STABILITY AND SHADOWING IN CONVEX DISCRETE-TIME SYSTEMS

PHIL DIAMOND AND ALEXEI POKROVSKII

ABSTRACT. In the realization of a dynamical system on a computer, all computational processes are of a discretization, where continuum state space is replaced by the finite set of machine arithmetic. When chaos is present, the discretized system often manifests collapsing effects to a fixed point or to short cycles. These phenomena exhibit a statistical structure which can be modelled by random mappings with an absorbing centre. This model gives results which are very much in line with computational experiments and there appears to be a type of universality summarized by an Arcsine Law. The effects are discussed with special reference to the family of mappings $f_\ell(x) = 1 - |1 - 2x|^\ell$, $x \in [0,1]$, $1 < \ell \leq 2$. Computer experiments display close agreement with the predictions of the model.

1. INTRODUCTION

Chaotic systems challenge scientific computation. Consider the simple scenario of a chaotic dynamical system generated by a mapping

$$f : X \to X, \quad X \subset \Re^n,$$

with trajectories $\{x_0, x_1, x_2, \ldots\}$, $x_{k+1} = f(x_k)$. Because of sensitivity to initial conditions, trajectories from nearby initial values diverge exponentially. So, whether the computed system is treated as a mapping on a finite set, or as a rounded off computation, a computational trajectory rapidly diverges from the theoretical orbit starting from the same initial value. The well known shadowing lemma [1, 11] is often interpreted as stating that every computed orbit is approximated by a true orbit, albeit from a different initial value, for some length of time. Too, unshadowable orbits exist. For these reasons, individual numerical trajectories cannot be treated as accurate system properties and in a simulation their *statistical* or ergodic properties are usually studied instead. For example, the invariant measure μ is often calculated from computed orbits by histogram methods.

However, collapsing effects exist which can distort the histogram as an estimator of μ [5, 6, 7, 8, 9, 4]. In computation, there is a tendency for computed trajectories to collapse onto fixed points of the realized f (like zero) or low order periodic numerical trajectories. In the first, the computed measure has an atom at the origin, while in the second it is a measure concentrated on the short cycles.

These collapsing effects are not readily described. They seem to have a random character, apparently generated from the structure of finite computer arithmetic rather than operations of roundoff and truncation. For this reason, we treat the

computational dynamical process as a discrete function on a finite set. This approach offers somewhat different insights from a more conventional description in terms of roundoff errors.

2. MAPPINGS ON LATTICES $\mathbf{L}_\nu$

As a prime example of chaotic behaviour, consider the dynamical system generated by the mapping

$$f_\ell(x) = 1 - |1 - 2x|^\ell, \qquad x \in [0, 1]$$

where $\ell \in (1, \infty)$ is a parameter. The logistic mapping $f_2(x) = 4x(1-x)$ corresponds to $\ell = 2$. Note that each mapping f_ℓ has fixed point zero and that the modal point $x = 1/2$ is the second preimage of this fixed point. These two properties play a major role in establishing the existence of an ergodic absolutely continuous invariant measure for f_ℓ. Let $\varphi_\nu^{(\ell)}$ be the discretization of the mapping f_ℓ on the lattice $\mathbf{L}_\nu = \{0, 1/\nu, \ldots, (\nu-1)/\nu, 1\}$. The relationship between the original system f_ℓ and discretizations φ_ν^ℓ has some unexpected features as $\nu \to \infty$.

Note one such key property for what follows. For each mapping φ of a finite set $\mathbf{L}$ into itself, let $\mathcal{A}(\varphi)$ be the totality of points of $\mathbf{L}$ which are eventually absorbed by fixed points of the mapping φ. Denote by $\mathcal{C}(\varphi)$ the proportion of absorbed points $\#(\mathcal{A}(\varphi))/\#(\mathbf{L})$. Here, $\#(\cdot)$ denotes the cardinality of a finite set. The sequence $\{\mathcal{C}(\varphi_\nu^{(\ell)})\}$ varies very irregularly with ν and has virtually zero autocorrelation. Nonetheless, for a long series of lattices $\mathbf{L}_N, \ldots, \mathbf{L}_{N+n}$, with N, n large, the average value

$$E_\ell(N, n) = \frac{1}{n} \sum_{\nu=N}^{N+n} \mathcal{C}(\varphi_\nu^{(\ell)})$$

approaches a nonzero limit as $n \to \infty$ for each $\ell \in (2, \infty)$. For $1 < \ell \leq 2$ this average decreases to zero as $n \to \infty$, but rather slowly. Figure 1 shows the proportion of points absorbed by zero for the logistic on a sequence of lattices.

