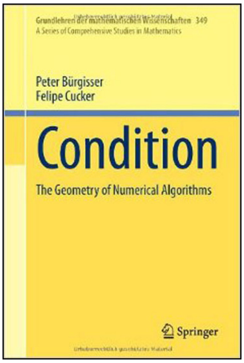
BOOK REVIEW

The Unavoidable Condition… A Report on the Book

Peter Bürgisser, Felipe Cucker: "Condition. The Geometry of Numerical Algorithms". Grundlehren der mathematischen Wissenschaften 349, Springer 2013, 554 pp.

Jean-Claude Yakoubsohn[1]

Published online: 30 May 2015



I remember my first numerical experiments using computers and the shout of joy after lots of attempts: *"it works"*. But disappointment came soon enough due to the inconsistency and lack of understanding of the numerical results. The first reaction was to believe that the computer code contains errors. But with time, coding errors disappear while numerical instabilities remain. Only theoretical conditions allow us to understand the mechanism of numerical errors. It was the book of Gastinel [5] that introduced me to the notion of condition. The terminology of condition number or conditioning is also used without distinction. Then, I understood that the computer was not the only one responsible for errors but the problem itself could contain the seeds of numerical disasters. The number of papers with *condition number* in the title is very huge and, it is very surprising that this book published in 2013 is the first book devoted entirely on this subject. It must be said that this book is a full success since it realizes a synthesis of ideas and works on the mathematical foundations on conditioning. First of all, I will say that I enjoyed the general spirit of the book explaining the points of view from which are discussed various aspects of condition. By the way, I will explain how Smale's 17th problem has stimulated research in this area and strengthened connections between algebraic geometry, integral geometry, probability and numerical analysis.

A problem of numerical analysis is a subset of the product of two sets called the inputs space and the outputs space. The condition of this problem measures the sensitivity of the output with respect to variations of the input. The problems considered in

✉ J.-C. Yakoubsohn
yak@math.univ-toulouse.fr

1 Toulouse, France

this book come from the resolution of linear systems, polynomial systems and linear programming.

In the case of the linear systems and/or polynomial systems the set of solutions appears implicitly under the form $F(input, output) = 0$ such that the implicit function theorem applies. Then the condition is defined as the norm of the derivative of this implicit function. Now comes the beautiful link between this measurement of the sensibility and the geometry. The geometry is governed by the locus of inputs where the implicit function does not exist, namely the singular variety. The condition number theorem, a very deep mathematical result, states that the condition of an input is the inverse of the distance from this input to the singular variety times the norm of this input. In the linear case, it is the celebrated Eckart-Young theorem. This characterization of the condition is the key for the probabilistic analysis that will be explained below.

For problems coming from linear optimization problems, it is not always possible to use differential calculus to define a condition concept. Then, it is an idea due to Renegar [6], to define the condition of an input as inversely proportional to the distance from this input to the set of ill-posed data.

This approach of condition only depends on the numerical problem and is totally disconnected from algorithms. But we can say that the numerical analysis of a class of algorithms computing approximately a solution of a given problem is mainly governed by the condition of this problem. The consequence for the computations using the floating point number system is that the loss of precision depends on the log of the condition. Therefore we need to estimate the size of the condition according to the size of the input. There are two probabilistic ways to study this issue.

The first probabilistic study is the average analysis which consists of choosing the input randomly in a certain space according to a certain probability distribution. This allows to investigate the condition as a random variable. More precisely the deal is to establish that the probability that the condition is greater than ϵ^{-1} is bounded by a polynomial in the input size and in a power of ϵ. In this way, we are done since it is very easy to deduce a bound for the expectation of the log of the condition. Finally this bound describes how many digits we need in average to perform the computation with a bounded loss of precision.

The second probabilistic way is the smoothed analysis, a theory born within this century and due to Spielman and Teng [14]. Concerning the condition, smoothed analysis consists in realizing the program of the average analysis but with a slight random perturbation of inputs.

In the case of the condition of a square matrix it is relatively easy to get polynomial complexity in the both cases of average and smoothed analyses for uniform distributions. I appreciated the simplicity and the elegance of the exposition of these analyses in Chapter 2 of this book. These analyses complete those of Azaïs and Wschebor [1] in the cases of Gaussian and non Gaussian distributions. For polynomial systems the things are technically much more tedious and the path found is due to Beltrán and Pardo [2]. To show polynomial complexity in the average analysis, the space of homogeneous polynomials is endowed with a unitarily invariant inner product (Weyl's inner product). The input space is the set product of the homogeneous polynomial systems and complex projective space. The numerical problem is the solution manifold, i.e., the couples of solutions *"(system, zero)"*. Moreover there is the singular variety characterized by the set of *"(system, zero)"* where the derivative has not full rank. This allows to define a notion of normalized condition number and to prove a geometric condition number theorem which expresses this normalized condition as the inverse distance to the singular variety. The distance here used is the Fubiny-Study distance. But to get good results of complexity we need to endow the solution manifold with a probability distribution that respects the nature of the input *"(system, zero)"* as a product of space. More precisely, if one has a probability distribution on the set of homogeneous systems this induces a probability for the projective space: this second probability distribution is classically named *pushforward density* of the initial one. Moreover, the famous co-area formula furnishes an explicit expression of the pushforward density. In this way, the solution variety is endowed with a probability distribution named standard probability distribution. It is the key for both average and smoothed analyses.

