J. Nonlinear Sci. Vol. 12: pp. 113–141 (2002)
DOI: 10.1007/s00332-001-0450-4

Journal of Nonlinear Science

Set-Valued Markov Chains and Negative Semitrajectories of Discretized Dynamical Systems

P. Diamond[1] and I. Vladimirov[2,*]

[1] Department of Mathematics, University of Queensland, Brisbane, QLD 4072, Australia
E-mail: pmd@maths.uq.edu.au

[2] Department of Mathematics, University of Queensland, Brisbane, QLD 4072, Australia
E-mail: igv@maths.uq.edu.au

Received January 9, 2001; accepted January 2, 2002
Online publication April 8, 2002
Communicated by E. Doedel

Summary. Computer simulation of dynamical systems involves a phase space which is the finite set of machine arithmetic. Rounding state values of the continuous system to this grid yields a spatially discrete dynamical system, often with different dynamical behaviour. Discretization of an invertible smooth system gives a system with set-valued negative semitrajectories. As the grid is refined, asymptotic behaviour of the semitrajectories follows probabilistic laws which correspond to a set-valued Markov chain, whose transition probabilities can be explicitly calculated. The results are illustrated for two-dimensional dynamical systems obtained by discretization of fractional linear transformations of the unit disc in the complex plane.

1. Introduction

Nonlinearity introduces serious challenges for both theoretical and applied research in science and engineering and is frequently studied by computation. Wide-ranging numerical experiments are carried out to visualise and understand the complicated dynamical behaviour which nonlinearities commonly produce. For example, chaotic systems are very often investigated in this manner. However, computer simulation has limitations imposed by the nature of digital arithmetic. Machine computation replaces the continuum of real numbers by the large but nevertheless finite set of computer arithmetic. It is important to realize that this discretization can have severe effects upon system behaviour. For example, again in chaos theory, all computed orbits are eventually periodic

* *Permanent address:* Institute for Information Transmission Problems, 19 Bolshoi Karetny Lane, GSP–4, 101447 Moscow, Russia

after a transient segment, which is in stark contrast to the theoretical time evolution of the system.

In fact, the overall behaviour of nonlinear systems can be fundamentally changed by a computer implementation. Artifacts and spurious behaviour can appear [7], [22], [37], even in linear systems [6]. Perhaps the most striking examples occur in computation of chaotic systems, where degenerate collapsing behaviour can occur. Long-term computed trajectories are attracted to fixed points or short cycles in a way that is quite at odds with the behaviour of the theoretical system [3], [4], [8], [10]–[17], [22], [37]. The underlying system is a differential equation or an iterated function T defined on a continuum $X \subset \mathbb{R}^n$. Its corresponding computer implementation is an iterative algorithm, represented by a function T_ε on a finite, discrete subset, the ε-grid $X_\varepsilon = X \bigcap (\varepsilon \mathbb{Z}^n)$ of X. For most initial points $x \in X$, the iterations $T(x), T^2(x), \ldots, T^k(x), \ldots$ wander through an invariant set, never repeating nor attracted to a single point. However, for many systems the sequence $T_\varepsilon(x), \ldots, (T^k)_\varepsilon(x), \ldots$ collapses to a fixed point or short cycle for a surprisingly large number of initial values $x \in X_\varepsilon$. Moreover, the proportion of collapsing points is extremely sensitive to the cardinality of X_ε, that is, to ε, varying in such an irregular way as to appear random. Various explanations have been given for this phenomenon, but despite the sophistication of the arguments there is, by and large, always a heuristic element present [3], [4], [20], [22], [36]–[39], [48].

There are good heuristic reasons to believe that this effect is very dependent upon the invariant measure of the original system [11], [15]. Some useful predictions were made by modeling long-term iterations by certain types of random graphs [10], [11], [16], [17], drawing ultimately on inspiration from a number of sources concerning random mappings [9], [21], [26], [34], [40], [43]. Nonetheless, this random graphical approach was still heuristic and, despite the closeness of prediction with extensive computational experiments, still lacked a rigorous justification based on the structure of computer arithmetic and its interaction with function iteration.

In a rather different direction, periodicity occurs in pseudorandom number generators. For those sequences z_n/m generated by congruences $z_{n+1} = (az_n + c) \bmod m$, aimed at mimicking uniformly distributed numbers on $[0, 1]$, period length is dependent on $a,\ m$ [32]. For sequences $z_{n+1} = f(z_n)$, where f maps $[0 \ldots m-1]$ to itself, if f is uniformly distributed over the m^m possible functions and z_0 uniformly distributed over $[0 \ldots m-1]$, the probability of computational collapse is roughly $1.25m^{-1/2}$ ([32], Ex. 11, Section 3.1 and solution, p. 518). But note that the sensitivity here is due to *the choice of* f and not to the grid size, although the probability depends on the latter, and the effect described is thus different from that which concerns this paper, although not unrelated to some aspects of random graph models of collapse. See also [34] for the expected length of periods under similarly random choice of the function f.

Nevertheless, certain number theoretic considerations arise, related to poor approximation by rationals and rational dependence, which are also important in dynamical systems and the problem of small divisors, and normal forms of functions on $\mathbb{R}^n$ [42], [31], [44]. Similar independence conditions are required for uniformity of the distribution of values of generalized polynomials involving the floor function [23], [24]. Note, however, that limiting asymptotic distributions of certain artifacts are not uniform [17]. Indeed, subtle changes in round-off result in marked changes of experimentally computed invariant measures of discretized functions [15]. The effects of uniformly distributed round-off errors are transformed by their exponential propagation along trajectories of chaotic systems and the distributions of artifacts resulting from round-off are themselves not uniform.

This paper is concerned with rigorous justification of how certain asymptotic distributions arise in computer implementations on models of fixed-point computer arithmetic X_ε. If T is a smooth invertible function on a bounded open set $X \subset \mathbb{R}^n$, its discretization induces a dynamical system on X_ε. The discretized function $T_\varepsilon\colon X_\varepsilon \to X_\varepsilon$ is neither injective nor surjective [19]. On the one hand, positive semitrajectories can collapse to a common orbit. On the other, there may be points in X_ε which are unreachable by the discretized system, apart from being themselves initial values. This appears to be the mechanism underlying possible computational collapse to short-period cycles. These cycles lie in the set of points in X_ε which are reachable in some given number $k > 0$ of iterations of T_ε.

Asymptotic probabilistic properties, as $\varepsilon \to 0$, of this phenomena are investigated below by considering negative semitrajectories of the discretized system. In contrast to the original invertible T, these are set-valued and are here studied using the techniques of compensating system and quantization errors developed in [46], [18], [47]. The compensating system is generated by discretizing T^{-1}. A negative semitrajectory of T_ε can be represented as the set-valued Minkowski sum of a positive semitrajectory of the compensating system (note $(T_\varepsilon)^{-1} \neq (T^{-1})_\varepsilon$).

The evolution of this set-valued semitrajectory is determined by the quantization errors of the compensating system. Moreover, with increasing refinement of the grid X_ε, as $\varepsilon \to 0$, the evolution satisfies asymptotic probabilistic laws corresponding to a set-valued nonhomogeneous Markov chain, in the sense of *frequency measurability* of subsets of $\varepsilon\mathbb{Z}^n$. That is, there exists an absolutely continuous asymptotic spatial distribution of its points, with density described by a frequency function. This notion is used to show that the sets of k-reachable points are all frequency measurable, with frequency functions which are the nonextinction probabilities of the set-valued Markov chain. A martingale property, for the asymptotic distribution of normalized cardinalities of the set-valued negative semitrajectory states of the discretized system, follows as a consequence.

Convergence to the limiting distribution holds under the condition that the *resonance set* [18] of the original diffeomorphism is of zero Lebesgue measure. That is, a number theoretic property, involving rational independence, holds almost everywhere in X.

Note that the set-valued Markov chains studied in this paper have nothing in common with the Markov set-chains considered in [27], which are Markov chains with uncertain transition probability matrices belonging to a set of stochastic matrices. In contrast, the set-valued Markov chains have a denumerable phase space consisting of finite subsets of the n-dimensional integer lattice $\mathbb{Z}^n$. A distinction should also be drawn from the Markov chains arising from the Conley decomposition of a dynamical system and its relation to ε-pseudo-orbits [28].

The results of the paper can be applied to give rigorous justification for the heuristic and often ad hoc models for computational collapse in chaotic systems; see for example [11]–[17], [22], [48]. It is important to emphasize that the asymptotic probability laws considered are, in general, not uniform distributions as in the classical and quite distinct context of [32] and [23], [24], although quantization errors are uniformly distributed.

Although most computation occurs in floating-point arithmetic, presently it is difficult to model complex behaviour of dynamical systems on these exponentially spaced grids. Nevertheless, floating-point is "locally" a regularly spaced grid, in the sense that for any fixed exponent the mantissa values are equally spaced. Consequently, although quantization floating-point effects are not *quantitatively* the same as in fixed-point arithmetic, they are still persistently present, and their limiting distributions are *qualitatively* similar. Further, many well-known and frequently cited results [32], [23], [24] are accomplished on sets which are regularly spaced or without any structure at all [34].

The paper is organized as follows. In Section 2, spatial discretization of smooth invertible dynamical systems and the notion of frequency measurability of sets are defined. In Section 3, the compensating system and its quantization errors are defined, and set-valued negative semitrajectories of the discretized system are represented in terms of these. In Section 4, the mappings which carry out the representation are shown to be continuously convergent as the stepsize of the grid approaches zero. Section 5 contains a key lemma on asymptotic independence and the uniform distribution of quantization errors. In Section 6, the main theorem is proved. This relates asymptotic behaviour of negative semitrajectories of the discretized system to an associated set-valued Markov chain. Section 7 provides a recursive equation which expresses the frequency functions of sets of reachable points in terms of the probabilities of nonextinction of the set-valued Markov chain. Section 8 establishes a martingale property for the cardinality of the set-valued states of semitrajectories. Section 9 develops some further properties of nonextinction probabilities of the Markov chain. Section 10 illustrates these results for two-dimensional dynamical systems generated by discretized Möbius transformations of the unit disc in the complex plane. In Section 11, an algorithm is outlined for computing the transition probabilities of the set-valued Markov chain in the two-dimensional case. Each section begins with a brief heuristic summary of its contents, to improve readability.

2. Preliminaries

This section introduces some basic definitions and much of the notation used throughout. These include the form of round-off, the regular grids in which arithmetical operations occur, discretized mappings, reachable points, and the notion of frequency measurability.

Let T be a continuously differentiable diffeomorphism of a Jordan measurable open set $X \subset \mathbb{R}^n$. Recall that Jordan measurability means that X is bounded and has boundary ∂X with $\mathrm{mes}_n\, \partial X = 0$, where $\mathrm{mes}_n(\,\cdot\,)$ is n-dimensional Lebesgue measure. This mapping, acting as a transition operator, generates an autonomous dynamical system with phase space X.

Suppose that the system is to be simulated in fixed-point computer arithmetic with accuracy $\varepsilon > 0$. Although in binary arithmetic with b significant digits after the binary point, $\varepsilon = 2^{-b}$, it is convenient for our purposes that ε be regarded as a small parameter not necessarily of this specific form.

In machine arithmetic with accuracy ε, the only points of X which appear in their exact form are those of the ε-*grid*

$$X_\varepsilon = X \bigcap (\varepsilon \mathbb{Z}^n), \tag{1}$$

where $\mathbb{Z}^n$ denotes the n-dimensional integer lattice. Applying the mapping T to X_ε can, in general, yield points not representable in the arithmetic. Therefore, what is actually implemented on a computer is some *ε-discretization* $T_\varepsilon\colon\ X_\varepsilon \to \varepsilon\mathbb{Z}^n$ of the transition operator T.