This paper highlights some peculiarities of scaling, with respect to ν, in the limiting behaviour of various statistics associated with collapsing effects when $1 < \ell \leq 2$. In particular, the mean $E_\ell(N, n)$ decreases to zero as $N^{1/2-1/\ell}$ for $N \gg n \gg 1$, whereas the corresponding median scales as $N^{1-2/\ell}$.

These scaling phenomena seem to be of a universal nature. In particular, our experiments indicate that either the Arcsine Law or straightforward modifications are valid for both one-dimensional mappings, such as the β-mapping, and two-dimensional maps like the Henon mapping, Lozi mapping, Belykh mapping and so on.

The paper is organized as follows. In the next section random mappings with an absorbing centre are introduced. This concept is the main technical tool that is used. Asymptotic results for these mappings are formulated in this section. Section 4 is devoted to the correspondence between properties of discretizations of chaotic dynamical systems and analogous properties of random mappings with a single absorbing centre. Results of numerical experiments with mappings f_ℓ are discussed. A slightly lengthy proof is relegated to the final section.

3. Random mappings with a single absorbing centre

Let Δ, $K > 0$ be positive integers and let

$$X(\Delta, K) = \{-\Delta + 1, \ldots, -1, 0, 1, \ldots, K\}.$$

Define the set $\Psi(\Delta, K)$ of all mappings $\psi : X(\Delta, K) \to X(\Delta, K)$ satisfying $\psi(i) = 0$ for $i \leq 0$. This collection is called a *random mapping, with an absorbing centre*. The set $\{\xi \in X(\Delta, K) : \xi \leq 0\}$ is the absorbing centre: once a trajectory of φ enters this set it cannot leave. If S is a subset of $\Psi(\Delta, K)$ associated with some given property A, then the proportion of elements of Ψ which belong to S will be called the *probability of the event A* and is denoted by $P(A)$.

These mappings are similar to, though differ from mappings with a single attracting centre [2, 3]. It is convenient to define $\Psi(0, K)$ as a *completely random mapping* on the set

$$X(0, K) = \{1, \ldots, K\},$$

that is, as the totality of all possible mappings $X(0, K) \to X(0, K)$. Properties of basins of attraction of the short cycles of such random mappings will be investigated in detail. Let $P(s; \Delta, K)$, $\Delta \leq s \leq \Delta + K$ denote the probability of the event that exactly s elements from the set $X(\Delta, K)$ are absorbed by fixed points of a mapping $\psi \in \Psi(\Delta, K)$. Most importantly, this includes the zero fixed point.

Write

$$p(s, d, k) = \frac{d}{s} \binom{k}{s - d} \left(\frac{s}{d + k - 1} \right)^{s - d} \left(\frac{k + d - s - 1}{d + k - 1} \right)^{k + d - s}$$

where d, k, s are non-negative integers with $d \leq s \leq d + k$.

Theorem 1.

$$P(s; \Delta, K)) = \sum_{m=0}^{s - \Delta} \alpha(m) p(s, \Delta + m, K - m),$$

where

$$\alpha(m) = \binom{K}{m} \left(\frac{\Delta + K - 1}{\Delta + K} \right)^{K - m}.$$

This theorem is not dissimilar to some well known formulas, in particular that of Burtin [3], p. 407 (3), but does not seem to have appeared in the literature.

