Available online at www.sciencedirect.com
SCIENCE DIRECT
ELSEVIER
Physics Reports 422 (2006) 167–222
PHYSICS REPORTS
www.elsevier.com/locate/physrep

Deterministic Brownian motion: The effects of perturbing a dynamical system by a chaotic semi-dynamical system

Michael C. Mackey[a], Marta Tyran-Kamińska[b,*]

[a]*Departments of Physiology, Physics & Mathematics and Centre for Nonlinear Dynamics, McGill University, 3655 Promenade Sir William Osler, Montreal, QC, Canada, H3G 1Y6*
[b]*Institute of Mathematics, Silesian University, ul. Bankowa 14, 40-007 Katowice, Poland*

Accepted 22 September 2005
Available online 4 November 2005
editor: D.K. Campbell

Abstract

Here we review and extend central limit theorems for chaotic deterministic semi-dynamical discrete time systems. We then apply these results to show how Brownian motion-like behavior can be recovered and how an Ornstein–Uhlenbeck process can be constructed within a totally deterministic framework. These results illustrate that under certain circumstances the contamination of experimental data by "noise" may be alternately interpreted as the signature of an underlying chaotic process.

PACS: 05.40.−a; 05.45.Ac; 05.45.Pq; 05.70; 47.52.+j; 74.40.+k; 95.10.Fh

Keywords: Brownian motion; Ornstein–Uhlenbeck process; Ergodic maps; Chaos

Contents

* Corresponding author. Tel./fax: +48 32 258 2976.
E-mail addresses: mackey@cnd.mcgill.ca (M.C. Mackey), mtyran@us.edu.pl (M. Tyran-Kamińska).

0370-1573/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.physrep.2005.09.002

1. Introduction

Almost anyone who has ever looked through a microscope at a drop of water has been intrigued by the seemingly erratic and unpredictable movement of small particles suspended in the water, e.g. dust or pollen particles. This phenomenon, noticed shortly after the invention of the microscope by many individuals, now carries the name of "Brownian motion" after the English botanist Robert Brown who wrote about his observations in 1828. Almost three-quarters of a century later Einstein (1905) gave a theoretical (and essentially molecular) treatment of this macroscopic motion that predicted the phenomenology of Brownian motion. (A very nice English translation of this, and other, work of Einstein on Brownian motion can be found in Fürth (1956).) The contribution of Einstein led to the development of much of the field of stochastic processes and to the notion that Brownian movement is due to the summated effect of a very large number of tiny impulsive forces delivered to the macroscopic particle being observed. This was also one of the most definitive arguments of the time for an atomistic picture of the microscopic world.

Other ingenious experimentalists used this conceptual idea to explore the macroscopic effects of microscopic influences. One of the more interesting is due to Kappler (1931), who devised an experiment in which a small mirror was suspended by a quartz fiber (cf. Mazo (2002) for an analysis of this experimental setup). Any rotational movement of the mirror would tend to be counterbalanced by a restoring torsional force due to the quartz fiber. The position of the mirror was monitored by shining a light on it and recording the reflected image some distance away (so small changes in the rotational position of the mirror were magnified). Air molecules striking the mirror caused a transient deflection that could be monitored, and the frequency of these collisions was controlled by changing the air pressure. Fig. 1.1, taken from Kappler (1931), shows two sets of data taken using this arrangement and offers a vivid macroscopic depiction of microscopic influences.

In trying to understand theoretically the basis for complicated and irreversible experimental observations, a number of physicists have supplemented the reversible laws of physics with various hypotheses about the irregularity of the physical world. One of the first of these, and arguably one of the most well known, is the so-called "molecular chaos" hypothesis of Boltzmann (1995). This hypothesis, which postulated a lack of correlation between the movement of molecules in a small collision volume, allowed the derivation of the Boltzmann equation from the Liouville equation and led to the celebrated H theorem. The origin of the loss of correlations was never specified. In an effort to understand the nature of turbulent flow, Ruelle (1978–1980) postulated a type of mixing dynamics to be necessary. More recently, several authors have made "chaotic hypotheses" about the nature of dynamics at the microscopic level. The most prominent of these is Gallavotti (1999), and virtually the entire book of Dorfman (1999) is predicated on the implicit assumption that microscopic dynamics have a chaotic (loosely defined, but usually taken to be mixing) nature. All of these hypotheses have been made despite the fact that none of the microscopic dynamics that we write down in physics actually display such properties.

Others have taken this suggestion (chaotic hypothesis) quite seriously, and attempted an experimental confirmation. Fig. 1.2 shows a portion of the data, taken from Gaspard et al. (1998), that was obtained in an examination of a microscopic system for the presence of chaotic behavior. Their data analysis showed a positive lower bound on the sum of Lyapunov exponents of the system composed of a macroscopic Brownian particle and the surrounding fluid. From their analysis, they argued that the Brownian motion was due to (or the signature of) deterministic microscopic chaos. However, Briggs et al. (2001) were more cautious in their interpretation, and Mazo (2002, Chapter 18) has explored the possible interpretations of experiments like these in some detail.

If true, the existence of deterministic chaos (whatever that means) would be an intriguing possibility since it could serve as an explanation of a host of unresolved problems in the sciences. Most notably, it could serve as an explanation

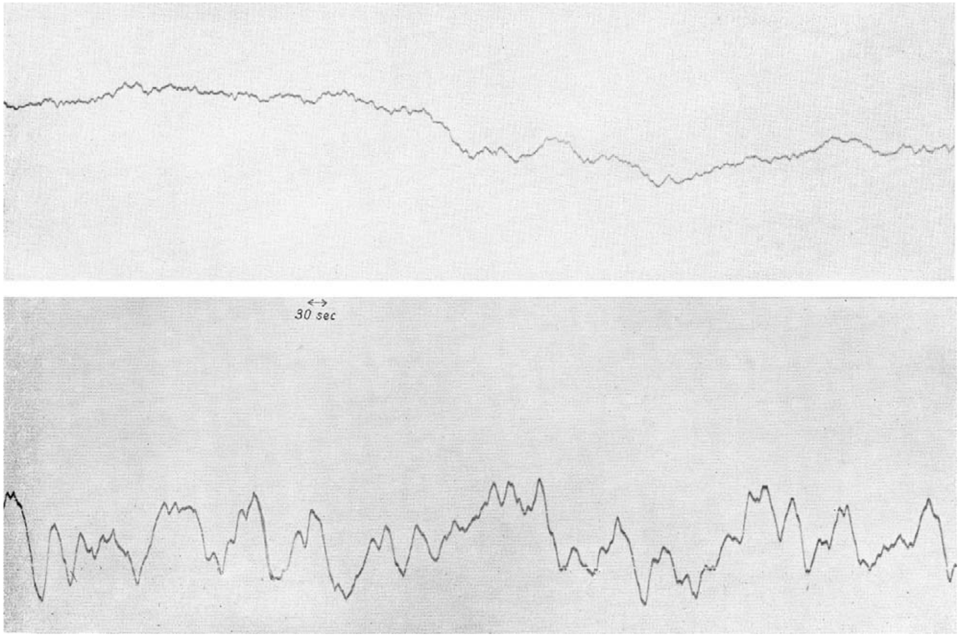

Fig. 1.1. The upper panel shows a recording of the movement of the mirror in the Kappler (1931) experiment over a period of about 30 minutes at atmospheric pressure (760 mm Hg). The bottom panel shows the same experiment at a pressure of 4×10^{-3} mm Hg. Both figures are from Kappler (1931). See the text for more detail. Published with permission of Annalen der Physik.

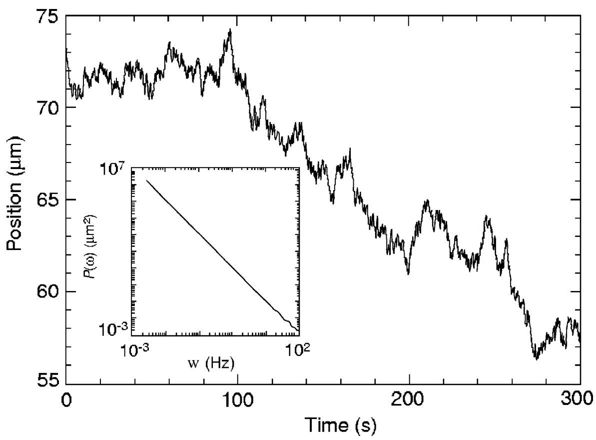

Fig. 1.2. The data shown here, taken from Gaspard et al. (1998), show the position of a 2.5 μm particle in water over a 300 second period with a sampling interval of $\frac{1}{60}$ sec (see Gaspard et al. (1998); Briggs et al. (2001) for the experimental details). The inset figure shows the power spectrum, which displays a typical decay (for Brownian motion) with ω^{-2}. Published with permission of Nature.

for the manifest irreversibility of our physical and biological world in the face of physical laws that fail to encompass irreversibility without the most incredulous of assumptions. In particular, it would clarify the foundations of irreversible statistical mechanics, e.g. the operation of the second law of thermodynamics (Dorfman, 1999; Gallavotti, 1999; Mackey, 1989, 1992; Schulman, 1997), and the implications of the second law for the physical and biological sciences.

In this paper, we have a rather more modest goal. We address a different facet of this chaotic hypothesis by studying how and when the characteristics of Brownian motion can be reproduced by deterministic systems. To motivate this,

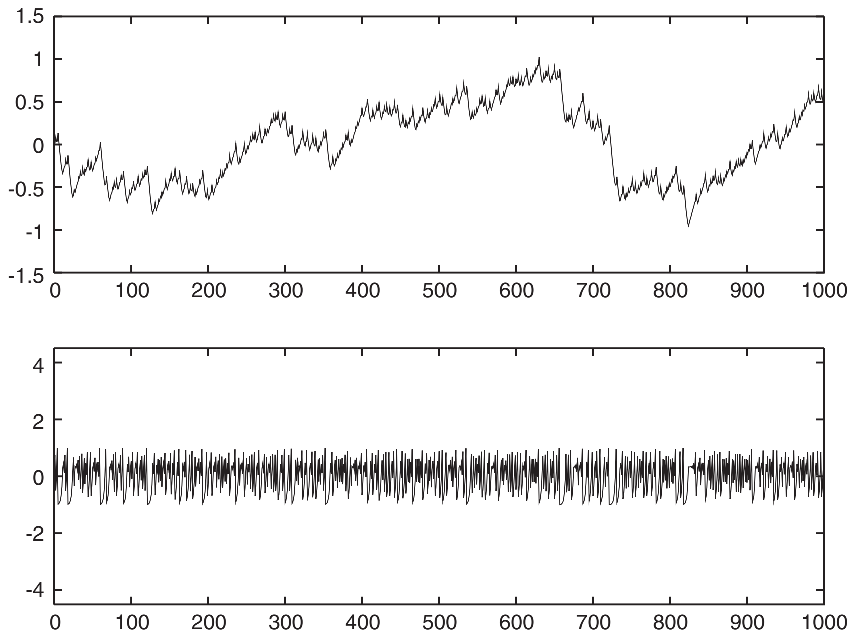

Fig. 1.3. The top panel shows the simulated position of a particle obeying Eqs. (1.1)–(1.4) using Eq. (3.28), while the bottom panel shows the velocity of the same particle computed with Eq. (3.21). The parameters used were: $\gamma = \Gamma/m = 10$, $\kappa = 1$, and $\tau = -\frac{1}{10}\ln(9 \times 10^{-4}) \simeq 0.932$ so $\lambda \equiv e^{-\gamma\tau} = 9 \times 10^{-4}$. The initial condition on the tent map given by Eq. (1.4) was $y_0 = 0.12562568$.

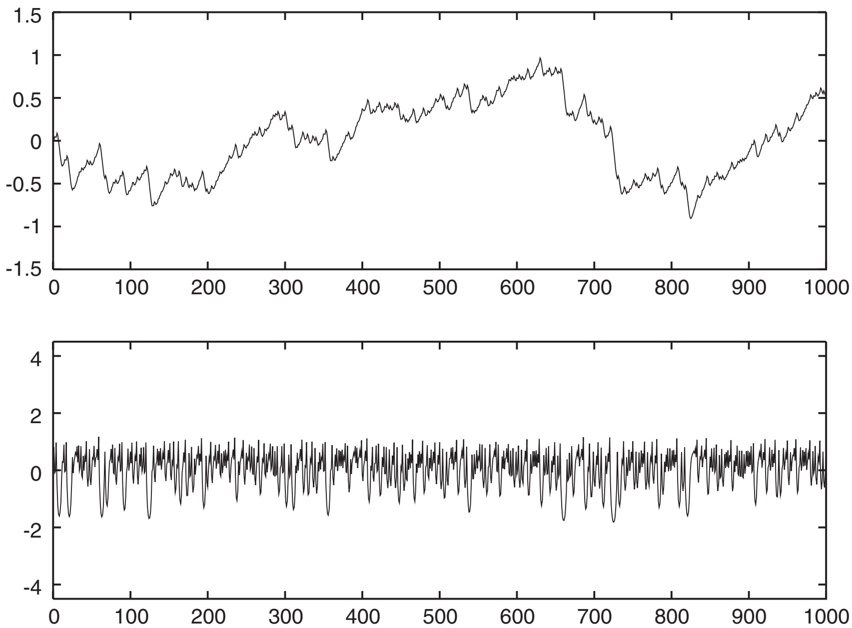

Fig. 1.4. As in Fig. 1.3 except that $\tau = \frac{1}{10}\ln(2) \simeq 0.069$ so $\lambda = \frac{1}{2}$.

in Figs. 1.3–1.6 we show the position (x) and velocity (v) of a particle of mass m whose dynamics are described by

$$\frac{dx}{dt} = v \ , \tag{1.1}$$

$$m\frac{dv}{dt} = -\Gamma v + \mathscr{F}(t) \ . \tag{1.2}$$

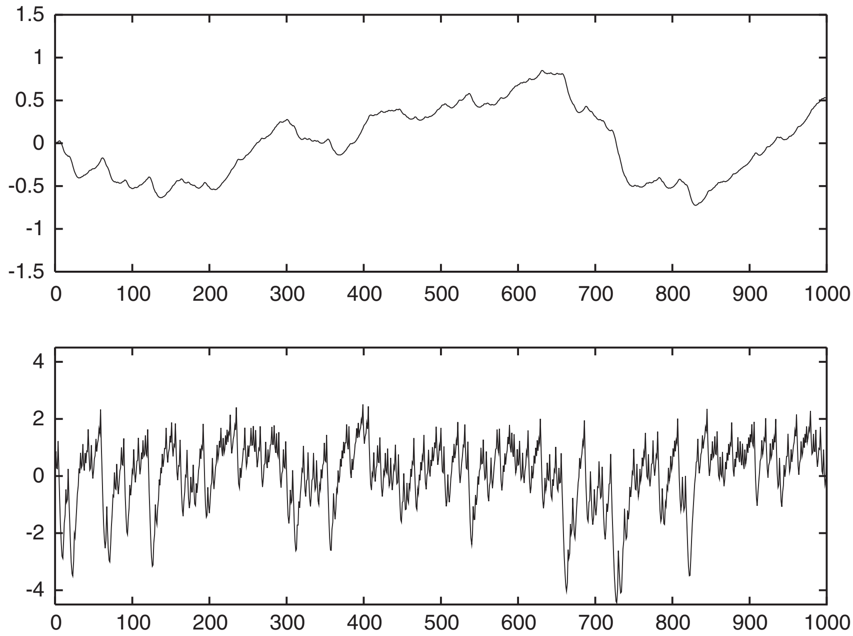

Fig. 1.5. As in Fig. 1.4 except that $\tau = -\frac{1}{10}\ln(0.9) \simeq 0.011$ so $\lambda = 0.9$.

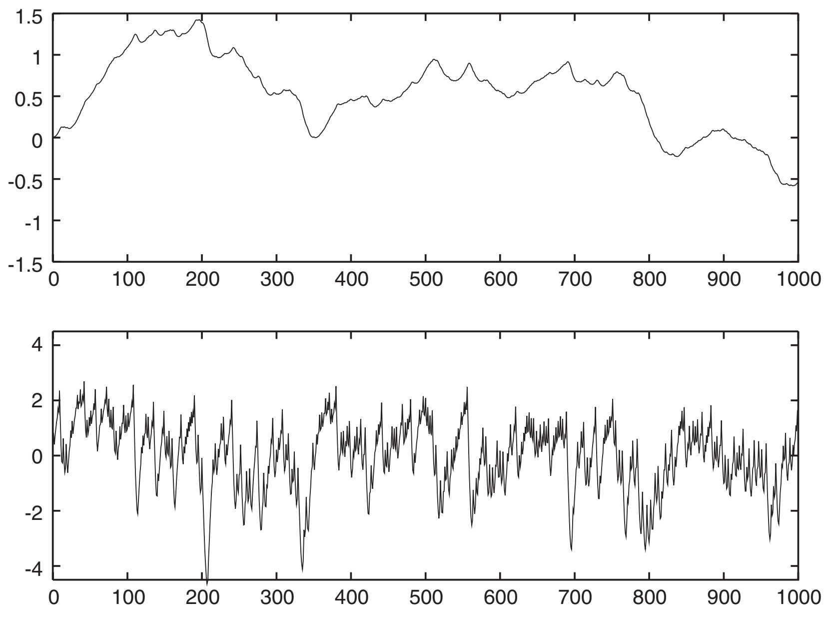

Fig. 1.6. As in Fig. 1.3, with the parameters of Fig. 1.5 and an initial condition on the tent map (1.4) of $y_0 = 0.1678549321$.

In Eqs. (1.1) and (1.2), we have taken $\mathscr{F}$ to be a fluctuating "force" consisting of a sequence of delta-function like impulses given by

$$\mathscr{F}(t) = m\kappa \sum_{n=0}^{\infty} \xi(t)\delta(t - n\tau) \ , \tag{1.3}$$

and ξ is a "highly chaotic" (exact, see Section 2) deterministic variable generated by $\xi_{t+\tau} = T(\xi_t)$ where T is the tent map on $[-1, 1]$ defined by

$$T(y) = \begin{cases} 2(y + \frac{1}{2}) & \text{for } y \in [-1, 0) , \\ 2(\frac{1}{2} - y) & \text{for } y \in [0, 1) . \end{cases} \tag{1.4}$$

When $\mathscr{F}(t)$ is a white noise, Eq. (1.2) is called the Langevin equation (Chandrasekhar, 1943).

In this paper we examine the behavior of systems described by equations like (1.1)–(1.4) and establish, analytically, the eventual limiting behavior of ensembles. In particular, *we address the question of how Brownian-like motion can arise from a purely deterministic dynamics*. We do this by studying the dynamics from a statistical, or ergodic theory, standpoint.

The outline of the paper is as follows. Section 2 gives required background and new material. This includes the definitions of a hierarchy of chaotic behaviors (ergodic, mixing, and exact) and a discussion of the evolution of densities under the action of transfer operators such as the Frobenius–Perron operator. We then go on to treat central limit theorems and functional central limit theorems. Section 3 returns to the specific problem that Eqs. (1.1)–(1.4) illustrate. We show when the particle velocity distribution may converge and the particle position may become asymptotically Gaussian if properly scaled. All of these considerations are for a more general class of maps than given by (1.4). In Section 3.3 we illustrate the application of the results from Section 3.2 using a specific chaotic map (the dyadic map) to act as a surrogate noise source. Section 4 considers the question when one can obtain Gaussian processes by studying appropriate scaling limits of the velocity and position variables, and the convergence of the velocity process to an Ornstein–Uhlenbeck process as the interval τ between chaotic perturbations approaches 0. The paper concludes with a brief discussion in Section 5. The Appendices collect and extend general central limit theorems from probability theory and weak convergence criteria that are used in the main results of the paper.

2. Ergodic theory and central limit theorems

We are going to examine the behavior illustrated in Section 1 using techniques from ergodic theory, and closely related concepts from probability theory, applied to the dynamics of semi-dynamical systems. Ergodic theory is ideally suited to this task as it treats the convergence properties of ensembles and their relation to trajectory dynamics. In this section we collect together the necessary machinery to do so. Much of this background material can be found in Lasota and Mackey (1994).

2.1. Density evolution operators

In looking at ensemble behavior, the natural framework is to look at the evolution of a density as the description of the temporal behavior of the ensemble. Thus we start by looking at the operators important for describing this density evolution.

Let $(Y_1, \mathscr{B}_1, \nu_1)$ and $(Y_2, \mathscr{B}_2, \nu_2)$ be two σ-finite measure spaces and let the transformation $T : Y_1 \to Y_2$ be measurable, i.e. $T^{-1}(\mathscr{B}_2) \subseteq \mathscr{B}_1$ where $T^{-1}(\mathscr{B}_2) = \{T^{-1}(A) : A \in \mathscr{B}_2\}$. Then we say that T is nonsingular (with respect to ν_1 and ν_2) if $\nu_1(T^{-1}(A)) = 0$ for all $A \in \mathscr{B}_2$ with $\nu_2(A) = 0$. Associated with the transformation T we have the Koopman operator U_T defined by

$$U_T g = g \circ T$$

for every measurable function $g : Y_2 \to \mathbb{R}$. We define the transfer operator $P_T : L^1(Y_1, \mathscr{B}_1, \nu_1) \to L^1(Y_2, \mathscr{B}_2, \nu_2)$ as follows. For any $f \in L^1(Y_1, \mathscr{B}_1, \nu_1)$, there is a unique element $P_T f$ in $L^1(Y_2, \mathscr{B}_2, \nu_2)$ such that

$$\int_A P_T f(y) \nu_2(\mathrm{d}y) = \int_{T^{-1}(A)} f(y) \nu_1(\mathrm{d}y) . \tag{2.1}$$

Eq. (2.1) simply gives an implicit relation between an initial density of states (f) and that density after the action of the map T, i.e. $P_T f$. The Koopman operator $U_T : L^\infty(Y_2, \mathscr{B}_2, \nu_2) \to L^\infty(Y_1, \mathscr{B}_1, \nu_1)$ and the transfer operator

P_T are adjoint, so

$$\int_{Y_2} g(y) P_T f(y) \nu_2(\mathrm{d}y) = \int_{Y_1} f(y) U_T g(y) \nu_1(\mathrm{d}y)$$

for $g \in L^\infty(Y_2, \mathscr{B}_2, \nu_2)$, $f \in L^1(Y_1, \mathscr{B}_1, \nu_1)$.

In some special cases Eq. (2.1) allows us to obtain an explicit form for P_T. Let $Y_2 = \mathbb{R}$, $\mathscr{B}_2 = \mathscr{B}(\mathbb{R})$ be the Borel σ-algebra, and ν_2 be the Lebesgue measure. Let Y_1 be an interval $[a, b]$ on the real line $\mathbb{R}$, $\mathscr{B}_1 = [a, b] \cap \mathscr{B}(\mathbb{R})$ and ν_1 be the Lebesgue measure restricted to $[a, b]$. We will simply write $L^1([a, b])$ when the underlying measure is the Lebesgue measure.

The transformation $T : [a, b] \to \mathbb{R}$ is called *piecewise monotonic* if

(i) there is a partition $a = a_0 < a_1 < \cdots < a_l = b$ of $[a, b]$ such that for each integer $i = 1, \ldots, l$ the restriction of T to (a_{i-1}, a_i) has a C^1 extension to $[a_{i-1}, a_i]$ and
(ii) $|T'(x)| > 0$ for $x \in (a_{i-1}, a_i)$, $i = 1, \ldots, l$.

If a transformation $T : [a, b] \to \mathbb{R}$ is piecewise monotonic, then for $f \in L^1([a, b])$ we have

$$P_T f(y) = \sum_{i=1}^{l} \frac{f(T_{(i)}^{-1}(y))}{|T'(T_{(i)}^{-1}(y))|} 1_{T((a_{i-1}, a_i))}(y) \ ,$$

where $T_{(i)}^{-1}$ is the inverse function for the restriction of T to (a_{i-1}, a_i). Note that we have equivalently

$$P_T f(y) = \sum_{x \in T^{-1}(\{y\})} \frac{f(x)}{|T'(x)|} \ . \tag{2.2}$$

Of course these formulas hold almost everywhere with respect to the Lebesgue measure.

The notion of a piecewise monotonic transformation on an interval can be extended to "piecewise smooth" transformations $T : Y \to Y$ with $Y \subset \mathbb{R}^k$. Therefore if T has, for example, finitely many inverse branches and the Jacobian matrix $DT(x)$ of T at x exists and $\det DT(x) \neq 0$ for almost every x, then the Frobenius–Perron operator is given by

$$P_T f(y) = \sum_{x \in T^{-1}(\{y\})} \frac{f(x)}{|\det DT(x)|}$$

for $f \in L^1(Y)$. If T is invertible then we have $P_T f(y) = f(T^{-1}(y)) |\det DT^{-1}(y)|$.

Let $(Y, \mathscr{B})$ be a measurable space and let $T : Y \to Y$ be a measurable transformation. The definition of the transfer operator for T depends on a given σ-finite measure on $\mathscr{B}$, which in turn gives rise to different operators for different underlying measures on $\mathscr{B}$. If ν is a probability measure on $\mathscr{B}$ which is invariant for T, i.e. $\nu(T^{-1}(A)) = \nu(A)$ for all $A \in \mathscr{B}$, then T is nonsingular. The transfer operator $P_T : L^1(Y, \mathscr{B}, \nu) \to L^1(Y, \mathscr{B}, \nu)$ is well defined and when we want to emphasize that the underlying measure ν in the transfer operator is invariant under the transformation T we will write $\mathscr{P}_{T,\nu}$. The Koopman operator U_T is also well defined for $f \in L^1(Y, \mathscr{B}, \nu)$ and is an isometry of $L^1(Y, \mathscr{B}, \nu)$ into $L^1(Y, \mathscr{B}, \nu)$, i.e. $\|U_T f\|_1 = \|f\|_1$ for all $f \in L^1(Y, \mathscr{B}, \nu)$. The following relation holds between the operators $U_T, \mathscr{P}_{T,\nu} : L^1(Y, \mathscr{B}, \nu) \to L^1(Y, \mathscr{B}, \nu)$

$$\mathscr{P}_{T,\nu} U_T f = f \quad \text{and} \quad U_T \mathscr{P}_{T,\nu} f = E(f | T^{-1}(\mathscr{B})) \tag{2.3}$$

for $f \in L^1(Y, \mathscr{B}, \nu)$, where $E(\cdot | T^{-1}(\mathscr{B})) : L^1(Y, \mathscr{B}, \nu) \to L^1(Y, T^{-1}(\mathscr{B}), \nu)$ denotes the operator of conditional expectation (see Appendix A). Both of these equations are based on the following change of variables (Billingsley, 1995, Theorem 16.13): $f \in L^1(Y, \mathscr{B}, \nu)$ if and only if $f \circ T \in L^1(Y, \mathscr{B}, \nu)$, in which case the following holds

$$\int_{T^{-1}(A)} f \circ T(y) \nu(\mathrm{d}y) = \int_A f(y) \nu(\mathrm{d}y), \quad A \in \mathscr{B} \ . \tag{2.4}$$

If the measure ν is finite, we have $L^p(Y, \mathscr{B}, \nu) \subset L^1(Y, \mathscr{B}, \nu)$ for $p \geqslant 1$. The operator $U_T : L^p(Y, \mathscr{B}, \nu) \to L^p(Y, \mathscr{B}, \nu)$ is also an isometry in this case. Note that if the conditional expectation operator $E(\cdot | T^{-1}(\mathscr{B})) : L^1(Y, \mathscr{B}, \nu) \to L^1(Y, \mathscr{B}, \nu)$ is restricted to $L^2(Y, \mathscr{B}, \nu)$, then this is the orthogonal projection of $L^2(Y, \mathscr{B}, \nu)$ onto $L^2(Y, T^{-1}(\mathscr{B}), \nu)$.

One can also consider any σ-finite measure m on $\mathscr{B}$ with respect to which T is nonsingular and the corresponding transfer operator $P_T : L^1(Y, \mathscr{B}, m) \to L^1(Y, \mathscr{B}, m)$. To be specific, let Y be a Borel subset of $\mathbb{R}^k$ with Lebesgue measure m and $\mathscr{B} = \mathscr{B}(Y)$ be the σ-algebra of Borel subsets of Y. Throughout this paper m will denote Lebesgue measure and $L^1(Y)$ will denote $L^1(Y, \mathscr{B}, m)$. The transfer operator $P_T : L^1(Y) \to L^1(Y)$ is usually known as the Frobenius–Perron operator. A measure v (on Y) is said to have a density g_* if $v(A) = \int_A g_*(y)\,\mathrm{d}y$ for all $A \in \mathscr{B}$, where $g_* \in L^1(Y)$ is nonnegative and $\int_Y g_*(y)\,\mathrm{d}y = 1$. A measure v is called *absolutely continuous* if it has a density. If the Frobenius–Perron operator P_T has a nontrivial fixed point in $L^1(Y)$, i.e. the equation $P_T f = f$ has a nonzero solution in $L^1(Y)$, then the transformation T has an absolutely continuous invariant measure v, its density g_* is a fixed point of P_T, and we call g_* an *invariant density* under the transformation T. The following relation holds between the operators P_T and $\mathscr{P}_{T,v}$

$$P_T(fg_*) = g_* \mathscr{P}_{T,v} f \quad \text{for } f \in L^1(Y, \mathscr{B}, v) \ . \tag{2.5}$$

In particular, if the density g_* is strictly positive, i.e. $g_*(y) > 0$ for almost every $y \in Y$, then the measures m and v are equivalent and we also have

$$P_T(f) = g_* \mathscr{P}_{T,v}\left(\frac{f}{g_*}\right) \quad \text{for } f \in L^1(Y) \ .$$

2.2. *Probabilistic and ergodic properties of density evolution*

Some of the most important concepts and results from ergodic theory are related to the convergence properties of sequences of densities (starting from an initial density) under the action of a transfer operator associated with the underlying dynamics. This section considers a hierarchy of convergence behaviors and illustrates them through concrete examples.