In what follows, attention is restricted to the very simple discretization scheme where the mapping T_ε is given by the composition

$$T_\varepsilon = R_\varepsilon \circ T|_{X_\varepsilon}. \tag{2}$$

Here, R_ε is the *round-off operator* which maps a vector $u = (u_k)_{1\le k\le n} \in \mathbb{R}^n$ to the nearest node of the cubic lattice $\varepsilon\mathbb{Z}^n$,

$$R_\varepsilon(u) = \varepsilon(\lfloor u_k/\varepsilon + 1/2\rfloor)_{1\le k\le n} = \varepsilon R_1(u/\varepsilon), \tag{3}$$

where $\lfloor\cdot\rfloor$ is the floor function. Clearly, R_1 commutes with the additive group of translations of $\mathbb{Z}^n$, that is, $R_1(u+z) = R_1(u) + z$ for all $u \in \mathbb{R}^n$, $z \in \mathbb{Z}^n$, and the full preimage of the zero vector under the mapping is the half-open cube

$$R_1^{-1}(0) = V = [-1/2, 1/2)^n. \tag{4}$$

Practical implementation of the original mapping can, in general, be a multistage procedure with not only the final but also intermediate quantities being rounded off. Nevertheless, (2) is a reasonably good model for the computer discretization of simple mappings T which are computed by a few operations like multiplication or division performed with double precision.

Note that, because of boundary effects, the ε-grid (1) is generally not invariant with respect to the mapping (2). That is, iterations of T_ε can map an initial value $x \in X_\varepsilon$ to the complement of X, in which case the corresponding positive semitrajectory cannot be subsequently defined. Nevertheless, this technical point can be ignored in practice because of the following two reasons.

First, since T is a smooth diffeomorphism of an open set X, for any $r \in \mathbb{N}$ and any compact set $K \subset X$, the r first images of $K \bigcap X_\varepsilon$ lie in X_ε for all sufficiently small $\varepsilon > 0$,

$$T_\varepsilon^k\left(K \bigcap X_\varepsilon\right) \subset X_\varepsilon, \qquad 1 \le k \le r.$$

That is, for any given r, and for any arbitrarily thin $L = X\backslash K$, at least r iterates of T_ε are well defined for all ε small enough and for all initial values $x \in X_\varepsilon\backslash L$ separated from the boundary ∂X.

Secondly, this study is principally concerned with asymptotic properties of a fixed but otherwise arbitrarily large number of iterates of T_ε as $\varepsilon \to +0$. For this purpose, the mapping T_ε is considered to be the transition operator of a spatially discrete autonomous dynamical system with phase space X_ε, which is used as a *model* for the computer implementation of the original dynamical system in fixed-point arithmetic.

Invertibility of the original diffeomorphism T is normally not inherited by the discretization T_ε, which is in general neither surjective nor injective on the grid X_ε. Because of the loss of injectivity, positive semitrajectories of T_ε starting at distinct points can eventually coalesce. On the other hand, there can be points to which no points of the grid are mapped by T_ε. Consequently, the set of *k-reachable* points,

$$X_{\varepsilon,k} = X_\varepsilon \bigcap T_\varepsilon^k(X_\varepsilon) = \left\{x \in X_\varepsilon\colon\ T_\varepsilon^{-k}(x) \neq \emptyset\right\}, \tag{5}$$

is nonincreasing in $k \in \mathbb{Z}_+$. Here, $X_{\varepsilon,0} = X_\varepsilon$, and $T_\varepsilon^{-k}(x) = \{y \in X_\varepsilon\colon\ T_\varepsilon^k(y) = x\}$ denotes the k-th full preimage of a point $x \in X_\varepsilon$ under T_ε.

Heuristically, points of X_ε under the action of T_ε can be thought of as acting like particles of a compressible fluid flow evolving in discrete time. In k steps of its time evolution, the fluid flow fills up the set $X_{\varepsilon,k} \subset X_\varepsilon$ and never again visits $X_\varepsilon \backslash X_{\varepsilon,k}$. Furthermore, the fluid is eventually absorbed by the set

$$C_\varepsilon = \bigcap_{k \geq 1} X_{\varepsilon,k} \tag{6}$$

consisting of limit cycles of the mapping T_ε. A surprising feature of this behaviour for small ε is that the total length of the cycles $\#C_\varepsilon$ is usually extremely small in comparison with the number of points $\#X_\varepsilon$ in the set X_ε, even if the original transition operator T has an absolutely continuous invariant measure with strictly positive density. This *degeneration* or *collapse* of T_ε in its phase space X_ε is an intrinsically discrete phenomenon which can be quantified by the ratios

$$\#X_{\varepsilon,k}/\#X_\varepsilon. \tag{7}$$

The denominator $\#X_\varepsilon$ of these grows asymptotically like $\varepsilon^{-n}\ \mathrm{mes}_n\ X$ as $\varepsilon \to +0$. Hence, the asymptotic behaviour of (7) for small ε is dictated by that of $\varepsilon^n \# X_{\varepsilon,k}$. Indeed, it is shown below that, under a nonresonance condition on the original diffeomorphism T, the set of k-reachable points (5) has an asymptotic absolutely continuous spatial distribution in the following sense.

A set $A_\varepsilon \subset X_\varepsilon$, parameterized by the stepsize of the grid $\varepsilon > 0$, is said to be *frequency measurable*, with *frequency function* $f\colon\ X \to [0, 1]$, if for any Jordan measurable set $G \subset X$,

$$\lim_{\varepsilon \to +0} \left(\varepsilon^n \#(A_\varepsilon \bigcap G)\right) = \int_G f(x)\, dx.$$

Note that frequency measurability is an asymptotic property of the set-valued mapping $0 < \varepsilon \mapsto A_\varepsilon \subset X_\varepsilon$ as $\varepsilon \to +0$, but not of its individual value for a given ε. Clearly, the grid X_ε itself is frequency measurable with constant frequency function 1.

In what follows, various subsets of X_ε arise which, although highly irregular in contrast to the grid X_ε, are still amenable to asymptotic statistical analysis in terms of frequency measurability as $\varepsilon \to 0+$.

3. Compensating System and Quantization Errors

Here, the key technical tools for the ensuing mathematical development are set out. The compensating system is induced by the discretization of the inverse of the underlying diffeomorphism. Quantization errors are normalized round-off errors of an iterated mapping and a recurrence is developed for set-valued perturbations which arise as a consequence.

Let $U = T^{-1}$ be the inverse of the diffeomorphism T. As in (2), define the ε-discretization of the diffeomorphism U as the mapping

$$U_\varepsilon = R_\varepsilon \circ U|_{X_\varepsilon}. \tag{8}$$

The dynamical system with transition operator U_ε will be called the *compensating system*. The significance of this is demonstrated by Lemma 1 below which shows that every negative semitrajectory of T_ε can be split into single-valued and set-valued parts, with the first a positive semitrajectory of U_ε and the second expressed in terms of appropriately defined quantization errors for the compensating system.

With the round-off operator (3), associate the *normalized round-off error* E_ε: $\mathbb{R}^n \to V$, taking values in the cube (4), defined by

$$E_\varepsilon(u) = (u - R_\varepsilon(u))/\varepsilon = E_1(u/\varepsilon). \tag{9}$$

Clearly, E_1 is unit periodic in each of its n variables and maps the cube V onto itself as the identity. That is, $E_1(u + z) = u$ for all $u \in V$, $z \in \mathbb{Z}^n$. Following [18], define the k-th *quantization error* of the compensating system as the mapping

$$E_{\varepsilon,k} = E_\varepsilon \circ U \circ U_\varepsilon^{k-1} = E_{\varepsilon,1} \circ U_\varepsilon^{k-1}. \tag{10}$$

To formulate Lemma 1, some additional notation is required. Write $\mathcal{Z}$ for the class of finite subsets of $\mathbb{Z}^n$, including the empty set, and endowed with the discrete topology,

$$\mathcal{Z} = \{A \subset \mathbb{Z}^n\colon \#A < +\infty\}. \tag{11}$$

Define the set-valued mapping F_ε: $X \times \mathcal{Z} \times V \to \mathcal{Z}$ by

$$F_\varepsilon(x, A, v) = \mathbb{Z}^n \bigcap (G_\varepsilon(x, A) + v), \tag{12}$$

where

$$G_\varepsilon(x, A) = (U(x + \varepsilon(A + V)) - U(x))/\varepsilon. \tag{13}$$

Here, $B + C = \{b + c\colon\ b \in B,\ c \in C\}$ denotes the Minkowski sum of subsets B and C of a vector space, and the convention that $\emptyset + C = \emptyset$ is used.

Lemma 1. *For any $k \in \mathbb{N}$, the k-th preimage of a point $x \in X_\varepsilon$ under the mapping T_ε has a representation*

$$T_\varepsilon^{-k}(x) = U_\varepsilon^k(x) + \varepsilon S_{\varepsilon,k}(x), \tag{14}$$

where the set-valued mappings $S_{\varepsilon,k}$: $X_\varepsilon \to \mathcal{Z}$ are defined by the recurrence

$$S_{\varepsilon,k+1}(x) = F_\varepsilon(U_\varepsilon^k(x), S_{\varepsilon,k}(x), E_{\varepsilon,k+1}(x)), \tag{15}$$

with initial condition

$$S_{\varepsilon,0}(x) = \{0\}. \tag{16}$$

Proof. Using (8)–(13) and the property that $\varepsilon\mathbb{Z}^n$ is an additive group, it is straightforward to verify that for any $x \in X_\varepsilon$ and any set $A \in \mathcal{Z}$,

$$\begin{aligned} T_\varepsilon^{-1}(x+\varepsilon A) &= (\varepsilon\mathbb{Z}^n)\bigcap U(x+\varepsilon(A+V)) \\ &= (\varepsilon\mathbb{Z}^n)\bigcap(\varepsilon G_\varepsilon(x,A)+U(x)) \\ &= (\varepsilon\mathbb{Z}^n)\bigcap(\varepsilon(G_\varepsilon(x,A)+E_{\varepsilon,1}(x))+U_\varepsilon(x)) \\ &= U_\varepsilon(x)+\varepsilon\left(\mathbb{Z}^n\bigcap(G_\varepsilon(x,A)+E_{\varepsilon,1}(x))\right) \\ &= U_\varepsilon(x)+\varepsilon F_\varepsilon(x,A,E_{\varepsilon,1}(x)). \end{aligned} \tag{17}$$

Proof of (14) now proceeds by induction on $k \in \mathbb{Z}_+$ as follows. For $k = 0$, the representation is clearly true by (16). Suppose that (14) holds for some $k \in \mathbb{Z}_+$. Then, applying (17) and using the definition (15), obtain

$$\begin{aligned} T_\varepsilon^{-(k+1)}(x) &= T_\varepsilon^{-1}(U_\varepsilon^k(x)+\varepsilon S_{\varepsilon,k}(x)) \\ &= U_\varepsilon(U_\varepsilon^k(x))+\varepsilon F_\varepsilon(U_\varepsilon^k(x),S_{\varepsilon,k}(x),E_{\varepsilon,1}(U_\varepsilon^k(x))) \\ &= U_\varepsilon^{k+1}(x)+\varepsilon S_{\varepsilon,k+1}(x), \end{aligned}$$

which completes the proof of the lemma. □

From Lemma 1, it follows that each of the mappings $S_{\varepsilon,k}$ can be expressed in terms of k first quantization errors of the compensating system by

$$S_{\varepsilon,k}(x) = H_{\varepsilon,k}\left(x, E_{\varepsilon,1}(x), \ldots, E_{\varepsilon,k}(x)\right), \tag{18}$$

where the mappings $H_{\varepsilon,k}\colon\ X \times V^k \to \mathcal{Z}$ are defined by the recurrence

$$H_{\varepsilon,k+1}(x, y_1, \ldots, y_{k+1}) = F_\varepsilon\left((R_\varepsilon \circ U)^k(x), H_{\varepsilon,k}(x, y_1, \ldots, y_k), y_{k+1}\right), \tag{19}$$

for all $k \in \mathbb{Z}_+$, $x \in X$ and $y_1, \ldots, y_{k+1} \in V$, with initial condition

$$H_{\varepsilon,0}(x) = \{0\}. \tag{20}$$

4. Continuous Convergence of Mappings

The ideas of this section govern the regularity, as gridsize goes to zero, of the set-valued representation arising in the previous section. The Jacobian matrix of the inverse diffeomorphism appears and plays an important theoretical role in this and subsequent sections.