The asymptotics which are used below follow in the usual way from Theorem 1 and Stirling's formula. Let S be a finite set of non-negative real numbers and define the *empirical distribution function of the set S*, $\mathcal{D}(\cdot\,; S) : [0, \infty) \to [0, 1]$, by

$$\mathcal{D}(x; S) = \frac{\#(\{s \in S : s \leq x\})}{\#(S)}, \qquad 0 \leq x < \infty.$$

The statistical distributions to be studied are

$$D(x; \Delta, K) = \mathcal{D}(x; \{\mathcal{C}(\psi) : \psi \in \Psi(\Delta, K)\}).$$

Observe, that by the definition $D(x; \Delta, K) \equiv 1$ for $x \geq 1$. Statistics of special interest are the mathematical expectation

$$E(\Delta, K) = \#(\Psi(\Delta, K))^{-1} \sum_{\psi \in \Psi(\Delta, K)} \mathcal{C}(\psi) = \int_0^1 (1 - D(x; \Delta, K))\, dx$$

65

and the median $\mathcal{M}(\Delta, K)$, defined for a given Δ, K by a number $a \geq 0$ such that
$$\#(\{\psi \in \Psi(\Delta, K) : \mathcal{C}(\psi) \leq a\}) = \#(\{\psi \in \Psi(\Delta, K) : \mathcal{C}(\psi) \geq a\}).$$
Denote the complementary error function by
$$\text{erfc}\,(t) = \frac{2}{\sqrt{\pi}} \int_t^\infty e^{-s^2}\,ds\,.$$

Corollary 1.

(a) : *For each $\varepsilon > 0$ there exists a (small) positive $\delta(\varepsilon)$ such that the inequality*

(1) $$\Delta^2 < \delta(\varepsilon)K$$

implies that
$$\left| \frac{\sqrt{K}}{\Delta} E(\Delta, K) - \sqrt{\frac{\pi}{2}} \right| < \varepsilon.$$

(b) : *For each $\varepsilon > 0$ there exists (small) $\delta(\varepsilon) > 0$ such that the relations*

(2) $$\delta(\varepsilon)^{-1} < \Delta^2 < \delta(\varepsilon)K,$$

imply that
$$\left| D\left(x\frac{\Delta^2}{K}; \Delta, K \right) - \text{erfc}\left(\frac{1}{\sqrt{2x}} \right) \right| \leq \varepsilon, \quad x > 0.$$

Observe that different scalings for the mean value and the median arise here. If Q is the median of the complementary error function, namely the unique positive number satisfying
$$\int_0^Q e^{-t^2}\,dt = \frac{\sqrt{\pi}}{4}.$$

Corollary 2. *For each $\varepsilon > 0$ there exists a positive $\delta(\varepsilon)$ such that the relations* (2) *imply that*
$$\left| \frac{K}{\Delta^2} \mathcal{M}(\Delta, K) - \frac{1}{2Q^2} \right| < \varepsilon.$$

That mean values and medians admit different scalings reflects the fact that a very small proportion of fixed points of all the mappings in the family $\Psi(\Delta, K)$ absorb a quite disproportionate share of all those points absorbed. A rigorous formulation of this observation is as follows. Write
$$\mathcal{C}_u(\psi) = \begin{cases} \mathcal{C}(\psi), & \text{if } \mathcal{C}(\psi) \leq u, \\ 0, & \text{otherwise.} \end{cases} \qquad 0 \leq u \leq 1,$$
and consider the function
$$H(u; \Delta, K) = \frac{\sum_{\psi \in \Psi(\Delta, K)} \mathcal{C}_u(\psi)}{\sum_{\psi \in \Psi(\Delta, K)} \mathcal{C}(\psi)}, \qquad 0 \leq u \leq 1,$$
that is,
$$H(u; \Delta, K) = E(\Delta, K)^{-1} \left(uD(u; \Delta, K) - \int_0^u D(s; \Delta, K)\,ds \right), \quad 0 \leq u \leq 1.$$

Corollary 3. (Arcsine Law)(Arcsine Law) *For each $\varepsilon > 0$ there exist a positive $\delta(\varepsilon)$ such that the inequality* (1) *implies that*

$$\left| H(u; \Delta, K) - \frac{2}{\pi} \arcsin(\sqrt{u}) \right| < \varepsilon.$$

Briefly consider the rate of convergence in each of these Corollaries. In Corollary 1(a) and in Corollary 3 the value $\delta(\varepsilon)$ can be chosen to be of the same order as ε. Corollary 1(b) is more interesting. As stated, for each $1 < x_0 \leq \infty$ there exists $\delta(\varepsilon, x_0)$ such that the inequalities $\delta(\varepsilon, x_0)^{-1} < \Delta^2 < \delta(\varepsilon, x_0)K$ imply

$$\left| D\left(x\frac{\Delta^2}{K}; \Delta, K \right) - \mathrm{erfc}\left(\frac{1}{\sqrt{2x}} \right) \right| \leq \varepsilon, \quad 0 \leq x \leq x_0.$$