Now it is time to speak about the numerical algorithms with respect to the condition number. The problem of polynomial system solving arises in many theories and applications and, the literature on this subject is very huge. The background of this problem is the Bézout theorem which states that *generically* the number of (complex) zeros of a system of homogeneous polynomials is equal to the product of the degrees of each polynomial of the system. Classical algorithms to approximate the zeros are homotopy or path-following method. The goal is to start with a pair (*system, zero*) and to follow numerically the solution curve defined by the classical linear homotopy until to get an approximate zero of the system. Moreover the starting system must have the same number of zeros than the given system. The homotopy curve is numerically followed by steps based on the projective Newton method. This needs quantitative results on the local behaviour of the (projective) Newton method in terms of condition. The size of homotopy steps depends on this analysis. All the program comes from the five Bézout papers written by Shub and Smale that give a new numerical point of view to consider the polynomial system solving. But the question of choice of the starting pair (*system, zero*) was not solved. In this context, Smale published at the end of the last century a list of eighteen unsolved mathematical problems. Smale's 17th problem asks *Can a zero of n complex polynomial equations in n unknowns be found approximately, on the average [for a suitable probability measure on the space of inputs], in polynomial time with a uniform algorithm?* A first partial resolution of Smale's 17th problem has been given by Beltrán and Pardo [2]. They provide a uniform probabilistic algorithm where the crucial point is the random choice of the pair (*system, zero*) to start the algorithm. The result is that if this choice is done according to the standard probability distribution then, this algorithm computes approximations of zeros for most homogeneous polynomial systems with a complexity which depends polynomially on the input size.

Thanks to a result due to Shub [12] which states that the number of steps of the linear homotopy is bounded in terms of the integral of the square of normalized condition number, Cucker and Bürgisser [3] give a constructive version of the randomized linear homotopy *à la Beltrán-Pardo*. Next they perform a smoothed analysis of this algorithm to prove polynomial complexity in terms of the input size and the inverse of the random perturbation.

To conclude this review, a few words about the general organization of this book. The unavoidable condition is presented as the conductor of a symphony in three acts *Linear Algebra*(Adagio), *Linear Optimization*(Andante) and *Polynomial Equation Solving*(Allegro con brio): I suggest to listen to the Mozart symphony 38 which has the same structure! The book is self contained and easy to read: the technical background of various topics is introduced progressively. A *Crash Course on Probability* appears when necessary and, the links with differential geometry, algebraic geometry and projective spaces are clearly stated. Moreover, notes referring to each chapter, in addition to containing further details and more interesting historical references, give links between various results which show how old and new problems are related : for instance the very interesting note of the chapter 21 about the volume of tubes. The book ends with the statement of eighteen open problems that shows that *Mr. Condition* has a bright future ahead of him.

References

The average study of condition with Gaussian and non Gaussian distributions:

1. Azaïs, J.-M., Wschebor, M.: Level Sets and Extrema of Random Processes and Fields. Wiley, New York (2009)

The first elegant response at the Smale's 17th problem which is detailed in the book... but to have the complete story, see also in the bibliography of the book, the papers of Beltrán-Pardo of years 2007, 2008 and 2009:

2. Beltrán, C., Pardo, L.M.: Fast linear homotopy to find approximate zeros of polynomial systems. Found. Comput. Math. **11**(1), 95–129 (2011)

Smoothed analysis of homotopy for polynomial systems:

3. Bürgisser, P., Cucker, F.: On a problem posed by Steve Smale. Ann. Math. **174**(3), 1785–1836 (2011)

An concise view of condition:

4. Dedieu, J.-P.: Condition operators, condition numbers and condition number theorem for the generalized eigenvalue problem. Linear Algebra Appl. **263**, 1–24 (1997)

A pioneer of condition:

5. Gastinel, N.: Linear Numerical Analysis. Academic Press, San Diego (1970)

A first notion of condition for linear programming:

6. Renegar, J.: Incorporating condition measures into the complexity theory of linear programming. SIAM J. Optim. **5**(3), 506–524 (1995)

The five precursor papers for condition at the dawn of the twenty-first century:

7. Shub, M., Smale, S.: Complexity of Bézout's theorem I: geometric aspects. J. Am. Math. Soc., 459–501 (1993)
8. Shub, M., Smale, S.: Complexity of Bézout's theorem II: volumes and probabilities. Computational Algebraic Geometry, 267–285 (1993)
9. Shub, M., Smale, S.: Complexity of Bézout's theorem III: condition number and packing. J. Complex. **9**(1), 4–14 (1993)
10. Shub, M., Smale, S.: Complexity of Bézout's theorem IV: probability of success; extensions. SIAM J. Numer. Anal. **33**(1), 128–148 (1996)
11. Shub, M., Smale, S.: Complexity of Bézout's theorem V: polynomial time. Theor. Comput. Sci. **133**(1), 141–164 (1994)

… Without forgetting the bound of number steps of linear homotopy to get approximate zero with respect the integral of square of condition:

12. Shub, M.: Complexity of Bézout's theorem VI: Geodesics in the condition (number) metric. Found. Comput. Math. **9**(2), 171–178 (2009)

The existence of a homotopy path bounded by the log of condition:

13. Beltrán, C., Shub, M.: Complexity of Bézout's theorem VII: distance estimates in the condition (number) metric. Found. Comput. Math. **9**(2), 179–195 (2009)

The pioneer paper of smoothed analysis. last version… but the better!

14. Spielman, D.A., Teng, S.-H.: Smoothed analysis of algorithms: why the simplex algorithm usually takes polynomial time. http://arxiv.org/pdf/cs/0111050.pdf