A mapping $\Phi_\varepsilon\colon\ \Omega \to \Upsilon$, parameterized by $\varepsilon > 0$, of a set $\Omega \subset \mathbb{R}^r$ to a separable metric space Υ is said to be *continuously convergent* to a mapping Φ as $\varepsilon \to +0$, written as $\Phi_\varepsilon \xrightarrow{\mathrm{c}} \Phi$, if

$$\lim_{(\varepsilon,y)\to(+0,x)} \Phi_\varepsilon(y) = \Phi(x), \quad \text{for } \mathrm{mes}_r\text{-almost all } x \in \Omega.$$

Lemma 2. *The mapping F_ε in (12) is continuously convergent to the mapping $F\colon X \times \mathcal{Z} \times V \to \mathcal{Z}$, as $\varepsilon \to +0$, given by*

$$F(x, A, v) = \mathbb{Z}^n \bigcap (U'(x)(A + V) + v), \tag{21}$$

where $U'(x)$ denotes the Jacobian matrix of the inverse diffeomorphism U.

Proof. Since $\mathcal{Z}$ is equipped with the discrete topology, the continuous convergence $F_\varepsilon \xrightarrow{c} F$ will be proved if it is shown that for any given $A \in \mathcal{Z}$,

$$\lim_{(\varepsilon, y, w) \to (+0, x, v)} F_\varepsilon(y, A, w) = F(x, A, v),$$
$$\text{for } \operatorname{mes}_{2n} - \text{almost all } (x, v) \in X \times V. \tag{22}$$

Note that $U'(x)(A + V)$ in (21) is the derivative of the smooth diffeomorphism U at a point $x \in X$ along the set $A + V$ [1]. Therefore, the definitions (12)–(13) easily imply that the set

$$\Gamma(x, A) = \{v \in V\colon\ F_\varepsilon(y, A, w) \not\to F(x, A, v), \quad \text{as } (\varepsilon, y, w) \to (+0, x, v)\}, \tag{23}$$

where $\not\to$ signifies lack of convergence, satisfies

$$\Gamma(x, A) \subset V \bigcap (-U'(x)\, \partial(A + V) + \mathbb{Z}^n).$$

The set on the right of this last inclusion is contained in a union of finitely many $(n-1)$-dimensional hyperplanes, and consequently,

$$\operatorname{mes}_n \Gamma(x, A) = 0. \tag{24}$$

Since $x \in X$ was chosen arbitrarily,

$$\operatorname{mes}_{2n}\{(x, v)\colon\ x \in X,\ v \in \Gamma(x, A)\} = \int_X \operatorname{mes}_n \Gamma(x, A)\, dx = 0,$$

which, by the definition (23), immediately yields (22), and the proof of the lemma is complete. □

Using the smoothness of U and the uniform boundedness of the normalized round-off error, it is easy to verify inductively that for any $k \in \mathbb{N}$,

$$\sup_{x \in X} \left|(R_\varepsilon \circ U)^k(x) - U^k(x)\right| \to 0, \quad \text{as } \varepsilon \to +0. \tag{25}$$

Hence, Lemma 2 makes it sensible to expect that the asymptotic behaviour of the mappings $H_{\varepsilon,k}$ in (19)–(20) can be described in terms of the mappings $H_k\colon X \times V^k \to \mathcal{Z}$ satisfying the recurrence

$$H_{k+1}(x, y_1, \ldots, y_{k+1}) = F\left(U^k(x), H_k(x, y_1, \ldots, y_k), y_{k+1}\right), \tag{26}$$

for all $k \in \mathbb{Z}_+$, $x \in X$ and $y_1, \ldots, y_{k+1} \in V$, with initial condition

$$H_0(x) = \{0\}. \tag{27}$$

Note that (26) is obtained by formal replacement of F_ε and $R_\varepsilon \circ U$ in (19) with F and U, respectively.

Lemma 3. *For any $k \in \mathbb{N}$,*

$$H_{\varepsilon,k} \xrightarrow{\mathrm{c}} H_k, \quad \text{as } \varepsilon \to +0. \tag{28}$$

Proof. Proof proceeds by induction on $k \in \mathbb{Z}_+$. For $k = 0$, the convergence (28) follows immediately from (20) and (27), from which $H_{\varepsilon,0} = H_0$. Suppose that the assertion of the lemma holds for some $k \in \mathbb{Z}_+$ or, equivalently, that the set

$$\Lambda_k = \left\{ u \in X \times V^k \colon \lim_{(\varepsilon,w)\to(+0,u)} H_{\varepsilon,k}(w) = H_k(u) \right\}$$

has full $(k+1)n$-dimensional Lebesgue measure. Then, by definition (23) and the uniform convergence (25), the set Λ_{k+1}, defined similarly for the mappings $H_{\varepsilon,k+1}$ and H_{k+1}, satisfies

$$\Lambda_{k+1} \supset \left\{ (x, y, v) \in X \times V^{k+1} \colon (x, y) \in \Lambda_k,\ v \in V \backslash \Gamma(U^k(x), H_k(x, y)) \right\}.$$

Hence, by (24),

$$\begin{aligned} \mathrm{mes}_{(k+2)n} \left((X \times V^{k+1}) \backslash \Lambda_{k+1} \right) &= \mathrm{mes}_{(k+2)n} \left((\Lambda_k \times V) \backslash \Lambda_{k+1} \right) \\ &\le \int_{\Lambda_k} \mathrm{mes}_n\, \Gamma \left(U^k(x), H_k(x, y) \right) (dx \times dy) = 0. \end{aligned}$$

This last relation implies that the set Λ_{k+1} is of full $(k+2)n$-dimensional Lebesgue measure, thereby finishing the inductive step for (28), and the proof of the lemma is complete. ☐

5. Asymptotic Distribution of Quantization Errors

Below, the idea of resonance is introduced. This is essentially a form of rational dependence. The results of the paper are true if the rows of Jacobian matrices of iterates of T are mutually nonresonant almost everywhere. Although technical, such ideas of independence arise frequently in quite different areas; see for example in functional equations [31], [44], celestial mechanics [42], and computer science [24], [32]. The number theoretic property is used to set up machinery for the existence of limiting distributions. Although the error distribution is asymptotically uniform, again it should be emphasized that the computer effects that are observed, like collapse and distortion of invariant measures, are not uniformly distributed.

Say that a matrix *resonates* if its rows are linearly dependent over the field of real rationals. Associate with the transition operator T its *resonance set*

$$\mathcal{R}(T) = \bigcup_{k\in\mathbb{N}} \mathcal{R}_k(T), \tag{29}$$

where

$$\mathcal{R}_k(T) = \left\{ x \in X \colon \begin{bmatrix} I_n \\ T'(x) \\ \vdots \\ (T^k)'(x) \end{bmatrix} \in \mathbb{R}^{(k+1)n\times n} \text{ resonates} \right\}, \tag{30}$$

where I_n denotes the identity matrix of order n. The mapping T will be called *iteratively nonresonant* if $\mathrm{mes}_n\, \mathcal{R}(T) = 0$.

Lemma 4. *Let the diffeomorphism T be iteratively nonresonant. Then for any $k \in \mathbb{N}$ and any bounded continuous function $f\colon X \times V^k \to \mathbb{R}$, the quantization errors (10) of the compensating system satisfy*

$$\lim_{\varepsilon \to +0} \left(\varepsilon^n \sum_{x \in X_\varepsilon} f(x, E_{\varepsilon,1}(x), \dots, E_{\varepsilon,k}(x)) \right) = \int_{X \times V^k} f(z)\, dz. \tag{31}$$

Proof. From

$$\begin{bmatrix} I_n \\ T'(U^k(x)) \\ \vdots \\ (T^k)'(U^k(x)) \end{bmatrix} (U^k)'(x) = \begin{bmatrix} (U^k)'(x) \\ \vdots \\ U'(x) \\ I_n \end{bmatrix},$$

it follows that the sets (30), considered for the inverse diffeomorphism U, satisfy $\mathcal{R}_k(U) = T^k(\mathcal{R}_k(T))$. Since T is a smooth diffeomorphism, the condition $\mathrm{mes}_n\, \mathcal{R}(T) = 0$ is equivalent to $\mathrm{mes}_n\, \mathcal{R}(U) = 0$, so U is also iteratively nonresonant. Therefore, applying [18, Theorem 1] to the compensating system U_ε gives the asymptotic independence and uniform distribution of the quantization errors on the cube V, in the sense of (31). □

6. Set-Valued Markov Chains

Here, another key tool is introduced: a Markov chain whose states are finite subsets of the integer lattice $\mathbb{Z}^n$. Even though quantization errors have, asymptotically, a uniform distribution, the transition probabilities of the chain obviously cannot. The chain describes the forward evolution of the discretization of T on the grid by going backwards via the inverse of the discretized diffeomorphism. The significance of this is that, starting from a given point x of the grid, if after k steps of the chain an empty state is reached, that is extinction, it means that the point x is not k-reachable. Thus the transition probabilities of the chain, and its extinction probabilities, provide a means of studying the asymptotic distributions of sets of reachable points.

Using the mapping F given by (21), associate with every point $x \in X$ a nonhomogeneous set-valued Markov chain $\sigma_x = (\sigma_{x,k})_{k \in \mathbb{Z}_+}$ with the denumerable state space $\mathcal{Z}$ in (11) and defined by the recurrence

$$\sigma_{x,k} = F(U^{k-1}(x),\ \sigma_{x,k-1},\ \omega_k), \tag{32}$$

where ω_k are independent random vectors distributed uniformly on the cube V. The transition probabilities of the chain at the k-th step of its evolution are described by the function $P_k\colon \mathcal{Z}^2 \times X \to [0, 1]$ given by

$$\begin{aligned} P_k(B \mid A, x) &= \mathbf{P}\left(\sigma_{x,k} = B \mid \sigma_{x,k-1} = A\right) \\ &= \mathrm{mes}_n \left\{ v \in V\colon\ F\left(U^{k-1}(x), A, v\right) = B \right\} \\ &= P_1\left(B \mid A, U^{k-1}(x)\right), \end{aligned} \tag{33}$$

where $\mathbf{P}(\,\cdot \mid \cdot\,)$ denotes the conditional probability on the underlying probability space. For any $r \in \mathbb{N}$, the conditional joint distribution of the r-th initial segment of σ_x is given by the probabilities

$$\mathbf{P}\left(\sigma_{x,1} = A_1,\ \ldots,\ \sigma_{x,r} = A_r \mid \sigma_{x,0} = A_0\right) = \prod_{k=1}^{r} P_k\left(A_k \mid A_{k-1}, x\right), \tag{34}$$

for all $A_0, \ldots, A_r \in \mathcal{Z}$. Note that each of the functions $P_k(B \mid A, x)$ is continuous in $x \in X$ for any fixed $A, B \in \mathcal{Z}$. Hence, so also are all the functions (34). Theorem 1 below shows that the set-valued mappings $S_{\varepsilon,k}(x)$ defined by (15)–(16) are asymptotically distributed like the elements $\sigma_{x,k}$ of the Markov chain σ_x.