Clearly,

$$D\left(x\frac{\Delta^2}{K}; \Delta, K \right) = 1, \qquad x \geq \frac{K}{\Delta^2}.$$

Hence, for x_0 of the order K/Δ^2, the discrepancy

$$\left| D\left(x\frac{\Delta^2}{K}; \Delta, K \right) - \mathrm{erfc}\left(\frac{1}{\sqrt{2x}} \right) \right|$$

is no less than $\mathrm{erf}\left(\Delta/\sqrt{K} \right) \sim \Delta/\sqrt{K}$. That is, $\delta(\varepsilon, x_0)$ can be no better than of $O(\varepsilon^2)$.

4. INTERPRETATION

Now consider the family of mappings f_ℓ, specifically for $1 < \ell \leq 2$. There is compelling experimental computational evidence to suggest that properties of the flow of discretizations $\varphi_\nu^{(\ell)}$, $\nu = 1, 2, \ldots$ *are statistically similar to the flow of mappings* ψ_ν *independently sampled from* $\Psi(\Delta_\ell(\nu), K_\ell(\nu))$ for appropriate $K_\ell(\nu), \Delta_\ell(\nu)$. This is strongly supported by heuristic and physical reasoning which also justifies appropriate choice of the parameters $\Delta_\ell(\nu)$ and $K_\ell(\nu)$, which are not *ad hoc* parameters merely fixed to give a best fit.

The statement is rather informal and it will be made more precise below by specifying the nature of the statistical relationship between the discretizations $\varphi_\nu^{(\ell)}$ and the random mappings $\Psi(\Delta_\ell(\nu), K_\ell(\nu))$; then providing arguments for the choice of $K_\ell(\nu)$, $\Delta_\ell(\nu)$; and finally providing experimental evidence.

First, we will very briefly sketch the reasons behind the choice of Δ_ℓ, K_ℓ and formulate a Hypothesis linking the behaviour of discretizations to that of the set of random graphs $\Psi(\Delta_\ell(\nu), K_\ell(\nu))$. The mapping f_ℓ has an absolutely continuous invariant measure μ_ℓ and so a cell $[(i-1)/\nu, i/\nu]$ of the lattice $\mathbf{L}_\nu$ may be given a natural weight $w_i = \mu_\ell([(i-1)/\nu, i/\nu])$. Similar reasoning to that of [10] suggests that $K(\nu)$ should increase at the same rate as $Y(\nu) = \left(\sum w_i^2 \right)^{-1}$. This gives $K(\nu) = b\nu$ when $1 < \ell < 2$ and $K(\nu) = b\nu/\ln \nu$ when $\ell = 2$, for a constant $b = b(\ell)$.

Turning to $\Delta(\nu)$, it is unnatural and, more importantly, inadequate to choose as a model the *completely random mapping* $\mathcal{G}(K_\ell(\nu))$, that is, the totality of mappings defined on the set $\{1, 2, \ldots, K_\ell(\nu)\}$ and endowed with the uniform measure. Indeed, one point of the lattice $\mathbf{L}_\nu$, namely the modal point $c = [1/2]\nu$, is quite untypical. The value $\varphi_\nu^{(\ell)}(c)$ has as preimage under $\varphi_\nu^{(\ell)}$ a number of order $O(\nu^{1-1/\ell})$ of the other points of the lattice. This is because of the flatness of f_ℓ near c so that the

images of nearby points are identified with 1 in $\mathbf{L}_\nu$. Besides, it is the second preimage of the zero fixed point of the mapping. Consequently, in these circumstances, it is more accurate to prescribe to one point of the lattice $\mathbf{L}_\nu$ a weight w of the order $\Delta_\ell(\nu) = O(\nu^{1-1/\ell})$ relative to other weights w_i and to regard this more heavily weighted point as absorbing. A natural way of realizing these rather informal considerations in a technical framework leads to mappings with an absorbing centre as defined in the previous section. Then, for $1 < \ell < 2$, $K_\ell(\nu) \sim b\nu$, $\Delta_\ell(\nu) \sim a\nu^{1-1/\ell}$, while for $\ell = 2$, $K_2(\nu) \sim b\nu/\ln\nu$, $\Delta_2(\nu) \sim a\sqrt{\nu}/\ln\nu$. See [5], 1995], pp.562-564, for a more detailed account.