Let $(Y, \mathscr{B}, v)$ be a normalized measure space and let $T : Y \to Y$ be a measurable transformation preserving the measure v. We can discuss the ergodic properties of T in terms of the convergence behavior of its transfer operator $\mathscr{P}_{T,v} : L^1(Y, \mathscr{B}, v) \to L^1(Y, \mathscr{B}, v)$. To this end, we note that the transformation T is defined to be

(i) *Ergodic* (with respect to v) if and only if every invariant set $A \in \mathscr{B}$ is such that $v(A) = 0$ or $v(Y \backslash A) = 0$. This is equivalent to: T is ergodic (with respect to v) if and only if for each $f \in L^1(Y, \mathscr{B}, v)$ the sequence $(1/n)\sum_{k=0}^{n-1} \mathscr{P}_{T,v}^k f$ is weakly convergent in $L^1(Y, \mathscr{B}, v)$ to $\int f(y)v(\mathrm{d}y)$, i.e. for all $g \in L^\infty(Y, \mathscr{B}, v)$

$$\lim_{n\to\infty} \int \frac{1}{n} \sum_{k=0}^{n-1} \mathscr{P}_{T,v}^k f(y) g(y) v(\mathrm{d}y) = \int f(y) v(\mathrm{d}y) \int g(y) v(\mathrm{d}y) \ .$$

(ii) *Mixing* (with respect to v) if and only if

$$\lim_{n\to\infty} v(A \cap T^{-n}(B)) = v(A)v(B) \quad \text{for } A, B \in \mathscr{B} \ .$$

Mixing is equivalent to: For each $f \in L^1(Y, \mathscr{B}, v)$ the sequence $P_T^n f$ is weakly convergent in $L^1(Y, \mathscr{B}, v)$ to $\int f(y)v(\mathrm{d}y)$, i.e.

$$\lim_{n\to\infty} \int \mathscr{P}_{T,v}^n f(y) g(y) v(\mathrm{d}y) = \int f(y) v(\mathrm{d}y) \int g(y) v(\mathrm{d}y) \quad \text{for } g \in L^\infty(Y, \mathscr{B}, v) \ .$$

(iii) *Exact* (with respect to v) if and only if

$$\lim_{n\to\infty} v(T^n(A)) = 1 \quad \text{for } A \in \mathscr{B} \quad \text{with} \quad T(A) \in \mathscr{B}, \ v(A) > 0 \ .$$

Exactness is equivalent to: For each $f \in L^1(Y, \mathscr{B}, v)$ the sequence $P_{T,v}^n f$ is strongly convergent in $L^1(Y, \mathscr{B}, v)$ to $\int f(y)v(\mathrm{d}y)$, i.e.

$$\lim_{n\to\infty} \int |\mathscr{P}_{T,v}^n f(y) - \int f(y) v(\mathrm{d}y)| v(\mathrm{d}y) = 0 \ .$$

The characterization of the ergodic properties of transformations through the properties of the evolution of densities requires that we know an invariant measure ν for T. Examples of ergodic, mixing, and exact transformations are given in the following.

Example 2.1. The transformation on [0, 1]

$$T(y) = y + \phi \quad \bmod 1 ,$$

known as rotation on the circle, is ergodic with respect to the Lebesgue measure when ϕ is irrational. The associated Frobenius–Perron operator is given by

$$P_T f(y) = f(y - \phi) .$$

Example 2.2. The baker map on $[0, 1] \times [0, 1]$

$$T(y, z) = \begin{cases} (2y, \frac{1}{2}z) & 0 \leqslant z \leqslant \frac{1}{2} , \\ (2y - 1, \frac{1}{2} + \frac{1}{2}z) & \frac{1}{2} < z \leqslant 1 \end{cases}$$

is mixing with respect to the Lebesgue measure. The Frobenius–Perron operator is given by

$$P_T f(y, z) = \begin{cases} f(\frac{1}{2}y, 2z) & 0 \leqslant z \leqslant \frac{1}{2} , \\ f(\frac{1}{2} + \frac{1}{2}y, 2z - 1) & \frac{1}{2} < z \leqslant 1 . \end{cases}$$

Example 2.3. The tent map on $[-1, 1]$ defined by Eq. (1.4) is exact with respect to the Lebesgue measure, and has a Frobenius–Perron operator given by

$$P_T f(y) = \tfrac{1}{2}[f(\tfrac{1}{2}y - \tfrac{1}{2}) + f(\tfrac{1}{2} - \tfrac{1}{2}y)] .$$

Example 2.4. A class of piecewise linear transformations on [0, 1] are given by

$$T_N(y) = \begin{cases} N\left(y - \frac{2n}{N}\right) & \text{for } y \in \left[\frac{2n}{N}, \frac{2n+1}{N}\right) \\ N\left(\frac{2n+2}{N} - y\right) & \text{for } y \in \left[\frac{2n+1}{N}, \frac{2n+2}{N}\right) , \end{cases} \tag{2.6}$$

where $n = 0, 1, \ldots, [(N-1)/2]$ and $[z]$ denotes the integer part of z. For $N \geqslant 2$, these piecewise linear maps generalize the tent map, are exact with respect to the Lebesgue measure, and have the invariant density

$$g_*(y) = 1_{[0,1]}(y) . \tag{2.7}$$

Example 2.5. The Chebyshev maps (Adler and Rivlin, 1964) on $[-1, 1]$ studied by Beck and Roepstorff (1987), Beck (1996) and Hilgers and Beck (1999) are given by

$$S_N(y) = \cos(N \arccos y), \quad N = 0, 1, \ldots \tag{2.8}$$

with $S_0(y) = 1$ and $S_1(y) = y$. They are conjugate to the transformation of Example 2.4, and satisfy the recurrence relation $S_{N+1}(y) = 2yS_N(y) - S_{N-1}(y)$. For $N \geqslant 2$ they are exact with respect to the measure with the density

$$g_*(y) = \frac{1}{\pi\sqrt{1 - y^2}} .$$

For $N = 2$ the Frobenius–Perron operator is given by

$$P_{S_2} f(y) = \frac{1}{2\sqrt{2y + 2}} \left[f\left(\sqrt{\frac{1}{2}y + \frac{1}{2}}\right) + f\left(-\sqrt{\frac{1}{2}y + \frac{1}{2}}\right) \right]$$

and the transfer operator by

$$\mathscr{P}_{S_2,\nu} f(y) = \frac{1}{2} \left[f\left(\sqrt{\frac{1}{2}y + \frac{1}{2}}\right) + f\left(-\sqrt{\frac{1}{2}y + \frac{1}{2}}\right) \right] .$$

Given a transfer operator under which the Lebesgue measure is invariant it is rather easy to construct a second transformation (and its associated transfer operator) under which a measure ν is invariant, and vice-versa, through conjugation. This construction of a transfer operator for a conjugate map is presented in the next theorem.

Theorem 2.6 (Lasota and Mackey (1994, Theorem 6.5.2)). *Let $T : [0, 1] \to [0, 1]$ be a measurable and nonsingular (with respect to the Lebesgue measure) transformation. Let $\nu : \mathscr{B}([a, b]) \to [0, \infty)$ be a probability measure with a strictly positive density g_*, that is $g_*(y) > 0$ for almost every y. Let a second transformation $S : [a, b] \to [a, b]$ be given by $S = G^{-1} \circ T \circ G$, where*

$$G(x) = \int_a^x g_*(y)\,\mathrm{d}y, \quad a \leqslant x \leqslant b \ .$$

Then the transfer operator $\mathscr{P}_{S,\nu}$ is given by

$$\mathscr{P}_{S,\nu} f = U_G P_T U_{G^{-1}} f \quad \textit{for } f \in L^1([a, b], \mathscr{B}([a, b]), \nu) \ , \tag{2.9}$$

where $U_G, U_{G^{-1}}$ are Koopman operators for G and G^{-1}, respectively, and P_T is the Frobenius–Perron operator for T. As a consequence, ν is invariant for S if and only if the Lebesgue measure is invariant for T.

Example 2.7. The dyadic map on $[-1, 1]$ is given by

$$T(y) = \begin{cases} 2y + 1, & y \in [-1, 0] \\ 2y - 1, & y \in (0, 1] \ , \end{cases} \tag{2.10}$$

and has the uniform invariant density

$$g_*(y) = \tfrac{1}{2} 1_{[-1,1]}(y) \ .$$

Like the tent map, it is exact with respect to the normalized Lebesgue measure on $[-1, 1]$. It has a Frobenius–Perron operator given by

$$P_T f(y) = \tfrac{1}{2}[f(\tfrac{1}{2}y - \tfrac{1}{2}) + f(\tfrac{1}{2}y + \tfrac{1}{2})] \ .$$

Example 2.8. Alexander and Yorke (1984) defined a generalized baker transformation (also known as a fat/skiny baker transformation) $S_\beta : [-1, 1] \times [-1, 1] \to [-1, 1] \times [-1, 1]$ by

$$S_\beta(x, y) = (\beta x + (1 - \beta)h(y), T(y))$$

where $0 < \beta < 1$, T is the dyadic map on $[-1, 1]$, and

$$h(y) = \begin{cases} 1, & y \geqslant 0 \ , \\ -1, & y < 0 \ . \end{cases}$$

For every $\beta \in (0, 1)$ the transformation S_β has an invariant probability measure on $[-1, 1] \times [-1, 1]$ and is mixing. The invariant measure is the product of a so-called infinitely convolved Bernoulli measure (see Section 3.3) and the normalized Lebesgue measure on $[-1, 1]$. If $\beta = \frac{1}{2}$, the transformation S_β is conjugated through a linear transform of the plane to the baker map of Example 2.2. If $\beta < \frac{1}{2}$, the transformation S_β does not have an invariant density (with respect to the planar Lebesgue measure).

Example 2.9. The continued fraction map

$$T(y) = \frac{1}{y} \mod 1 \quad y \in (0, 1] \tag{2.11}$$

has an invariant density

$$g_*(y) = \frac{1}{(1 + y)\ln 2} \tag{2.12}$$

and is exact. The Frobenius–Perron operator is given by

$$P_T f(y) = \sum_{k=1}^{\infty} \frac{1}{(y+k)^2} f\left(\frac{1}{y+k}\right)$$

and the transfer operator by

$$\mathscr{P}_{T,v} f(y) = \sum_{k=1}^{\infty} \frac{y+1}{(y+k)(y+k+1)} f\left(\frac{1}{y+k}\right) .$$

Example 2.10. The quadratic map is given by

$$T_\beta(y) = 1 - \beta y^2, \quad y \in [-1, 1]$$

where $0 < \beta \leqslant 2$. It is known that there exists a positive Lebesgue measure set of parameter values β such that the map T_β has an absolutely continuous (with respect to Lebesgue measure) invariant measure v_β (Jakobson, 1981; Benedics and Carleson, 1985). Let $\alpha > 0$ be a very small number and let

$$\Delta_\epsilon = \{\beta \in [2-\epsilon, 2] : |T_\beta^n(0)| \geqslant \mathrm{e}^{-\alpha n} \text{ and } |(T_\beta^n)'(T_\beta(0))| \geqslant (1.9)^n \;\; \forall n \geqslant 0\}$$

for $\epsilon > 0$. Young (1992) proved that for sufficiently small ϵ and for every $\beta \in \Delta_\epsilon$ the transformation T_β is exact with respect to v_β and this measure is supported on $[T_\beta^2(0), T_\beta(0)]$.

Example 2.11. The Manneville–Pomeau map $T_\beta : [0, 1] \to [0, 1]$ is given by

$$T_\beta(y) = y + y^{1+\beta} \quad \mathrm{mod}\, 1 ,$$

where $\beta \in (0, 1)$. The map has an absolutely continuous invariant probability measure v_β with density satisfying

$$\frac{c_1}{y^\beta} \leqslant g_*(y) \leqslant \frac{c_2}{y^\beta}$$

for some constants $c_2 \geqslant c_1 > 0$ (cf. Thaler, 1980), and is exact.

Finally, we discuss the notion of Sinai–Ruelle–Bowen measure or SRB measure of T which was first conceived in the setting of Axiom A diffeomorphisms on compact Riemannian manifolds. This notion varies from author to author (Alexander and Yorke, 1984; Eckmann and Ruelle, 1985; Tsujii, 1996; Young, 2002; Hunt et al., 2002). Let (Y, ρ) be a compact metric space with a reference measure m, e.g. a compact subset of $\mathbb{R}^k$ and m the Lebesgue or a compact Riemannian manifold and m the Riemannian measure on Y. If $T : Y \to Y$ is a continuous map, then by the Bogolyubov–Krylov theorem there always exists at least one invariant probability measure for T. When there is more than one measure, the question arises which invariant measure is "interesting", and has led to attempts to give a good definition of "physically" relevant invariant measures. Though this seems to be a rather vague and poorly defined concept, loosely speaking one would expect that one criteria for a physically relevant invariant measure would be whether or not it was observable in the context of some laboratory or numerical experiment.

An invariant measure v for T is called a natural or physical measure if there is a positive Lebesgue measure set $Y_0 \subset Y$ such that for every $y \in Y_0$ and for every continuous observable $f : Y \to \mathbb{R}$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i(y)) = \int f(z) v(\mathrm{d}z) . \tag{2.13}$$

In other words the average of f along the trajectory of T starting in Y_0 is equal to the average of f over the space Y.

Observe that if v is ergodic then from the individual Birkhoff ergodic theorem it follows that for every $f \in L^1(Y, \mathscr{B}, v)$ Condition 2.13 holds for almost all $y \in Y$, i.e. except for a subset of Y of (v) measure zero. Thus, if T has an ergodic absolutely continuous invariant measure v with density g_* then every continuous function f is integrable with respect v and Condition 2.13 holds for almost every point from the set $\{y \in Y : g_*(y) > 0\}$. Therefore such v is a physical measure for T. Not only absolutely continuous measures are physical measures. Consider, for example, the generalized baker transformation S_β of Example 2.8. Alexander and Yorke (1984) showed that there is a unique physical measure v_β for each $\beta \in (0, 1)$. This measure is mixing and hence ergodic. Although for $\beta > \frac{1}{2}$ the transformation expands areas, the measure v_β might not be absolutely continuous for certain values of the parameter β (e.g. $\beta = \frac{-1+\sqrt{5}}{2}$) in which case the Birkhoff ergodic theorem only implies Condition 2.13 on a zero Lebesgue measure set. Therefore a completely different argument was needed in the proof of the physical property.

In the context of smooth invertible maps having an Axiom A attractor the existence of a unique physical measure on the attractor was first proved for Anosov diffeomorphisms by Sinai (1972) and later generalized by Ruelle (1976) and Bowen (1975). Roughly speaking, these are maps having uniformly expanding and contracting directions and their physical invariant measures have densities with respect to the Lebesgue measure in the expanding directions (being usually singular in the contracting directions). This property led then to the characterization of a Sinai–Ruelle–Bowen measure. In a recent attempt to go beyond maps having an Axiom A attractor, Young (2002) additionally requires that T has a positive Lyapunow exponent almost everywhere. The precise definition strongly relies on the smoothness and invertibility of the map T. Note that the generalized baker transformation S_β has Lyapunov exponents equal to $\ln 2$ and $\ln \beta$ and the measure v_β is absolutely continuous along all vertical directions.

2.3. Brownian motion from deterministic perturbations

We are now ready to start developing the machinery to examine the question posed in the Introduction, i.e. how Brownian-like motion can arise from a purely deterministic dynamics. We first, in Section 2.3.1, give some background material and review the content of the central limit theorem and introduce a stronger result from Billingsley (1968) known as a functional central limit theorem. Section 2.3.2 then goes on to develop a number of essential results using the central limit theorem and functional central limit theorem for invertible maps.

2.3.1. Preliminaries

We follow the terminology of Billingsley (1968). If ζ is a measurable mapping from a probability space $(\Omega, \mathscr{F}, \Pr)$ into a measurable space $(Z, \mathscr{A})$, we call ζ a Z-valued random variable. The distribution of ζ is the normalized measure $\mu = \Pr \circ \zeta^{-1}$ on $(Z, \mathscr{A})$, i.e.

$$\mu(A) = \Pr(\zeta^{-1}(A)) = \Pr\{\omega : \xi(\omega) \in A\} = \Pr\{\xi \in A\}, \quad A \in \mathscr{A} .$$

The random variables ζ and ξ are, by definition, (statistically) independent if for all measurable A and B

$$\Pr\{\zeta \in A, \xi \in B\} = \Pr\{\zeta \in A\} \Pr\{\xi \in B\} ,$$

i.e. the distribution of the pair (ζ, ξ) is the product of the distribution of ζ with that of ξ.

Let (Z, ρ) be a metric space and $\mathscr{B}(Z)$ be the σ-algebra of Borel subsets of Z. A sequence (μ_n) of normalized measures on $(Z, \mathscr{B}(Z))$ is said to converge weakly to a normalized measure μ if

$$\lim_{n\to\infty} \int_Z f(z)\mu_n(\mathrm{d}z) = \int_Z f(z)\mu(\mathrm{d}z)$$

for every continuous bounded function $f : Z \to \mathbb{R}$. As an example, in Eq. (2.13) the measures $(1/n)\sum_{i=0}^{n-1} \delta_{T^i(y)}$ are weakly convergent to v for each $y \in Y_0$.

Note that the integrals $\int_Z f(z)\mu(\mathrm{d}z)$ completely determine μ, thus the sequence (μ_n) cannot converge weakly to two different limits. Note also that weak convergence depends only on the topology of Z, not on the specific metric that generates it; thus two equivalent metrics give rise to the same notion of weak convergence. If we have a family $\{\mu_\tau : \tau \geqslant 0\}$ of normalized measures instead of a sequence, we can also speak of weak convergence of μ_τ to μ when τ goes to ∞ or some finite value τ_0 in a continuous manner. This then means that μ_τ converges weakly to μ as $\tau \to \tau_0$ if and only if μ_{τ_n} converges weakly to μ for each sequence (τ_n) such that $\tau_n \to \tau_0$ as $n \to \infty$.

If $Z = \mathbb{R}^k$ and F and F_n are, respectively, the distribution functions of μ and μ_n, i.e. $F(z) = \mu\{y : y_i \leqslant z_i, i \leqslant k\}$ and $F_n(z) = \mu_n\{y : y_i \leqslant z_i, i \leqslant k\}$ for $z = (z_1, \ldots, z_k) \in \mathbb{R}^k$, then (μ_n) converges weakly to μ if and only if

$$\lim_{n\to\infty} F_n(z) = F(z) \quad \text{at continuity points } z \text{ of } F .$$

The characteristic function φ_μ of a normalized measure μ on $\mathbb{R}^k$ is defined by

$$\varphi_\mu(r) = \int \exp(\mathrm{i}\langle r, z\rangle)\mu(\mathrm{d}z) ,$$

where $\mathrm{i} = \sqrt{-1}$ and $\langle r, z\rangle = \sum_{j=1}^{k} r_j z_j$ denotes the inner product in $\mathbb{R}^k$. The continuity theorem (Billingsley, 1968, Theorem 7.6) gives us the following: (μ_n) converges weakly to μ if and only if

$$\lim_{n\to\infty} \varphi_{\mu_n}(r) = \varphi_\mu(r) \quad \text{for each } r \in \mathbb{R}^k . \tag{2.14}$$

A sequence ζ_n of Z-valued random variables converges in distribution, or weakly, to a normalized measure μ on $(Z, \mathscr{B}(Z))$, if the corresponding distributions of ζ_n converge weakly to μ. This is denoted by

$$\zeta_n \to^{\mathrm{d}} \mu .$$

If μ is the distribution of a random variable ζ, we write $\zeta_n \to^{\mathrm{d}} \zeta$. Note that the underlying probability spaces for the random variables $\zeta, \zeta_1, \zeta_2 \ldots$ may be all distinct.

A sequence ζ_n of Z-valued random variables converges in probability to a Z-valued random variable ζ if

$$\lim_{n\to\infty} \Pr(\rho(\zeta_n, \zeta) > \epsilon) = 0 \quad \text{for all } \epsilon > 0 \tag{2.15}$$

and almost surely (a.s.) to ζ if

$$\lim_{n\to\infty} \zeta_n(\omega) = \zeta(\omega) \quad \text{for almost every } \omega .$$

This is denoted, respectively, by

$$\zeta_n \to^{P} \zeta \quad \text{and} \quad \zeta_n \to \zeta \text{ a.s.}$$

Here all the random variables are defined on the same probability space. Note that if Condition 2.15 holds then

$$\lim_{n\to\infty} \Pr_0(\rho(\zeta_n, \zeta) > \epsilon) = 0 \quad \text{for all } \epsilon > 0$$

for every probability measure $\Pr_0$ on $(\Omega, \mathscr{F})$ which is absolutely continuous with respect to Pr. In other words convergence in probability is preserved by an absolutely continuous change of measure. We will also frequently use the following result from Billingsley (1968, Theorem 4.1): If (Z, ρ) is a separable metric space, and

$$\text{if} \quad \tilde{\zeta}_n \to^{\mathrm{d}} \zeta \quad \text{and} \quad \rho(\zeta_n, \tilde{\zeta}_n) \to^{P} 0, \quad \text{then } \zeta_n \to^{\mathrm{d}} \zeta . \tag{2.16}$$

We will write $\mathrm{N}(0, \sigma^2)$ for either a real-valued random variable which is Gaussian distributed with mean 0 and variance σ^2, or the measure on $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ with density

$$\frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{x^2}{2\sigma^2}\right) . \tag{2.17}$$

The characteristic function of $\mathrm{N}(0, \sigma^2)$ is $\phi(r) = \exp(-\frac{\sigma^2}{2} r^2)$, $r \in \mathbb{R}$. Since $\sigma\mathrm{N}(0, 1) = \mathrm{N}(0, \sigma^2)$ when $\sigma > 0$, we can always write $\sigma\mathrm{N}(0, 1)$ for $\sigma \geqslant 0$, which in the case $\sigma = 0$ reduces to the point measure δ_0 at 0.

Let $(\zeta_j)_{j\geqslant 1}$ be a sequence of real-valued zero-mean random variables with finite variance. If there is $\sigma > 0$ such that

$$\frac{\sum_{j=1}^{n} \zeta_j}{\sqrt{n}} \to^{\mathrm{d}} \sigma\mathrm{N}(0, 1) , \tag{2.18}$$

then $(\zeta_j)_{j\geqslant 1}$ is said to satisfy the central limit theorem (CLT). Note that if the random variables ζ_j are defined on the same probability space $(\Omega, \mathscr{F}, \Pr)$, we have the equivalent formulation of (2.18) in the case of $\sigma > 0$

$$\lim_{n\to\infty} \Pr\left\{\omega \in \Omega : \frac{\sum_{j=1}^{n} \zeta_j(\omega)}{\sigma\sqrt{n}} \leqslant z\right\} = \Phi(z), \quad z \in \mathbb{R} ,$$

where

$$\Phi(z)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z}\exp\left(-\frac{1}{2}t^2\right)\mathrm{d}t \tag{2.19}$$

is the standard Gaussian distribution function. In the case of $\sigma=0$

$$\frac{\sum_{j=1}^{n}\zeta_j}{\sqrt{n}}\to^{P} 0\ .$$

Let $\{w(t), t\in[0,\infty)\}$ be a standard Wiener process (Brownian motion), i.e. $\{w(t), t\in[0,\infty)\}$ is a family of real-valued random variables on some probability space $(\Omega,\mathscr{F},\Pr)$ satisfying the following properties:

(i) the process starts at zero: $w(0)=0$ a.e.;
(ii) for $0\leqslant t_1<t_2$ the random variable $w(t_2)-w(t_1)$ is Gaussian distributed with mean 0 and variance t_2-t_1;
(iii) for times $t_1<t_2<\cdots<t_n$ the increments $w(t_2)-w(t_1),\ldots,w(t_n)-w(t_{n-1})$ are independent random variables.

The existence of the Wiener process $\{w(t), t\in[0,1]\}$ is equivalent to the existence of the Wiener measure W on the space $C[0,1]$ of continuous functions on $[0,1]$ with uniform convergence, in a topology which makes $C[0,1]$ a complete separable metric space. Then, simply, W is the distribution of a random variable $W:\Omega\to C[0,1]$ defined by $W(\omega): t\mapsto w(t)(\omega)$.

Let $D[0,1]$ be the space of right continuous real valued functions on $[0,1]$ with left-hand limits. We endow $D[0,1]$ with the Skorohod topology which is defined by the metric

$$\rho_S(\psi,\widetilde{\psi})=\inf_{s\in S}\left(\sup_{t\in[0,1]}|\psi(t)-\widetilde{\psi}(s(t))|+\sup_{t\in[0,1]}|t-s(t)|\right),\quad \psi,\widetilde{\psi}\in D[0,1]\ ,$$

where S is the family of strictly increasing, continuous mappings s of $[0,1]$ onto itself such that $s(0)=0$ and $s(1)=1$ (Billingsley, 1968, Section 14). The metric space $(D[0,1],\rho_S)$ is separable and is not complete, but there is an equivalent metric on $D[0,1]$ which turns $D[0,1]$ with the Skorohod topology into a complete separable metric space. Since the Skorohod topology and the uniform topology on $C[0,1]$ coincide, W can be considered as a measure on $D[0,1]$.

A stronger result than the CLT is a weak invariance principle, also called a functional central limit theorem (FCLT). Let $(\zeta_j)_{j\geqslant 1}$ be a sequence of real-valued zero-mean random variables with finite variance. Let $\sigma>0$ and define the process $\{\psi_n(t), t\in[0,1]\}$ by

$$\psi_n(t)=\frac{1}{\sigma\sqrt{n}}\sum_{j=1}^{[nt]}\zeta_j\quad \text{for } t\in[0,1]\ ,$$

(where the sum from 1 to 0 is set equal to 0). Note that ψ_n is a right continuous step function, a random variable of $D[0,1]$ and $\psi_n(0)=0$. If

$$\psi_n\to^{\mathrm{d}} w$$

(here the convergence in distribution is in $D[0,1]$), then $(\zeta_j)_{j\geqslant 1}$ is said to satisfy the FCLT.

If for every $k\geqslant 1$ and every vector $(t_1,\ldots,t_k)\in[0,1]^k$ with $t_1<\cdots<t_k$ the joint distribution of the vector $(\psi_n(t_1),\ldots,\psi_n(t_k))$ converges to the joint distribution of $(w(t_1),\ldots,w(t_k))$, then we say that the finite dimensional distributions of ψ_n converge to those of w. For one dimensional distributions this convergence is equivalent to the central limit theorem.

The convergence of all finite-dimensional distributions of ψ_n to those of w is not sufficient to conclude that $\psi_n\to^{\mathrm{d}} w$ in $D[0,1]$. According to Theorems 15.1 and 15.5 of Billingsley (1968) if, additionally, for each positive ϵ

$$\lim_{\delta\to 0}\limsup_{n\to\infty}\Pr\left(\sup_{|t-s|\leqslant\delta}|\psi_n(s)-\psi_n(t)|>\epsilon\right)=0\ , \tag{2.20}$$

then ψ_n converges in distribution to the Wiener process w.

The term "functional central limit theorem" comes from the mapping theorem (Billingsley, 1968, Theorem 5.1), according to which for any functional $f : D[0,1] \to \mathbb{R}$, measurable and continuous on a set of Wiener measure 1, the distribution of $f(\psi_n)$ converges weakly to the distribution of $f(w)$. This applies in particular to the functional $f(\psi) = \sup_{0 \leqslant s \leqslant 1} \psi(s)$. Instead of real-valued functionals one can also consider mappings with values in a metric space. For example, this theorem applies for any $f : D[0,1] \to D[0,1]$ of the form $f(\phi)(t) = \sup_{s \leqslant t} \phi(s)$ or $f(\phi)(t) = \int_0^t \phi(s)\,\mathrm{d}s$.

2.3.2. *CLT and FCLT for maps with zero auto-correlation*

Obtaining Brownian-like motion is intimately connected with central limit theorems (CLT) for maps and various invariance principles. We start to investigate this question in the present section.