Theorem 1. *Let the diffeomorphism T be iteratively nonresonant. Then for any $r \in \mathbb{N}$ and any sets $A_1, \ldots, A_r \in \mathcal{Z}$, the set*

$$\left\{x \in X_\varepsilon\colon\ S_{\varepsilon,1}(x) = A_1,\ \ldots,\ S_{\varepsilon,r}(x) = A_r\right\}$$

is frequency measurable with frequency function

$$\mathbf{P}\left(\sigma_{x,1} = A_1,\ \ldots,\ \sigma_{x,r} = A_r \mid \sigma_{x,0} = \{0\}\right)$$

given by (34).

Proof. For any $\varepsilon > 0$ and $r \in \mathbb{N}$, define a countably additive measure $\lambda_{\varepsilon,r}$ on Borel sets $B \subset X \times V^r$ by

$$\lambda_{\varepsilon,r}(B) = \varepsilon^n \,\#\left\{x \in X_\varepsilon\colon\ \left(x,\ E_{\varepsilon,1}(x),\ \ldots,\ E_{\varepsilon,r}(x)\right) \in B\right\}. \tag{35}$$

By Lemma 4, the measure converges weakly to $\mathrm{mes}_{(r+1)n}$,

$$\lambda_{\varepsilon,r} \xrightarrow{\mathrm{w}} \mathrm{mes}_{(r+1)n}, \quad \text{as } \varepsilon \to +0, \tag{36}$$

(see [5] for the general definition of weak convergence of measures). Assembling the mappings $H_{\varepsilon,k}$ given by (19)–(20), define the mapping $W_{\varepsilon,r}\colon X \times V^r \to X \times \mathcal{Z}^r$ by

$$W_{\varepsilon,r}(x, y_1, \ldots, y_r) = \left(x,\ H_{\varepsilon,1}(x, y_1),\ \ldots,\ H_{\varepsilon,r}(x, y_1, \ldots, y_r)\right). \tag{37}$$

Lemma 3 easily implies the continuous convergence

$$W_{\varepsilon,r} \xrightarrow{\mathrm{c}} W_r, \quad \text{as } \varepsilon \to +0, \tag{38}$$

where the limiting mapping W_r is given by

$$W_r(x, y_1, \ldots, y_r) = (x,\ H_1(x, y_1),\ \ldots,\ H_r(x, y_1, \ldots, y_r)). \tag{39}$$

Consider the countably additive measure $\mu_{\varepsilon,r}$ defined on Borel subsets of $X \times \mathcal{Z}^r$ by

$$\mu_{\varepsilon,r} = \lambda_{\varepsilon,r} \circ W_{\varepsilon,r}^{-1}.$$

Applying the results of [45] (see also [5, Theorem 5.5 on p. 34]), from (36) and (38), obtain that

$$\mu_{\varepsilon,r} \xrightarrow{\mathrm{w}} \mathrm{mes}_{(r+1)n} \circ W_r^{-1} = \mu_r. \tag{40}$$

The definitions (26)–(27), (33), (34), and (39) imply that for any Borel set $G \subset X$ and any sets $A_1, \ldots, A_r \in \mathcal{Z}$, the value of the limiting measure μ_r on the set

$$B = G \times (A_1, \ldots, A_r) \tag{41}$$

is given by

$$\begin{aligned} \mu_r(B) &= \int_G \mathrm{mes}_{rn}\{(v_1, \ldots, v_r) \in V^r\colon\ F(U^{k-1}(x), A_{k-1}, v_k) = A_k,\ 1 \le k \le r\}\, dx \\ &= \int_G \prod_{k=1}^{r} P_k\,(A_k \mid A_{k-1}, x)\ dx \\ &= \int_G \mathbf{P}\left(\sigma_{x,1} = A_1,\ \ldots,\ \sigma_{x,r} = A_r \mid \sigma_{x,0} = \{0\}\right)\, dx, \end{aligned} \tag{42}$$

where $A_0 = \{0\}$. Therefore, if G is a Jordan measurable subset of X, the set (41) is μ_r-continuous in the sense that $\mu_r(\partial B) = 0$ and applying to (40) the well-known criterion for weak convergence of measures yields

$$\lim_{\varepsilon \to +0} \mu_{\varepsilon,r}(B) = \mu_r(B). \tag{43}$$

Now, note that from the representation (18) and (35), (37), it easily follows that for the set (41),

$$\mu_{\varepsilon,r}(B) = \varepsilon^n\, \#\{x \in G \bigcap X_\varepsilon\colon\ S_{\varepsilon,k}(x) = A_k,\ 1 \le k \le r\}. \tag{44}$$

Assembling (42)–(44), obtain that

$$\begin{aligned} &\lim_{\varepsilon \to +0} (\varepsilon^n \#\{x \in G \bigcap X_\varepsilon\colon\ S_{\varepsilon,1}(x) = A_1,\ \ldots,\ S_{\varepsilon,r}(x) = A_r\}) \\ &\qquad = \int_G \mathbf{P}(\sigma_{x,1} = A_1,\ \ldots,\ \sigma_{x,r} = A_r \mid \sigma_{x,0} = \{0\})\, dx, \end{aligned}$$

for any Jordan measurable set $G \subset X$, thereby completing the proof of the theorem. $\square$

7. Frequency Functions of Reachable Points

Enough technical machinery has now been set up to investigate the asymptotic distribution of reachable points. The frequency functions of the sets of grid points reachable in k iterations of the discretized map is below shown to satisfy a recurrence whose coefficients depend on the transition probabilities of the set-valued Markov chain. Although the proof is technical and involves the nonresonance condition, the recurrence does give a computational procedure for evaluating the distribution.

Theorem 2. *Let the diffeomorphism T be iteratively nonresonant. Then for any $k \in \mathbb{N}$, the set of k-reachable points $X_{\varepsilon,k}$ in (5) is frequency measurable with frequency function*

$$q_k(x) = Q_k(x, \{0\}), \tag{45}$$

where the functions $Q_k\colon X \times \mathcal{Z} \to [0, 1]$ satisfy the recurrence

$$Q_{k+1}(x, A) = \sum_{B \in \mathcal{Z}} Q_k(U(x), B)\, P_1(B \mid A, x), \tag{46}$$

for all $k \in \mathbb{Z}_+$, $x \in X$, and $A \in \mathcal{Z}$, with initial condition

$$Q_0(x, A) = \begin{cases} 0 & \text{for } A = \emptyset \\ 1 & \text{otherwise.} \end{cases} \tag{47}$$

Proof. Since $P_k(\emptyset \mid \emptyset, x) = 1$ for all $k \in \mathbb{N}$ and $x \in X$, the empty set is an absorbing state of the Markov chain σ_x. The random event $\{\sigma_{x,k} = \emptyset\}$ will be interpreted as *extinction* of the chain during the k first steps of its evolution. For any $A \in \mathcal{Z}$, denote the conditional probability of the complementary event by

$$Q_k(x, A) = \mathbf{P}(\sigma_{x,k} \neq \emptyset \mid \sigma_{x,0} = A). \tag{48}$$

$Q_k\colon X \times \mathcal{Z} \to [0, 1]$ will be called the k-th *nonextinction probability function*. Since the set of k-reachable points can be written as $X_{\varepsilon,k} = \{x \in X_\varepsilon\colon S_{\varepsilon,k}(x) \neq \emptyset\}$, from Theorem 1 it follows that for any Jordan measurable set $G \subset X$,

$$\lim_{\varepsilon \to +0} \left(\varepsilon^n \,\#\left(G \bigcap X_{\varepsilon,k}\right)\right) = \int_G Q_k(x, \{0\})\, dx. \tag{49}$$

It now remains to derive the recurrence for the functions Q_k defined in (48). Clearly, Q_0 satisfies (47). For any $k \in \mathbb{N}$, the function Q_k is just

$$Q_k\,(x, A_0) = \sum_{A_1,\ldots,A_k \in \mathcal{Z}\colon\ A_k \neq \emptyset}\ \prod_{j=1}^{k} P_j\left(A_j \mid A_{j-1}, x\right),$$

for all $x \in X$ and $A_0 \in \mathcal{Z}$. Hence, using the property that $P_{k+1}(B \mid A, x) = P_k(B \mid A, U(x))$, which immediately follows from the definition of the transition probabilities (33), obtain that for any $x \in X$ and any $B \in \mathcal{Z}$,

$$Q_k(U(x), B) = \mathbf{P}\left(\sigma_{x,k+1} \neq \emptyset \mid \sigma_{x,1} = B\right). \tag{50}$$

Therefore,

$$\begin{aligned} Q_{k+1}(x, A) &= \mathbf{P}\left(\sigma_{x,k+1} \neq \emptyset \mid \sigma_{x,0} = 0\right) \\ &= \sum_{B \in \mathcal{Z}} \mathbf{P}\left(\sigma_{x,k+1} \neq \emptyset \mid \sigma_{x,1} = B\right) P_1(B \mid A, x) \\ &= \sum_{B \in \mathcal{Z}} Q_k(U(x), B) P_1(B \mid A, x). \end{aligned}$$

This last implies that the functions Q_k satisfy (46)–(47) which, together with (49), completes the proof of the theorem. □

In particular, Theorem 2 implies the following convergence of the ratios (7):

$$\lim_{\varepsilon\to+0}\frac{\#X_{\varepsilon,k}}{\#X_\varepsilon}=\frac{1}{\mathrm{mes}_n\, X}\int_X q_k(x)\,dx. \tag{51}$$

From the definition of the nonextinction probability functions (48), it follows that $Q_k(x, A)$ is nonincreasing in $k \in \mathbb{Z}_+$. Moreover, using the continuity of the functions (33) and the recurrence (46)–(47), it can be shown inductively that $Q_k(x, A)$ are all continuous in $x \in X$. Consequently, the frequency functions (45) are continuous and are also a nonincreasing sequence, $q_{k+1}(x) \le q_k(x)$ for all $k \in \mathbb{N}$, $x \in X$. The monotonicity implies the existence of the limiting functions $q_\infty\colon X \to [0, 1]$ and $Q_\infty\colon X \times \mathcal{Z} \to [0, 1]$ satisfying

$$\begin{aligned} q_\infty(x) &= \lim_{k\to+\infty} q_k(x) = Q_\infty(x, \{0\}), && (52)\\ Q_\infty(x, A) &= \lim_{k\to+\infty} Q_k(x, A) \\ &= \sum_{B\in\mathcal{Z}} Q_\infty(U(x), B)\, P_1(B \mid A, x), \qquad Q_\infty(x, \emptyset) = 0. && (53)\end{aligned}$$

Applying the Lebesgue Dominated Convergence Theorem, from (51) and (52) obtain that the total length of the limit cycles in (6) satisfies

$$\limsup_{\varepsilon\to+0}\frac{\#C_\varepsilon}{\#X_\varepsilon}\le\frac{1}{\mathrm{mes}_n\, X}\int_X q_\infty(x)\,dx. \tag{54}$$

The first of the functions (45) gives an explicit representation in terms of the Jacobian matrix of the inverse diffeomorphism U,

$$\begin{aligned} q_1(x) &= 1 - P_1(\emptyset \mid \{0\}, x)\\ &= \mathrm{mes}_n\{v \in V\colon\ F(x, \{0\}, v) \ne \emptyset\}\\ &= \mathrm{mes}_n \left(V \bigcap (U'(x)V + \mathbb{Z}^n)\right)\\ &= \mathrm{mes}_n\, E_1(U'(x)V), && (55)\end{aligned}$$

where the relations

$$\left\{v \in V\colon\ \mathbb{Z}^n \bigcap (G + v) \ne \emptyset\right\} = V \bigcap (-G + \mathbb{Z}^n) = E_1(-G),$$

for any $G \subset \mathbb{R}^n$, have been used. From (55) it follows that q_1 takes strictly positive values. Moreover,

$$\max\left(1, \left\|T'(U(x))\right\|\right)^{-n} \le q_1(x) \le |\det U'(x)|,$$

where $\|M\| = \max_{1\le j\le n}\sum_{k=1}^n |m_{jk}|$ denotes the maximal absolute row sum norm of a matrix $M = (m_{jk})_{1\le j,k\le n}$ induced by the ℓ_∞ vector norm in $\mathbb{R}^n$. Further properties of the functions q_k and Q_k are studied in Section 9.