For a positive integer N and $\gamma > 0$, denote by $\Phi_\gamma^{(\ell)}(N)$ the set of discretized mappings

$$\{\varphi_\nu^{(\ell)} : N \leq \nu \leq (1+\gamma)N\}.$$

Compare the finite set of discretizations $\Phi_\gamma^{(\ell)}(N)$ with independently sampled elements

(3) $$\psi_\nu \in \Psi(a\nu^{1/\ell}, b\nu^{2/\ell}), \qquad N \leq \nu \leq (1+\gamma)N,$$

for increasing N, where $\gamma > 0$ is small. Then, because γ is small, ν and $\nu^{1-1/\ell}$ do not vary significantly in $[N, (1+\gamma)N]$ and so the distributions $D(x; a\nu^{1-1/\ell}, b\nu)$, $N \leq \nu \leq (1+\gamma)N$, are all close to $D(x; aN^{1-1/\ell}, bN)$. On the other hand, the length of the interval $[N, (1+\gamma)N]$ increases sufficiently fast with N so that from standard estimates it can be deduced that, with probability 1, the distribution from sampling as in (3) will be close to the distribution $D(x; aN^{1-1/\ell}, bN)$ *for all sufficiently* large N. Consequently, the distributions

$$D_{\gamma,N}^{(\ell)}(x) = \mathcal{D}\left(x; \left\{\nu^{2/\ell-1}\mathcal{C}(\varphi_\nu^\ell) : N \leq \nu \leq (1+\gamma)N\right\}\right), \quad 1 < \ell \leq 2,$$

associated with the flow of discretizations should be qualitatively and quantitatively similar to the function

(4) $$D(xN^{2/\ell-1}; aN^{1-1/\ell}, bN).$$

On the other hand, for sufficiently large N, all functions (4) satisfy the relation

$$D(xN^{2/\ell-1}; aN^{1-1/\ell}, bN) \approx \mathrm{erfc}\left(\frac{a}{\sqrt{2bx}}\right)$$

by Corollary 1(b). That is, for all for all $\gamma > 0$ and for all sufficiently large N,

(5) $$D_{\gamma,N}^{(\ell)}(x) \approx \mathrm{erfc}\left(\frac{a}{\sqrt{b}}\frac{1}{\sqrt{2x}}\right).$$

Similar reasoning using Corollary 1(a) leads to the conclusion that, for the means

$$E_{\gamma,N}^{(\ell)} = \frac{1}{\gamma N} \sum_{\nu=N}^{(1+\gamma)N} \nu^{1/\ell-1/2}\mathcal{C}(\varphi_\nu^{(\ell)})$$

the asymptotic formula is

$$E_{\gamma,N}^{(\ell)} \approx \frac{a}{\sqrt{b}}\sqrt{\frac{\pi}{2}},$$

for $\gamma > 0$ and for all sufficiently large N. This last relation and (5) can be combined as

$$D_{\gamma,N}^{(\ell)}(x) \approx \mathrm{erfc}\left(\frac{c}{\sqrt{2x}}\right) \qquad \text{where} \qquad c = \sqrt{\frac{2}{\pi}}\, E_{\gamma,N}^{(\ell)}.$$

Arguing in the same way from the Arcsine Law,

$$H_{\gamma,N}^{(\ell)}(u) \approx \frac{2}{\pi} \arcsin(\sqrt{u})$$

where

$$H_{\gamma,N}^{(\ell)}(u) = \frac{\sum_{\nu=N}^{(1+\gamma)N} C_u(\varphi_\nu^{(\ell)})}{\sum_{\nu=N}^{(1+\gamma)N} C(\varphi_\nu^{(\ell)})}, \qquad 0 \le u \le 1.$$

These arguments motivate the following:

Hypothesis 1. *Suppose that $1 < \ell < 2$. For each $\gamma > 0$ and for sufficiently large N, the following asymptotic relations hold.*