Many CLT results and invariance principles for maps have been proved, see e.g. the survey of Denker (1989). These results extend back over some decades, including contributions by Ratner (1973), Boyarsky and Scarowsky (1979), Wong (1979), Keller (1980), Jabłoński and Malczak (1983a), Jabłoński (1991), Liverani (1996) and Viana (1997).

First, however, remember that if we have a time series $y(j)$ and a bounded integrable function $h : X \to R$, then the auto-correlation of h is defined as

$$R_h(n) = \lim_{N\to\infty} \frac{1}{N} \sum_{j=0}^{N-1} h(y(j+n))h(y(j)) \ .$$

If the time series is generated by a measurable transformation $T : Y \to Y$ operating on a normalized measure space $(Y, \mathscr{B}, v)$, and if v is an invariant measure under T and T is ergodic, then we can rewrite the auto-correlation as

$$R_h(n) = \lim_{N\to\infty} \frac{1}{N} \sum_{j=0}^{N-1} h(y(j+n))h(y(j)) = \int h(y)h(T^n(y))v(\mathrm{d}y) \ .$$

The average $\langle h \rangle$ is just

$$\langle h \rangle = \lim_{N\to\infty} \frac{1}{N} \sum_{j=0}^{N-1} h(y(j)) = \int h(y)v(\mathrm{d}y) \ .$$

In this section we give CLT and FCLT results for systems in which $R_h(n) = 0$ for $n \geqslant 1$. This restriction on the auto-correlation holds for a number of the examples we discuss, e.g. the tent map of Example 2.3, the Chebyshev maps of Example 2.5, and the quadratic map of Example 2.10 with $\beta = 2$. Note, however, that the dyadic map introduced in Example 2.7 and used extensively as an example does not satisfy this restriction.

Let $(Y, \mathscr{B}, v)$ be a normalized measure space, and $T : Y \to Y$ be a measurable transformation such that v is T-invariant. $(Y, \mathscr{B}, v)$ will serve as our probability space $(\Omega, \mathscr{F}, \Pr)$. Let $h \in L^2(Y, \mathscr{B}, v)$ be such that $\int h(y)v(\mathrm{d}y) = 0$. The random variables $\zeta_j = h \circ T^{j-1}$, $j \geqslant 1$ are real-valued, have zero-mean and finite variance equal to $\|h\|_2^2 = \int h^2(y)v(\mathrm{d}y)$. Thus the terminology from Section 2.3.1 applies.

The explicit formulae for the Frobenius–Perron operators in Section 2.2 show that the equations $P_T h = 0$ or $\mathscr{P}_{T,v}h = 0$ can be easily solved. For instance, in Example 2.3, every function h which is odd is a solution of these equations. In particular, considering h with $\mathscr{P}_{T,v}h = 0$ turns out to be very fruitful. Statistical properties of the sequence $(h \circ T^j)_{j \geqslant 0}$ when $\mathscr{P}_{T,v}h = 0$ are summarized in the following

Theorem 2.12 (*CLT*). *Let $(Y, \mathscr{B}, v)$ be a normalized measure space and $T : Y \to Y$ be ergodic with respect to v. If $h \in L^2(Y, \mathscr{B}, v)$ is such that $\mathscr{P}_{T,v}h = 0$, then*

(i) $\int h(y)v(\mathrm{d}y) = 0$ *and* $\int h(y)h(T^n(y))v(\mathrm{d}y) = 0$ *for all* $n \geqslant 1$.
(ii) $\sum_{j=0}^{n-1} h \circ T^j/\sqrt{n} \to^{\mathrm{d}} \sigma \mathrm{N}(0,1)$ *and* $\sigma = \|h\|_2$.
(iii) *If* $\sigma > 0$ *then* $(h \circ T^j)_{j \geqslant 0}$ *satisfies the CLT and FCLT.*

(iv) *If* $h \in L^\infty(Y, \mathscr{B}, \nu)$ *and* $\sigma > 0$ *then all moments of* $\sum_{j=0}^{n-1} h \circ T^j/\sqrt{n}$ *converge to the corresponding moments of* $\sigma\mathrm{N}(0,1)$, *i.e. for each* $k \geqslant 1$ *we have*

$$\lim_{n\to\infty} \int \left(\frac{\sum_{j=0}^{n-1} h(T^j(y))}{\sqrt{n}}\right)^{2k} \nu(\mathrm{d}y) = \frac{(2k)!\sigma^k}{k!2^k} ,$$

$$\lim_{n\to\infty} \int \left(\frac{\sum_{j=0}^{n-1} h(T^j(y))}{\sqrt{n}}\right)^{2k-1} \nu(\mathrm{d}y) = 0 .$$

Proof. First note that the transfer operator $\mathscr{P}_{T,\nu}$ preserves the integral, i.e. $\int \mathscr{P}_{T,\nu}h(y)\nu(\mathrm{d}y) = \int h(y)\nu(\mathrm{d}y)$. Hence $\int h(y)\nu(\mathrm{d}y) = 0$. Now let $n \geqslant 1$. Since ν is a finite measure, the Koopman and transfer operators are adjoint on the space $L^2(Y, \mathscr{B}, \nu)$. This implies

$$\begin{aligned}\int h(y)h(T^n(y))\nu(\mathrm{d}y) &= \int h(y)U_T^n h(y)\nu(\mathrm{d}y) \\ &= \int \mathscr{P}_{T,\nu}h(y)U_T^{n-1}h(y)\nu(\mathrm{d}y) = 0\end{aligned}$$

and completes the proof of (i). Part (ii) follows from Lemma A.7 and Theorem A.2 in the appendix, since for each $n \geqslant 1$ we have

$$\frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} h \circ T^j = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} h \circ T^{n-j} .$$

For Part (iii), if $\sigma > 0$, then the CLT is a consequence of part (ii), while the FCLT follows from Lemma A.8 and A.9 in Appendix A. Finally, the existence and convergence of moments in Part (iv) follows from Theorem 5.3 and 5.4 of Billingsley (1968) and from Lemma A.7 in Appendix A. □

Remark 2.13. Note in particular that if T is invertible then the equation $\mathscr{P}_{T,\nu}(h) = 0$ has only a zero solution, so the theorem does not apply.

A central assumption of Theorem 2.12 is that $\mathscr{P}_{T,\nu}(h) = 0$, so we now address the question of solvability of the equation $\mathscr{P}_{T,\nu}(h) = 0$ and give a sufficient condition for this to hold.

Proposition 2.14. *Let* $(Y, \mathscr{B}, \nu)$ *be a normalized measure space and* $T : Y \to Y$ *be a measurable map preserving the measure* ν. *Let* $Y = Y_1 \cup Y_2$ *with* $Y_1, Y_2 \in \mathscr{B}$ *and* $\nu(Y_1 \cap Y_2) = 0$ *and let a bijective map* $\varphi : Y_1 \to Y_2$ *be such that both* φ *and* φ^{-1} *are measurable and preserve the measure* ν. *Assume that for every* $A \in \mathscr{B}$ *there is* $B \in \mathscr{B}$ *such that* $B \subseteq Y_1$ *and*

$$T^{-1}(A) = B \cup \varphi(B) . \tag{2.21}$$

If $h(y) + h(\varphi(y)) = 0$ *for almost every* $y \in Y_1$ *then* $\mathscr{P}_{T,\nu}h = 0$.

Proof. From Condition 2.21 it follows that

$$\int_{T^{-1}(A)} h(y)\nu(\mathrm{d}y) = \int_B h(y)\nu(\mathrm{d}y) + \int_{\varphi(B)} h(y)\nu(\mathrm{d}y) .$$

Since

$$\int_{\varphi(B)} h(y)\nu(\mathrm{d}y) = \int_{\varphi(B)} h(\varphi^{-1}(\varphi(y)))\nu(\mathrm{d}y) ,$$

the last integral is equal to

$$\int_B h(\varphi(y))(\nu \circ \varphi)(\mathrm{d}y) = \int_B h(\varphi(y))\nu(\mathrm{d}y)$$

by the change of variables applied to φ^{-1} and finally by the invariance of ν for φ together with the definition of $\mathscr{P}_{T,\nu}$. $\square$

Remark 2.15. The above proposition can be easily generalized. For example, we can have

$$T^{-1}(A) = B \cup \varphi_1(B) \cup \varphi_2(B)$$

with the sets B, $\varphi_1(B)$, $\varphi_2(B)$ pairwise disjoint.

Note that if Y is an interval then it is enough to check Condition 2.21 for intervals of the form $[a, b)$.

Example 2.16. For an even transformation T on $[-1, 1]$ with an even invariant density we can take $Y_1 = [-1, 0]$ and $\varphi(y) = -y$. In this case $\mathscr{P}_{T,\nu}h = 0$ for every odd function on $[-1, 1]$. In particular, this applies to the tent map of Example 2.3 and to the Chebyshev maps S_N of Example 2.5 with N even. We also have $\mathscr{P}_{S_N,\nu}h = 0$ when $h(y) = y$ and S_N is the Chebyshev map with N odd. Indeed, first observe that by Theorem 2.6 we have $\mathscr{P}_{S_N,\nu}h = 0$ if $P_{T_N} f = 0$ where T_N is given by 2.6 and $f(y) = \cos(\pi y)$ for $y \in [0, 1]$ and then note that $P_{T_N} f = 0$ follows because the expression

$$f(y) + \sum_{n=1}^{(N-1)/2} \left(f\left(\frac{2n}{N} + y\right) + f\left(\frac{2n}{N} - y\right)\right)$$

reduces to

$$\cos(\pi y)\left(1 + 2\sum_{n=1}^{(N-1)/2} \cos\left(\frac{2n\pi}{N}\right)\right)$$

which is equal to 0. For the dyadic map, we can take $\varphi(y) = y + 1$. Then any function satisfying $h(y) + h(y + 1) = 0$ gives a solution to $\mathscr{P}_{T,\nu}h = 0$.

The next example shows that the assumption of ergodicity in Theorem 2.12 is essential.

Example 2.17. Let $T : [0, 1] \to [0, 1]$ be defined by

$$T(y) = \begin{cases} 2y, & y \in [0, \frac{1}{4}) , \\ 2y - \frac{1}{2}, & y \in [\frac{1}{4}, \frac{3}{4}) , \\ 2y - 1, & y \in [\frac{3}{4}, 1] . \end{cases}$$

The Frobenius–Perron operator is given by

$$P_T f(y) = \frac{1}{2} f\left(\frac{1}{2} y\right) 1_{[0,\frac{1}{2})}(y) + \frac{1}{2} f\left(\frac{1}{2} y + \frac{1}{4}\right) + \frac{1}{2} f\left(\frac{1}{2} y + \frac{1}{2}\right) 1_{[\frac{1}{2},1]}(y) .$$

Observe that the Lebesgue measure on $([0, 1], \mathscr{B}([0, 1]))$ is invariant for T and that T is not ergodic since $T^{-1}([0, \frac{1}{2}]) = [0, \frac{1}{2}]$ and $T^{-1}([\frac{1}{2}, 1]) = [\frac{1}{2}, 1]$. Consider the following functions:

$$h_1(y) = \begin{cases} 1, & y \in [0, \frac{1}{4}), \\ -1, & y \in [\frac{1}{4}, \frac{3}{4}), \\ 1, & y \in [\frac{3}{4}, 1], \end{cases} \qquad h_2(y) = \begin{cases} 1, & y \in [0, \frac{1}{4}) , \\ -1, & y \in [\frac{1}{4}, \frac{1}{2}) , \\ -2, & y \in [\frac{1}{2}, \frac{3}{4}) , \\ 2, & y \in [\frac{3}{4}, 1] . \end{cases}$$

We see at once that $P_T h_1 = P_T h_2 = 0$, $P_T h_1^2 = h_1^2$, and $P_T h_2^2 = h_2^2$. It is immediate that for every $y \in [0, 1]$ we have $h_1^2(T(y)) = h_1^2(y) = 1$ and $h_2^2(T(y)) = h_2^2(y)$. Lemma A.7 and Theorem A.3 in the appendix show that

$$\frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} h_1 \circ T^j \to^{\mathrm{d}} \mathrm{N}(0, 1) ,$$

while

$$\frac{1}{\sqrt{n}}\sum_{j=0}^{n-1} h_2\circ T^j \to^{\mathrm{d}} \zeta\ ,$$

where ζ has the characteristic function of the form

$$\phi_\zeta(r)=\int_0^1 \exp\left(-\frac{r^2}{2}h_2^2(y)\right)\mathrm{d}y\ .$$

The density of ζ is equal to

$$\frac{1}{2}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right)+\frac{1}{2}\frac{1}{\sqrt{8\pi}}\exp\left(-\frac{x^2}{8}\right),\quad x\in\mathbb{R}\ .$$

Consequently, the sequence $(h_2\circ T^j)_{j\geqslant 0}$ does not satisfy the CLT.

We now turn to a discussion of the problem of changing the underlying probability space for the sequence $(h\circ T^j)_{j\geqslant 0}$. The random variables $h\circ T^j$ in Theorem 2.12 are defined on the probability space $(Y,\mathscr{B},\nu)$. Since ν is invariant for T, they have the same distribution and constitute a stationary sequence. We shall show that the result (iii) of Theorem 2.12 remains true if the transformation T is exact and $h\circ T^j$ are random variables on $(Y,\mathscr{B},\nu_0)$ where ν_0 is an arbitrary normalized measure absolutely continuous with respect to ν. In other words we can consider random variables $h(T^j(\xi_0))$ with ξ_0 distributed according to ν_0. Now these random variables are not identically distributed and constitute a non-stationary sequence. For example, consider the tent map $T(y)=1-2|y|$ on $[-1,1]$ and $h(y)=y$. Then Theorem 2.12 applies if ξ_0 is uniformly distributed. In the next theorem we will show that we can also consider ξ_0 having a density with respect to the normalized Lebesgue measure on $[-1,1]$.

Theorem 2.18. *Let $(Y,\mathscr{B},\nu)$ be a normalized measure space and $T:Y\to Y$ be exact with respect to ν. Let $h\in L^2(Y,\mathscr{B},\nu)$ be such that $\mathscr{P}_{T,\nu}h=0$ and $\sigma=\|h\|_2>0$. If ξ_0 is distributed according to a normalized measure ν_0 on $(Y,\mathscr{B})$ which is absolutely continuous with respect to ν, then $(h\circ T^j(\xi_0))_{j\geqslant 0}$ satisfies both the CLT and FCLT.*

Proof. Let g_0 be the density of the measure ν_0 with respect to ν. On the probability space $(Y,\mathscr{B})$ define the random variables ζ_n by

$$\zeta_n(y)=\frac{1}{\sigma\sqrt{n}}\sum_{j=0}^{n-1}h(T^j(y)),\quad y\in Y\ .$$

To prove the CLT we shall use the continuity theorem, and to do so we must show that

$$\lim_{n\to\infty}\int \exp(\mathrm{i}r\zeta_n(y))g_0(y)\nu(\mathrm{d}y)=\exp\left(-\frac{r^2}{2}\right),\quad r\in\mathbb{R}\ .$$

Fix $\epsilon>0$. Since T is exact, there exists $m\geqslant 1$ such that

$$\int|\mathscr{P}^m_{T,\nu}g_0(y)-1|\nu(\mathrm{d}y)\leqslant\epsilon\ .$$

Define

$$\tilde{\zeta}_n(y)=\frac{1}{\sigma\sqrt{n}}\sum_{j=0}^{n-m-1}h(T^j(y)),\quad n>m$$

and observe that

$$\zeta_n=\frac{1}{\sigma\sqrt{n}}\sum_{j=0}^{m-1}h\circ T^j+\tilde{\zeta}_n\circ T^m$$

for sufficiently large n. Since for every y the sequence $\frac{1}{\sigma\sqrt{n}}\sum_{j=0}^{m-1}h(T^j(y))$ converges to 0 as $n\to\infty$, we obtain

$$\lim_{n\to\infty}\left|\int \exp(\mathrm{i}r\zeta_n(y))g_0(y)\nu(\mathrm{d}y)-\int \exp(\mathrm{i}r\tilde{\zeta}_n(T^m(y)))g_0(y)\nu(\mathrm{d}y)\right|=0\ .$$

The equality $\exp(\mathrm{i}r\tilde{\zeta}_n\circ T^m)=U_T^m(\exp(\mathrm{i}r\tilde{\zeta}_n))$ implies that

$$\int \exp(\mathrm{i}r\tilde{\zeta}_n(T^m(y)))g_0(y)\nu(\mathrm{d}y)=\int \exp(\mathrm{i}r\tilde{\zeta}_n(y))\mathscr{P}_{T,\nu}^m g_0(y)\nu(\mathrm{d}y)\ ,$$

since the operators U_T and $\mathscr{P}_{T,\nu}$ are adjoint. This gives

$$\left|\int \exp(\mathrm{i}r\tilde{\zeta}_n(T^m(y)))g_0(y)\nu(\mathrm{d}y)-\int \exp(\mathrm{i}r\tilde{\zeta}_n(y))\nu(\mathrm{d}y)\right|\leqslant\int |\mathscr{P}_{T,\nu}^m g_0(y)-1|\nu(\mathrm{d}y)\leqslant\epsilon\ .$$

From Theorem 2.12 and the continuity theorem, it follows that

$$\lim_{n\to\infty}\int \exp(\mathrm{i}r\tilde{\zeta}_n(y))\nu(\mathrm{d}y)=\exp\left(-\frac{r^2}{2}\right)\ .$$

Consequently,

$$\limsup_{n\to\infty}\left|\int \exp(\mathrm{i}r\zeta_n(y))g_0(y)\nu(\mathrm{d}y)-\exp\left(-\frac{r^2}{2}\right)\right|\leqslant\epsilon\ ,$$

which leads to the desired conclusion, as ϵ was arbitrary.

Similar arguments as above, the multidimensional version of the continuity theorem, and Lemma A.8 in Appendix A allow us to show that the finite dimensional distributions of

$$\psi_n(t)=\frac{1}{\sigma\sqrt{n}}\sum_{j=0}^{[nt]-1}h(T^j(\xi_0))$$

converge to the finite dimensional distributions of the Wiener process w. By Lemma A.9 in Appendix A, Condition 2.20 holds with $\Pr=\nu$. Since ν_0 is absolutely continuous with respect to ν, it is easily seen that this condition also holds with $\Pr=\nu_0$. □

2.3.3. CLT and FCLT for maps with not necessarily zero auto-correlation

Does the CLT still hold when h does not satisfy the equation $\mathscr{P}_{T,\nu}h=0$? The answer to this question is positive provided that h can be written as a sum of two functions in which one satisfies the assumptions of Theorem 2.12 while the other is irrelevant for the central limit theorem to hold. This idea goes back to Gordin (1969), and allows us to develop CLT results for maps in which the auto-correlation $R_h(n)$ is not necessarily zero for $n\geqslant 1$.

Theorem 2.19. *Let* $(Y,\mathscr{B},\nu)$ *be a normalized measure space,* $T:Y\to Y$ *be ergodic with respect to* ν, *and* $h\in L^2(Y,\mathscr{B},\nu)$ *be such that* $\int h(y)\nu(\mathrm{d}y)=0$. *If there exists* $\tilde{h}\in L^2(Y,\mathscr{B},\nu)$ *such that* $\mathscr{P}_{T,\nu}\tilde{h}=0$ *and the sequence* $(1/\sqrt{n})\sum_{j=0}^{n-1}(h-\tilde{h})\circ T^j$ *is convergent in* $L^2(Y,\mathscr{B},\nu)$ *to* 0, *then*

$$\lim_{n\to\infty}\frac{\|\sum_{j=0}^{n-1}h\circ T^j\|_2^2}{n}=\|\tilde{h}\|_2^2 \tag{2.22}$$

and

$$\frac{1}{\sqrt{n}}\sum_{j=0}^{n-1}h\circ T^j\to^{\mathrm{d}}\|\tilde{h}\|_2\mathrm{N}(0,1)\ .$$

Moreover, if the series $\sum_{j=1}^{\infty}\int h(y)h(T^j(y))\nu(\mathrm{d}y)$ *is convergent, then*

$$\|\tilde{h}\|_2^2=\int h^2(y)\nu(\mathrm{d}y)+2\sum_{n=1}^{\infty}\int h(y)h(T^n(y))\nu(\mathrm{d}y)\ . \tag{2.23}$$

Proof. We have

$$\frac{\sum_{j=0}^{n-1}h\circ T^j}{\sqrt{n}}=\frac{\sum_{j=0}^{n-1}\tilde{h}\circ T^j}{\sqrt{n}}+\frac{\sum_{j=0}^{n-1}(h-\tilde{h})\circ T^j}{\sqrt{n}}\ . \tag{2.24}$$

Since $\mathscr{P}_{T,\nu}\tilde{h}=0$, we obtain $\int\tilde{h}(y)\nu(\mathrm{d}y)=0$ and $\int\tilde{h}(T^i(y))\tilde{h}(T^j(y))\nu(\mathrm{d}y)=0$ for $i\neq j$ by Condition (i) of Theorem 2.12. Hence

$$\|\tilde{h}\|_2^2=\frac{1}{n}\sum_{j=0}^{n-1}\|\tilde{h}\circ T^j\|_2^2=\left\|\frac{1}{\sqrt{n}}\sum_{j=0}^{n-1}\tilde{h}\circ T^j\right\|_2^2$$

and therefore Eq. (2.22) holds. Since the sequence

$$\frac{1}{\sqrt{n}}\sum_{j=0}^{n-1}(h-\tilde{h})\circ T^j$$

is convergent to 0 in $L^2(Y,\mathscr{B},\nu)$, it is also convergent to 0 in probability. Eq. (2.24), Condition (ii) of Theorem 2.12 applied to $\tilde{h}$, and property (2.16) complete the proof of the first part.

It remains to show that

$$\lim_{n\to\infty}\frac{\|\sum_{j=0}^{n-1}h\circ T^j\|_2^2}{n}=\int h^2(y)\nu(\mathrm{d}y)+2\sum_{n=1}^{\infty}\int h(y)h(T^n(y))\nu(\mathrm{d}y)\ .$$

Since ν is T-invariant, we have

$$\frac{1}{n}\int\left(\sum_{j=0}^{n-1}h(T^j(y))\right)^2\nu(\mathrm{d}y)=\|h\|_2+2\frac{1}{n}\sum_{j=1}^{n-1}\sum_{l=1}^{j}\int h(y)h(T^l(y))\nu(\mathrm{d}y)\ .$$

By assumption the sequence $(\sum_{j=1}^{n}\int h(y)h(T^j(y))\nu(\mathrm{d}y))_{n\geqslant 1}$ is convergent to $\sum_{j=1}^{\infty}\int h(y)h(T^j(y))\nu(\mathrm{d}y)$, which implies Eq. (2.23). $\square$

Remark 2.20. Note that in the above proof of the CLT we only used the weaker condition that the sequence $(1/\sqrt{n})\sum_{j=0}^{n-1}(h-\tilde{h})\circ T^j$ is convergent to 0 in probability. The stronger assumption that this sequence is convergent in $L^2(Y,\mathscr{B},\nu)$ was used to derive Eq. (2.22). Note also that all of the computations are useless when $\|\tilde{h}\|_2=0$ and the most interesting situation is when $\tilde{h}$ is nontrivial, i.e. $\|\tilde{h}\|_2>0$.

Strengthening the assumptions of the last theorem leads to the functional central limit theorem.

Theorem 2.21. *Let* $(Y,\mathscr{B},\nu)$ *be a normalized measure space*, $T:Y\to Y$ *be ergodic with respect to* ν, *and* $h\in L^2(Y,\mathscr{B},\nu)$ *be such that* $\int h(y)\nu(\mathrm{d}y)=0$. *If there exists a nontrivial* $\tilde{h}\in L^2(Y,\mathscr{B},\nu)$ *such that* $\mathscr{P}_{T,\nu}\tilde{h}=0$ *and the sequence* $(1/\sqrt{n})\sum_{j=0}^{n-1}(h-\tilde{h})\circ T^j$ *is a.s. convergent to* 0, *then* $(h\circ T^j)_{j\geqslant 0}$ *satisfies the FCLT.*

Proof. Since every sequence convergent ν almost everywhere is convergent in probability, the CLT follows by Remark 2.20 and Theorem 2.19. To derive the FCLT define

$$\tilde{\psi}_n(t)=\frac{1}{\sigma\sqrt{n}}\sum_{j=0}^{[nt]-1}\tilde{h}\circ T^j\quad\text{and}\quad\psi_n(t)=(1/\sigma\sqrt{n})\sum_{j=0}^{[nt]-1}h\circ T^j,\quad t\in[0,1]\ ,$$

where $\sigma = \|\tilde{h}\|_2$. Then by (iii) of Theorem 2.12 we have

$$\widetilde{\psi}_n \to^{\mathrm{d}} w \ .$$

By property (2.16) it remains to show that

$$\rho_S(\psi_n, \widetilde{\psi}_n) \to^P 0 \ .$$

To this end observe that

$$\rho_S(\psi_n, \widetilde{\psi}_n) \leqslant \sup_{0 \leqslant t \leqslant 1} |\psi_n(t) - \widetilde{\psi}_n(t)| \leqslant (1/\sigma\sqrt{n}) \max_{1 \leqslant k \leqslant n} \left| \sum_{j=0}^{k-1} (h - \tilde{h}) \circ T^j \right| .$$

Since the sequence $(1/\sqrt{n}) \sum_{j=0}^{n-1} (h - \tilde{h}) \circ T^j$ is a.s. convergent to 0, the same holds for the sequence $(1/\sigma\sqrt{n}) \max_{1 \leqslant k \leqslant n} |\sum_{j=0}^{k-1} (h - \tilde{h}) \circ T^j|$ by an elementary analysis. □

Remark 2.22. If T is exact, then Theorems 2.19 and 2.21 hold for the sequence $(h \circ T^j(\xi_0))_{j \geqslant 0}$ provided that ξ_0 is distributed according to a normalized measure ν_0 on (Y, ν) which is absolutely continuous with respect to ν. Indeed, since convergence to zero in probability is preserved by an absolutely continuous change of measure, we can apply the above arguments again, with Theorem 2.12 replaced by Theorem 2.18.

One situation when all assumptions of the two preceding theorems are met is described in the following.

Theorem 2.23. *Let $(Y, \mathscr{B}, \nu)$ be a normalized measure space, $T : Y \to Y$ be ergodic with respect to ν, and $h \in L^2(Y, \mathscr{B}, \nu)$. Suppose that the series*

$$\sum_{n=0}^{\infty} \mathscr{P}^n_{T,\nu} h$$

is convergent in $L^2(Y, \mathscr{B}, \nu)$. Define $f = \sum_{n=1}^{\infty} \mathscr{P}^n_{T,\nu} h$ and $\tilde{h} = h + f - f \circ T$.

Then $\tilde{h} \in L^2(Y, \mathscr{B}, \nu)$, $\mathscr{P}_{T,\nu}\tilde{h} = 0$, $\int h(y)\nu(\mathrm{d}y) = 0$, and the sequence $((1/\sqrt{n}) \sum_{j=0}^{n-1} (h - \tilde{h}) \circ T^j)_{n \geqslant 1}$ is convergent to 0 both in $L^2(Y, \mathscr{B}, \nu)$ and a.s.

In particular, $\|\tilde{h}\|_2 = 0$ if and only if $h = f \circ T - f$ for some $f \in L^2(Y, \mathscr{B}, \nu)$.

Proof. Since $\mathscr{P}_{T,\nu}(h + f) = f$, we have by Eq. (2.3)

$$\mathscr{P}_{T,\nu}\tilde{h} = \mathscr{P}_{T,\nu}(h + f) - \mathscr{P}_{T,\nu}(f \circ T) = f - \mathscr{P}_{T,\nu} U_T f = 0.$$

Thus it remains to study the behavior of the sequence $(1/\sqrt{n}) \sum_{j=0}^{n-1} (h - \tilde{h}) \circ T^j$, which, with our notation, reduces to $(1/\sqrt{n})(f \circ T^n - f)$. This sequence is obviously convergent to 0 in $L^2(Y, \mathscr{B}, \nu)$ because $\|f \circ T^n - f\|_2 \leqslant 2\|f\|_2$. It is also a.s. convergent to 0 which follows from the Borel-Cantelli lemma and the fact that for every $\epsilon > 0$ the series $\sum_{n=1}^{\infty} \nu(f^2 \circ T^n \geqslant n\epsilon) = \sum_{n=1}^{\infty} \nu(f^2 \geqslant n\epsilon)$ is convergent since $f \in L^2(Y, \mathscr{B}, \nu)$. □

Summarizing our results for general h we arrive at the following sufficient conditions for the CLT and FCLT to hold.

Corollary 2.24. *Let $(Y, \mathscr{B}, \nu)$ be a normalized measure space, $T : Y \to Y$ be ergodic with respect to ν, and $h \in L^2(Y, \mathscr{B}, \nu)$. If*

$$\sum_{n=0}^{\infty} \|\mathscr{P}^n_{T,\nu} h\|_2 < \infty \ , \tag{2.25}$$

then $\sigma \geqslant 0$ *given by*

$$\sigma^2 = \int h^2(y)\nu(\mathrm{d}y) + 2\sum_{n=1}^{\infty}\int h(y)h(T^n(y))\nu(\mathrm{d}y)$$

is finite and $(h \circ T^j)_{j \geqslant 0}$ *satisfies the CLT and FCLT provided that* $\sigma > 0$.