8. Martingale Property of the Set-Valued Markov Chain

Although the structure of the Markov chain may appear complicated, not least for having set-valued states, it does have some nice properties. For example, the number of points in the states at the k-th step of the chain forms a sequence of random variables which, when suitably normalized, has the martingale property. The normalization factor depends on the Jacobian of the k-th iterate of the inverse diffeomorphism.

For each $x \in X$ and each corresponding set-valued Markov chain σ_x in (32), define a sequence of integer-valued random variables $\nu_{x,k}$ with values in $\mathbb{Z}_+$ given by

$$\nu_{x,k} = \#\sigma_{x,k}. \tag{56}$$

For any $r \in \mathbb{N}$, the conditional joint distribution of $\nu_{x,1}, \ldots, \nu_{x,r}$, given $\sigma_{x,0} = \{0\}$, is given by $\Pi_{r,\cdot}(x)\colon \mathbb{Z}_+^r \to [0,1]$, where

$$\begin{aligned} \Pi_{r,m_1,\ldots,m_r}(x) &= \mathbf{P}(\nu_{x,1} = m_1, \ldots, \nu_{x,r} = m_r \mid \sigma_{x,0} = \{0\}) \\ &= \sum_{\#A_1 = m_1, \ldots, \#A_r = m_r} \prod_{k=1}^{r} P_k(A_k \mid A_{k-1}, x), \qquad A_0 = \{0\}. \end{aligned} \tag{57}$$

From the representation (14) and Theorem 1, it follows that the functions $N_{\varepsilon,k}\colon X_\varepsilon \to \mathbb{Z}_+$, defined by

$$N_{\varepsilon,k}(x) = \#T_\varepsilon^{-k}(x) = \#S_{\varepsilon,k}(x),$$

which give the number of points in states of the set-valued negative semitrajectories of T_ε, are asymptotically distributed as the random variables $\nu_{x,k}$. That is, for any $r \in \mathbb{N}$ and any $m_1, \ldots, m_r \in \mathbb{Z}_+$, the set

$$\left\{x \in X_\varepsilon\colon\ N_{\varepsilon,1}(x) = m_1,\ \ldots,\ N_{\varepsilon,r}(x) = m_r\right\}$$

is frequency measurable with the frequency function $\Pi_{r,m_1,\ldots,m_r}\colon X \to [0,1]$ given in (57).

The following lemma shows that, when appropriately normalized, the random variables (56) form a martingale (for the general definition of martingales and related properties of conditional expectations with respect to sub-σ-algebras, see for example, [41, Sections VII.1, II.7]).

Lemma 5. *For any fixed $x \in X$, the random variables*

$$\rho_{x,k} = \frac{\nu_{x,k}}{|\det(U^k)'(x)|} \tag{58}$$

form a martingale. That is, for any $k \in \mathbb{Z}_+$, $\mathbf{E}(\rho_{x,k+1} \mid \rho_{x,0}, \ldots, \rho_{x,k}) = \rho_{x,k}$ holds almost surely (here, $\mathbf{E}(\cdot \mid \cdot)$ denotes the conditional expectation).

Proof. Since the random variables $\rho_{x,k}$ are deterministic functions of the corresponding elements $\sigma_{x,k}$ of the Markov chain σ_x, the equations

$$\begin{aligned} \mathbf{E}\left(\rho_{x,k+1} \mid \rho_{x,0}, \ldots, \rho_{x,k}\right) &= \mathbf{E}\left(\mathbf{E}\left(\rho_{x,k+1} \mid \sigma_{x,0}, \ldots, \sigma_{x,k}\right) \mid \rho_{x,0}, \ldots, \rho_{x,k}\right) \\ &= \frac{\mathbf{E}\left(\mathbf{E}\left(\nu_{x,k+1} \mid \sigma_{x,k}\right) \mid \rho_{x,0}, \ldots, \rho_{x,k}\right)}{|\det(U^{k+1})'(x)|} \end{aligned}$$

hold almost surely. Hence, since $(U^{k+1})'(x) = U'(U^k(x))\,(U^k)'(x)$, it is sufficient to show that

$$\mathbf{E}\left(\nu_{x,k+1} \mid \sigma_{x,k} = A\right) = \left|\det U'\left(U^k(x)\right)\right| \#A, \tag{59}$$

for any $x \in X$, $k \in \mathbb{Z}_+$, and $A \in \mathcal{Z}$. To show this, recall the following result from the theory of geometric probability: For any Lebesgue measurable set $G \subset \mathbb{R}^n$,

$$\int_V \#\left(\mathbb{Z}^n \bigcap (G+v)\right) dv = \sum_{z\in\mathbb{Z}^n} \int_V \mathcal{I}_{G+v}(z)\, dv \tag{60}$$

$$= \sum_{z\in\mathbb{Z}^n} \mathrm{mes}_n \left(G \bigcap (V+z)\right) = \mathrm{mes}_n\, G, \tag{61}$$

where $\mathcal{I}_M(\,\cdot\,)$ is the indicator function of a set M. The property that translations $V+z$ of the cube V by vectors $z \in \mathbb{Z}^n$ form a partitioning of $\mathbb{R}^n$ has been used here. In particular, this property implies that $\mathrm{mes}_n(A+V) = \#A$ for any set $A \in \mathcal{Z}$. Using (61), (21), (32), and (56), obtain that for any set $A \in \mathcal{Z}$,

$$\begin{aligned}\mathbf{E}\left(\nu_{x,k+1} \mid \sigma_{x,k} = A\right) &= \int_V \#F\left(U^k(x), A, v\right) dv \\ &= \mathrm{mes}_n \left(U'\left(U^k(x)\right)(A+V)\right) = \left|\det U'\left(U^k(x)\right)\right| \#A,\end{aligned}$$

which is just (59). □

Using Lemma 5 and applying the Doob Martingale Convergence Theorem (see, for example, [25, Theorem 2.5 on p. 17] or [41, Theorem 1 and Corollary 3 on pp. 508–509]), obtain that for any fixed $x \in X$, the nonnegative martingale $\{\rho_{x,k}\}_{k\in\mathbb{N}}$ in (58) converges almost surely,

$$\mathbf{P}\left(\lim_{k\to+\infty} \rho_{x,k} = \rho_{x,\infty} \mid \sigma_{x,0} = \{0\}\right) = 1. \tag{62}$$

The probability distribution of the limiting random variable $\rho_{x,\infty}$ is parameterized by $x \in X$ and satisfies the inequality $\mathbf{E}(\rho_{x,\infty} \mid \sigma_{x,0} = \{0\}) \le 1$. Moreover,

$$\mathbf{P}(\rho_{x,\infty} = 0 \mid \sigma_{x,\infty} = \{0\}) \ge 1 - q_\infty(x),$$

where the function q_∞ is defined by (52). Note that if the transition operator T has an absolutely continuous invariant probability measure, then its density $p\colon X \to \mathbb{R}_+$ satisfies the functional equation

$$p(x) = p(U(x))\,|\det U'(x)|.$$

Hence, applying Lemma 5, it can be shown inductively that, for any fixed $x \in X$, the random sequence $\{p(U^k(x))\,\nu_{x,k}\}_{k\in\mathbb{Z}_+}$ is also a nonnegative martingale.

9. Nonextinction Probabilities

By studying the nonextinction probabilities of the Markov chain, measures of the probability of k-reachability, and hence of computational collapse in k iterations from x, are obtained. This section considers the rate of convergence of the probability of nonextinction, as $k \to \infty$, that is in long runs of computer iterations and shows it to be no faster than geometric.

Define a partial order $\prec$ on the class $\mathcal{Z}$ by $A \prec B$ if $A \subset B + z$ for some $z \in \mathbb{Z}^n$. This is weaker than the partial order induced by set-inclusion, since $A \subset B$ implies $A \prec B$. Clearly, $A \prec B$ and $B \prec A$ if and only if $A = B + z$ for some $z \in \mathbb{Z}^n$, for which write $A \cong B$.

Lemma 6. *For every $x \in X$, each of the functions $Q_k(x, \cdot)\colon\ \mathcal{Z} \to [0, 1]$ is*

(a) *translation invariant,*

$$A \cong B \ \Rightarrow\ Q_k(x, A) = Q_k(x, B); \tag{63}$$

(b) *subadditive,*

$$Q_k\left(x, A \bigcup B\right) \le Q_k(x, A) + Q_k(x, B); \tag{64}$$

(c) *monotonic with respect to $\prec$,*

$$A \prec B \ \Rightarrow\ Q_k(x, A) \le Q_k(x, B). \tag{65}$$

Proof. Property (a) is proved by induction on $k \in \mathbb{Z}_+$. For $k = 0$, (63) follows immediately from (47). Suppose that (63) holds for some $k \in Z_+$. From (21), it is straightforward to obtain that for any $A \in \mathcal{Z}$, $z \in \mathbb{Z}^n$, and $v \in V$,

$$\begin{aligned} F(x, A + z, v) &= \mathbb{Z}^n \bigcap (U'(x)(A + V) + v + U'(x)z) \\ &= R_1(U'(x)z + v) + \mathbb{Z}^n \bigcap (U'(x)(A + V) + E_1(U'(x)z + v)) \\ &= R_1(U'(x)z + v) + F(x, A, E_1(U'(x)z + v)). \end{aligned} \tag{66}$$

Note that if a random vector ω is distributed uniformly on the cube V, then so also is the random vector $\widetilde{\omega} = E_1(U'(x)z + \omega)$, since the uniform distribution on the n-dimensional torus is the Haar measure for the additive group of shifts on the torus. Combining this property with (66) and defining the random vector $\zeta = R_1(U'(x)z + \omega)$ with values in $\mathbb{Z}^n$, use (33) to rewrite the recurrence (46) to give

$$\begin{aligned} Q_{k+1}(x, A + z) &= \mathbf{E}\, Q_k(U(x), F(x, A + z, \omega)) \\ &= \mathbf{E}\, Q_k(U(x), F(x, A, \widetilde{\omega}) + \zeta) \\ &= \mathbf{E}\, Q_k(U(x), F(x, A, \widetilde{\omega})) = Q_{k+1}(x, A). \end{aligned}$$

Since $A \in \mathcal{Z}$ and $z \in \mathbb{Z}^n$ are arbitrary, this last equation completes the proof of (a).