> **(a) :** $\quad D_{\gamma,N}^{(\ell)}(x) \approx \operatorname{erfc}\left(\frac{c}{\sqrt{2x}}\right), \qquad$ where $\qquad c = \sqrt{\frac{2}{\pi}} \, E_{\gamma,N}^{(\ell)}.$
>
> **(b) :** $\qquad H_{\gamma,N}^{(\ell)}(u) \approx \frac{2}{\pi} \arcsin(\sqrt{u}).$

This Hypothesis can be tested by straightforward, if lengthy computational experiments. In these, we took $\ell = 3/2$, $N = 10^6$, $\gamma = 0.01$. The quantity $E_{\gamma,N}^{(\ell)} \approx 5.46$ computationally, giving an estimate $c \approx 0.42$. The function $D_{\gamma,N}^{(3/2)}(x)$ was constructed and was compared with $\operatorname{erfc}\left(\frac{c}{\sqrt{2x}}\right)$ for $0 \le x \le 14$. Results are graphed in Figure 2. Agreement is excellent: the two curves are virtually coincident. Figure 3 graphs the function $H_{\gamma,N}^{(3/2)}(u)$ against $\frac{2}{\pi} \arcsin(\sqrt{u})$ and agreement is again excellent, although not as spectularly so as in Figure 2.

A similar statement can be drafted for $\ell = 2$, the logistic mapping. However, since $\Delta^2/K \sim \ln(\nu)^{-1}$ in this case, to obtain good approximations without taking a very large number of lattices to average over, a higher order asymptotic term is added to the asymptotics derived from Theorem 1. Without going into cumbersome details, we simply state the more accurate formulas and test by experiment.

Define the functions

$$g_N(x;c) = \operatorname{erfc}\left(c\sqrt{\frac{1 - \frac{x}{\ln(N)}}{2x}}\right), \qquad 0 \le x \le \ln(N),$$

and

$$h_N(u,c) = \int_0^u \frac{1}{\sqrt{s(1-s)}} e^{\frac{c^2}{\ln(N)}(1-s)/2s} \, ds, \qquad 0 \le u \le 1,$$

where c is a parameter. For each γ, N denote by $\widehat{c} = c(\gamma, N)$ the solution in c of the equation

$$\frac{1}{\ln(N)} \int_0^{\ln(N)} (1 - g_N(x,c)) \, dx = E_{\gamma,N}^{(2)}.$$

Hypothesis 2. *For each $\gamma > 0$ and for sufficiently large N the following asymptotic relations hold.*

> **(a) :** $\quad D_{\gamma,N}^{(2)}(x) \approx g_N(x;\widehat{c}), \qquad 0 \le x \le \ln(N).$
>
> **(b) :** $\quad H_{\gamma,N}^{(2)}(u) \approx h_N(u;\widehat{c})/h_N(1;\widehat{c}), \qquad 1 \le u \le 1.$

Note that as $N \to \infty$, $g_N(x;c)$ tends to $\operatorname{erfc}\left(\frac{c}{\sqrt{2x}}\right)$ and $h_N(u;c)/h_N(1;c)$ tends to $\frac{2}{\pi} \arcsin(u)$. Therefore, for very large N, the second hypothesis captures the Arcsine

Law. The role of the higher order asymptotics is to give accuracy for middle values of N in the slow growth of $\ln N$. In particular, for $N = 10^7, \gamma = 0.001$ as before, the parameter $\widehat{c} = c(0.001, 10^7) \approx 0.9$. Figure 4 graphs $H^{(2)}_{0.001, 10^7}(u)$ against $h_{10^7}(u; \widehat{c})/h_N(1; \widehat{c})$. The agreement between experiment and theory is very good indeed, as is also that for $D^{(}2)_{0.001, 10^7}(x)$ (see [9] for further details).