Proof. Since the operators $\mathscr{P}_{T,\nu}$ and U_T are adjoint on the space $L^2(Y, \mathscr{B}, \nu)$, we have

$$\int h(y)h(T^n(y))\nu(\mathrm{d}y) = \int \mathscr{P}^n_{T,\nu}h(y)h(y)\nu(\mathrm{d}y) \ .$$

Thus

$$\left|\int h(y)h(T^n(y))\nu(\mathrm{d}y)\right| \leqslant \int |\mathscr{P}^n_{T,\nu}h(y)h(y)|\nu(\mathrm{d}y) \leqslant \|\mathscr{P}^n_{T,\nu}h\|_2\|h\|_2$$

by Schwartz's inequality. Hence

$$\sum_{n=1}^{\infty}\int |h(y)h(T^n(y))|\mathrm{d}\nu \leqslant \|h\|_2 \sum_{n=1}^{\infty}\|\mathscr{P}^n_{T,\nu}h\|_2 \ ,$$

which shows that the series $\sum_{n=1}^{\infty}\int h(y)h(T^n(y))\nu(\mathrm{d}y)$ is convergent. Since assumption (2.25) implies that the series $\sum_{n=0}^{\infty}\mathscr{P}^n_{T,\nu}h$ is absolutely convergent in $L^2(Y, \mathscr{B}, \nu)$, the assertions follow from Theorems 2.19, 2.21, and 2.23. □

Remark 2.25. Note that if Condition 2.25 holds then

$$\lim_{n\to\infty} \|\mathscr{P}^n_{T,\nu}h\|_2 = 0 \ .$$

Since ν is finite, we have

$$\lim_{n\to\infty} \|\mathscr{P}^n_{T,\nu}h\|_1 = 0 \ .$$

Therefore the validity of Condition 2.25 on a dense subset of $\{h \in L^1(Y, \mathscr{B}, \nu) : \int h(y)\nu(\mathrm{d}y) = 0\}$ implies that T is exact.

We now give the following generalizations of Theorem 2.23 which allow us to have the CLT and FCLT for maps with polynomial decay of correlations.

Theorem 2.26 (Tyran-Kamińska (2005)). *Let* $(Y, \mathscr{B}, \nu)$ *be a normalized measure space*, $T : Y \to Y$ *be ergodic with respect to* ν, *and* $h \in L^2(Y, \mathscr{B}, \nu)$ *be such that* $\int h(y)\nu(\mathrm{d}y) = 0$. *Suppose that*

$$\sum_{n=1}^{\infty} n^{-\frac{3}{2}}\left\|\sum_{k=0}^{n-1}\mathscr{P}^k_{T,\nu}h\right\|_2 < \infty \ . \tag{2.26}$$

Then there exists $\tilde{h} \in L^2(Y, \mathscr{B}, \nu)$ *such that* $\mathscr{P}_{T,\nu}\tilde{h} = 0$ *and the sequence* $((1/\sqrt{n})\sum_{j=0}^{n-1}(h-\tilde{h})\circ T^j)_{n\geqslant 1}$ *is convergent in* $L^2(Y, \mathscr{B}, \nu)$ *to zero. If*

$$\left\|\sum_{k=0}^{n-1}\mathscr{P}^k_{T,\nu}h\right\|_2 = \mathrm{O}(n^{\alpha}) \quad \textit{with } \alpha < \frac{1}{2}$$

then $((1/\sqrt{n})\sum_{j=0}^{n-1}(h-\tilde{h})\circ T^j)_{n\geqslant 1}$ *is convergent a.s. to* 0.

Now we give a simple result that derives CLT and FCLT from a bound on the rate of decay of correlations.

Corollary 2.27 (Tyran-Kamińska (2005)). *Let $(Y, \mathscr{B}, \nu)$ be a normalized measure space, $T : Y \to Y$ be ergodic with respect to ν, and let $h \in L^\infty(Y, \mathscr{B}, \nu)$ be such that $\int h(y)\nu(\mathrm{d}y) = 0$. If there are $\alpha > 1$ and $c > 0$ such that*

$$\left|\int h(y)g(T^n(y))\nu(\mathrm{d}y)\right| \leqslant \frac{c}{n^\alpha} \|g\|_\infty \tag{2.27}$$

for all $g \in L^\infty(Y, \mathscr{B}, \nu)$ and $n \geqslant 1$, then $\sigma \geqslant 0$ given by

$$\sigma^2 = \int h^2(y)\nu(\mathrm{d}y) + 2\sum_{n=1}^{\infty} \int h(y)h(T^n(y))\nu(\mathrm{d}y)$$

is finite and $(h \circ T^j)_{j \geqslant 0}$ satisfies the CLT and FCLT provided that $\sigma > 0$.

2.3.4. Examples

The material of the previous sections has been dense and difficult, but absolutely necessary for us to understand how systems like Eqs. (1.1)–(1.4) can generate processes with the properties of white noise. As we show in the next two sections, the answer to this question is intimately related to an extension of a familiar concept from probability theory (the central limit theorem) to situations like we are considering where the $\xi(t)$ are not statistically independent. Before doing so, however, we turn to a few examples to illustrate some of the material of the previous sections.

Assume that Y is an interval $[a, b]$ in $\mathbb{R}$ for some a, b. Recall that a function $h : [a, b] \to \mathbb{R}$ is said to be of bounded variation if

$$\bigvee_a^b h = \sup \sum_{i=1}^{n} |h(y_{i-1}) - h(y_i)| < \infty ,$$

where the supremum is taken over all finite partitions, $a = y_0 < y_1 < \cdots < y_n = b$, $n \geqslant 1$, of Y.

Let $V([0, 1])$ denote the space of all integrable functions with bounded variation over $[0, 1]$ such that $\int_0^1 h(y)\,\mathrm{d}y = 0$. We have

$$|h(y)| \leqslant \bigvee_0^1 h \quad \text{for } h \in V([0, 1]),\ \ y \in [0, 1] . \tag{2.28}$$

Example 2.28. For the continued fraction map there exists a positive constant $c < 1$ such that for every function h of bounded variation over $[0, 1]$ we have (Iosifescu, 1992, Corollary, p. 904)

$$\bigvee_0^1 \mathscr{P}^n_{T,\nu} h \leqslant c^n \bigvee_0^1 h \quad \text{for all } n \geqslant 1 ,$$

where ν is Gauss's measure with density g_* as in Example 2.9. From this and Condition 2.28 it follows that for every $h \in V([0, 1])$

$$\|\mathscr{P}^n_{T,\nu} h\|_2 \leqslant \sup_{y\in[0,1]} |\mathscr{P}^n_{T,\nu} h(y)| \leqslant c^n \bigvee_0^1 h .$$

Consequently, Condition 2.25 is satisfied and Corollary 2.24 applies.

By definition, the Frobenius–Perron operator is a linear operator from $L^1([0, 1])$ to $L^1([0, 1])$, but for sufficiently smooth piecewise monotonic maps it can be defined as a pointwise map of $V([0, 1])$ into $V([0, 1])$. Since functions of bounded variation have only countably many points of discontinuity, redefining P_T at those points does not change its L^1 properties. If, moreover, one is able to give an estimate for the iterates of P_T in the bounded variation norm

$$\|h\|_{BV} = \bigvee_0^1 h + \int_0^1 |h(y)|\,\mathrm{d}y ,$$

then obviously one is able to estimate the norm of $P_T^n f$ in all $L^p([0,1])$ spaces. In many cases there exist $c_1, c_2 > 0$ and $r \in (0,1)$ such that

$$\|P_T^n h\|_{BV} \leqslant c_1 r^n \left(\bigvee_0^1 h + c_2 \|h\|_1 \right), \quad h \in V([0,1]) . \tag{2.29}$$

We now describe two classes of chaotic maps for which one can easily show that Condition 2.25 holds for every $h \in V([0,1])$.

Example 2.29. Consider a transformation $T : [0,1] \to [0,1]$ having the following properties

(i) there is a partition $0 = a_0 < a_1 < \cdots < a_l = 1$ of $[0,1]$ such that for each integer $i = 1, \ldots, l$ the restriction of T to $[a_{i-1}, a_i)$ is continuous and convex,
(ii) $T(a_{i-1}) = 0$ and $T'(a_{i-1}) > 0$,
(iii) $T'(0) > 1$.

For such transformation the Frobenius–Perron operator has a unique fixed point g_*, where g_* is of bounded variation and a decreasing function of y (Lasota and Mackey, 1994). Moreover it is bounded from below when, for example, $T([0, a_1]) = [0,1]$. It is known (Jabłoński and Malczak, 1983b) that the estimate in Eq. (2.29) holds for the Frobenius–Perron operator P_T for transformations with these three properties. Suppose that $g_*(y) > 0$ for almost every $y \in [0,1]$. Since $\|\mathscr{P}_{T,v}\|_\infty \leqslant \|f\|_\infty$ for all $f \in L^\infty([0,1], \mathscr{B}, v)$, we have for all $h \in f \in L^\infty([0,1], \mathscr{B}, v)$

$$\|\mathscr{P}_{T,v}^n h\|_2 \leqslant \|h\|_\infty^{1/2} \|\mathscr{P}_{T,v}^n h\|_1^{1/2} \tag{2.30}$$

If h is of bounded variation with $\int h(y) g_*(y)\, \mathrm{d}y = 0$, then $h g_* \in V([0,1])$ and $\|\mathscr{P}_{T,v}^n h\|_1 = \|P_T^n (h g_*)\|_1$. Thus

$$\|\mathscr{P}_{T,v}^n h\|_2 = \mathrm{O}(r^{n/2})$$

by Eq. (2.29) and Corollary 2.24 applies.

Example 2.30. Let a transformation $T : [0,1] \to [0,1]$ be piecewise monotonic, the function $\frac{1}{|T'(y)|}$ be of bounded variation over $[0,1]$ and $\inf_{y \in [0,1]} |T'(y)| > 1$. For such transformations the Frobenius–Perron operator has a fixed point g_* and g_* is of bounded variation (Lasota and Mackey, 1994). Suppose that P_T has a unique invariant density g_* which is strictly positive. Then the transformation T is ergodic and g_* is bounded from below. There exists k such that T^k is exact, the estimate in Eq. (2.29) is valid for the Frobenius–Perron operator P_{T^k} corresponding to T^k, and the following holds (Jabłoński et al., 1985): There exists $c_1 > 0$ and $r \in (0,1)$ such that

$$|P_{T^k}^n f(y)| \leqslant c_1 r^n \left(\bigvee_0^1 f + \int_0^1 |f(y)| \mathrm{d}y \right), \quad y \in [0,1], \quad f \in V([0,1]) .$$

Hence Corollary 2.24 applies to T^k and every h of bounded variation with $\int h(y) g_*(y)\, \mathrm{d}y = 0$. One can relax the assumption that g_* is strictly positive and have instead $g_* \geqslant c$ for almost every $y \in Y_* = \{y \in [0,1] : g_*(y) > 0\}$. Then the above estimate is valid for $y \in Y_*$ and f with $\operatorname{supp} f = \{y \in [0,1] : f(y) \neq 0\} \subset Y_*$ and Corollary 2.24 still applies to T^k.

Example 2.31. The conclusions of Corollary 2.24 can be extended to conjugated maps. If the Lebesgue measure on $[0,1]$ is invariant with respect to T, then Theorem 2.6 offers the following.

Let $T : [0,1] \to [0,1]$ be a transformation for which the Lebesgue measure on $[0,1]$ is invariant and for which Eq. (2.29) holds. Let g_* be a positive function and $S : [a,b] \to [a,b]$ be given by $S = G^{-1} \circ T \circ G$, where

$$G(x) = \int_a^x g_*(y)\, \mathrm{d}y, \quad a \leqslant x \leqslant b .$$

By Theorem 2.6 we have

$$\mathscr{P}_{S,v} f = U_G P_T U_{G^{-1}} f \quad \text{for } f \in L^1([a,b], \mathscr{B}([a,b]), v) \ .$$

The operator $U_G : L^2([0,1]) \to L^2([a,b], \mathscr{B}([a,b]), v)$ is an isometry, thus

$$\|\mathscr{P}^n_{S,v} h\|_2 = \|P^n_T U_{G^{-1}} h\|_{L^2([0,1])} \ .$$

Since G is increasing, G^{-1} is a function of bounded variation. As a result, if $h : [a,b] \to \mathbb{R}$ is a function of bounded variation with $\int h(y) g_*(y)\,\mathrm{d}y = 0$, then $U_{G^{-1}} h = h \circ G^{-1}$ is of bounded variation over $[0,1]$. Consequently, by Eq. (2.29)

$$\sum_{n=0}^{\infty} \|\mathscr{P}^n_{S,v} h\|_2 < \infty \ ,$$

where $\|\cdot\|_2$ denotes the norm in $L^2([a,b], \mathscr{B}([a,b]), v)$ and v is the measure with density g_*.

Example 2.32. Finally we discuss the case of quadratic maps. We follow the formulation in Viana (1997). Consider the quadratic map T_β, $\beta \in (0,2)$, of Example 2.10 and assume that for the critical point $c=0$ there are constants $\lambda_c > 1$ and $\alpha > 0$ such that $\lambda_c > \mathrm{e}^{2\alpha}$ and

(i) $|(T^n_\beta)'(T(c))| \geqslant \lambda^n_c$ for every $n \geqslant 1$;
(ii) $|T^n_\beta(c) - c| \geqslant \mathrm{e}^{-\alpha n}$ for every $n \geqslant 1$;
(iii) T_β is topologically mixing on the interval $[T^2_\beta(c), T_\beta(c)]$, i.e. for every interval $I \subset [-1,1]$ there exists k such that $T^k_\beta(I) \supset [T^2_\beta(c), T_\beta(c)]$.

Young (1992) shows that there is a set of parameters β close to 2 for which conditions (i), (ii), (iii) hold with $\lambda_c = 1.9$ and $\alpha = 10^{-6}$ and that the central limit theorem holds for $(h \circ T^j_\beta)$ with h of bounded variation. Viana (1997) (Section 5) proves that conditions (i), (ii), (iii) imply all assumptions of our Corollary 2.24 for the transformation T_β and functions h of bounded variation with $\int h(y) v_\beta(\mathrm{d}y) = 0$.

Remark 2.33. Only recently was the FCLT established by Pollicott and Sharp (2002) for maps such as the Manneville–Pomeau map of Example 2.11 and for Hölder continuous functions h with $\int h(y) v(\mathrm{d}y) = 0$ under the hypothesis that $0 < \beta < \frac{1}{3}$. When $0 < \beta < \frac{1}{2}$ the CLT was proved by Young (1999), where it was shown that in this case condition 2.27 holds. Thus our Corollary 2.27 gives both the CLT and the FCLT for maps satisfying the following:

(i) $T(0) = 0$, $T'(0) = 1$, T is increasing and piecewise C^2 and onto $[0,1]$,
(ii) $\inf_{\epsilon \leqslant y \leqslant 1} |T'(y)| > 1$ for every $\epsilon > 0$,
(iii) $\lim_{y \to 0} y^{1-\beta} T''(y) \neq 0$.

as the "tower method" of (Young, 1999) gives us the estimate in Eq. (2.27) with $\alpha = \frac{1}{\beta} - 1$ for all Hölder continuous h and $g \in L^\infty([0,1], \mathscr{B}, v)$ with the constant c dependent only on h.

3. 'Langevin equations' with deterministic perturbations

The original problem we posed in Section 1 was "...how Brownian-like motion can arise from a purely deterministic dynamics." Having developed the necessary machinery to tackle this in the previous section, in this section we commence our examination of this problem. In a sense, the conclusions of the section are negative since we will show what does not work. However, the results of this investigation will serve as a guide to a successful resolution of the question in Section 4.

We consider the position (x) and velocity (v) of a dynamical system defined by

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = v(t) , \tag{3.1}$$

$$\frac{\mathrm{d}v(t)}{\mathrm{d}t} = b(v(t)) + \eta(t) , \tag{3.2}$$

with a perturbation η in the velocity. We further assume that $\eta(t)$ consists of a series of delta-function-like perturbations that occur at times $t_0, t_1, t_2, \ldots$. These perturbations have an amplitude $h(\xi(t))$, and $\eta(t)$ takes the explicit form

$$\eta(t) = \kappa \sum_{n=0}^{\infty} h(\xi(t))\delta(t - t_n) . \tag{3.3}$$

We further assume that ξ is generated by a dynamical system that at least has an invariant measure for the results of Section 3.1 to hold, or is at least ergodic for the central limit theorem to hold as in Section 3.2.

In practice, we will illustrate our results assuming that ξ is the trace of a highly chaotic semi-dynamical system that is, indeed, even exact in the sense of Lasota and Mackey (1994) (cf. Section 2). ξ could, for example, be generated by the differential delay equation

$$\delta \frac{\mathrm{d}\xi}{\mathrm{d}t} = -\xi(t) + T(\xi(t-1)) , \tag{3.4}$$

where the nonlinearity T has the appropriate properties to generate chaotic solutions (Mackey and Glass, 1977; an der Heiden and Mackey, 1982; Dorizzi et al., 1987; Berre et al., 1990). The parameter δ controls the time scale for these oscillations, and in the limit as $\delta \to 0$ we can approximate the behavior of the solutions through a map of the form

$$\xi_{n+1} = T(\xi_n) . \tag{3.5}$$

Thus, we can think of the map T as being generated by the sampling of a chaotic continuous time signal $\xi(t)$ as, for example, by the taking a Poincaré section of a semi-dynamical system operating in a high dimensional phase space.

Let $(Y, \mathscr{B}, \nu)$ be a normalized measure space. Let $b : \mathbb{R}^k \to \mathbb{R}^k$, $h : Y \to \mathbb{R}^k$ be measurable transformations, and let $(t_n)_{n \geqslant 0}$ be an increasing sequence of real numbers. Assume that $\xi : \mathbb{R}_+ \times Y \to Y$ is such that $\xi(t_{n+1}) = T(\xi(t_n))$ for $n \geqslant 0$, where $T : Y \to Y$ is a measurable transformation preserving the measure ν. Combining (3.2) with (3.3) we have

$$\frac{\mathrm{d}v(t)}{\mathrm{d}t} = b(v(t)) + \kappa \sum_{n=0}^{\infty} h(\xi(t))\delta(t - t_n) . \tag{3.6}$$

We say that $v(t)$, $t \geqslant t_0$, is a solution of Equation (3.6) if, for each $n \geqslant 0$, $v(t)$ is a solution of the Cauchy problem

$$\begin{aligned} &\frac{\mathrm{d}v(t)}{\mathrm{d}t} = b(v(t)), \quad t \in (t_n, t_{n+1}) , \\ &v(t_n) = v(t_n^-) + \kappa h(\xi(t_n)) , \end{aligned} \tag{3.7}$$

where v_0 is an arbitrary point of $\mathbb{R}^k$ and $v(t_n^-) = \lim_{t \to t_n^-} v(t)$ for $n \geqslant 1$.

Let $\pi : \mathbb{R}_+ \times \mathbb{R}^k \to \mathbb{R}^k$ be the semigroup generated by the Cauchy problem (if b is a Lipschitz map then π is well defined)

$$\begin{aligned} &\frac{\mathrm{d}\tilde{v}(t)}{\mathrm{d}t} = b(\tilde{v}(t)), \quad t > 0 , \\ &\tilde{v}(0) = \tilde{v}_0 , \end{aligned} \tag{3.8}$$

i.e. for every $\tilde{v}_0 \in \mathbb{R}^k$ the unique solution of (3.8) is given by $\tilde{v}(t) = \pi(t, \tilde{v}_0)$ for $t \geqslant 0$. As a result, the solution of (3.6) is given by

$$v(t) = \pi(t - t_n, v(t_n)) \quad \text{for } t \in [t_n, t_{n+1}),\ n \geqslant 0 .$$

After integration, for $t \in [t_n, t_{n+1})$ we have

$$x(t) - x(t_n) = \int_{t_n}^{t} v(s)\,\mathrm{d}s = \int_{t_n}^{t} \pi(s - t_n, v(t_n))\,\mathrm{d}s = \int_{0}^{t-t_n} \pi(s, v(t_n))\,\mathrm{d}s \ .$$

Consequently, the solutions of Eqs. (3.1) and (3.2) are given by

$$x(t) = x(t_n) + \int_{0}^{t-t_n} \pi(s, v(t_n))\,\mathrm{d}s \ , \tag{3.9}$$

$$v(t) = \pi(t - t_n, v(t_n)), \quad \text{for } t \in [t_n, t_{n+1}),\ n \geqslant 0 \ . \tag{3.10}$$

Observe that $x(t)$ is continuous in t, while $v(t)$ is only right continuous, with left-hand limits, and $v(t_n) = \lim_{t \to t_n^+} v(t)$.

We are interested in the variables $v(t_n)$, $v_n := v(t_n^-)$, $\xi_n := \xi(t_n)$, and $x_n := x(t_n)$ which appear in the definition of the solutions $v(t)$ and $x(t)$. We have

$$v(t_n) = v_n + \kappa h(\xi_n) \ , \tag{3.11}$$

$$v_{n+1} = \pi(t_{n+1} - t_n, v(t_n)) \ , \tag{3.12}$$

$$\xi_{n+1} = T(\xi_n) \ , \tag{3.13}$$

$$x_{n+1} = x_n + \int_{0}^{t_{n+1}-t_n} \pi(s, v(t_n))\,\mathrm{d}s \ . \tag{3.14}$$

We are going to examine the dynamics of these variables from a statistical point of view.

We start by supposing that v_0 has a distribution μ, ξ_0 has a distribution ν, and that the random variables are independent. The next section considers what we can say about the long-term behavior of the distribution of the random variables $v(t_n)$ or v_n.

3.1. Weak convergence of $v(t_n)$ and v_n

To simplify the presentation and easily use Proposition B.2 of Appendix B, assume that the differences $t_{n+1} - t_n$ do not depend on n, so that $t_n = n\tau$ for $n \geqslant 0$. Define $\Lambda : \mathbb{R}^k \to \mathbb{R}^k$ by

$$\Lambda(v) = \pi(\tau, v), \quad v \in \mathbb{R}^k \ , \tag{3.15}$$

where π describes the solutions of the unperturbed system as defined by Eq. (3.8). In particular, adding chaotic deterministic perturbations to any exponentially stable system produces a stochastically stable system as stated in the following

Corollary 3.1. *Let $\Lambda : \mathbb{R}^k \to \mathbb{R}^k$ be a Lipschitz map with a Lipschitz constant $\lambda \in (0, 1)$. Let $T : Y \to Y$ be a transformation preserving the measure ν, and $h : Y \to \mathbb{R}^k$ be such that $\int_Y |h(y)|\nu(\mathrm{d}y) < \infty$. Assume that the random variables v_0 and ξ_0 are independent and that ξ_0 has a distribution ν. Then $v(n\tau)$ converges in distribution to a probability measure μ_* on $\mathbb{R}^k$ and μ_* is independent of the distribution of the initial random variable v_0. Moreover, v_n converges in distribution to the probability measure $\mu_* \circ \Lambda^{-1}$.*

Proof. From Eqs. (3.11) and (3.12) it follows that

$$v((n+1)\tau) = \Lambda(v(n\tau)) + \kappa h(\xi_{n+1}), \quad n \geqslant 0 \ .$$

Define the transformation $R_n : \mathbb{R}^k \times Y \to \mathbb{R}^k$ recursively:

$$\begin{aligned} &R_0(v, y) = v + \kappa h(y) \ , \\ &R_{n+1}(v, y) = \Lambda(R_n(v, y)) + \kappa h(T^{n+1}(y)), \quad v \in \mathbb{R}^k,\ \ y \in Y,\ \ n \geqslant 0 \ . \end{aligned} \tag{3.16}$$

Then $v(n\tau) = R_n(v_0, \xi_0)$. One can easily check by induction that all assumptions of Proposition B.2 are satisfied. Thus $v(n\tau)$ converges in distribution to a unique probability measure μ_* on $\mathbb{R}^k$. Since $v_{n+1} = \Lambda(v(n\tau))$ and Λ is a continuous transformation, it follows from the definition of weak convergence that the distribution of v_{n+1} converges weakly to $\mu_* \circ \Lambda^{-1}$. □

We call the measure μ_* the limiting measure for $v(n\tau)$. Note that μ_* may depend on v.

Remark 3.2. Although Corollary 3.1 shows that there is a unique limiting measure, we cannot conclude in general that this measure has a density absolutely continuous with respect to the Lebesgue measure. See Example 2.8 and Remark 3.12.

3.2. The linear case in one dimension

We now consider Eq. (3.2) when $b(v) = -\gamma v$ and $\gamma \geqslant 0$. In this situation, we are considering a frictional force linear in the velocity, so Eqs. (3.1) and (3.2) become

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = v(t) \;, \qquad \frac{\mathrm{d}v(t)}{\mathrm{d}t} = -\gamma v(t) + \kappa \sum_{n=0}^{\infty} h(\xi(t))\delta(t - t_n) \;. \tag{3.17}$$

To make the computations of the previous section completely transparent, multiply Eq. (3.17) by the integrating factor $\exp(\gamma t)$, rearrange, and integrate from $(t_n - \epsilon)$ to $(t_{n+1} - \epsilon)$, where $0 < \epsilon < \min_{n \geqslant 0} (t_{n+1} - t_n)$, to give

$$\begin{aligned} v(t_{n+1} - \epsilon)\mathrm{e}^{\gamma(t_{n+1}-\epsilon)} - v(t_n - \epsilon)\mathrm{e}^{\gamma(t_n-\epsilon)} &= \kappa \sum_{n=0}^{\infty} \int_{t_n-\epsilon}^{t_{n+1}-\epsilon} h(\xi(z))\delta(z - t_n)\,\mathrm{d}z \\ &= \kappa \mathrm{e}^{\gamma(t_n-\epsilon)} h(\xi(t_n - \epsilon)) \;. \end{aligned} \tag{3.18}$$

Taking the $\lim_{\epsilon \to 0}$ in Eq. (3.18) and remembering that $v(t_n^-) = v_n$ and $\xi(t_n) = \xi_n$, we have

$$v_{n+1} = \lambda_n v_n + \kappa \lambda_n h(\xi_n) \;, \tag{3.19}$$

where $\tau_n \equiv t_{n+1} - t_n$ and $0 \leqslant \lambda_n \equiv \mathrm{e}^{-\gamma \tau_n} < 1$.

We simplify this formulation by taking $t_{n+1} - t_n \equiv \tau > 0$ so the perturbations are assumed to be arriving periodically. As a consequence, $\lambda_n \equiv \lambda$ with

$$\lambda = \mathrm{e}^{-\gamma\tau} \;. \tag{3.20}$$

Then, Eq. (3.19) becomes

$$v_{n+1} = \lambda v_n + \kappa \lambda h(\xi_n) \;. \tag{3.21}$$

This result can also be arrived at from other assumptions.[1]

[1] Alternately but, as it will turn out, equivalently, we can think of the perturbations as constantly applied. In this case we write an Euler approximation to the derivative in Eq. (3.17) so with an integration step size of τ we have

$$v(t + \tau) = (1 - \gamma\tau)v(t) + \tau\kappa h(\xi(t)) \;. \tag{3.22}$$

Measuring time in units of τ so $t_{n+1} = t_n + \tau$ we then can write this in the alternate equivalent form

$$v_{n+1} = \lambda v_n + \kappa_1 \lambda h(\xi_n) \;, \tag{3.23}$$

where, now, $\lambda = 1 - \gamma\tau$ and $\kappa_1 = \kappa\tau\lambda^{-1}$. Again, by induction we obtain Eq. (3.24).