Now, (b) and (c) are also proved by induction on $k \in \mathbb{Z}_+$. For $k = 0$, both relationships (64) and (65) follow again from (47). Suppose that they hold for some $k \in Z_+$. Note that the mapping (21) preserves set-theoretical operations over its second argument. In particular, $F\left(x, A \bigcup B, v\right) = F(x, A, v) \bigcup F(x, B, v)$ for all $x \in X$, $A, B \in \mathcal{Z}$, and $v \in V$. Consequently,

$$\begin{aligned} Q_{k+1}\left(x, A \bigcup B\right) &= \mathbf{E}\, Q_k(x, F(x, A, \omega) \bigcup F(x, B, \omega)) \\ &\leq \mathbf{E}\, Q_k(x, F(x, A, \omega)) + \mathbf{E}\, Q_k(x, F(x, B, \omega)) \\ &= Q_{k+1}(x, A) + Q_{k+1}(x, B), \end{aligned} \tag{67}$$

where ω is a random vector distributed uniformly on V. This last inequality completes the inductive step for the assertion (b). On the other hand, the leftmost equality in (67) implies that $Q_{k+1}(x, A \bigcup B) \geq \mathbf{E}\, Q_k(x, F(x, A, \omega)) = Q_{k+1}(x, A)$, and hence that

$$A \subset B \ \Rightarrow\ Q_{k+1}(x, A) \leq Q_{k+1}(x, B).$$

Combining this last implication with (a) gives $A \prec B \ \Rightarrow\ Q_{k+1}(x, A) \leq Q_{k+1}(x, B)$, completing the inductive step for (c), and the proof of the lemma is complete. ☐

As can be seen, the proof of (a) in Lemma 6 contains a stronger result, namely that the conditional joint distribution of $\nu_{x,1}, \ldots, \nu_{x,k}$, given $\sigma_{x,0} = A + z$, does not depend on $z \in \mathbb{Z}^n$.

From (45) and from Lemma 6, it follows that

$$q_k(x) = Q_k(x, \{0\}) \leq Q_k(x, A) \leq q_k(x)\#A, \quad \text{for } A \neq \emptyset. \tag{68}$$

In fact, a stronger property holds: For any $k \in \mathbb{Z}_+$, any $x \in X$, and any sets $A, B \in \mathcal{Z}$, $A \neq \emptyset$,

$$Q_k(x, B) \leq Q_k(x, A)\, d(A, B),$$

where

$$d(A, B) = \min\{\#C\colon\ C \in \mathcal{Z},\ B \subset A + C\}.$$

Clearly, $d(A, B) \leq 1$ if and only if $B \prec A$. Furthermore, if A is a singleton set, then $d(A, B) = \#B$ for all $B \in \mathcal{Z}$. However, in general, $\lceil \#B/\#A \rceil \leq d(A, B) \leq \#B$, where $\lceil \cdot \rceil$ is the ceiling function. Also note that $d(A, C) \leq d(A, B) d(B, C)$ for any sets $A, B, C \in \mathcal{Z}$, $A,\ B \neq \emptyset$r. That is, the logarithm of the function d, evaluated at nonempty finite subsets of $\mathbb{Z}^n$, satisfies the triangle inequality.

Lemma 7. *For any $j, k \in \mathbb{N}$ and any $x \in X$, the frequency function (45) satisfies the inequality*

$$q_{j+k}(x) \geq q_j(x)\, q_k(U^j(x)). \tag{69}$$

Proof. In a similar fashion to (50), $\mathbf{P}(\sigma_{x,j+k} \neq \emptyset \mid \sigma_{x,j} = A) = Q_k(U^j(x), A)$ for all $x \in X$, $j, k \in \mathbb{N}$, and $A \in \mathcal{Z}$. Hence, using the leftmost inequality in (68), obtain

$$\begin{aligned} q_{j+k}(x) &= \sum_{A \in \mathcal{Z}\colon\ A \neq \emptyset} \mathbf{P}(\sigma_{x,j+k} \neq \emptyset \mid \sigma_{x,j} = A)\, \mathbf{P}(\sigma_{x,j} = A \mid \sigma_{x,0} = \{0\}) \\ &= \sum_{A \in \mathcal{Z}\colon\ A \neq \emptyset} Q_k(U^j(x), A)\, \mathbf{P}(\sigma_{x,j} = A \mid \sigma_{x,0} = \{0\}) \end{aligned}$$

$$\geq\ q_k(U^j(x)) \sum_{A\in\mathcal{Z}:\ A\neq\emptyset} \mathbf{P}(\sigma_{x,j} = A \mid \sigma_{x,0} = \{0\})$$
$$=\ q_j(x) q_k(U^j(x)),$$

which was to be proved. $\square$

From (69) it follows that $q_k(x) \geq \prod_{j=0}^{k-1} q_1(U^j(x))$. Hence, by the strict positiveness of q_1, all the functions q_k are strictly positive. Therefore, if T has an invariant probability measure P, then

$$\liminf_{k\to+\infty} \left(k^{-1} \int_X \ln q_k(x) P(dx) \right) \geq \int_X \ln q_1(x)\ P(dx).$$

Intuitively, this last relationship shows that if $q_k(x) \to 0$ as $k \to +\infty$, then on average (in the sense of P-measure), the convergence cannot be faster than that of a geometric progression with parameter $\exp(\int_X \ln q_1(x)\ P(dx))$.

10. Illustrative Example: Möbius Transformations

The ideas and results of previous sections are demonstrated for the case of a rational transformation. Such transformations have been studied extensively because of their Julia sets and associated fractal properties. The simplest such mapping is considered here. If it appears a little artificial and perhaps too simplistic, its use is nonetheless justified for two reasons.

- The nonresonance property is easy to state, but difficult to check. It cannot really be checked on a computer, because in computer arithmetic all vectors are rationally dependent at some accuracy. So, although it would appear almost certain that many interesting mappings, associated with chaotic motions and computational collapse, satisfy the nonresonance condition, in most the calculations become extremely unwieldy. This is not the case here.
- In the example discussed below, the nonresonance condition has a very simple and intuitive expression in terms of the irrationality of the rotation number θ/π.

Let $M = (m_{jk})_{1\leq j,k\leq 2} \in \mathbb{C}^{2\times 2}$ be a J-unitary matrix. That is, $M^* J M = J = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, where M^* is the complex conjugate transpose of M. Each such M induces a Möbius transformation $T_M\colon\ X \to X$ of the open unit disc in the complex plane [2] $X = \{z \in \mathbb{C}\colon\ |z| < 1\}$ defined by

$$T_M(z) = \frac{m_{11}z + m_{12}}{m_{21}z + m_{22}}. \tag{70}$$

It is well known that T_M is conformal on X. Hence, by the natural bijection between $\mathbb{C}$ and $\mathbb{R}^2$, the mapping T_M can be identified with the two-dimensional diffeomorphism of the disc X,

$$T(x) = (\mathrm{Re}\, T_M(z),\ \mathrm{Im}\, T_M(z)), \quad x = (x_1, x_2), \quad z = x_1 + i x_2, \tag{71}$$

where $\mathrm{Re}(\cdot)$ and $\mathrm{Im}(\cdot)$ are the real and imaginary parts of a complex number. Note that $M \mapsto T_M$ is a homomorphism of the group of J-unitary matrices to the group of fractional linear conformal mappings (70), since $T_{M_1} \circ T_{M_2} = T_{M_1 M_2}$. When considered on the whole complex plane, the function T_M is meromorphic with a simple pole at $-1/h(M)$ where

$$h(M) = \frac{m_{21}}{m_{22}} \in X. \tag{72}$$

Since (70) is invariant under multiplication of M by nonzero complex numbers, the J-unitary matrix M can be parameterized by three numbers $\varphi, \alpha, \beta \in \mathbb{R}$ as

$$M = \begin{bmatrix} \cosh\varphi\, \exp(i\alpha) & \sinh\varphi\, \exp(-i\beta) \\ \sinh\varphi\, \exp(i\beta) & \cosh\varphi\, \exp(-i\alpha) \end{bmatrix}. \tag{73}$$

This representation of an arbitrary J-unitary matrix is defined up to multiples $\exp(i\gamma)$ on the right of (73). Because of the multiplicative invariance of T_M, without loss of generality, the $\exp(i\gamma)$ can be omitted. It is straightforward to show that if

$$|\cosh\varphi \cos\alpha| < 1, \tag{74}$$

the eigenvalues λ_1 and λ_2 of M lie on the unit circle ∂X and are given by

$$\lambda_{1,2} = \cosh\varphi \cos\alpha \ \pm i\sqrt{1 - (\cosh\varphi \cos\alpha)^2},$$

which implies that $\lambda_1/\lambda_2 = \exp(i\theta)$ with

$$\theta = -i \ln \frac{\lambda_1}{\lambda_2} = 2 \arccos(\cosh\varphi \cos\alpha). \tag{75}$$

Lemma 8. *Suppose that, in the representation (73), $\varphi \neq 0$, and that the inequality (74) holds, with θ/π irrational. Then the the diffeomorphism (71) is iteratively nonresonant.*

Proof. For a fixed but otherwise arbitrary $k \in \mathbb{N}$, consider the component $\mathcal{R}_k(T)$ (30) of the resonance set corresponding to the diffeomorphism T given by (71). It is straightforward to show that $\mathcal{R}_k(T)$ consists of those points z in the unit disc X for which the $k+1$ complex numbers $(T_M^j)'(z)$, $0 \le j \le k$, are linearly dependent over the field of complex rationals, where $T_M^0(z) \equiv z$. Hence, if $\mathrm{mes}_2\, \mathcal{R}_k(f) > 0$, then, by the uniqueness theorem for analytic functions, there exist complex rationals $c_0, \ldots, c_k$, not all zero, such that $\sum_{j=0}^k c_j (T_M^j)'(z) = 0$ everywhere in $\mathbb{C}$ except at the poles $-1/h(M^j)$ of the functions $T_M^j = T_{M^j}$, $1 \le j \le k$. Therefore, if the complex numbers $h(M^k) \in X$ are all pairwise distinct, then $\mathrm{mes}_2\, \mathcal{R}_k(T) = 0$ for any $k \in \mathbb{N}$ and, consequently, so also is the resonance set $\mathcal{R}(T)$ a null set. It only remains to prove that, under the assumptions of the lemma, the $h(M^k)$ are pairwise distinct for all $k \in \mathbb{N}$. Using the representation (73), rewrite the matrix M as

$$M = U \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} U^{-1}, \tag{76}$$

where

$$U = \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} = \begin{bmatrix} \sinh\varphi \exp(-i\beta) & i\delta \\ -i\delta & \sinh\varphi \exp(i\beta) \end{bmatrix} \tag{77}$$

is a matrix of eigenvectors of M, with

$$\delta = \cosh\varphi \sin\alpha - \sqrt{1-(\cosh\varphi\cos\alpha)^2}\,. \tag{78}$$

From (72), (75), and (76), it follows that for each $k \in \mathbb{N}$,

$$h(M^k) = h\left(U \begin{bmatrix} \exp(ik\theta) & 0 \\ 0 & 1 \end{bmatrix} U^{-1}\right) = T_L(\exp(ik\theta)), \tag{79}$$

where

$$L = \left(U^T\right)^{-1} \begin{bmatrix} u_{21} & 0 \\ 0 & u_{22} \end{bmatrix}.$$

From (78) it is easy to see that $\varphi \neq 0$ implies $\delta \neq 0$, and consequently, by (77), the matrix L is nonsingular. Therefore, the corresponding fractional linear mapping T_L is injective on $\mathbb{C}$ except for its pole $-1/h(L)$, where

$$h(L) = -\frac{u_{12}u_{21}}{u_{11}u_{22}} = -\left(\frac{\delta}{\sinh\varphi}\right)^2 = \frac{-\delta}{\cosh\varphi\sin\alpha + \sqrt{1-(\cosh\varphi\cos\alpha)^2}}\,.$$

Hence, $|h(L)| \neq 1$ and so the pole of T_L does not lie on the unit circle ∂X. Thus, the circle is mapped by T_L injectively into $\mathbb{C}$. Since θ/π is irrational, the numbers $\exp(ik\theta)$ are all pairwise distinct which, together with the injectivity of T_L on ∂X, implies the same property for the numbers $h(M^k)$. The proof is complete. □

Under the assumptions of Lemma 8, the diffeomorphism T in (71) is a nonlinear mapping of the disc X, similar to the aperiodic rotation by the angle (75) since $T_M(z) = T_U(\exp(i\theta)T_U^{-1}(z))$, where T_U is the Möbius transformation corresponding to the matrix (77). Hence, T has a unique fixed point $T_U(0)$, and that point is neutrally stable. The rest of the disc X is split into one-dimensional invariant manifolds $\{z \in X\colon |T_U^{-1}(z)| = r\}$, $r \in (0,1)$, each diffeomorphic to the unit circle.