5. Proof of Theorem 1

Let S be a subset of $\{1, \dots, K\}$. Denote by $\Psi(\Delta, K, S)$ the subset of $\Psi(\Delta, K)$ consisting of mappings ψ, each of whose fixed point set coincides with S. Let $G_0(\Delta, K)$ be the set of all mappings on $X(\Delta, K)$ with no fixed points at all. For each $\psi \in \Psi(\Delta, K, S)$ denote by $\mathcal{G}_0(\psi)$ the set of all mappings $g \in G_0(\Delta, K)$ which coincide with ψ on $\{1, \dots, K\} \setminus S$

Let $\widehat{g}^{-1}(S)$ be the transitive closure of the set S under g^{-1}, that is,

$$\widehat{g}^{-1}(S) = \bigcup_{i=0}^{K} g^{-i}(S).$$

For each $\psi \in \Psi(\Delta, K, S)$, and $g \in \mathcal{G}_0(\psi)$, the transitive closure $\widehat{g}^{-1}(S)$ coincides with the set $\mathcal{A}(\psi)$ of all those points eventually absorbed by fixed points. Observe that each set $\mathcal{G}_0(\psi)$ contains the same number of elements, namely $(\Delta + K - 1)^{\Delta + \#(S)}$. Therefore, writing $P(A \,|\, B)$ as the probability of A conditioned on the event B,

$$P\left(\mathcal{C}(\psi) = s\right) \,|\, \psi \in \Psi(\Delta, K, S)) = P\left(\#(\widehat{g}^{-1}(X_- \bigcup S)) = s \,|\, g \in G_0(\Delta, K)\right).$$

For $\psi \in \Psi(\Delta, K)$ let $\mathrm{fp}(\psi)$ be the number of strictly positive fixed points of ψ. From the last equality it follows that

$$P(s; \Delta, K)) = \sum_{m=0}^{K} \alpha_m p_m(s),$$

where

$$\alpha_m = P(\mathrm{fp}(\psi) = m) = \frac{\#(\{\psi \in \Psi(\Delta, K) : \mathrm{fp}(\psi) = m\})}{\Psi(\Delta, K)}$$

and

$$p_m(s) = P\left(\#(\widehat{g}^{-1}(\{-\Delta, \dots, -1, 0, 1, \dots, m\})) = s \,\big|\, g \in \mathcal{G}_0(\Delta, K)\right).$$

Clearly,

$$\alpha_m = \binom{K}{m} \left(\frac{\Delta + K - 1}{\Delta + K}\right)^{K-m}.$$

That is, α_m is the same as $\alpha(m)$ in Theorem 1.

It remains to establish that

$$(6) \qquad p_m(s) = p(s, \Delta + m, K - m) = \binom{K - m}{s - \Delta - m} q_m(s),$$

where

$$(7) \qquad q_m(s) = \frac{\Delta + m}{s} \left(\frac{s}{\Delta + K - 1}\right)^{s - \Delta - m} \left(\frac{\Delta + K - s - 1}{\Delta + K - 1}\right)^{\Delta + K - s}.$$

Let $A_m = \{-\Delta - 1, \ldots, m\}$ and let B be some subset from $X(\Delta, K) \backslash A_m$ consisting from $s - \Delta - m$ elements. Following the proof of Theorem 38, [2], p.372, write

$$(8) \qquad P(\widehat{g}^{-1}(A_m) = A_m \bigcup B) = P_1 P_2 P_3,$$

where P_1 is the probability of the event $g(B) \subset A_m \bigcup B$, P_2 is the probability of the event $g(X \backslash (A \bigcup B)) \subset X \backslash (A \bigcup B)$, and P_2 is the proportion of elements from $\mathcal{G}(\Delta, K)$ such that for every $b \in B$ there exists an oriented path beginning at b and ending in A. Then

$$P_1 = \left(\frac{s-1}{\Delta + K - 1} \right)^{s - \Delta - m},$$

$$P_2 = \left(\frac{K + \Delta - s - 1}{\Delta + K - 1} \right)^{K + \Delta - s},$$

$$P_3 = \frac{\Delta + m}{s} \left(\frac{s}{s-1} \right)^{s - \Delta - m}.$$

The first two equalities are clear, and the third one follows immediately from Burtin's Proposition 1, case $\varepsilon_i \equiv 1$, see [3], page 404. Hence

$$P_1 P_2 P_3 = \left(\frac{s-m}{\Delta + K - 1} \right)^{s - \Delta - m} \left(\frac{K + \Delta - s - 1}{\Delta + K - 1} \right)^{K + \Delta - s}$$

and, by (8),

$$P\left(\widehat{g}^{-1}(A_m) = A_m \bigcup B \right) = q_m(s),$$

where $q_m(s)$ is defined by (7). The set B may be chosen in $\begin{pmatrix} K - m \\ s - \Delta - m \end{pmatrix}$ ways. Therefore,

$$p_m(s) = \begin{pmatrix} K - m \\ s - \Delta - m \end{pmatrix} q_m(s),$$

which coincides with (6). The theorem is proved. $\qquad \square$

References

[1] Bowen, R., "Equilibrium States and the Ergodic Theory of Axiom A Diffeomorphisms," Lecture Notes in Mathematics, No. 470, Springer-Verlag, New York, 1975.