3.2.1. Behavior of the velocity variable

For a given initial v_0 we have, by induction,

$$\begin{aligned} v_n &= \lambda^n v_0 + \kappa\lambda \sum_{j=0}^{n-1} \lambda^{n-1-j} h(\xi_j)\ , \\ &= \lambda^n v_0 + \kappa\lambda \sum_{j=0}^{n-1} \lambda^{n-1-j} h(T^j(\xi_0))\ . \end{aligned} \tag{3.24}$$

We now calculate the asymptotic behavior of the variance of v_n when ξ_0 is distributed according to ν, the invariant measure for T. Assume for simplicity that $v_0 = 0$, set $\sigma^2 = \int h^2(y)\nu(\mathrm{d}y)$ and assume that the auto-correlation $\int h(y)h(T^n(y))\nu(\mathrm{d}y) = 0$ for $n \geqslant 1$. Then we have

$$\int v_n^2 \nu(\mathrm{d}y) = \kappa^2 \int \left(\sum_{j=0}^{n-1} \lambda^{n-j} h(T^j(y)) \right)^2 \nu(\mathrm{d}y)\ .$$

Since the sequence $h \circ T^j$ is uncorrelated by our assumption,

$$\int \left(\sum_{j=0}^{n-1} \lambda^{n-j} h(T^j(y)) \right)^2 \nu(\mathrm{d}y) = \sum_{j=0}^{n-1} \lambda^{2n-2j} \int h(T^j(y))^2 \nu(\mathrm{d}y) = \frac{1-\lambda^{2n}}{1-\lambda^2}\sigma^2\ .$$

Thus

$$\int v_n^2 \nu(\mathrm{d}y) = \kappa^2\sigma^2 \left(\frac{1-\lambda^{2n}}{1-\lambda^2} \right)\ . \tag{3.25}$$

Since $b(v) = -\gamma v$, we have $\pi(t, v) = \mathrm{e}^{-\gamma t} v$. Eq. (3.21), in conjunction with Eqs. (3.12) and (3.11), leads to $v_{n+1} = \lambda v(n\tau)$ where $v(n\tau) = v_n + \kappa h(\xi_n)$. Thus

$$v(n\tau) = \lambda^n v_0 + \kappa \sum_{j=0}^{n} \lambda^{n-j} h(T^j(\xi_0))\ . \tag{3.26}$$

Case 1: If $\lambda < 1$ and ξ_0 is distributed according to ν, then by Corollary 3.1 there exists a unique limiting measure μ_* for $v(n\tau)$ provided that h is integrable with respect to ν. The sequence v_n also converges in distribution. Since the function Λ defined by Eq. (3.15) is linear, $\Lambda(v) = \lambda v$, both sequences $v(n\tau)$ and v_n are either convergent or divergent in distribution.

What can happen if the random variable ξ_0 in Eq. (3.26) is distributed according to a different measure?

Proposition 3.3. *Let $\lambda < 1$, the transformation T be exact with respect to ν, and $h \in L^1(Y, \mathscr{B}, \nu)$. If the random variable ξ_0 is distributed according to a normalized measure ν_0 on $(Y, \mathscr{B})$ which is absolutely continuous with respect to ν, then*

$$\kappa \sum_{j=0}^{n} \lambda^{n-j} h(T^j(\xi_0)) \to^{\mathrm{d}} \mu_* \tag{3.27}$$

and μ_* does not depend on ν_0.

Proof. Recall that by the continuity theorem

$$\kappa \sum_{j=0}^{n} \lambda^{n-j} h(T^j(\xi_0)) \to^{\mathrm{d}} \mu_*$$

if and only if for every $r \in \mathbb{R}$

$$\lim_{n\to\infty} \int_Y \exp\left(\mathrm{i} r\kappa \sum_{j=0}^{n} \lambda^{n-j} h(T^j(y)) \right) v_0(\mathrm{d}y) = \int_{\mathbb{R}} \exp(\mathrm{i}rx)\mu_*(\mathrm{d}x) \ .$$

We know that the last equation holds when $v_0 = v$. Since for every $m \geqslant 1$ the sequence

$$\kappa \sum_{j=0}^{m-1} \lambda^{n-j} h(T^j(y))$$

is convergent to 0 as $n \to \infty$ and

$$\sum_{j=m}^{n} \lambda^{n-j} h \circ T^j = \sum_{j=0}^{n} \lambda^{n-m-j} h \circ T^j \circ T^m \ ,$$

an analysis similar to that in the proof of Theorem 2.18 completes the demonstration. □

Case 2: If $\lambda = 1$ we have $v_n = v_0 + \kappa \sum_{j=0}^{n-1} h(T^j(\xi_0))$. Since v_0 and ξ_0 are independent random variables, v_0 and $\kappa \sum_{j=0}^{n-1} h(T^j(\xi_0))$ are also independent. Hence v_n converges if and only if $\kappa \sum_{j=0}^{n-1} h(T^j(\xi_0))$ does. Moreover, if there is a limiting measure μ_* for v_n, then the sequence $v(n\tau)$ converges in distribution, say to v_*, and μ_* is a convolution of the distribution of v_0 and v_*. As a result, μ_* depends on the distribution of v_0. However, if the map T and function h satisfy the CLT, then

$$\frac{\sum_{j=0}^{n-1} h(T^j(\xi_0))}{\sigma\sqrt{n}} \to^{\mathrm{d}} \mathrm{N}(0,1) \ .$$

Hence $\sum_{j=0}^{n-1} h(T^j(\xi_0))$ is not convergent in distribution since the density is spread out on the entire real line.

3.2.2. Behavior of the position variable

For the position variable we have, for $t \in [n\tau, (n+1)\tau)$,

$$x(t) - x(n\tau) = \int_{n\tau}^{t} v(s)\,\mathrm{d}s = \int_{n\tau}^{t} \mathrm{e}^{-\gamma(s-n\tau)} v(n\tau)\,\mathrm{d}s = \frac{1-\mathrm{e}^{-\gamma(t-n\tau)}}{\gamma} v(n\tau) \ .$$

With $x(n\tau) = x_n$ we have

$$x_{n+1} - x_n = \frac{1-\mathrm{e}^{-\gamma\tau}}{\gamma} v(n\tau) = \frac{1-\lambda}{\gamma} v(n\tau) \ . \tag{3.28}$$

Summing from 0 to n gives

$$x_{n+1} = x_0 + \frac{1-\lambda}{\gamma} \sum_{j=0}^{n} v(j\tau) \ . \tag{3.29}$$

From this and Eq. (3.26) we obtain

$$\begin{aligned} x_{n+1} &= x_0 + \frac{1-\lambda}{\gamma} \sum_{j=0}^{n} \left(\lambda^j v_0 + \kappa \sum_{i=0}^{j} \lambda^{j-i} h(T^i(\xi_0)) \right) \\ &= x_0 + \frac{(1-\lambda^{n+1})}{\gamma} v_0 + \frac{(1-\lambda)\kappa}{\gamma} \sum_{j=0}^{n} \sum_{i=0}^{j} \lambda^{j-i} h(T^i(\xi_0)) \ . \end{aligned}$$

Changing the order of summation in the last term gives

$$\sum_{j=0}^{n}\sum_{i=0}^{j}\lambda^{j-i}h(T^i(\xi_0)) = \sum_{i=0}^{n}\sum_{j=i}^{n}\lambda^{j-i}h(T^i(\xi_0)) = \sum_{i=0}^{n}\frac{1-\lambda^{n-i+1}}{1-\lambda}h(T^i(\xi_0))$$
$$= \frac{1}{1-\lambda}\sum_{i=0}^{n}h(T^i(\xi_0)) - \frac{\lambda}{1-\lambda}\sum_{i=0}^{n}\lambda^{n-i}h(T^i(\xi_0)) \ .$$

Consequently

$$x_{n+1} = x_0 + \frac{(1-\lambda^{n+1})}{\gamma}v_0 + \frac{\kappa}{\gamma}\sum_{i=0}^{n}h(T^i(\xi_0)) - \frac{\lambda\kappa}{\gamma}\sum_{i=0}^{n}\lambda^{n-i}h(T^i(\xi_0)) \ .$$

In conjunction with Eq. (3.28), this gives

$$x_n = x_0 + \frac{(1-\lambda^{n})}{\gamma}v_0 + \frac{\kappa}{\gamma}\sum_{i=0}^{n}h(T^i(\xi_0)) - \frac{\kappa}{\gamma}\sum_{i=0}^{n}\lambda^{n-i}h(T^i(\xi_0)) \ . \tag{3.30}$$

Next we calculate the asymptotic behavior of the variance of x_n. Assume as before that $x_0 = v_0 = 0$, and that $\int h(y)h(T^j(y))\nu(\mathrm{d}y) = 0$ and $\sigma^2 = \int h^2(y)\nu(\mathrm{d}y)$. We have

$$\left(\sum_{i=0}^{n}h(T^i(\xi_0)) - \sum_{i=0}^{n}\lambda^{n-i}h(T^i(\xi_0)\right)^2 = \left(\sum_{i=0}^{n}h(T^i(\xi_0))\right)^2 + \left(\sum_{i=0}^{n}\lambda^{n-i}h(T^i(\xi_0))\right)^2$$
$$- 2\sum_{i=0}^{n}h(T^i(\xi_0))\sum_{i=0}^{n}\lambda^{n-i}h(T^i(\xi_0)) \ .$$

Since, by assumption, the sequence $h \circ T^i$ is again uncorrelated we have

$$\int\left(\sum_{i=0}^{n}h(T^i(y))\right)^2\nu(\mathrm{d}y) = \sum_{i=0}^{n}\sum_{j=0}^{n}\int h(T^i(y))h(T^j(y))\nu(\mathrm{d}y)$$
$$= \sum_{i=0}^{n}\int h(T^i(y))^2\nu(\mathrm{d}y) = (n+1)\sigma^2 \ .$$

Analogous to the computation for the velocity variance

$$\int\left(\sum_{i=0}^{n}\lambda^{n-i}h(T^i(y))\right)^2\nu(\mathrm{d}y) = \sum_{i=0}^{n}\lambda^{2n-2i}\int h(T^i(y))^2\nu(\mathrm{d}y) = \frac{1-\lambda^{2n+2}}{1-\lambda^2}\sigma^2$$

and

$$\int\sum_{i=0}^{n}h(T^i(y))\sum_{j=0}^{n}\lambda^{n-j}h(T^j(y))\nu(\mathrm{d}y) = \sum_{i=0}^{n}\sum_{j=0}^{n}\lambda^{n-j}\int h(T^i(y))h(T^j(y))\nu(\mathrm{d}y)$$
$$= \sum_{i=0}^{n}\lambda^{n-i}\int h(T^i(y))^2\nu(\mathrm{d}y) = \frac{1-\lambda^{n+1}}{1-\lambda}\sigma^2 \ .$$

Consequently, if $x_0 = v_0 = 0$ then

$$\int x_n^2\nu(\mathrm{d}y) = \frac{\kappa^2\sigma^2}{\gamma^2}\left(n+1+\frac{1-\lambda^{2n+2}}{1-\lambda^2} - 2\,\frac{1-\lambda^{n+1}}{1-\lambda}\right) \ . \tag{3.31}$$

Theorem 3.4. *Let $(Y, \mathcal{B}, \nu)$ be a normalized measure space, $T : Y \to Y$ be a measurable map such that T preserves the measure ν, let $\sigma > 0$ be a constant, and $h \in L^2(Y, \mathcal{B}, \nu)$ be such that $\int h(y)\nu(\mathrm{d}y) = 0$. Then*

$$\frac{\sum_{i=0}^{n} h(T^i(\xi_0))}{\sqrt{n}} \to^{\mathrm{d}} \mathrm{N}(0, \sigma^2) \tag{3.32}$$

if and only if

$$\frac{x_n}{\sqrt{n}} \to^{\mathrm{d}} \mathrm{N}\left(0, \frac{\kappa^2\sigma^2}{\gamma^2}\right) . \tag{3.33}$$

Proof. Assume that Condition 3.32 holds. From Eq. (3.30) we obtain

$$\frac{x_n}{\sqrt{n}} = \frac{x_0}{\sqrt{n}} + \frac{(1-\lambda^n)}{\gamma}\frac{v_0}{\sqrt{n}} + \frac{\kappa}{\gamma\sqrt{n}}\sum_{i=0}^{n} h(T^i(\xi_0)) - \frac{\kappa}{\gamma\sqrt{n}}\sum_{i=0}^{n} \lambda^{n-i}h(T^i(\xi_0)) .$$

By assumption

$$\frac{\kappa}{\gamma\sqrt{n}}\sum_{i=0}^{n} h(T^i(\xi_0)) \to^{\mathrm{d}} \frac{\kappa}{\gamma}\mathrm{N}(0, \sigma^2) .$$

Thus the result will follow when we show that the remaining terms are convergent in probability to zero. The first term

$$\frac{x_0}{\sqrt{n}} + \frac{(1-\lambda^n)}{\gamma}\frac{v_0}{\sqrt{n}}$$

is convergent a.s. to zero and consequently in probability. The sequence $\sum_{i=0}^{n} \lambda^{n-i}h(T^i(\xi_0))$ is convergent in distribution and $\frac{\kappa}{\gamma\sqrt{n}} \to 0$ as $n \to \infty$. Consequently, the sequence

$$\frac{\kappa}{\gamma\sqrt{n}}\sum_{i=0}^{n} \lambda^{n-i}h(T^i(\xi_0))$$

is convergent in probability to zero which completes the proof. The proof of the converse is analogous. □

Remark 3.5. Observe that if the transformation T is exact and ξ_0 is distributed according to a measure absolutely continuous with respect to ν, then the conclusion of Theorem 3.4 still holds.

Theorem 3.4 generalizes the results of Chew and Ting (2002). In Section 2.3.2 we have discussed when Condition 3.32 holds for a given ergodic transformation.

Remark 3.6. Note that if we multiply Gaussian distributed random variable $\mathrm{N}(0, 1)$ by a positive constant c, then it becomes Gaussian distributed $\mathrm{N}(0, c^2)$ with variance c^2. Thus if we multiply both sides of Eq. (3.33) by $\frac{1}{\sqrt{\tau}}$, we obtain

$$\frac{x(n\tau)}{\sqrt{n\tau}} \to^{\mathrm{d}} \mathrm{N}\left(0, \frac{\kappa^2\sigma^2}{\tau\gamma^2}\right) .$$

So if $\kappa = \sqrt{\gamma m\tau}$, as in Chew and Ting (2002), then

$$\frac{\kappa^2\sigma^2}{\tau\gamma^2} = \frac{m\sigma^2}{\gamma} .$$

What we have demonstrated in this section for the case of a linear frictional term in the velocity equation is quite interesting. Namely, the velocity will converge in distribution (but not necessarily to a Gaussian) and the position variable will not. However, given the proper scaling, as we have shown in Theorem 3.4, there will be convergence of the position variable.

3.3. Identifying the limiting velocity distribution

In Remark 3.2 we noted that although Corollary 3.1 guaranteed the existence of a unique limiting invariant measure, we could not conclude in general that it had a density absolutely continuous with respect to Lebesgue measure. In this section we consider a concrete example in which the perturbations ξ are taken from iterates of the dyadic map of Example 2.7 and analytically construct the density of the limiting measure. We further illustrate the temporal approach of the densities to this invariant density graphically.

Let $Y \subset \mathbb{R}$ be an interval, $\mathscr{B} = \mathscr{B}(Y)$, and $T : Y \to Y$ be a transformation preserving a normalized measure ν on $(Y, \mathscr{B}(Y))$. Recall from Section 3.1 that μ_* is the limiting measure for the sequence of random variables $(v(n\tau))$ starting from $v_0 \equiv 0$, i.e.

$$v(n\tau) = \kappa \sum_{i=0}^{n} \lambda^{n-i} h(\xi_i) ,$$

where $h : Y \to \mathbb{R}$ is a given integrable function, $0 < \lambda < 1$, $\xi_i = T^i(\xi_0)$, and ξ_0 is distributed according to ν.

Proposition 3.7. *Let $Y = [a, b]$ and $h : Y \to \mathbb{R}$ be a bounded function. Then the limiting measure μ_* has moments of all order given by*

$$\int x^k \mu_*(\mathrm{d}x) = \lim_{n\to\infty} \int v(n\tau)^k \nu(\mathrm{d}y) \tag{3.1}$$

and the characteristic function of μ_* is of the form

$$\phi_*(r) = \sum_{k=0}^{\infty} \frac{(\mathrm{i}r)^k}{k!} \int x^k \mu_*(\mathrm{d}x), \quad r \in \mathbb{R} .$$

Moreover,

$$\mu_*\left(\left[-\frac{\kappa c}{1-\lambda}, \frac{\kappa c}{1-\lambda}\right]\right) = 1 ,$$

where $c = \sup_{y\in Y} |h(y)|$.

Proof. Since h is bounded, we have

$$|v(n\tau)|^k \leqslant \left(\frac{\kappa c}{1-\lambda}\right)^k, \quad n, \ k \geqslant 0 .$$

The existence and convergence of moments now follow from Theorems 5.3 and 5.4 of Billingsley (1968). Since $v(n\tau)$ has all its values in the interval

$$\left[-\frac{\kappa c}{1-\lambda}, \frac{\kappa c}{1-\lambda}\right]$$

we obtain

$$\mu_n\left(\left[-\frac{\kappa c}{1-\lambda}, \frac{\kappa c}{1-\lambda}\right]\right) = 1 ,$$

where μ_n is the distribution of $v(n\tau)$. Convergence in distribution (Billingsley, 1968, Theorem 2.1) implies that

$$\limsup_{n\to\infty} \mu_n(F) \leqslant \mu_*(F)$$

for all closed sets. Therefore

$$\mu_*\left(\left[-\frac{\kappa c}{1-\lambda}, \frac{\kappa c}{1-\lambda}\right]\right) = 1 .$$

The formula for the characteristic function is a consequence of the other statements. □

Remark 3.8. Note that if the characteristic function of μ_* is integrable, then μ_* has a continuous and bounded density which is given by

$$f_*(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp(-\mathrm{i}xr)\phi_*(r)\,\mathrm{d}r, \quad x \in \mathbb{R} .$$

On the other hand if μ_* has a density then $\phi_*(r) \to 0$ as $|r| \to \infty$.

Note also that if a density exists then it must be zero outside the interval

$$\left[-\frac{\kappa c}{1-\lambda}, \frac{\kappa c}{1-\lambda}\right] .$$

Let $Y = [a, b]$ be an interval and $h(y) = y$ for $y \in Y$. Thus h is bounded and for this choice of h all moments of the corresponding limiting distribution μ_* exist by Proposition 3.7. However, the calculation might be quite tedious. We are going to determine the measure μ_* for a specific example of the transformation T by using a different method.

Let $h_n : Y \to \mathbb{R}$ be defined by

$$h_n(y) = \sum_{i=0}^{n} \lambda^{n-i} T^i(y), \quad y \in Y, \ n \geqslant 0 . \tag{3.2}$$

Then $v(n\tau) = \kappa h_n(\xi_0)$ and $v_n = \kappa\lambda h_{n-1}(\xi_0)$. Thus knowing the limiting distribution for these sequences is equivalent to knowing the limiting distribution for $h_n(\xi_0)$.

3.3.1. Dyadic map

To give a concrete example for which many of the preceding considerations can be completely illustrated, consider the generalized dyadic map defined by Eq. (2.10):

$$T(y) = \begin{cases} 2y+1, & y \in [-1, 0] \\ 2y-1, & y \in (0, 1] . \end{cases}$$

Proposition 3.9. *Let ξ, ξ_0 be random variables uniformly distributed on $[-1, 1]$. Let (α_k) be a sequence of independent random variables taking values drawn from $\{-1, 1\}$ with equal probability. Assume that ξ is statistically independent of the sequence (α_k). Then for every $\lambda \in (0, 1)$*

$$h_n(\xi_0) \to^{\mathrm{d}} \frac{1}{2-\lambda}\left(\xi + \sum_{k=0}^{\infty} \lambda^k \alpha_{k+1}\right) . \tag{3.3}$$

Proof. The random variable

$$\xi_0 = \sum_{k=1}^{\infty} \frac{\alpha_k}{2^k}$$

is uniformly distributed on $[-1, 1]$. It is easily seen that for the dyadic map

$$\xi_i = T^i(\xi_0) = \sum_{k=1}^{\infty} \frac{\alpha_{k+i}}{2^k} \quad \text{for } i \geqslant 1 . \tag{3.4}$$

Using this representation we obtain

$$\sum_{i=0}^{n-1} \lambda^{n-1-i}\xi_i = \sum_{i=0}^{n-1} \lambda^{n-1-i} \sum_{k=1}^{n-i} \frac{\alpha_{k+i}}{2^k} + \sum_{i=0}^{n-1} \lambda^{n-1-i} \sum_{k=n-i+1}^{\infty} \frac{\alpha_{k+i}}{2^k} .$$

Changing the order of summation leads to

$$\sum_{k=1}^{n}\left(\sum_{i=1}^{k}\frac{\lambda^{n-1-k+i}}{2^i}\right)\alpha_k+\sum_{i=0}^{n-1}\frac{\lambda^{n-1-i}}{2^{n-i}}\sum_{k=1}^{\infty}\frac{\alpha_{k+n}}{2^k}$$
$$=\frac{1}{2-\lambda}\sum_{k=1}^{n}\lambda^{n-k}\left(1-\left(\frac{\lambda}{2}\right)^k\right)\alpha_k+\frac{1-\left(\frac{\lambda}{2}\right)^n}{2-\lambda}\xi_n\ .$$

Consequently

$$\sum_{i=0}^{n-1}\lambda^{n-1-i}\xi_i=\frac{1}{2-\lambda}\sum_{k=1}^{n}\lambda^{n-k}\alpha_k+\frac{1}{2-\lambda}\xi_n-\lambda^n w_n\ ,$$

where

$$w_n=\frac{1}{2-\lambda}\left[\sum_{k=1}^{n}\left(\frac{1}{2}\right)^k\alpha_k+\left(\frac{1}{2}\right)^n\xi_n\right]\ .$$

This gives

$$h_{n-1}(\xi_0)+\lambda^n w_n=\frac{1}{2-\lambda}\left(\sum_{k=1}^{n}\lambda^{n-k}\alpha_k+\xi_n\right)\ . \tag{3.5}$$

Note that for every n we have $|w_n|\leqslant 2$. Therefore $\lambda^n w_n$ is a.s. convergent to 0 as $n\to\infty$. Since $h_n(\xi_0)$ converges in distribution, say to $\widetilde{\mu}_*$, we have $h_{n-1}(\xi_0)+\lambda^n w_n\to^{\mathrm{d}}\widetilde{\mu}_*$ and the random variables on the right-hand side of Eq. (3.5) converge in distribution to $\widetilde{\mu}_*$. Since the random variables α_k are independent, the random variables $\sum_{k=1}^{n}\lambda^{n-k}\alpha_k$ and ξ_n are also independent for every n. The same is true for $\sum_{k=0}^{n-1}\lambda^k\alpha_{k+1}$ and ξ. Moreover, $\xi_n+\sum_{k=1}^{n}\lambda^{n-k}\alpha_k$ and $\xi+\sum_{k=0}^{n-1}\lambda^k\alpha_{k+1}$ have identical distributions. Thus

$$\frac{1}{2-\lambda}\left(\xi+\sum_{k=0}^{n-1}\lambda^k\alpha_{k+1}\right)\to^{\mathrm{d}}\mu_*\ .$$

On the other hand $\sum_{k=0}^{n-1}\lambda^k\alpha_{k+1}\to\sum_{k=0}^{\infty}\lambda^k\alpha_{k+1}$ almost surely as $n\to\infty$, but this implies convergence in distribution. $\square$

Before stating our next result, we review some of the known properties of the random variable which appears in Eq. (3.3). For every $\lambda\in(0,1)$ let

$$\zeta_\lambda=\sum_{k=0}^{\infty}\lambda^k\alpha_{k+1}\ , \tag{3.6}$$

and let ϱ_λ be the distribution function of ζ_λ, $\varrho_\lambda(x)=\Pr\{\zeta_\lambda\leqslant x\}$ for $x\in\mathbb{R}$. Explicit expressions for ϱ_λ are, in general, not known. The measure induced by the distribution ϱ_λ is called an infinitely convolved Bernoulli measure (see Peres et al., 2000 for the historical background and recent advances).

It is known (Jessen and Wintner, 1935) that ϱ_λ is continuous and it is either absolutely continuous or singular. Recall that x is a point of increase of ϱ_λ if $\varrho_\lambda(x-\epsilon)<\varrho_\lambda(x+\epsilon)$ for all $\epsilon>0$. The set of points of increase of ϱ_λ (Kershner and Wintner, 1935) is either the interval $[-\frac{1}{1-\lambda},\frac{1}{1-\lambda}]$ when $\lambda\geqslant\frac{1}{2}$ or a Cantor set K_λ of zero Lebesgue measure contained in this interval when $\lambda<\frac{1}{2}$, ϱ_λ is always singular for $\lambda\in(0,\frac{1}{2})$ and the Cantor set K_λ satisfies $K_\lambda=(\lambda K_\lambda+1)\cup(\lambda K_\lambda-1)$ and $\frac{1}{1-\lambda},-\frac{1}{1-\lambda}\in K_\lambda$. Wintner (1935) noted that ϱ_λ has the uniform density $\rho_\lambda(x)=\frac{1}{4}1_{[-2,2]}(x)$ for $\lambda=\frac{1}{2}$ and that it is absolutely continuous for the kth roots of $\frac{1}{2}$. Thus it was suspected that ϱ_λ is absolutely continuous for all $\lambda\in[\frac{1}{2},1)$.

However, Erdös (1939a) showed that ϱ_λ is singular for

$$\lambda = \frac{\sqrt{5}-1}{2}$$

and for the reciprocal of the so-called Pisot numbers in $(1, 2)$. Later Erdös (1939b) showed that there is a $\beta < 1$ such that for almost all $\lambda \in (\beta, 1)$ the measure ϱ_λ is absolutely continuous. Only recently, Solomyak (1995) showed that $\beta = \frac{1}{2}$.

Proposition 3.10. *Let $\kappa > 0$ and $\lambda \in (0, 1)$. Define*

$$a_\lambda = \frac{\kappa}{1-\lambda} \quad \textit{and} \quad b_\lambda = \frac{\kappa\lambda}{(2-\lambda)(1-\lambda)} \tag{3.7}$$

Then the density f_*^λ of the limiting measure μ_*^λ of $v(n\tau)$ satisfies

$$f_*^\lambda(v) = 0 \textit{ if and only if } |v| \geqslant a_\lambda \ .$$

Moreover, on the interval $(-a_\lambda, a_\lambda)$ we have

$$f_*^\lambda(v) = \begin{cases} \frac{2-\lambda}{2\kappa}\, \varrho_\lambda\left(\frac{2-\lambda}{\kappa}\, v + 1\right), & -a_\lambda < v < -b_\lambda \ , \\ \frac{2-\lambda}{2\kappa}\left[\varrho_\lambda\left(\frac{2-\lambda}{\kappa} v + 1\right) - \varrho_\lambda\left(\frac{2-\lambda}{\kappa} v - 1\right)\right], & -b_\lambda \leqslant v \leqslant b_\lambda \ , \\ \frac{2-\lambda}{2\kappa}\left[1 - \varrho_\lambda\left(\frac{2-\lambda}{\kappa} v - 1\right)\right], & b_\lambda < v < a_\lambda \ , \end{cases} \tag{3.8}$$

where ϱ_λ is the distribution function of ζ_λ defined by Eq. (3.6).

Proof. Recall that we have $v(n\tau) = h_n(\xi_0)$. By Proposition 3.9, the sequence $(2-\lambda)h_n(\xi_0)$ converges in distribution to $\xi + \zeta_\lambda$ and the random variables ξ and ζ_λ are statistically independent. Since ξ has the uniform density on $[-1, 1]$ and ϱ_λ is continuous, the density of $\xi + \zeta_\lambda$ is given by

$$\begin{aligned}\int_{-\infty}^{\infty} \frac{1}{2} 1_{[-1,1]}(x-z)\, \mathrm{d}\rho_\lambda(z) &= \int_{-\infty}^{\infty} \frac{1}{2} 1_{[x-1,x+1]}(z)\, \mathrm{d}\rho_\lambda(z) \\ &= \int_{x-1}^{x+1} \frac{1}{2}\, \mathrm{d}\rho_\lambda(z) = \frac{1}{2}\left(\varrho_\lambda(x+1) - \rho_\lambda(x-1)\right) \ .\end{aligned}$$

Since $v(n\tau)$ converges in distribution to $\kappa/(2-\lambda)(\xi + \zeta_\lambda)$, it follows that μ_*^λ is the distribution of $\kappa/(2-\lambda)(\xi + \zeta_\lambda)$. Thus μ_*^λ has a density given by

$$f_*^\lambda(v) = \frac{2-\lambda}{2\kappa}\left(\varrho_\lambda\left(\frac{2-\lambda}{\kappa} v + 1\right) - \varrho_\lambda\left(\frac{2-\lambda}{\kappa} v - 1\right)\right), \quad v \in \mathbb{R} \ .$$

Consequently, $f_*^\lambda(v) = 0$ if and only if

$$\varrho_\lambda\left(\frac{2-\lambda}{\kappa} v + 1\right) = \varrho_\lambda\left(\frac{2-\lambda}{\kappa} v - 1\right) \ . \tag{3.9}$$

Since ϱ_λ is nondecreasing, it must be constant outside the set of points on which it is increasing, which is contained in the interval $[-\frac{1}{1-\lambda}, \frac{1}{1-\lambda}]$. Hence $\varrho_\lambda(x) = 0$ for $x \leqslant -\frac{1}{1-\lambda}$ and $\varrho_\lambda(x) = 1$ for $x \geqslant \frac{1}{1-\lambda}$. Therefore if $|v| \geqslant \frac{\kappa}{1-\lambda}$, then $f_*^\lambda(v) = 0$ and Eq. (3.8) follows. If $\lambda \geqslant \frac{1}{2}$ the function ϱ_λ is increasing on the interval $[-\frac{1}{1-\lambda}, \frac{1}{1-\lambda}]$, thus f_*^λ is positive on $[-\frac{\kappa}{1-\lambda}, \frac{\kappa}{1-\lambda}]$. Now let $\lambda < \frac{1}{2}$. Since $\frac{1}{1-\lambda}, -\frac{1}{1-\lambda} \in K_\lambda$, we also have $\frac{2\lambda-1}{1-\lambda}, \frac{1-2\lambda}{1-\lambda} \in K_\lambda$ and they divide the interval $[-\frac{1}{1-\lambda}, \frac{1}{1-\lambda}]$ into three intervals of length 2, $2\frac{1-2\lambda}{1-\lambda}$, and 2 respectively. Since the middle interval has length less than 2 and the distance between the points $\frac{2-\lambda}{\kappa} v + 1$ and $\frac{2-\lambda}{\kappa} v - 1$ is always 2, the result follows in this case as well. □

Remark 3.11. If $\lambda = \frac{1}{2}$ then $\varrho_\lambda(x) = \frac{1}{2}(x+2)$, $2-\lambda = \frac{3}{2}$ and the density is equal to

$$f_*^\lambda(v) = \begin{cases} \frac{9}{32\kappa}(v+2), & -2\kappa < v < -\frac{2}{3}\kappa\ , \\ \frac{3}{8\kappa}, & |v| \leqslant \frac{2}{3}\kappa\ , \\ \frac{9}{32\kappa}(2-v), & \frac{2}{3}\kappa < v < 2\kappa\ . \end{cases} \tag{3.10}$$

Remark 3.12. The invariant measure for the baker transformation S_β of Example 2.8 is the product of the distribution of $(1-\lambda)\zeta_\lambda$ and the normalized Lebesgue measure. Thus, in this example the limiting measure for v_n is the distribution of $(1-\lambda)\zeta_\lambda$, which may be either singular or absolutely continuous.