As Lemma 8 shows, the nonresonance property of the Möbius transformation T_M reduces to the irrationality of θ/π. Note that $\pi^{-1}\arccos u$ is irrational for many u, for example, for any rational $u \in (-1,1)\backslash\{0, \pm 1/2\}$. Therefore, if the J-unitary matrix M has all rational entries with $|\operatorname{Re} m_{11}| \in (0,1)\backslash\{1/2\}$ and $m_{12} \neq 0$, it automatically satisfies the assumptions of the lemma. A simple such example is the matrix

$$M = \begin{bmatrix} 3/4 + i & 3/4 \\ 3/4 & 3/4 - i \end{bmatrix},$$

which is exactly representable in any practical binary arithmetic, with at least two digits after the binary point. For discretizations of the corresponding diffeomorphism (71), the frequency functions q_k of the sets of k-reachable points $X_{\varepsilon,k}$, $k = 1, 2$, computed by the recurrence (45)–(47), are graphed in Figure 1. In particular, for the fixed point of the

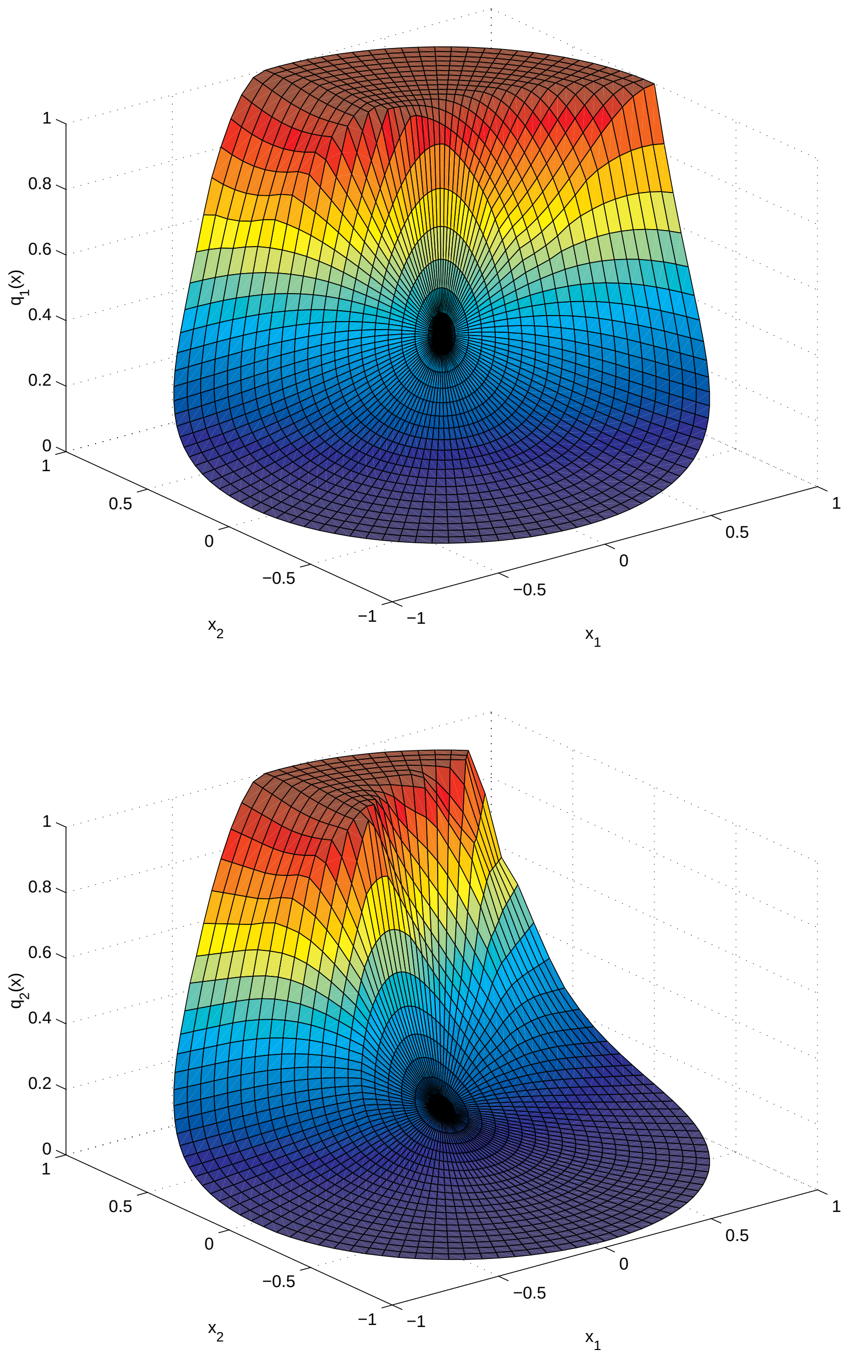

Fig. 1. The frequency function q_k of the sets of k-reachable points $X_{\varepsilon,k}$, $k = 1, 2$, for the discretized Möbius transformation.

diffeomorphism, $x_* = (0,\ (4 - \sqrt{7})/3)$, the values of the first four frequency functions are given by

k	1	2	3	4
$q_k(x_*)$	0.9863	0.9729	0.9599	0.9473

The experimental results are presented in Figure 2, where the sets $X_{\varepsilon,k}$, $k = 1, 2$, are shown for $\varepsilon = 0.01$. The table below contains the relative proportions of k-reachable points and their theoretically predicted limiting values (51) calculated by numerical integration of the frequency functions over the unit disc X:

k	$\#X_{0.01,k}/\#X_{0.01}$	$\int_X q_k(x)\,dx/\pi$	Rel. gap, %
1	0.4651	0.4635	0.3452
2	0.2717	0.2864	5.1327

This comparison provides compelling experimental evidence for the efficacy of Theorems 1 and 2 as models of digital simulation, especially if it is taken into account that ε is here not very small.

11. Algorithm for Computing Transition Probabilities

Numerical implementation of the recurrence (46) reduces principally to computation of the transition probabilities in (33). This can be carried out by an algorithm which, for any given set $A \in \mathcal{Z}$, gives the list of all sets $B \in \mathcal{Z}$ satisfying $P_1(B \mid A, x) > 0$, along with the appropriate transition probabilities. Such an algorithm was developed and programmed in Lahey-Fujitsu Fortran 95-5.5 for the two-dimensional case $n = 2$ and is outlined below. Although the principal ideas are also applicable to higher dimensions, the implementation is rather complicated for $n \geq 3$.

Given a point $x \in X$ and a fixed nonempty finite set $A \subset \mathbb{Z}^2$, for notational convenience simplify (21) and (33) by writing

$$P(B) = \mathrm{mes}_2\{v \in V\colon\ F(v) = B\}, \tag{80}$$

where

$$F(v) = \mathbb{Z}^2 \bigcap (M(A + V) + v). \tag{81}$$

Here, B is a finite subset of $\mathbb{Z}^2$, $M = U'(x)$ is a nonsingular (2×2)-matrix, and $V = [-1/2, 1/2)^2$. No special structure for M is assumed. So, the following applies not only to the case where T corresponds to a conformal mapping. That the square V is half-open is nonessential for what follows and will be ignored.

The sets $\{v \in V\colon\ F(v) = B\}$ in (80), taken over all $B \in \mathcal{Z}$, form a partition of V. They cannot be arcwise connected and their connected components are concave polygons in general. This complicates direct application of standard schemes, like triangulation, used to compute the area of elements of the partition.

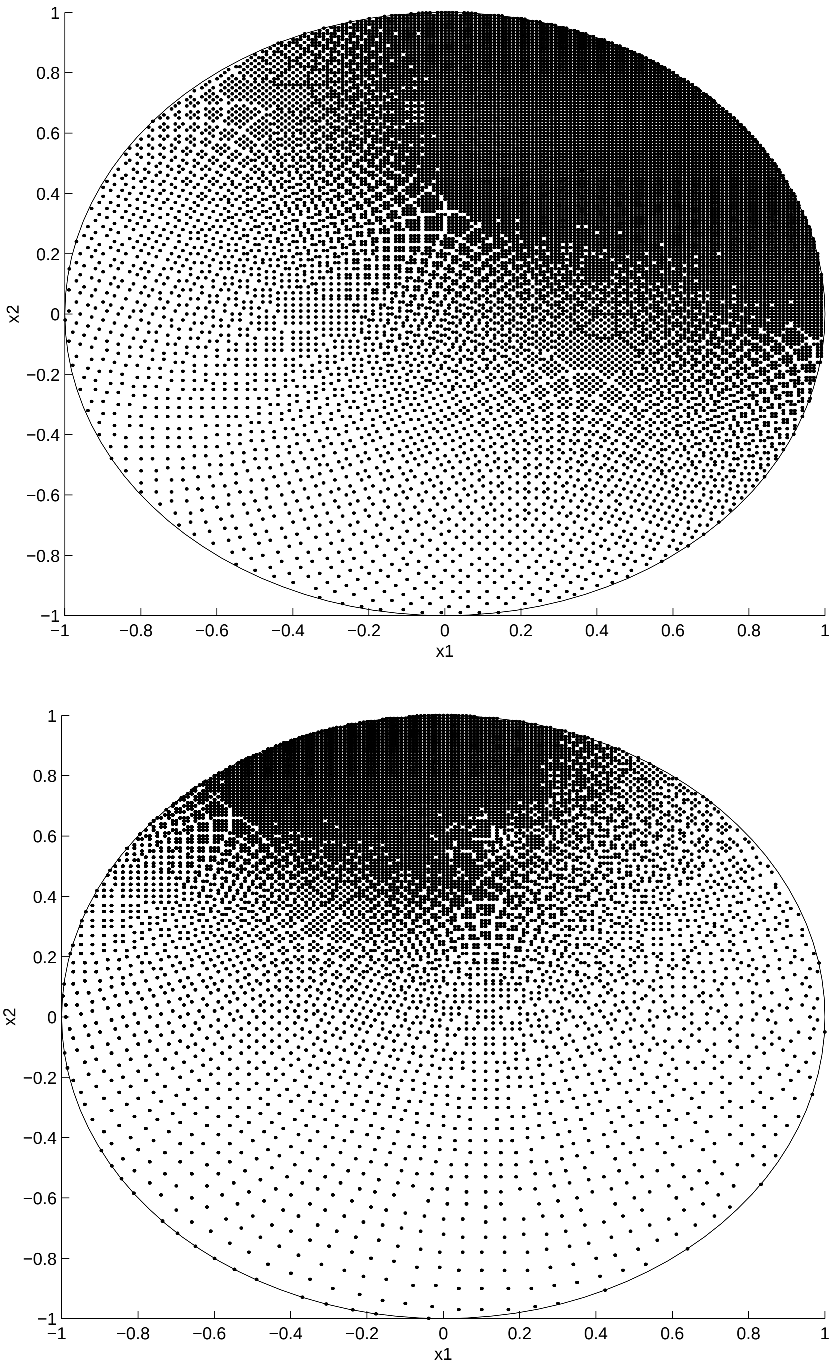

Fig. 2. The sets of k-reachable points $X_{0.01,k}$, $k = 1, 2$, for the discretized Möbius transformation.