3.3.2. Graphical illustration of the velocity density evolution with dyadic map perturbations

What is the probability density function of $h_n(\xi_0)$ defined by Eq. (3.2) when ξ_0 is distributed according to v? For many maps, including the dyadic map example being considered here, this can be calculated analytically. This is the subject of this section.

Let Y be an interval and let v have a density g_* with respect to Lebesgue measure. Then the distribution of $h_n(\xi_0)$ is given by

$$\Pr\{h_n(\xi_0) \in A\} = \Pr\{\xi_0 \in h_n^{-1}(A)\} = \int_{h_n^{-1}(A)} g_*(y)\,\mathrm{d}y, \quad A \in \mathscr{B}(\mathbb{R})\ .$$

To obtain the density of $h_n(\xi_0)$ with respect to the Lebesgue measure, one has to write the last integral as $\int_A g_n(x)\,\mathrm{d}x$ for some nonnegative function g_n. If the map $h_n : Y \to \mathbb{R}$ is nonsingular with respect to Lebesgue measure, then the Frobenius–Perron operator $P_{h_n} : L^1([-1,1]) \to L^1(\mathbb{R})$ for h_n exists and $g_n = P_{h_n} g_*$.

Let $Y = [-1,1]$ and let T be the dyadic map. Remember that P_T has the uniform invariant density $g_*(y) = \frac{1}{2}1_{[-1,1]}(y)$. Since h_n is a linear function on each interval $(\frac{k}{2^n}, \frac{k+1}{2^n}]$ with constant derivative, say h'_n, we have

$$g_n(v) = P_{h_n} g_*(v) = \frac{1}{2h'_n} \sum_{k=-2^n}^{2^n-1} 1_{h_n((\frac{k}{2^n}, \frac{k+1}{2^n}])}(v), \quad v \in \mathbb{R}\ . \tag{3.11}$$

The derivative h'_n satisfies the recurrence equation $h'_n = \lambda^{n-1} + 2h'_{n-1}$, $n \geqslant 1$ and $h'_0 = 0$ and is equal to

$$\frac{(\lambda-3)2^n + \lambda^n}{\lambda - 2}$$

for each n.

In Fig. 3.1 we show the evolution of the velocity densities g_n when T is the dyadic map for two different values of λ. For $\lambda = \frac{1}{2}$ (left hand panels) the density rapidly (by $n = 8$) approaches the analytic form given in Eq. (3.10). On the right hand side, for $\lambda = 0.8$ the velocity densities have, by $n = 8$, approached a Gaussian-like form but supported on a finite interval. In both cases the support of the limiting densities is in agreement with Proposition 3.10. Fig. 3.2 shows $g_8(v)$ for six different values of λ.

3.3.3. r-dyadic map

Let $r \geqslant 2$ be an integer. Consider the r-dyadic transformation on the interval [0, 1]

$$T(y) = ry \pmod 1, \quad y \in [0,1]\ .$$

The proof of Proposition 3.9 carries over to this transformation when (α_k) is a sequence of independent random variables taking values in $\{0, 1, \ldots, r-1\}$ with equal probabilities, i.e. $\Pr(\alpha_k = i) = \frac{1}{r}$, $i = 0, 1, \ldots, r-1$, and ξ is a random variable uniformly distributed on [0,1] and independent of the sequence of random variables (α_k). Then the limiting measure for v_n is the distribution of the random variable

$$\frac{\kappa}{r-\lambda}\left(\xi + \sum_{k=0}^{\infty} \lambda^k \alpha_{k+1}\right)\ .$$

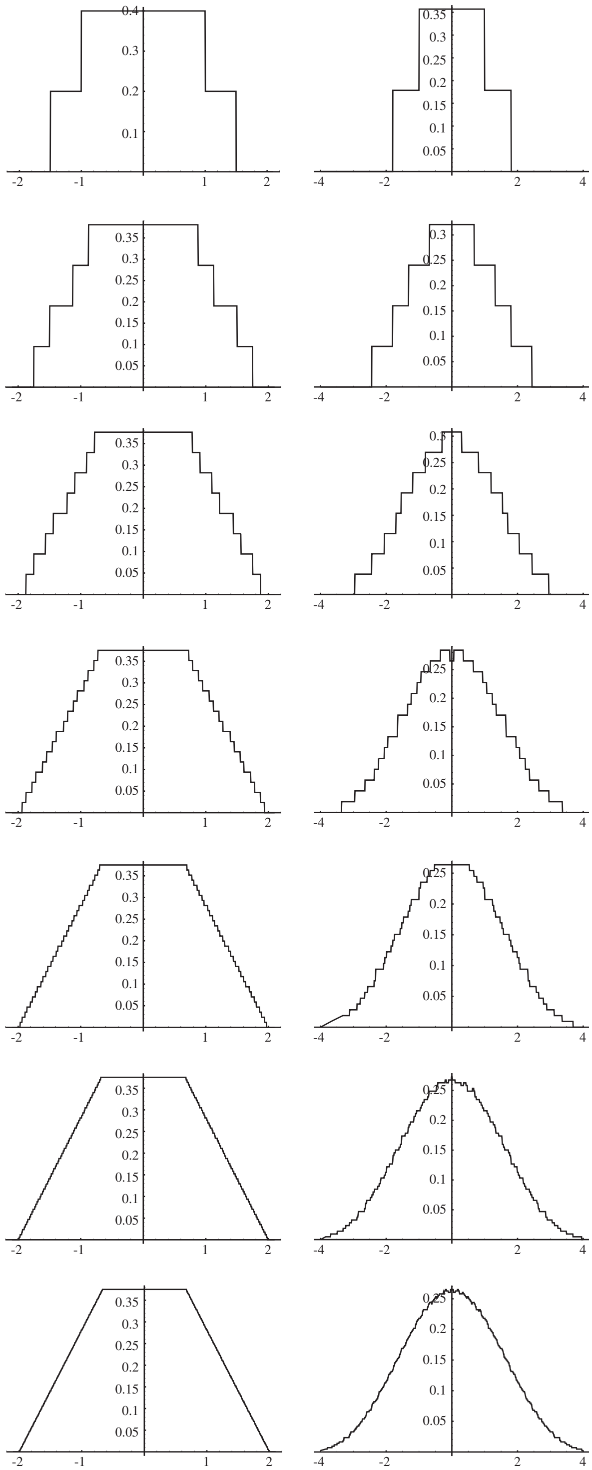

Fig. 3.1. This figure shows the successive evolution of the densities g_n, $n = 2, \ldots, 8$ of the velocity under perturbations from the dyadic map. In the left hand series of densities, $\lambda = \frac{1}{2}$, while on the right $\lambda = 0.8$. When $\lambda = \frac{1}{2}$ the densities g_n rapidly approach the limiting analytic form $f_*^{\frac{1}{2}}$ given in Eq. (3.10). In both cases, $\kappa = 1$.

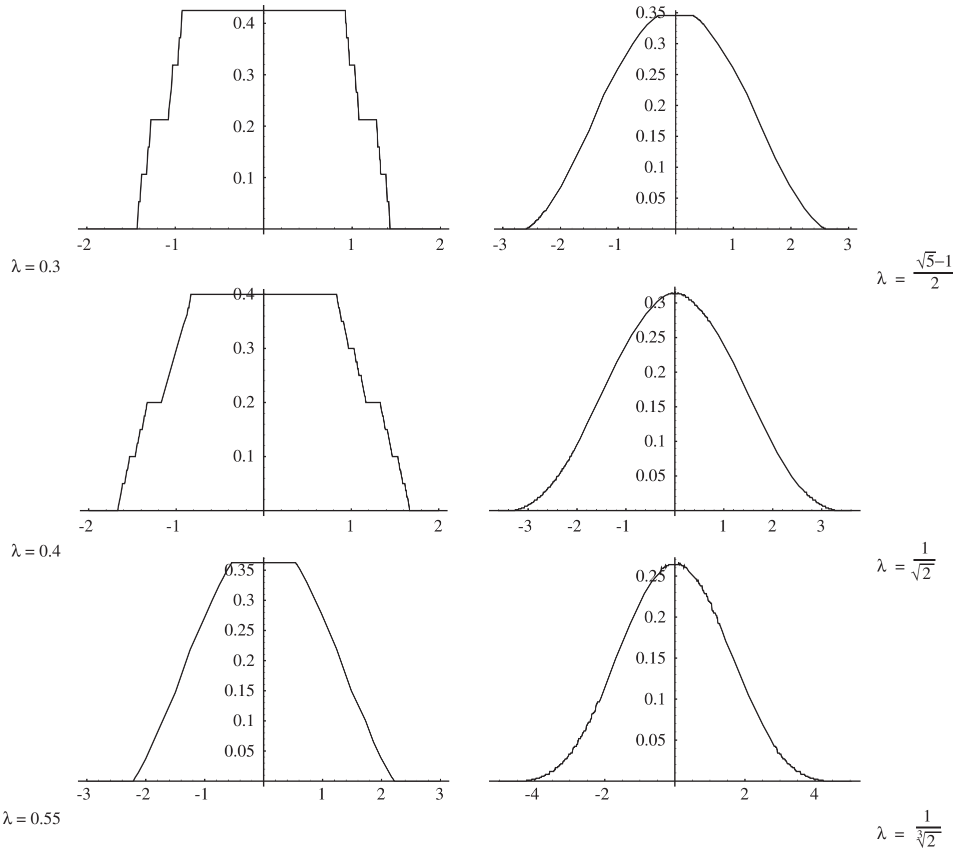

Fig. 3.2. This figure illustrates the form of the density $g_8(v)$ for perturbations coming from the dyadic map, as computed from Eq. (3.11), and various values of λ as indicated, with $\kappa = 1$. In every case the initial velocity density was the uniform invariant density of the dyadic map.

4. Gaussian behavior in the limit $\tau \to 0$

In the last section we have shown for the concrete example of the dyadic map that when κ is independent of τ then the density of the limiting invariant measure is not Gaussian though it may, on occasion, have a Gaussian-like appearance (cf. Figs. 3.1 and 3.2).

In this section we consider the possibility that κ depends on τ, $\kappa(\tau)$, and explore the type of dependence required to ensure a true limiting invariant measure with a Gaussian density. That is, we show how to truly converge to an Ornstein–Uhlenbeck process and thus reproduce Brownian motion from a deterministic process.

In a series of papers Beck and Roepstorff (1987), Beck (1990b), Beck (1990a), Beck (1996) and Hilgers and Beck (1999), motivated by questions related to alternative interpretations of Brownian motion, have numerically examined the dynamic character of the iterates of dynamical systems of the form

$$u_{n+1} = \lambda u_n + \sqrt{\tau} y_n, \quad \lambda \equiv \mathrm{e}^{-\gamma\tau} \tag{4.1}$$

$$y_{n+1} = T(y_n) , \tag{4.2}$$

in which T is a 'chaotic' mapping and τ is a small temporal scaling parameter. They refer to these systems as linear Langevin systems, and point out that they arise from the following

$$\dot{u} = -\gamma u + \sqrt{\tau} \sum_{i=1}^{\infty} y_{i-1}\delta(t - \mathrm{i}\tau) .$$

Integrating this equation one obtains Eq. (4.1) with $u_n = u(n\tau)$.

For situations in which the map T is selected from the class of Chebyshev maps [cf. Eq. (2.8) and Adler and Rivlin (1964)], Hilgers and Beck (2001) have provided abundant numerical evidence that the density of the distribution of the sequence of iterates $\{u_i\}_{i=0}^N$, for N quite large, may be approximately normal, or Gaussian, as $\lambda \to 1$, and Hilgers and Beck (1999) have provided some of the same type of numerical evidence for perturbations coming from the dyadic and tent maps. Our results provide the analytic basis for these observations.

In this section we consider and answer the question when one can obtain Gaussian processes by studying appropriate scaling limits of the velocity and position variables. We first recall what is meant by a Gaussian process.

An $\mathbb{R}$-valued stochastic process $\{\zeta(t); t \in (0,\infty)\}$ is called Gaussian if, for every integer $l \geqslant 1$ and real numbers $0 < t_1 < t_2 < \cdots < t_l < \infty$ the random vector $(\zeta(t_1), \ldots, \zeta(t_l))$ has a joint normal distribution or equivalently, for all $d_j \in \mathbb{R}$, $j = 1, \ldots, k$, the random variable $\sum_{j=1}^{l} d_j \zeta(t_j)$ is Gaussian. The finite dimensional distributions of a Gaussian process are completely determined by its first moment $m(t) = E(\zeta(t))$ and its covariance function

$$K_\zeta(t,s) = E(\zeta(t) - m(t))(\zeta(s) - m(s)), \quad s, t > 0.$$

If $m(t) \equiv 0$, $t > 0$ we say that ζ is a zero-mean Gaussian process. The initial random variable $\zeta(0)$ can be either identically equal to zero or can be any other random variable independent of the process $\{\zeta(t); t \in (0,\infty)\}$.

Now we recall the Ornstein–Uhlenbeck theory of Brownian motion for a free particle. The Ornstein–Uhlenbeck velocity process is a solution of the stochastic differential equation

$$\mathrm{d}V(t) = -\gamma V(t)\,\mathrm{d}t + \sigma_0\,\mathrm{d}w(t) ,$$

where w is a standard Wiener process, and the solution of this equation is

$$V(t) = \mathrm{e}^{-\gamma t} V(0) + \sigma_0 \int_0^t \mathrm{e}^{-\gamma(t-s)}\,\mathrm{d}w(s) .$$

In other words, V is an Ornstein–Uhlenbeck velocity process if ζ defined by $\zeta(t) = V(t) - \mathrm{e}^{-\gamma t} V(0)$, $t \geqslant 0$, is a zero-mean Gaussian process with covariance function

$$K_\zeta(t,s) = \frac{\sigma_0^2}{2\gamma}(\mathrm{e}^{2\gamma \min(t,s)} - 1)\mathrm{e}^{-\gamma(t+s)} . \tag{4.3}$$

If the initial random variable $V(0)$ has a normal distribution with mean zero and variance $\sigma_0^2/2\gamma$, then V itself is a stationary, zero-mean Gaussian process with covariance function

$$K_V(t,s) = \frac{\sigma_0^2}{2\gamma}\,\mathrm{e}^{-\gamma|t-s|} .$$

Let $X(t)$ denote the position of a Brownian particle at time t. Then

$$X(t) = X(0) + \int_0^t V(s)\,\mathrm{d}s .$$

In other words, X is an Ornstein–Uhlenbeck position process if η defined by $\eta(t) = X(t) - X(0) - (1 - \mathrm{e}^{-\gamma t})/\gamma V(0)$ is a zero-mean Gaussian process with covariance function

$$K_\eta(t,s) = \frac{\sigma_0^2}{2\gamma^3}(2\gamma \min(t,s) - 2 + 2\mathrm{e}^{-\gamma t} + 2\mathrm{e}^{-\gamma s} - \mathrm{e}^{-\gamma|t-s|} - \mathrm{e}^{-\gamma(t+s)}) .$$

In particular the variance of $\eta(t)$ is equal to $\sigma_0^2/2\gamma^3(2\gamma t - 3 + 4\mathrm{e}^{-\gamma t} - \mathrm{e}^{-2\gamma t})$.

Let $h \in L^2(Y, \mathscr{B}, \nu)$ be such that $\int h(y)\nu(\mathrm{d}y) = 0$. Assume that x_0, v_0, and ξ_0 are independent random variables on $(Y, \mathscr{B}, \nu)$ and ξ_0 is distributed according to ν. The solution of Eq. (3.17) is of the form

$$v(t) = \mathrm{e}^{-\gamma(t-n\tau)} v(n\tau), \quad t \in [n\tau, (n+1)\tau), \; n \geqslant 0 .$$

We indicate the dependence of $x(t)$ and $v(t)$ on τ by writing $x_\tau(t)$ and $v_\tau(t)$ respectively. Let $n=[\frac{t}{\tau}]$, where the notation $[\cdot]$ indicates the integer value of the argument, for $t \in [n\tau, (n+1)\tau)$, substitute $\lambda = \mathrm{e}^{-\gamma\tau}$, and use Eq. (3.26) to obtain

$$v_\tau(t) = \mathrm{e}^{-\gamma t} v_0 + \kappa(\tau)\mathrm{e}^{-\gamma t} \sum_{j=0}^{[\frac{t}{\tau}]} \mathrm{e}^{\gamma\tau j} h(T^j(\xi_0)), \quad t \geqslant 0 , \tag{4.4}$$

and Eq. (3.30) to obtain

$$x_\tau(t) = x_0 + \frac{1-\mathrm{e}^{-\gamma t}}{\gamma} v_0 + \frac{\kappa(\tau)}{\gamma} \left(\sum_{j=0}^{[\frac{t}{\tau}]} h(T^j(\xi_0)) - \mathrm{e}^{-\gamma t} \sum_{j=0}^{[\frac{t}{\tau}]} \mathrm{e}^{\gamma\tau j} h(T^j(\xi_0)) \right) . \tag{4.5}$$

Observe that the first moment of $v_\tau(t)$ is equal to

$$\int v_\tau(t) \nu(\mathrm{d}y) = \mathrm{e}^{-\gamma t} \int v_0(y)\nu(\mathrm{d}y) ,$$

since the random variables $h(T^j(\xi_0))$ have a first moment equal to 0. Assume for simplicity that $\int h(y)h(T^j(y))\nu(\mathrm{d}y) = 0$ for $j \geqslant 1$ and set $\sigma^2 = \int h^2(y)\nu(\mathrm{d}y)$. Since the random variables v_0 and $h(T^j(\xi_0))$ are independent, the second moment of $v_\tau(t)$ takes the form

$$\begin{aligned} \int v_\tau(t)^2 \nu(\mathrm{d}y) &= \mathrm{e}^{-2\gamma t} \int v_0^2(y)\nu(\mathrm{d}y) + [\kappa(\tau)]^2 \mathrm{e}^{-2\gamma t} \sum_{j=0}^{[\frac{t}{\tau}]} \mathrm{e}^{2\gamma\tau j} \int h^2(T^j(y))\nu(\mathrm{d}y) \\ &= \mathrm{e}^{-2\gamma t} \int v_0^2(y)\nu(\mathrm{d}y) + \sigma^2 [\kappa(\tau)]^2 \mathrm{e}^{-2\gamma t} \frac{1-\mathrm{e}^{2\gamma\tau([\frac{t}{\tau}]+1)}}{1-\mathrm{e}^{2\gamma\tau}} . \end{aligned}$$

If σ and γ do not depend on τ, we have

$$\lim_{\tau\to 0} \sigma^2 \mathrm{e}^{-2\gamma t} \left(\mathrm{e}^{2\gamma\tau([\frac{t}{\tau}]+1)} - 1 \right) \frac{\tau}{\mathrm{e}^{2\gamma\tau}-1} = \frac{\sigma^2}{2\gamma} (1 - \mathrm{e}^{-2\gamma t}) .$$

Hence the limit of the second moment of $v_\tau(t)$ as $\tau \to 0$ is finite and positive if and only if $\kappa(\tau)$ depends on τ in such a way that

$$\lim_{\tau\to 0} \frac{[\kappa(\tau)]^2}{\tau}$$

is finite and positive.

Beck and Roepstorff (1987) take $\kappa(\tau) = \sqrt{\tau}$ from the outset, and claim that in the limit $\tau \to 0$ the process $v_\tau(t)$ converges to the Ornstein–Uhlenbeck velocity process when the sequence $(h \circ T^j)$ has a so-called ϕ-mixing property[2] on the probability space $(Y, \mathscr{B}, \nu)$. In fact, the following result can be proved.

Theorem 4.1. *Let $(Y, \mathscr{B}, \nu)$ be a normalized measure space and $T : Y \to Y$ be ergodic with respect to ν. Let $h \in L^2(Y, \mathscr{B}, \nu)$ be such that $\sum_{n=0}^{\infty} \|\mathscr{P}_{T,\nu}^n h\|_2 < \infty$ and let*

$$\sigma = \left(\int h(y)^2 \nu(\mathrm{d}y) + 2 \sum_{n=1}^{\infty} \int h(y)h(T^n(y))\nu(\mathrm{d}y) \right)^{1/2}$$

---

[2] A sequence of random variables $\{\xi_j : j \geqslant 0\}$ is called ϕ-mixing if

$$\lim_{n\to\infty} \sup \left\{ \frac{|\Pr(A \cap B) - \Pr(A)\Pr(B)|}{\Pr(A)} : A \in \mathscr{F}_1^k, b \in \mathscr{F}_{k+n}^\infty, k \geqslant 1 \right\} = 0 ,$$

where $\mathscr{F}_1^k$ and $\mathscr{F}_{k+n}^\infty$ denote the σ-algebra generated by the random variables $\xi_1, \ldots, \xi_k$ and $\xi_{k+n}, \xi_{k+n+1}, \ldots$, respectively.

be positive. Assume that $\gamma > 0$ *and*

$$\lim_{\tau\to 0} \frac{[\kappa(\tau)]^2}{\tau} = 1 \ .$$

Then for each v_0 *the finite dimensional distributions of the velocity process* v_τ *given by Eq.* (4.4) *converge weakly as* $\tau \to 0$ *to the finite dimensional distributions of the Ornstein–Uhlenbeck velocity process V for which* $V(0) = v_0$ *and* $\sigma_0 = \sigma$.

Proof. By Theorem 2.23 we have $\mathscr{P}_{T,v}(\tilde{h}) = 0$ where $\tilde{h} = h + f - f \circ T$ and $f = \sum_{n=1}^{\infty} \mathscr{P}^n_{T,v} h$. For $t \geqslant 0$ and $\tau > 0$ define

$$\zeta_\tau(t) = \kappa(\tau) e^{-\gamma t} \sum_{j=0}^{[\frac{t}{\tau}]} e^{\gamma\tau j} \tilde{h} \circ T^j, \quad \tilde{\zeta}_\tau(t) = \kappa(\tau) e^{-\gamma t} \sum_{j=0}^{[\frac{t}{\tau}]} e^{\gamma\tau j} (f \circ T^{j+1} - f \circ T^j) \ .$$

Then

$$v_\tau(t) = e^{-\gamma t} v_0 + \zeta_\tau(t) + \tilde{\zeta}_\tau(t) \ .$$

Observe that

$$\tilde{\zeta}_\tau(t) = \kappa(\tau) e^{-\gamma t} (e^{\gamma\tau[\frac{t}{\tau}]} f \circ T^{[\frac{t}{\tau}]+1} - f) + \kappa(\tau) e^{-\gamma t} (e^{-\gamma\tau} - 1) \sum_{j=1}^{[\frac{t}{\tau}]} e^{\gamma\tau j} f \circ T^j \ .$$

Hence

$$\|\tilde{\zeta}_\tau(t)\|_2 \leqslant 2|\kappa(\tau)| e^{-\gamma t} (e^{\gamma\tau[\frac{t}{\tau}]} + 1) \|f\|_2 \ ,$$

and consequently

$$\|\tilde{\zeta}_\tau(t)\|_2 \leqslant 4|\kappa(\tau)| \|f\|_2, \quad t \geqslant 0, \quad \tau > 0 \ .$$

This and Lemma A.14 imply that the finite dimensional distributions of $v_\tau(t) - e^{-\gamma t} v_0$ converge weakly to the corresponding finite dimensional distributions of a zero mean Gaussian process ζ with $\zeta(0) = 0$ and the covariance function $K_\zeta(t, s)$ given by Eq. (4.3) where $\sigma_0^2 = \|\tilde{h}\|_2^2$. $\square$

For the corresponding position process we have the following.

Theorem 4.2. *Under the assumptions of Theorem* 4.1, *let* $V(0) = v_0$. *Then for each* x_0 *the finite dimensional distributions of the position process* x_τ *given by Eq.* (4.5) *converge weakly as* $\tau \to 0$ *to the finite dimensional distributions of the Ornstein–Uhlenbeck position process X for which* $X(0) = x_0$.

Proof. This follows from Lemma A.15 similarly as the preceding theorem follows from Lemma A.14. $\square$

Example 4.3. Let us apply Theorem 4.1 to a transformation $T : [-1, 1] \to [-1, 1]$ and $h(y) = y$. We have $\mathscr{P}_{T,v} h = 0$ when T is the tent map, Eq. (1.4). Then $\sigma^2 = 1/3$. Thus all assumptions of Theorem 4.1 are satisfied. When T is one of the Chebyshev maps (2.8) S_N we also have $\mathscr{P}_{T,v} h = 0$ by Example 2.16 and $\sigma^2 = 1/2$. For the dyadic map (2.10), the series $\sum_{n=1}^{\infty} \mathscr{P}^n_{T,v} h$ is absolutely convergent in $L^2([-1, 1], \mathscr{B}([-1, 1]), v)$ and is equal to h. This implies that $\sigma = \|2h - h \circ T\|_2$, thus in this case $\sigma = 1$, which can be easily calculated. Thus all of the numerical examples and studies of Beck and co-workers cited above are covered by this example.

5. Discussion

In this paper we were motivated by the strong statistical properties of discrete dynamical systems to consider when Brownian-motion-like behavior could emerge in a simple toy system. To illustrate this possibility, in the Introduction we have numerically illustrated how a simple system can produce behavior that has the visual appearance of an Ornstein–Uhlenbeck process.

To deal analytically with this problem, in Section 2 we have reviewed a number of different types of 'chaotic' temporal behaviors that dynamical systems may display, and also given the necessary background material from ergodic theory to allow one to examine these within the context of ensemble behavior. We have also reviewed and significantly extended a class of central limit theorems for discrete time maps in Section 2.3, and illustrated how a Wiener process can be produced from a deterministic system. Section 2.3.4 illustrates a number of the points developed in the previous parts of Section 2 with concrete examples.

In Section 3 we have turned to a consideration of the Langevin-like equations

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = v(t) \ ,$$
$$\frac{\mathrm{d}v(t)}{\mathrm{d}t} = -\gamma v(t) + \eta(t)$$

in which the perturbation $\eta(t)$ need not be a Gaussian noise but may be substituted by

$$\eta(t) = \kappa \sum_{n=0}^{\infty} h(\xi(t))\delta(t - n\tau)$$

with a highly irregular deterministic function $\xi(n\tau)$. When the variables $h(\xi(n\tau))$ are uncorrelated Gaussian distributed (thus in fact independent) random variables then the limiting distribution of $v(n\tau)$ is Gaussian. This need not be the case for the deterministic noise produced by perturbations derived from highly chaotic semi-dynamical systems. Specifically, in Section 3.1 we have shown that although the velocity will converge in distribution it will not necessarily be a Gaussian. Section 3.2.2 shows that the position variable will never converge. However, given the proper scaling of κ, as we have shown in Theorem 3.4, convergence of the position variable can be achieved. In Section 3.3 we have illustrated all of our results of Section 3.2 for the specific case of perturbations derived from the exact dyadic map.

In Section 4, in the limit $\tau \to 0$ the deterministic perturbations give precisely the same result as is customarily obtained using Gaussian distributed white noise, i.e. an Ornstein–Uhlenbeck process. Specifically, with respect to the question we posed in the Introduction (how Brownian-like motion can arise from a purely deterministic dynamics), the central and most important results of this paper are contained in Theorems 4.1 and 4.2. These two theorems give, for the first time we believe, sufficient conditions on the scaling of the noise amplitude κ with respect to the interval τ between perturbations for deterministic processes to reproduce an Ornstein–Uhlenbeck process in both velocity and position. Thus these results give the theoretical framework in which previous numerical results, e.g. by Beck and Roepstorff (1987); Beck (1990b, 1990a, 1996); Hilgers and Beck (1999); Chew and Ting (2002), can be understood.

The significance of these considerations is rather broad. It is the norm in experimental observations that any experimental variable that is recorded will be "contaminated" by "noise". Sometimes the distribution of this noise is approximately Gaussian, sometimes not. The considerations here illustrate quite specifically that the origins of the noise observed experimentally need not be due to the operation of a random process (random in the sense that there is no underlying physical law allowing one to predict exactly the future of the process based on the past). Rather, the results we present strongly suggest, as an alternative, that the fluctuations observed experimentally might well be the signature of an underlying deterministically chaotic process.

Acknowledgements

This work was supported by MITACS (Canada) and the Natural Sciences and Engineering Research Council (NSERC grant OGP-0036920, Canada). The first draft of this paper was written while the second author was visiting McGill University, whose hospitality and support are gratefully acknowledged. We thank Radosław Kamiński for providing the computations on which Fig. 1.3 was based. We would also like to acknowledge the very helpful comments of a referee, which substantially improved the final version of this paper.

Appendix A. Limit theorems for dependent random variables

This appendix reviews known general central limit theorems from probability theory which we can use directly in the context of dynamical systems. With their help we then prove several results which we have used in Sections 2.3 and 4.

Consider random variables arranged in a double array

$$\begin{array}{l} \zeta_{1,1}, \zeta_{1,2}, \ldots, \zeta_{1,k_1} \\ \zeta_{2,1}, \zeta_{2,2}, \ldots, \zeta_{2,k_2} \\ \qquad\vdots \\ \zeta_{n,1}, \zeta_{n,2}, \ldots, \zeta_{n,k_n} \\ \qquad\vdots \end{array} \tag{A.1}$$

with $k_n \to \infty$ as $n \to \infty$. We shall give conditions which imply that the row sums converge in distribution to a Gaussian random variable $\sigma\mathrm{N}(0,1)$, that is

$$\sum_{i=1}^{k_n} \zeta_{n,i} \to^{\mathrm{d}} \sigma\mathrm{N}(0,1)\ . \tag{A.2}$$

We shall require the Lindeberg condition

$$\lim_{n\to\infty} \sum_{i=1}^{k_n} E(\zeta_{n,i}^2 1_{\{|\zeta_{n,i}|>\epsilon\}}) = 0 \quad \text{for every } \epsilon > 0\ . \tag{A.3}$$

If independence in each row is allowed, then we have the classical Lindeberg–Feller theorem.

Theorem A.1 (Chung (2001, Theorem 7.2.1)). *Let the random variables* $\zeta_{n,1}, \ldots, \zeta_{n,k_n}$ *be independent for each n. Assume that* $E(\zeta_{n,i}) = 0$ *and* $\sum_{i=1}^{k_n} E(\zeta_{n,i}^2) = 1, n \geqslant 1$. *Then the Lindeberg condition holds if and only if*

$$\sum_{i=1}^{k_n} \zeta_{n,i} \to^{\mathrm{d}} \mathrm{N}(0,1)\ ,$$

and

$$for\ all\ \delta > 0, \qquad \max_{1\leqslant i\leqslant k_n} \Pr\{|\zeta_{n,i}| > \delta\} \to 0\ . \tag{A.4}$$

If one considers a map T on a probability space $(Y, \mathscr{B}, \nu)$ which preserves the measure ν, and defines $\zeta_{n,i}$ to be $(1/\sqrt{n}) f \circ T^{i-1}$ for $i = 1, \ldots, n$ with f measurable, then Condition A.4 holds. This is because

$$\Pr\{|\zeta_{n,i}| > \delta\} = \nu(\{y \in Y : |f(T^{i-1}(y))| > \delta\sqrt{n}\})\ ,$$

and $\{y \in Y : |f(T^{i-1}(y))| > \delta\sqrt{n}\} = T^{-i+1}(\{y \in Y : |f(y)| > \delta\sqrt{n}\})$. Thus by the invariance of ν this leads to

$$\max_{1\leqslant i\leqslant k_n} \Pr\{|\zeta_{n,i}| > \delta\} = \nu(\{y \in Y : |f(y)| > \delta\sqrt{n}\}) \to 0\ .$$

Similarly, if one takes a square integrable f, then the Lindeberg condition A.3 is satisfied. Indeed, we have

$$E(|\zeta_{n,i}|^2 1_{\{|\zeta_{n,i}|>\epsilon\}}) = \frac{1}{\sqrt{n}^2} \int_{\{z:|f(T^{i-1}(z))|\geqslant\sqrt{n}\epsilon\}} f^2(T^{i-1}(y))\nu(\mathrm{d}y)$$

and by the change of variables applied to T^{i-1} this reduces to

$$\frac{1}{n} \int_{\{|f|\geqslant\sqrt{n}\epsilon\}} f^2(y)\nu(\mathrm{d}y)\ .$$

Thus

$$\sum_{i=1}^{n} E(|\zeta_{n,i}|^2 1_{\{|\zeta_{n,i}|>\epsilon\}}) = \int_{\{|f|\geqslant\sqrt{n}\epsilon\}} f^2(y)\nu(\mathrm{d}y) \tag{A.5}$$

converges to 0 by the Lebesgue dominated convergence theorem and the assumption that f^2 is integrable.

Since our random variables are dependent, we cannot apply the above theorem. Instead, we use the notion of martingale differences for which there is a natural generalization of Theorem A.1. Moreover, additional assumptions are needed, as one can easily check that if T is the identity map, then

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} f\circ T^{i-1} = \frac{n}{\sqrt{n}} f$$

which cannot be convergent to a Gaussian random variable.

We first recall the definition of conditional expectation. Let a probability space $(\Omega, \mathscr{F}, \Pr)$ be given and let $\mathscr{G}$ be a sub-σ-algebra of $\mathscr{F}$. For $\zeta \in L^1(\Omega, \mathscr{F}, \Pr)$ there exists a random variable $E(\zeta|\mathscr{G})$, called the conditional expectation of ζ given $\mathscr{G}$, having the following properties: it is $\mathscr{G}$ measurable, integrable and satisfies the equation

$$\int_A E(\zeta|\mathscr{G})(\omega)\Pr(\mathrm{d}\omega) = \int_A \zeta(\omega)\Pr(\mathrm{d}\omega), \quad A\in\mathscr{G} .$$

The existence and uniqueness of $E(\zeta|\mathscr{G})$ for a given ζ follows from the Radon–Nikodym theorem. The transformation $\zeta \mapsto E(\zeta|\mathscr{G})$ is a linear operator between the spaces $L^1(\Omega, \mathscr{F}, \Pr)$ and $L^1(\Omega, \mathscr{G}, \Pr)$, so sometimes it is called an operator of conditional expectation.

Let $\{\zeta_{n,i} : 1\leqslant i\leqslant k_n, n\geqslant 1\}$ be a family of random variables defined on a probability space $(\Omega, \mathscr{F}, \Pr)$. For each $n\geqslant 1$, let a family $\{\mathscr{F}_{n,i} : i\geqslant 0\}$ of sub-σ-algebras of $\mathscr{F}$ be given. Consider the following set of conditions:

(1) $E(\zeta_{n,i}) = 0$ and $E(\zeta_{n,i}^2) < \infty$,
(2) $\mathscr{F}_{n,i-1} \subseteq \mathscr{F}_{n,i}$,
(3) $\zeta_{n,i}$ is $\mathscr{F}_{n,i}$ measurable,
(4) $E(\zeta_{n,i}|\mathscr{F}_{n,i-1}) = 0$ for each $1\leqslant i\leqslant k_n$, $n\geqslant 1$.

A family $\{\mathscr{F}_{n,i}, \zeta_{n,i} : 1\leqslant i\leqslant k_n, n\geqslant 1\}$ satisfying conditions (i)–(iv) is called a (square integrable) martingale differences array.

Theorem A.2 (Billingsley (1995, Theorem 35.12)). *Let* $\{\zeta_{n,i} : 1\leqslant i\leqslant k_n, n\geqslant 1\}$ *be a martingale difference array satisfying Lindeberg condition* A.3. *If*

$$\sum_{i=1}^{k_n} E(\zeta_{n,i}^2|\mathscr{F}_{n,i-1}) \to^P \sigma^2 , \tag{A.6}$$

where σ *is a nonnegative constant*, *then*

$$\sum_{i=1}^{k_n} \zeta_{n,i} \to^{\mathrm{d}} \sigma\mathrm{N}(0, 1) . \tag{A.7}$$

If the limit in Condition A.6 is a random variable instead of the constant σ^2, we obtain convergence to mixtures of normal distributions.

Theorem A.3 (Eagleson (1975, Corollary p. 561)). *Let* $\{\zeta_{n,i} : 1\leqslant i\leqslant k_n, n\geqslant 1\}$ *be a martingale difference array satisfying the Lindeberg condition* A.3. *If there exists an* $\mathscr{F}_\infty = \bigcap_{n=1}^{\infty}\mathscr{F}_{n,0}$*-measurable, a.s. positive and finite random variable* η *such that*

$$\sum_{i=1}^{k_n} E(\zeta_{n,i}^2|\mathscr{F}_{n,i-1}) \to^P \eta , \tag{A.8}$$

then $\sum_{i=1}^{k_n} \zeta_{n,i}$ *is convergent in distribution to a measure whose characteristic function is* $\varphi(r) = E(\exp(-\frac{1}{2}r^2\eta))$.

The above result shows that to obtain a normal distribution in the limit a specific normalization is needed and we have the following.

Theorem A.4 (Gaenssler and Joos (1992, Theorem 3.6)). *Let* $\{\zeta_{n,i} : 1 \leqslant i \leqslant k_n, n \geqslant 1\}$ *be a martingale difference array and let* η *be a real-valued random variable such that* $\Pr(0 < \eta < \infty) = 1$. *Suppose that*

$$\lim_{n\to\infty} E\left(\max_{1\leqslant i\leqslant k_n} |\zeta_{n,i}|\right) = 0 \tag{A.9}$$

and

$$\sum_{i=1}^{k_n} \zeta_{n,i}^2 \to^P \eta \ . \tag{A.10}$$

If η *is* $\mathscr{F}_{n,i}$*-measurable for each* n *and for each* $1 \leqslant i \leqslant k_n$, *then*

$$\frac{\sum_{i=1}^{k_n} \zeta_{n,i}}{\sqrt{\sum_{i=1}^{k_n} \zeta_{n,i}^2}} \to^{\mathrm{d}} \mathrm{N}(0, 1) \ .$$

Remark A.5. If η in Condition A.10 is constant, $\eta = \sigma^2$, then the conclusion of Theorem A.4 is equivalent to

$$\sum_{i=1}^{k_n} \zeta_{n,i} \to^{\mathrm{d}} \sigma\mathrm{N}(0, 1) \ .$$

Note also that Condition A.9 is implied by the Lindeberg condition.

The next result gives conditions for convergence of moments in Theorem A.2.

Theorem A.6 (Hall (1978, Theorem), Teicher (1988, Theorem 3)). *Let* $\{\zeta_{n,i} : 1 \leqslant i \leqslant k_n, n \geqslant 1\}$ *be a martingale difference array with* $\sum_{i=1}^{k_n} E(\zeta_{n,i}^2) = \sigma^2$ *where* $\sigma > 0$. *Suppose that for* $p > 1$

$$\sum_{i=1}^{k_n} E|\zeta_{n,i}|^{2p} \to 0 \quad \textit{and} \quad E\left|\sum_{i=1}^{k_n} E(\zeta_{n,i}^2|\mathscr{F}_{n,i-1}) - \sigma^2\right|^p \to 0 \ . \tag{A.11}$$

Then

$$\lim_{n\to\infty} E\left|\sum_{i=1}^{k_n} \zeta_{n,i}\right|^{2p} = E|\mathrm{N}(0, \sigma^2)|^{2p} \ .$$

Let $(Y, \mathscr{B}, \nu)$ be a normalized measure space and $T : Y \to Y$ be a measurable map such that T preserves the measure ν. Recall from Section 2 relation 2.3 between the transfer operator $\mathscr{P}_{T,\nu}$, the Koopman operator and the operator of conditional expectation which gives

$$\mathscr{P}_{T,\nu} \circ U_T f = f \quad \text{and} \quad U_T \circ \mathscr{P}_{T,\nu} f = E(f|T^{-1}(\mathscr{B})), \quad f \in L^1(Y, \mathscr{B}, \nu) \ .$$

Lemma A.7. *Let* $(Y, \mathscr{B}, \nu)$ *be a normalized measure space, and* $T : Y \to Y$ *be a measurable map such that* T *preserves the measure* ν. *Let* $\{c_{n,i} : 1 \leqslant i \leqslant k_n, n \geqslant 1\}$ *be a family of real numbers and* $h \in L^2(Y, \mathscr{B}, \nu)$. *Suppose that* $\mathscr{P}_{T,\nu} h = 0$. *Then*

$$\zeta_{n,i} = c_{n,i} h \circ T^{k_n - i}, \quad 1 \leqslant i \leqslant k_n, \qquad \zeta_{n,i} = 0, \quad i > k_n \ ,$$

with

$$\mathscr{F}_{n,i} = T^{-k_n+i}(\mathscr{B}), \quad 0 \leqslant i \leqslant k_n \quad and \quad \mathscr{F}_{n,i} = \mathscr{B}, \quad i > k_n, \quad n \geqslant 1$$

is a martingale difference array.

Moreover, if $c_{n,i} = \frac{1}{\sqrt{k_n}}$, $1 \leqslant i \leqslant k_n$ *then the following hold*

(i) *Lindeberg condition* A.3;
(ii) *Conditions* A.8 *and* A.10 *with* $\eta = E(h^2|\mathscr{I})$ *where* $\mathscr{I}$ *is the* σ*-algebra of all* T*-invariant sets*;
(iii) *Condition* A.6 *with* $\sigma^2 = \int h^2 d\nu$ *provided that* T *is ergodic*;
(iv) *Condition* A.11 *for every* $p > 1$ *provided that* T *is ergodic and* $h \in L^\infty(Y, \mathscr{B}, \nu)$.

Proof. To check conditions (2), and (3) of the definition of a martingale difference array, observe that $T^{-j-1}(\mathscr{B}) \subset T^{-j}(\mathscr{B})$ and $h \circ T^j$ is $T^{-j}(\mathscr{B})$ measurable. The Koopman and transfer operators for the iterated map T^j are just the jth iterates of the operators U_T and $\mathscr{P}_{T,\nu}$. From this and Eq. (2.3) we have $\mathscr{P}^j_{T,\nu} U^j_T h = h$ and

$$E(h \circ T^j | T^{-j-1}(\mathscr{B})) = U_T^{j+1} \mathscr{P}_{T,\nu}^{j+1} (h \circ T^j) = U_T^{j+1} \mathscr{P}_{T,\nu} h \ .$$

Since $\mathscr{P}_{T,\nu} h = 0$, we see that $E(h \circ T^j | T^{-j-1}(\mathscr{B})) = 0$ for $j \geqslant 0$ which proves condition (4).

Now assume that $c_{n,i} = \frac{1}{\sqrt{k_n}}$, $1 \leqslant i \leqslant k_n$. Then (i) follows from Eq. (A.5). To obtain Condition A.8 use Eq. (2.3), change the order of summation

$$\begin{aligned}
\sum_{i=1}^{k_n} E(\zeta_{n,i}^2 | \mathscr{F}_{n,i-1}) &= \frac{1}{k_n} \sum_{i=1}^{k_n} E(h^2 \circ T^{k_n - i} | T^{-k_n+i-1}(\mathscr{B})) \\
&= \frac{1}{k_n} \sum_{i=1}^{k_n} U_T^{k_n-i+1} \mathscr{P}_{T,\nu}^{k_n-i+1} U_T^{k_n-i}(h^2) \\
&= \frac{1}{k_n} \sum_{i=1}^{k_n} U_T^{k_n-i+1} \mathscr{P}_{T,\nu}(h^2) = \frac{1}{k_n} \sum_{i=0}^{k_n-1} U_T^{i+1} \mathscr{P}_{T,\nu}(h^2)
\end{aligned}$$

and apply Birkhoff's ergodic theorem to the integrable function $U_T \mathscr{P}_{T,\nu}(h^2)$ to conclude that this sequence is convergent to $E(U_T \mathscr{P}_{T,\nu}(h^2)|\mathscr{I})$ almost everywhere (with respect to ν), and consequently in probability. Since $U_T \mathscr{P}_{T,\nu}(h^2) = E(h^2 | T^{-1}(\mathscr{B}))$ and $\mathscr{I} \subseteq T^{-1}(\mathscr{B})$, we have $E(U_T \mathscr{P}_{T,\nu}(h^2)|\mathscr{I}) = E(h^2|\mathscr{I})$. Similarly, Condition A.10 follows from the Birkhoff ergodic theorem, which completes the proof of (ii). Assume, additionally, that T is ergodic. Then η is constant and is equal to $\int h^2 \, d\nu$, so (iii) holds. Since $h^2 \in L^p(Y, \mathscr{B}, \nu)$ and $p > 1$, we have

$$\sum_{i=1}^{k_n} E|\zeta_{n,i}|^{2p} = \frac{1}{n^p} \int h^{2p}(y) \nu(dy) \to 0$$

and $U_T \mathscr{P}_{T,\nu}(h^2) \in L^p(Y, \mathscr{B}, \nu)$. By the ergodic theorem in L^p spaces we get

$$\sum_{i=1}^{k_n} E(\zeta_{n,i}^2 | \mathscr{F}_{n,i-1}) = \frac{1}{k_n} \sum_{i=0}^{k_n-1} U_T^i U_T \mathscr{P}_{T,\nu}(h^2) \to \int U_T \mathscr{P}_{T,\nu}(h^2)(y) \nu(dy)$$

in $L^p(Y, \mathscr{B}, \nu)$, but $\int U_T \mathscr{P}_{T,\nu}(h^2)(y)\nu(dy) = \int h^2(y)\nu(dy)$ and the proof of (iv) is complete. □

We now turn to the FCLT for $(h \circ T^i)_{i \geqslant 0}$. Let $(Y, \mathscr{B}, \nu)$ be a normalized measure space and $T : Y \to Y$ be ergodic with respect to ν. Let $h \in L^2(Y, \mathscr{B}, \nu)$ and $\sigma = \|h\|_2 > 0$. We define a random function

$$\psi_n(t) = \frac{1}{\sigma\sqrt{n}} \sum_{i=0}^{[nt]-1} h \circ T^i \quad \text{for } t \in [0, 1]$$

(where the sum from 0 to -1 is set to be 0). Note that ψ_n is a right continuous step function, a random variable of $D[0,1]$ and $\psi_n(0)=0$.

Lemma A.8. *If $\mathscr{P}_{T,v}h=0$, then the finite dimensional distributions of ψ_n converge to those of the Wiener process w.*

Proof. To show that the finite dimensional distributions of ψ_n converge to the corresponding finite dimensional distributions of w we use the Cramér–Wold technique. If the $c_1,\dots,c_k$ are arbitrary real numbers and $t_0=0<t_1<\dots<t_k\leqslant 1$, we put

$$\zeta_{n,i}=\begin{cases}\frac{c_j}{\sigma\sqrt{n}}h\circ T^{n-i}, & n-[nt_j]<i\leqslant n-[nt_{j-1}],\ \ j=1,\dots,k\ ,\\ 0, & 1\leqslant i\leqslant n-[nt_k]\quad\text{and}\quad t_k<1\ ,\\ 0, & i>n\ .\end{cases}$$

Observe that

$$\sum_{i=1}^{n}\zeta_{n,i}=\sum_{j=1}^{k}c_j(\psi_n(t_j)-\psi_n(t_{j-1}))\ .$$

By Lemma A.7 the $\zeta_{n,i}$ array is a martingale differences array. We will verify the conditions of Theorem A.2. The Lindeberg condition reads as

$$\sum_{i=1}^{n}E(\zeta_{n,i}^2 1_{\{|\zeta_{n,i}|>\epsilon\}})=\sum_{j=1}^{k}\frac{c_j^2}{\sigma^2 n}([nt_j]-[nt_{j-1}])E(h^2 1_{\{|h|>\epsilon\sigma\sqrt{n}c_j^{-1}\}})$$

and the right-hand side is a finite sum of sequences converging to 0, so it is convergent to 0. Condition A.6 is a consequence of ergodicity of T and

$$\sum_{i=1}^{n}E(\zeta_{n,i}^2|T^{-n+i-1}(\mathscr{B}))=\sum_{j=1}^{k}\frac{c_j^2}{\sigma^2 n}\sum_{i=[nt_{j-1}]+1}^{[nt_j]}U_T^i P_T(h^2)\to^P\sum_{j=1}^{k}c_j^2(t_j-t_{j-1})\ .$$

Therefore, we obtain by Theorem A.2

$$\sum_{i=1}^{n}\zeta_{n,i}\to^{\mathrm{d}}\sqrt{\sum_{j=1}^{k}c_j^2(t_j-t_{j-1})}\,\mathrm{N}(0,1)\ .$$

Thus $\sum_{j=1}^{k}c_j(\psi_n(t_j)-\psi_n(t_{j-1}))$ converges to the Gaussian distributed random variable with mean 0 and variance $\sum_{j=1}^{k}c_j^2(t_j-t_{j-1})$, but this is the distribution of $\sum_{j=1}^{k}c_j(w(t_j)-w(t_{j-1}))$. $\square$

Lemma A.9. *If $\mathscr{P}_{T,v}h=0$, then Condition* 2.20 *holds for each positive ϵ.*

Proof. Define a martingale difference array

$$\zeta_{n,i}=\begin{cases}\frac{1}{\sigma\sqrt{n}}h\circ T^{n-i}, & 1\leqslant i\leqslant n\ ,\\ 0, & i>n,\ \ n\geqslant 1\ .\end{cases}$$

Let also $\zeta_{n,0}=0$ and $\widetilde{\psi}_n(t)=\sum_{i=0}^{[nt]}\zeta_{n,i}$, $t\in[0,1]$, $n\geqslant 1$. Since

$$\psi_n(t)=\widetilde{\psi}_n(1)-\widetilde{\psi}_n(1-t)\ ,$$

we obtain

$$\begin{aligned}\sup_{|t-s|\leqslant\delta}|\psi_n(s)-\psi_n(t)|&\leqslant\sup_{|t-s|\leqslant\delta}|\widetilde{\psi}_n(s)-\widetilde{\psi}_n(t)|\\ &\leqslant 4\sup_k\sup_{k\delta<t\leqslant(k+1)\delta}|\widetilde{\psi}_n(t)-\widetilde{\psi}_n(k\delta)|\ .\end{aligned}$$

This gives

$$\nu\left(\sup_{|t-s|\leqslant\delta}|\psi_n(s)-\psi_n(t)|>\epsilon\right)\leqslant\sum_{k\delta<1}\nu\left(\sup_{k\delta<t\leqslant(k+1)\delta}|\widetilde{\psi}_n(t)-\widetilde{\psi}_n(k\delta)|>\frac{\epsilon}{4}\right) .$$

Now arguments similar to those of Brown (1971, pp. 64–65) and Lemma A.8 complete the proof. □

For the next results we need the following.

Lemma A.10. *Let $(z_i)_{i\geqslant 1}$ be a sequence of real numbers such that*

$$\lim_{n\to\infty}\frac{\sum_{i=1}^{n}z_i}{n}=z .$$

Then

$$\lim_{n\to\infty}\frac{\sum_{i=1}^{k_n}a_n^i z_i}{\sum_{i=1}^{k_n}a_n^i}=z$$

for every sequence of integers $k_n\geqslant 1$ and every sequence of real numbers a_n satisfying

$$\lim_{n\to\infty}k_n=\infty,\quad \lim_{n\to\infty}a_n=1\quad\textit{and}\quad \lim_{n\to\infty}a_n^{k_n}\neq 1 \tag{A.12}$$

and either $a_n>1$ or $0<a_n<1$ for all $n\geqslant 1$.

Proof. Suppose that $a_n>1$ for all $n\geqslant 1$ (the case of $0<a_n<11$ for all $n\geqslant 1$ is proved analogously). Using the following method of summation:

$$\sum_{i=k}^{m}c_i b_i=c_m\sum_{i=k}^{m}b_i-\sum_{i=k}^{m-1}(c_{i+1}-c_i)\sum_{j=k}^{i}b_i ,$$

we can write

$$\sum_{i=1}^{k_n}a_n^i z_i-z\sum_{i=1}^{k_n}a_n^i=-\sum_{i=1}^{k_n-1}(a_n^{i+1}-a_n^i)\left(\sum_{j=1}^{i}z_j-iz\right)+a_n^{k_n}\left(\sum_{i=1}^{k_n}z_i-k_n z\right) .$$

Fix $\epsilon>0$ and let n_0 be such that

$$\left|\frac{\sum_{i=1}^{m}z_i}{m}-z\right|\leqslant\epsilon\quad\text{for } m\geqslant n_0 .$$

Combining these yields, for $k_n>n_0$,

$$\left|\frac{\sum_{i=1}^{k_n}a_n^i z_i}{\sum_{i=1}^{k_n}a_n^i}-z\right|\leqslant\frac{\sum_{i=1}^{n_0-1}(a_n^{i+1}-a_n^i)|\sum_{j=1}^{i}z_j-iz|}{\sum_{i=1}^{k_n}a_n^i}+\epsilon\,\frac{\sum_{i=n_0}^{k_n-1}(a_n^{i+1}-a_n^i)i+a_n^{k_n}k_n}{\sum_{i=1}^{k_n}a_n^i} .$$

Letting $n\to\infty$ we see that the first term on the right goes to zero, while the second term goes to ϵ times a constant not depending on ϵ. □

Theorem A.11. *Let $(Y,\mathscr{B},\nu)$ be a normalized measure space and $T:Y\to Y$ be ergodic with respect to ν. Let (k_n), (a_n) be sequences satisfying Condition* A.12. *Let $c\in\mathbb{R}$ and let (c_n) be a sequence of real numbers such that*

$$\lim_{n\to\infty}k_n c_n^2=c^2 .$$

If $h \in L^2(Y, \mathscr{B}, \nu)$ *is such that* $\mathscr{P}_{T,\nu}h = 0$, *then*

$$c_n \sum_{i=1}^{k_n} a_n^i h \circ T^i \to^{\mathrm{d}} \sigma \mathrm{N}(0, 1) \ ,$$

where

$$\sigma = \sqrt{\frac{c^2(a^2-1)}{\ln a^2}} \|h\|_2 \quad \textit{and} \quad a = \lim_{n\to\infty} a_n^{k_n} \ .$$

Proof. We shall apply Theorem A.4. From Lemma A.7 it follows that $\zeta_{n,i} = c_n a_n^{k_n+1-i} h \circ T^{k_n+1-i}$ is a martingale difference array and that

$$\lim_{n\to\infty} \frac{1}{\sqrt{k_n}} E\left(\max_{1\leqslant j\leqslant k_n} |h \circ T^j|\right) = 0 \ .$$

We have

$$\max_{1\leqslant i\leqslant k_n} |\zeta_{n,i}| \leqslant |c_n| \max(a_n^{k_n}, 1) \max_{1\leqslant i\leqslant k_n} |h \circ T^i| \ ,$$

so

$$E\left(\max_{1\leqslant i\leqslant k_n} |\zeta_{n,i}|\right) \leqslant \sqrt{k_n}|c_n| \max(a_n^{k_n}, 1) E\left(\frac{1}{\sqrt{k_n}} \max_{0\leqslant i\leqslant k_n} |h \circ T^i|\right) \ ,$$

which gives Condition A.9. Condition A.10 is a consequence of

$$\sum_{i=1}^{k_n} \zeta_{n,i}^2 = c_n^2 \sum_{i=1}^{k_n} a_n^{2i} h^2 \circ T^i = c_n^2 \sum_{i=1}^{k_n} a_n^{2i} \frac{\sum_{i=1}^{k_n} a_n^{2i} h^2 \circ T^i}{\sum_{i=1}^{k_n} a_n^{2i}} \ ,$$

the fact that

$$\lim_{n\to\infty} c_n^2 \sum_{i=1}^{k_n} a_n^{2i} = \frac{c^2(a^2-1)}{\ln a^2} \ ,$$

Birkhoff's ergodic theorem, and Lemma A.10. □

Corollary A.12. *Under the assumptions of Theorem* A.11, *if for* $h \in L^2(Y, \mathscr{B}, \nu)$ *the series* $\sum_{n=0}^{\infty} \mathscr{P}_{T,\nu}^n h$ *is convergent in* $L^2(Y, \mathscr{B}, \nu)$, *then*

$$c_n \sum_{i=1}^{k_n} a_n^i h \circ T^i \to^{\mathrm{d}} \sigma \mathrm{N}(0, 1) \ ,$$

where

$$\sigma = \sqrt{\frac{c^2(a^2-1)}{\ln a^2}} \|h + f - f \circ T\|_2 \quad \textit{and} \quad f = \sum_{n=1}^{\infty} \mathscr{P}_{T,\nu}^n h \ .$$

Proof. Theorem 2.23 implies $\mathscr{P}_{T,\nu}(h + f - f \circ T) = 0$. Thus

$$c_n \sum_{i=1}^{k_n} a_n^i (h + f - f \circ T) \circ T^i \to^{\mathrm{d}} \sigma \mathrm{N}(0, 1)$$

by Theorem A.11. Therefore it remains to prove that

$$c_n \sum_{i=1}^{k_n} a_n^i (f \circ T - f) \circ T^i \to^P 0 . \tag{A.13}$$

Observe that the left-hand side of Eq. (A.13) is equal to

$$c_n(a_n^{k_n} f \circ T^{k_n+1} - f \circ T) + c_n(a_n^{-1} - 1) \sum_{i=1}^{k_n} a_n^i f \circ T^i .$$

Since $c_n \to 0$ as $n \to \infty$, the first term converges to 0 in probability. From Lemma A.10 and Birkhoff's ergodic theorem it follows that

$$\frac{\sum_{i=1}^{k_n} a_n^i f \circ T^i}{\sum_{i=1}^{k_n} a_n^i} \to^P \int f(y)v(\mathrm{d}y) .$$

Therefore the sequence

$$c_n(a_n^{-1} - 1) \sum_{i=1}^{k_n} a_n^i f \circ T^i = c_n(1 - a_n^{k_n}) \frac{\sum_{i=1}^{k_n} a_n^i f \circ T^i}{\sum_{i=1}^{k_n} a_n^i}$$

is also convergent to 0 in probability. □

Remark A.13. Note that we can conclude from Theorem A.11 that

$$c_n \sum_{i=m}^{k_n} a_n^i h \circ T^i \to^{\mathrm{d}} \sigma \mathrm{N}(0, 1) ,$$

where $m \geqslant 0$ is any fixed integer, because c_n goes to zero and a_n to 1 as $n \to \infty$, so the difference

$$c_n \sum_{i=1}^{k_n} a_n^i h \circ T^i - c_n \sum_{i=m}^{k_n} a_n^i h \circ T^i ,$$

which is either equal to $c_n h$ or $c_n \sum_{i=1}^{m} a_n^i h \circ T^i$, converges in probability to zero.