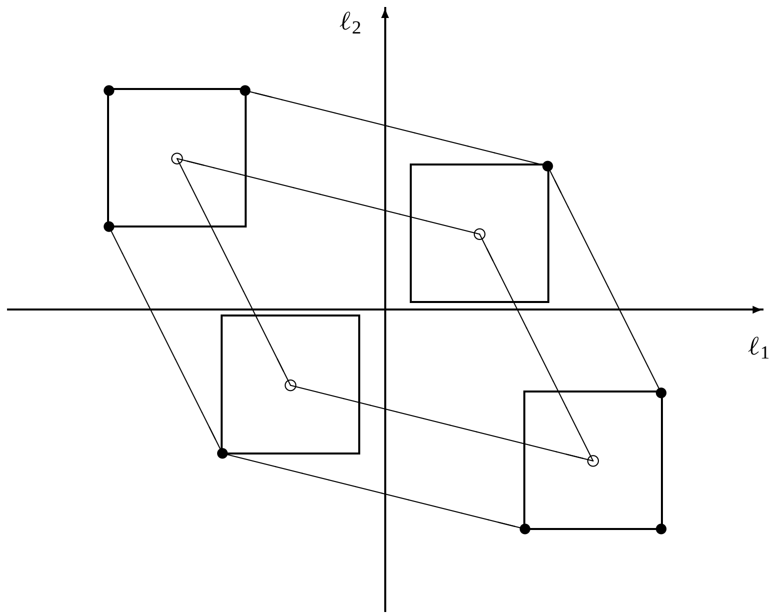

Fig. 3. The outermost contour is the boundary of the octagon L. The $\circ$'s and $\bullet$'s represent the vertices of the parallelogram MV and of the octagon, respectively.

However, there is a *generating* partition of V into finitely many convex polygons (intersections of the square with parallelograms whose sides are determined by column-vectors of the matrix M) so that each element of the original partition with nonzero area is the union of some elements of the generating partition.

The set-valued mapping (81) is piecewise constant on the square V, with values subsets of

$$F(V) = \bigcup_{v\in V} F(v) = \bigcup_{x\in A} Y(x),$$

where

$$Y(x) = \mathbb{Z}^2 \bigcap (Mx + L),$$

and

$$L = (MV) + V.$$

Note that L is an octagon, centrally symmetric about the origin (see Fig. 3). The discontinuity set of the mapping F satisfies the inclusion

$$\operatorname{discont} F \subset \bigcup_{x\in A} \bigcup_{y\in Y(x)} (y - Mx + \Lambda), \tag{82}$$

where

$$\Lambda = M\left((\{\pm 1/2\} \times \mathbb{R}) \bigcup (\mathbb{R} \times \{\pm 1/2\})\right)$$

is the union of two pairs of parallel lines which contains $M\partial V$. Therefore, the set on the right of (82) is the union of two collections of parallel lines which divide the square V into at most $(2\sum_{x\in A}\#Y(x)+1)^2$ sets. These last form a generating partition since each of these sets C satisfies the following properties: (i) the mapping F is constant on C, and (ii) the set is a convex polygon, since it is obtained from the intersection of V with a parallelogram.

The extreme points of each set C of the generating partition are found and their centre of mass $m(C)$ is calculated. The extreme points are then reordered counter clockwise about $m(C)$ and triangulation is applied to compute the area of C. The value $B = F(m(C))$ of F on the convex polygon C is computed and $\mathrm{mes}_2\, C$ contributes to the transition probability $P(B)$ in (80). Finite subsets $\mathbb{Z}^2$ are implemented as integer matrices with two rows and lexicographically ordered columns.

Acknowledgments

The work is supported by the Australian Research Council Grant A 4970 2246. One of the authors (I.V.) is grateful to Dr. Alex Klimenko for helpful discussions of the results of the paper. The authors thank an anonymous referee for recommending some additional references, as well as pointing out possible points of confusion and suggesting how these might be remedied.

References

[1] Aubin, J.-P., and Frankowska, H., *Set-Valued Analysis*, Birkhäuser, Boston, 1990.
[2] Beardon, A., *Iteration of Rational Functions*, Springer-Verlag, New York, 1991.
[3] Beck, C., Scaling behaviour of random maps, *Phys. Lett.*, **A136** (1989), 121–125.
[4] Beck, C., and Roepstorff, G., Effects of Phase Space Discretization on the Long-time Behaviour of Dynamical Systems, *Physica*, **D25** (1987), 173–180.
[5] Billingsley, P., *Convergence of Probability Measures*, John Wiley and Sons, New York, 1968.
[6] Binder, P. M., Machine iteration of a linear function: Local behaviour, *Comput. Math. Appl.* **21** (1991), 133–140.
[7] Binder, P. M., Limit cycles in a quadratic discrete iteration, *Physica*, **D57** (1992), 31–38.
[8] Blank, M., *Discreteness and Continuity in Problems of Chaotic Dynamics*, Translations of Mathematical Monographs, Vol. 161, American Mathematical Society, Providence, Rhode Island, 1997.
[9] Burtin, Y. D., On a simple formula for random mappings and its applications, *J. Appl. Prob.*, **17** (1980), 403–414.
[10] Diamond, P., Kloeden, P. E., Kozyakin, V. S., and Pokrovskii, A. V., A model for roundoff and collapse in computation of chaotic dynamical systems, *Mathematics and Computers in Simulation*, **44** (1997), 163–185.
[11] Diamond, P., Kloeden, P., Pokrovskii, A., and Vladimirov, A., Collapsing effects in numerical simulation of a class of chaotic dynamical systems and random mappings with a single attracting centre, *Physica D* **D86** (1995), 559–571.
[12] Diamond, P., Kloeden, P., Klemm, A., and Pokrovskii, A., Basin of attraction of measurable systems and their discretizations, *J. Statist. Phys.*, **84** (1996), 713–733.
[13] Diamond, P., Kloeden, P., Kozyakin, V., and Pokrovskii, A., Boundedness and dissipativity of truncated rotations on uniform planar lattices, *Math. Nachr.*, **171** (1995), 95–110.

[14] Diamond, P., Kloeden, P., Kozyakin, V., and Pokrovskii, A., Monotone dynamical systems under spatial discretization, *Proc. Am. Math. Soc.*, **126** (1998), 2169–2174.

[15] Diamond, P., Kloeden, P., and Pokrovskii, A., Interval stochastic matrices, a combinatorial lemma and the computation of invariant measures of dynamical systems, *J. Dyn. and Diff. Eq.*, **7** (1995), 341–364.

[16] Diamond, P., Kloeden, P., Pokrovskii, A., and Suzuki, M., Statistical properties of discretizations of a class of chaotic dynamical systems, *Comput. Math. Appl.*, **31** (1996), 83–95.

[17] Diamond, P., and Pokrovskii, A., Statistical laws for computational collapse of discretized chaotic mappings, *Int. J. Bifurcation Chaos*, **6** (1996), 2389–2399.

[18] Diamond, P., and Vladimirov, I., Asymptotic independence and uniform distribution of quantization errors for spatially discretized dynamical systems, *Int. J. Bifurcation Chaos*, **8** (1998), 1479–1490.

[19] Diamond, P., and Vladimirov, I., Loss of invertibility of smooth diffeomorphisms under spatial discretization, submitted for publication. See www.maths.uq.edu.au/∼pmd/reports.html

[20] Erber, T., and Gavelek, D., The iterative evolution of complex systems, *Physica* **A177** (1991), 394–400.

[21] Flajolet, P., and Odlyzko, A. M., Random mappings statistics, *Advances in Cryptology*, Springer Lecture Notes in Computer Sciences, **434**, 329–355, Springer-Verlag, Berlin, 1990.

[22] Grebogi, E., Ott, E., and Yorke, J. A., Roundoff–induced periodicity and the correlation dimension of chaotic attractors, *Phys. Rev.* **A34** (1988), 3688–3692.

[23] Håland, I. J., Uniform distribution of generalized polynomials, *J. Number Theory*, **45** (1993), 327–366.

[24] Håland, I. J., and Knuth, D. E., Polynomials involving the floor function, *Math. Scand.*, **76** (1995), 194–200.

[25] Hall, P., and Heyde, C. C., *Martingale Limit Theory and Its Application*, Academic Press, New York, 1980.

[26] Harris, B., Probability distributions related to random mapping, *Annals of Mathematical Statistics*, **31** (1960), 1045–1062.

[27] Hartfiel, D. J., *Markov Set-Chains*. Lecture Notes in Mathematics **1695**, Springer-Verlag, Heidelberg, 1998.

[28] Hunt, F. Y., Finite precision representation of the Conley decomposition, *J. Dyn. Diff. Eq.*, **13** (2001), 87–105.

[29] Izmailov, R., and Pokrovskii, A., Asymptotic analysis of aliasing structures, *J. Appl. Math. Stoch. Anal.*, **5** (1992), 193–204.

[30] Izmailov, R., Pokrovskii, A., and Vladimirov, A., Visualization of polynomials, *Comput. Graphics* **20** (1996), 95–105.

[31] Karlin, S., and McGregor, J., Iteration of analytic functions of several variables, *Problems in Analysis*, Gunning, R. C., Ed., Priceton University Press, Princeton, N.J., 1970, 81–92.

[32] Knuth, D. E., *The Art of Computer Programming. Vol. 2 Seminumerical Algorithms*, 2nd ed., Addison-Wesley, Reading, Mass., 1981.

[33] Kozyakin, V., Kuznetsov, N., Pokrovskii, A., and Vladimirov, I., Some problems in analysis of discretizations of continuous dynamical systems. Proceedings of the Second World Congress of Nonlinear Analysts, Part 2 (Athens, 1996), *Nonlin. Anal.* **30** (1997), 767–778.

[34] Kruskal, M. D., The expected number of components under a random mapping function, *Amer. Math. Monthly*, **61** (1954), 392–397.

[35] Kuznetsov, N., and Kloeden, P., The problem of information stability in computer studies of continuous systems, *Math. Comput. Simulation*, **43** (1997), 143–158.

[36] Lanford III, O. E., Informal remarks on the orbit structure of discrete approximations to chaotic maps, *Experiment. Math.*, **7** (1998), 317–324.

[37] Levy, Y. E., Some remarks about computer studies of dynamical systems, *Phys. Lett.*, **A88** (1982), 1–3.

[38] Nusse, H. E., and Yorke, J. A., Is every approximate trajectory of some process near an exact trajectory of a nearby process? *Commun. Math. Phys.*, **114** (1988), 363–379.

[39] Percival, I., and Vivaldi, F., Arithmetical properties of strongly chaotic motions, *Physica* **D25** (1987), 105–130.

[40] Pittel, B., On distributions related to transitive closures of random finite mappings, *Annals of Probability*, **11** (1983), 428–441.
[41] Shiryaev, A., *Probability*, 2nd edition, Springer, New York, 1996.
[42] Siegel, C. L., and Moser, J. K., *Lectures on Celestial Mechanics*, Springer-Verlag, Berlin, 1971.
[43] Stepanov, V. E., Random mappings with a single attracting centre, *Theory of Probability and Its Applications*, **16** (1971), 155–161.
[44] Sternberg, S., Local contractions and a theorem of Poincaré, *Am. J. Math.*, **79** (1957), 809–823.
[45] Topsøe, F., Preservation of weak convergence under mappings, *Ann. Math. Statist.*, **38** (1967), 1661–1665.
[46] Vladimirov, I., Quantized linear systems on integer lattices: Frequency-based approach, Parts I, II, *Center for Applied Dynamical Systems and Environmental Modeling, CADSEM Reports*, **96-032**, **96-033**, Deakin University, Geelong, Australia, 1996.
[47] Vladimirov, I., Kuznetsov, N., and Diamond, P., Frequency measurability, algebras of quasiperiodic sets and spatial discretizations of smooth dynamical systems, *Math. Comput. Simulation*, **52** (2000), 251–272.
[48] Yuan, G., and Yorke, J. A., Collapsing of chaos in one-dimensional maps, *Physica*, **D136** (2000), 18–30.