Lemma A.14. *Let $(Y, \mathscr{B}, v)$ be a normalized measure space, $T : Y \to Y$ be ergodic with respect to v, and $\gamma \neq 0$ be a constant. Let $\kappa(\tau)$, $\tau > 0$, be such that*

$$\lim_{\tau \to 0} \frac{[\kappa(\tau)]^2}{\tau} = 1 .$$

If $h \in L^2(Y, \mathscr{B}, v)$ is such that $\mathscr{P}_{T,v} h = 0$, then the finite dimensional distributions of the process ζ_τ defined by

$$\zeta_\tau(t) = \kappa(\tau)\mathrm{e}^{-\gamma t} \sum_{j=0}^{[\frac{t}{\tau}]} \mathrm{e}^{\gamma \tau j} h \circ T^j, \quad t \geqslant 0, \quad \tau > 0$$

converge weakly as $\tau \to 0$ to the corresponding finite dimensional distributions of the zero-mean Gaussian process ζ for which $\zeta(0) = 0$ and

$$E\zeta(t)\zeta(s) = \frac{\|h\|_2^2}{2\gamma} (\mathrm{e}^{2\gamma \min(t,s)} - 1)\mathrm{e}^{-\gamma(t+s)}, \quad t, s > 0 .$$

Proof. To prove the convergence of the finite dimensional distributions of ζ_τ to the corresponding finite dimensional distributions of the Gaussian process ζ, it is enough to prove that for any $l \geqslant 1$, any real numbers $0 < t_1 < \cdots < t_l < \infty$

and $d_1,\dots,d_l$ the distribution of $\sum_{j=1}^l d_j\zeta(t_j)$ is Gaussian and that $\sum_{j=1}^l d_j\zeta_\tau(t_j)$ converges in distribution to $\sum_{j=1}^l d_j\zeta(t_j)$ as $\tau\to 0$.

We consider first the case of $l=1$. It follows from Theorem A.11 that for $t>0$ the distribution of $\zeta_\tau(t)$ converges weakly as $\tau\to 0$ to a Gaussian random variable. To see this let τ_n be a sequence going to zero as $n\to\infty$. Take $k_n=[\frac{t}{\tau_n}]$, $a_n=\mathrm{e}^{\gamma\tau_n}$, $c_n=\kappa(\tau_n)\mathrm{e}^{-\gamma t}$, and observe that

$$\lim_{n\to\infty} k_n=\infty,\quad \lim_{n\to\infty} a_n=1,\quad \lim_{n\to\infty} a_n^{k_n}=\mathrm{e}^{\gamma t}\quad\text{and}\quad \lim_{n\to\infty} k_n c_n^2=t\mathrm{e}^{-2\gamma t}\ ,$$

and

$$\frac{t\mathrm{e}^{-2\gamma t}(\mathrm{e}^{2\gamma t}-1)}{\ln \mathrm{e}^{2\gamma t}}=\frac{1-\mathrm{e}^{-2\gamma t}}{2\gamma}\ .$$

The theorem then implies that

$$\kappa(\tau_n)\mathrm{e}^{-\gamma t}\sum_{j=0}^{[\frac{t}{\tau_n}]}\mathrm{e}^{\gamma\tau_n j}h(T^j(\xi_0))\to \mathrm{N}\left(0,\frac{\sigma_0^2}{2\gamma}(1-\mathrm{e}^{-2\gamma t})\right)\ ,$$

where $\sigma_0^2=\int h^2(y)\nu(\mathrm{d}y)$ and ξ_0 is distributed according to ν. Consequently $\zeta_\tau(t)\to^{\mathrm{d}}\zeta(t)$ as $\tau\to 0$, where $\zeta(t)$ is a Gaussian distributed random variable with mean 0 and variance given by

$$\frac{\|h\|_2^2}{2\gamma}(1-\mathrm{e}^{-2\gamma t}),\quad t>0\ .$$

Note that $\zeta_\tau(0)=\kappa(\tau)h$. Since $\lim_{\tau\to 0}\kappa(\tau)=0$, we also have $\zeta_\tau(0)\to 0$ as $\tau\to 0$.

We next consider the case of $l=2$. Let $t_1<t_2$ and d_1,d_2 be given. Let τ_n be a sequence going to zero as $n\to\infty$. Set $k_{n,1}=[\frac{t_1}{\tau_n}]$, $k_{n,2}=[\frac{t_2}{\tau_n}]$, $k_n=k_{n,2}+1$, and observe that $k_{n,1}<k_{n,2}$ for all n sufficiently large. Define

$$\eta_{n,j}=\begin{cases} d_2\mathrm{e}^{-\gamma t_2}\kappa(\tau_n)\mathrm{e}^{\gamma\tau_n(k_n-j)}h\circ T^{k_n-j}, & 0<j\leqslant k_{n,2}-k_{n,1}\ ,\\ (d_2\mathrm{e}^{-\gamma t_2}+d_1\mathrm{e}^{-\gamma t_1})\kappa(\tau_n)\mathrm{e}^{\gamma\tau_n(k_n-j)}h\circ T^{k_n-j}, & k_{n,2}-k_{n,1}<j\leqslant k_n\ ,\\ 0, & \text{otherwise}\ .\end{cases}$$

Then we have

$$d_1\zeta_{\tau_n}(t_1)+d_2\zeta_{\tau_n}(t_2)=\sum_{j=1}^{k_n}\eta_{n,j}\ .$$

Observe that

$$\sum_{j=1}^{k_n}\eta_{n,j}^2=d_2^2\mathrm{e}^{-2\gamma t_2}[\kappa(\tau_n)]^2\sum_{j=0}^{k_{n,2}}\mathrm{e}^{2\gamma\tau_n j}h^2\circ T^j$$
$$+(2d_2d_1\mathrm{e}^{-\gamma(t_2+t_1)}+d_1^2\mathrm{e}^{-2\gamma t_1})[\kappa(\tau_n)]^2\sum_{j=0}^{k_{n,1}}\mathrm{e}^{2\gamma\tau_n j}h^2\circ T^j\ .$$

As in the proof of Theorem A.11, we check that Theorem A.4 applies to $\{\eta_{n,j}:1\leqslant j\leqslant k_n,n\geqslant 1\}$ and conclude that

$$d_1\zeta_{\tau_n}(t_1)+d_2\zeta_{\tau_n}(t_2)\to^{\mathrm{d}}\sigma\mathrm{N}(0,1)\ ,$$

where $\sigma^2=\frac{\|h\|_2^2}{2\gamma}(d_2^2(1-\mathrm{e}^{-2\gamma t_2})+2d_2d_1\mathrm{e}^{-\gamma(t_2+t_1)}(\mathrm{e}^{2\gamma t_1}-1)+d_1^2(1-\mathrm{e}^{-2\gamma t_1}))$. Since $\sigma\mathrm{N}(0,1)$ is the distribution of $d_1\zeta(t_1)+d_2\zeta(t_2)$ and $E(\zeta(t_1)\zeta(t_2))=\frac{\|h\|_2^2}{2\gamma}\mathrm{e}^{-\gamma(t_2+t_1)}(\mathrm{e}^{2\gamma t_1}-1)$. The case of arbitrary l is deduced analogously from Theorem A.4. □

Lemma A.15. *Let* $(Y, \mathcal{B}, \nu)$ *be a normalized measure space,* $T : Y \to Y$ *be ergodic with respect to* ν*, and* $\gamma \neq 0$ *be a constant. Let* $\kappa(\tau)$, $\tau > 0$, *be such that*

$$\lim_{\tau\to 0} \frac{[\kappa(\tau)]^2}{\tau} = 1 \ .$$

If $h \in L^2(Y, \mathcal{B}, \nu)$ *is such that* $\mathcal{P}_{T,\nu}h = 0$, *then the finite dimensional distributions of the process* η_τ *defined by*

$$\eta_\tau(t) = \frac{\kappa(\tau)}{\gamma} \sum_{j=0}^{[\frac{t}{\tau}]} (1 - \mathrm{e}^{\gamma(\tau j - t)}) h \circ T^j, \quad t \geqslant 0, \quad \tau > 0$$

converge weakly as $\tau \to 0$ *to the corresponding finite dimensional distributions of the zero-mean Gaussian process* η *for which* $\eta(0) = 0$ *and*

$$E\eta(t)\eta(s) = \frac{\|h\|_2^2}{\gamma^3}(2\gamma \min(t, s) - 2 + 2\mathrm{e}^{-\gamma t} + 2\mathrm{e}^{-\gamma s} - \mathrm{e}^{-\gamma|t-s|} - \mathrm{e}^{-\gamma(t+s)})$$

for $t, s > 0$.

The lemma follows from Theorem A.4 in a similar fashion as the preceding lemma.

Appendix B. Weak convergence criteria

Let $(X, |\cdot|)$ be a phase space which is either $\mathbb{R}^k$ or a separable Banach space, and denote by $\mathcal{M}_1$ the space of all probability measures defined on the σ-algebra $\mathcal{B}(X)$ of Borel subsets of X. For a real-valued measurable bounded function f, and $\mu \in \mathcal{M}_1$, we introduce the scalar product notation

$$\langle f, \mu \rangle = \int_X f(x)\mu(\mathrm{d}x) \ .$$

One way to characterize weak convergence in $\mathcal{M}_1$ is to use the Fortet–Mourier metric in $\mathcal{M}_1$, which is defined by

$$d_{FM}(\mu_1, \mu_2) = \sup\{|\langle f, \mu_1\rangle - \langle f, \mu_2\rangle| : f \in \mathcal{F}_{FM}\} \quad \text{for } \mu_1, \mu_2 \in \mathcal{M}_1 \ ,$$

where

$$\mathcal{F}_{FM} = \left\{ f : X \to \mathbb{R} : \sup_{x\in X} |f(x)| \leqslant 1, |f|_L \leqslant 1 \right\}$$

and $|f|_L = \sup_{x\neq y} \frac{|f(x)-f(y)|}{|x-y|}$. This defines a complete metric on $\mathcal{M}_1$, and we have $\mu_n \to \mu$ weakly if and only if $d_{FM}(\mu_n, \mu) \to 0$ (cf. Dudley, 1989, Chapter 3).

We further introduce a distance on $\mathcal{M}_1$ by

$$d(\mu_1, \mu_2) = \sup\{|\langle f, \mu_1\rangle - \langle f, \mu_2\rangle| : |f|_L \leqslant 1\} \quad \text{for } \mu_1, \mu_2 \in \mathcal{M}_1 \ .$$

This quantity is always defined, but for some measures it may be infinite. It is easy to check that the function d is finite for elements of the set

$$\mathcal{M}_1^1 = \left\{ \mu \in \mathcal{M}_1 : \int_X |x|\mu(\mathrm{d}x) < \infty \right\} \ ,$$

and defines a metric on this set. Moreover, $\mathcal{M}_1^1$ is a dense subset of $(\mathcal{M}_1, d_{FM})$ and

$$d_{FM}(\mu_1, \mu_2) \leqslant d(\mu_1, \mu_2) \ .$$

Let $(Y, \mathscr{B}, \nu)$ be a normalized measure space and let $R_n : X \times Y \to X$ be a measurable transformation for each $n \in \mathbb{N}$. We associate with each transformation R_n an operator $P_n : \mathscr{M}_1 \to \mathscr{M}_1$ defined by

$$P_n\mu(A) = \int_X \int_Y 1_A(R_n(x, y))\nu(\mathrm{d}y)\mu(\mathrm{d}x) \tag{B.1}$$

for $\mu \in \mathscr{M}_1$, where

$$1_A(x) = \begin{cases} 1 & x \in A\ , \\ 0 & x \notin A \end{cases}$$

is the indicator function of a set A. Write

$$U_n f(x) = \int_Y f(R_n(x, y))\nu(\mathrm{d}y)$$

for measurable functions $f : X \to \mathbb{R}$, for which the integral is defined. The operators U_n and P_n satisfy the identity $\langle U_n f, \mu\rangle = \langle f, P_n\mu\rangle$. Note that if $\mu = \delta_x$, where δ_x is the point measure at x defined by

$$\delta_x(A) = \begin{cases} 1 & x \in A \\ 0 & x \notin A\ , \end{cases} \tag{B.2}$$

then $U_n f(x) = \langle f, P_n\delta_x\rangle$.

Remark B.1. Note that if $R_n(x, y)$ does not depend on y, then $U_n f = U_{R_n} f$ where U_{R_n} is the Koopman operator corresponding to $R_n : X \to X$. The following relation holds between the Frobenius–Perron operator P_{R_n} on $L^1(X, \mathscr{B}(X), m)$ and the operator P_n: If μ has a density f with respect to m, then $P_{R_n} f$ is a density of $P_n\mu$.

On the other hand if $R_n(x, y)$ does not depend on x, then $U_n f$ is equal to $\int U_{R_n} f(y)\nu(\mathrm{d}y)$, where U_{R_n} is the Koopman operator corresponding to $R_n : Y \to X$. The operator P_n has the same value $\nu \circ R_n^{-1}$ for every $\mu \in \mathscr{M}_1$.

Assume that for each $n \in \mathbb{N}$ the transformation $R_n : X \times Y \to X$ satisfies the following conditions:

(A1). There exists a measurable function $L_n : Y \to \mathbb{R}_+$ such that

$$|R_n(x, y) - R_n(\bar{x}, y)| \leqslant L_n(y)|x - \bar{x}| \quad \text{for } x, \bar{x} \in X,\ \ y \in Y\ .$$

(A2). The series

$$\sum_{n=1}^{\infty} \int_Y |R_n(0, T(y)) - R_{n+1}(0, y)|\nu(\mathrm{d}y)$$

is convergent, where $T : Y \to Y$ is a transformation preserving the measure ν.

(A3). The integral $\int_Y |R_n(0, y)|\nu(\mathrm{d}y)$ is finite for at least one n.

Proposition B.2. *Let the transformations R_n satisfy conditions* (A1)–(A3). *If*

$$\lim_{n\to\infty} \int_Y L_n(y)\nu(\mathrm{d}y) = 0,$$

then there exists a unique measure $\mu_* \in \mathscr{M}_1$ such that $(P_n\mu)$ converges weakly to μ_* for each measure $\mu \in \mathscr{M}_1$.

Proof. Assumptions (A2) and (A3) imply that $P_n(\mathscr{M}_1^1) \subset \mathscr{M}_1^1$. By the definition of the metric d we have

$$\mathrm{d}(P_n\delta_0, P_{n+1}\delta_0) = \sup\{|U_n f(0) - U_{n+1} f(0)| : |f|_L \leqslant 1\}\ .$$

Since the transformation T preserves the measure ν, we can write

$$U_n f(0) = \int_Y f(R_n(0, y)\nu(\mathrm{d}y) = \int_Y f(R_n(0, T(y))\nu(\mathrm{d}y)$$

for any f with $|f|_L \leqslant 1$. Hence

$$|U_n f(0) - U_{n+1} f(0)| \leqslant \int_Y |R_n(0, T(y)) - R_{n+1}(0, y)| \nu(\mathrm{d}y) \ .$$

Consequently

$$\mathrm{d}(P_n \delta_0, P_{n+1}\delta_0) \leqslant \int_Y |R_n(0, T(y)) - R_{n+1}(0, y)| \nu(\mathrm{d}y) \ ,$$

and

$$d_{FM}(P_n\delta_0, P_{n+1}\delta_0) \leqslant \mathrm{d}(P_n\delta_0, P_{n+1}\delta_0) \ .$$

From Condition (A2), the sequence $(P_n\delta_0)$ is a Cauchy sequence. Since the space $(\mathscr{M}_1, d_{FM})$ is complete, $(P_n\delta_0)$ is weakly convergent to a $\mu_* \in \mathscr{M}_1$. From (A1) it follows that

$$\mathrm{d}(P_n\mu_1, P_n\mu_2) \leqslant \int_Y L_n(y)\nu(\mathrm{d}y)\mathrm{d}(\mu_1, \mu_2) \ .$$

Hence $(P_n\mu)$ is weakly convergent for each $\mu \in \mathscr{M}_1^1$ and has the limit μ_*. Since, for sufficiently large n, each operator P_n satisfies

$$d_{FM}(P_n\mu_1, P_n\mu_2) \leqslant d_{FM}(\mu_1, \mu_2)$$

and the set $\mathscr{M}_1^1$ is dense in (M_1, d_{FM}). □

References

Adler, R., Rivlin, T., 1964. Ergodic and mixing properties of Chebyshev polynomials. Proc. Amer. Math. Soc. 15, 794–796.

Alexander, J.C., Yorke, J.A., 1984. Fat baker's transformations. Erg. Theory Dyn. System 4, 1–23.

an der Heiden, U., Mackey, M.C., 1982. The dynamics of production and destruction: analytic insight into complex behaviour. J. Math. Biol. 16, 75–101.

Beck, C., 1990a. Brownian motion from deterministic dynamics. Physica A 169, 324–336.

Beck, C., 1990b. Ergodic properties of a kicked damped particle. Commun. Math. Phys. 130, 51–60.

Beck, C., 1996. Dynamical systems of Langevin type. Physica A 233, 419–440.

Beck, C., Roepstorff, G., 1987. From dynamical systems to the Langevin equation. Physica A 145, 1–14.

Benedics, M., Carleson, L., 1985. On iterations of $1 - ax^2$ on $(-1, 1)$. Ann. Math. 122, 1–25.

Berre, M.L., Ressayre, E., Tallet, A., Pomeau, Y., 1990. Dynamic system driven by a retarded force acting as colored noise. Phys. Rev. A 41.

Billingsley, P., 1968. Convergence of Probablility Measures. Wiley, New York.

Billingsley, P., 1995. Probablility and Measure. Wiley, New York.

Boltzmann, L., 1995. Lectures on Gas Theory. Dover, Mineola, NY.

Bowen, R., 1975. Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms. Springer Lecture Notes in Mathematics, vol. 470. Springer, Berlin, New York.

Boyarsky, A., Scarowsky, M., 1979. On a class of transformations which have unique absolutely continuous invariant measures. Trans. Amer. Math. Soc. 255, 243–262.

Briggs, M., Sengers, J., Francis, M., Gaspard, P., Gammon, R., Dorfman, J., Calabrese, R., 2001. Tracking a colloidal particle for the measurement of dynamic entropies. Physica A 296, 42–59.

Brown, B.M., 1971. Martingale central limit theorems. Ann. Math. Statist. 42, 59–66.

Chandrasekhar, S., 1943. Stochastic problems in Physics and Astronomy. Rev. Mod. Phys. 15 (1), 1–89.

Chew, L., Ting, C., 2002. Microscopic chaos and Gaussian diffusion processes. Physica A 307, 275–296.

Chung, K.L., 2001. A Course in Probability Theory. Third ed. Academic Press, New York.

Denker, M., 1989. The central limit theorem for dynamical systems. In: Dynamical Systems and Ergodic Theory (Warsaw, 1986), vol. 23. Banach Center Publications, PWN, Warsaw, pp. 33–62.

Dorfman, J., 1999. An Introduction to Chaos in Nonequilibrium Statistical Mechanics, Cambridge Lecture Notes in Physics, vol. 14. Cambridge University Press, Cambridge, New York.

Dorizzi, B., Grammaticos, B., Berre, M., Pomeau, Y., Ressayre, E., Tallet, A., 1987. Statistics and dimension of chaos in differential delay equations. Phys. Rev. A 35.

Dudley, R.M., 1989. Real Analysis and Probability. Wadsworth, Belmont, CA.

Eagleson, G.K., 1975. Martingale convergence to mixtures of infinitely divisible laws. Ann. Probab. 3, 557–562.

Eckmann, J.-P., Ruelle, D., 1985. Ergodic theory of chaos and strange attractors. Rev. Mod. Phys. 57 (3, Part 1), 617–656.

Einstein, A., 1905. Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. Ann. d. Physik 17, 549–560.

Erdös, P., 1939a. On a family of symmetric Bernoulli convolutions. Amer. J. Math. 61, 974–976.

Erdös, P., 1939b. On the smoothness properties of a family of Bernoulli convolutions. Amer. J. Math. 62, 180–186.

Fürth, R. (Ed.) (1956). Investigations on the Theory of the Brownian Movement. Dover, New York.

Gaenssler, P., Joos, K., 1992. Another view on martingale central limit theorems. Stochast. Process. Appl. 40 (2), 181–197.

Gallavotti, G., 1999. Statistical Mechanics, A Short Treatise. Springer, Berlin, New York.

Gaspard, P., Briggs, M., Francis, M., Sengers, J., Gammon, R., Dorfman, J., Calabrese, R., 1998. Experimental evidence for microscopic chaos. Nature 394, 865–868.

Gordin, M.I., 1969. The central limit theorem for stationary processes. Dokl. Akad. Nauk SSSR 188, 739–741.

Hall, P., 1978. The convergence of moments in the martingale central limit theorem. Z. Wahrsch. Verw. Gebiete 44 (3), 253–260.

Hilgers, A., Beck, C., 1999. Approach to Gaussian stochastic behavior for systems driven by deterministic chaotic forces. Phys. Rev. E 60, 5385–5393.

Hilgers, A., Beck, C., 2001. Higher-order correlations of Tchebyscheff maps. Physica D 156, 1–18.

Hunt, B.R., Kennedy, J.A., Li, T.-Y., Nusse, H.E., 2002. SLYRB measures: natural invariant measures for chaotic systems. Phys. D 170 (1), 50–71.

Iosifescu, M., 1992. A very simple proof of a generalization of the Gauss–Kuzmin–Levy theorem on continued fractions, and related questions. Rev. Roumaine Math. Pures Apll. 37, 901–914.

Jabłoński, M., 1991. A central limit theorem for processes generated by a family of transformations. Dissertationes Mathematicae 307, 1–62.

Jabłoński, M., Malczak, J., 1983a. A central limit theorem for piecewise convex mappings of the unit interval. Tôhoku Math. J. 35, 173–180.

Jabłoński, M., Malczak, J., 1983b. The rate of convergence of iterates of the Frobenius–Perron operator for piecewise convex transformations of the unit interval. Bull. Pol. Ac.: Math. 31, 249–254.

Jabłoński, M., Kowalski, Z., Malczak, J., 1985. The rate of convergence of iterates of the Frobenius–Perron operator for Lasota-Yorke transformations. Univ. Jagel. Acta Math. 25, 189–193.

Jakobson, M., 1981. Absolutely continuous invariant measure for one-parameter families of one-dimensional maps. Commun. Math. Phys. 81, 39–88.

Jessen, B., Wintner, A., 1935. Distribution functions and the Riemann zeta function. Trans. Amer. Math. Soc. 38, 48–88.

Kappler, E., 1931. Versuche zur Messung der Avogadro-Loschmidtschen Zeit aus der Brownschen bewegung einer Drehwaage. Ann. Physik 11, 233–256.

Keller, G., 1980. Un théorème de la limite centrale pour une classe de transformations monotones par morceaux. C.R. Acad. Sc. Paris 291, 155–158.

Kershner, R., Wintner, A., 1935. On symmetric Bernoulli convolutions. Amer. J. Math. 57, 541–548.

Lasota, A., Mackey, M.C., 1994. Chaos, Fractals and Noise: Stochastic Aspects of Dynamics. Springer, Berlin, New York, Heidelberg.

Liverani, C., 1996. Central limit theorem for deterministic systems. In: International Conference on Dynamical Systems (Montevideo, 1995), Pitman Research Notes in Mathematics Series, vol. 362. Longman, Harlow, pp. 56–75.

Mackey, M.C., 1989. The dynamic origin of increasing entropy. Rev. Mod. Phys. 61, 981–1016.

Mackey, M.C., 1992. Time's Arrow, The Origins of Thermodynamic Behaviour. Springer, Berlin, New York, Heidelberg.

Mackey, M.C., Glass, L., 1977. Oscillation and chaos in physiological control systems. Science 197, 287–289.

Mazo, R., 2002. Brownian Motion: Fluctuations, Dynamics, and Applications. Clarendon Press, Oxford.

Peres, Y., Schlag, W., Solomyak, B., 2000. Sixty years of Bernoulli convolutions. In: Fractal Geometry and Stochastics, II (Greifswald/Koserow, 1998) Progress in Probability, vol. 46, Birkhäuser, Basel, pp. 39–65.

Pollicott, M., Sharp, R., 2002. Invariance principles for interval maps with an indifferent fixed point. Commun. Math. Phys. 229, 337–346.

Ratner, M., 1973. The central limit theorem for geodesic flows on n-dimensional manifolds of negative curvature. Israel J. Math. 16, 181–197.

Ruelle, D., 1976. A measure associated with Axiom A attractors. Am. J. Math. 98, 619–654.

Ruelle, D., 1978. Sensitive dependence on initial condition and turbulent behavior of dynamical systems. Ann. N.Y. Acad. Sci. 316, 408–416.

Ruelle, D., 1979. Microscopic fluctuations and turbulence. Phys. Lett. 72A, 81–82.

Ruelle, D., 1980. Measures describing a turbulent flow. Ann. N.Y. Acad. Sci. 357, 1–9.

Schulman, L.S., 1997. Time's Arrows and Quantum Measurement. Cambridge University Press, Cambridge.

Sinai, Y.G., 1972. Gibbs measure in ergodic theory. Russian Math. Surveys 27, 21–69.

Solomyak, B., 1995. On the random series $\sum \pm\lambda^n$ (an Erdös problem). Annals of Math. 142, 611–625.

Teicher, H., 1988. Distribution and moment convergence of martingales. Probab. Theory Rel. Fields 79 (2), 303–316.

Thaler, M., 1980. Estimates of the invariant densities of endomorphisms with indifferent fixed points. Israel J. Math. 37, 303–314.

Tsujii, M., 1996. On continuity of Bowen-Ruelle-Sinai measures in families of one-dimensional maps. Commun. Math. Phys. 177 (1), 1–11.

Tyran-Kamińska, M., 2005. An invariance principle for maps with polynomial decay of correlations. Commun. Math. Phys. 260, 1–15.

Viana, M., 1997. Stochastic dynamics of deterministic systems. Col. Bras. de Matemática 21, 197.

Wintner, A., 1935. On convergent Poisson convolutions. Amer. J. Math. 57, 827–838.

Wong, S., 1979. A central limit theorem for piecewise monotonic mappings of the unit interval. Ann. Probab. 7, 500–514.

Young, L.-S., 1992. Decay of correlations for certain quadratic maps. Commun. Math. Phys. 146, 123–138.

Young, L.-S., 1999. Recurrence times and rates of mixing. Israel J. Math. 110, 153–188.

Young, L.-S., 2002. What are SRB measures, and which dynamical systems have them? J. Statist. Phys. 108, 733–754.