[2] Bollobas, B., *Random graphs*, Academic Press, London, 1985.

[3] Burtin, Y.D., On a simple formula for random mappings and is applications, *J. Appl. Prob.*, **17** (1980), 403–414.

[4] Diamond, P., Kloeden, P., and Pokrovskii, A., Interval stochastic matrices, a combinatorial lemma and the computation of invariant measures, *J. Dynamics & Diff. Eqns.*, **7** (1995), 341–364.

[5] Diamond, P., Kloeden, P., Pokrovskii, A., & Vladimirov, A., Collapsing Effects in Numerical simulation of a class of chaotic dynamical Systems and Random Mappings with a Single Attracting Centre, *Physica D*, **86** (1995), 559–571.

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

[7] Diamond, P., Klemm, A., Kloeden, P. & Pokrovskii, A.. Basin of attraction of cycles of discretizations of dynamical systems with SRB invariant measures, *J. of Statistical Physics*, **84** (1996), 713–733.

[8] Diamond, P., Kloeden, P., Kozyakin, V. and Pokrovskii, A., A model for roundoff and collapse in computation of chaotic dynamical systems, submitted for publication.

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

[10] Grebogi, E., Ott, E., and Yorke, J.A., Roundoff–induced periodicity and the correlation dimension of chaotic attractors, *Phys. Rev. A*, **$4** (1988), 3688–3692.
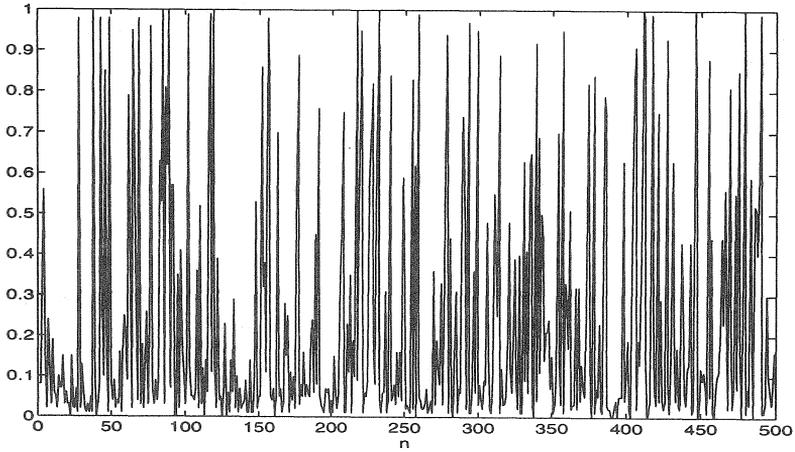[11] Shub, M., "Global Stability of Dynamical Systems", Springer-Verlag, New York, 1987.

Figure 1. Proportion of elements of $\mathbf{L}_\nu$ collapsing to zero, $2^{27} \leq \nu \leq 2^{27} + 500$, for the logistic $f_2(x) = 4x(1-x)$.
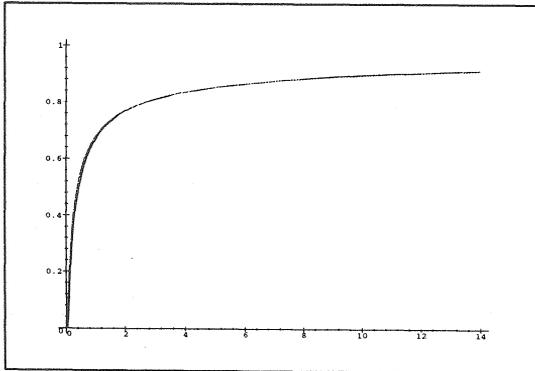

Figure 2. Experimental results $D^{(3/2)}_{0.001,\,10^6}(x)$ against the theoretical prediction erfc $\left(\frac{0.42}{\sqrt{2x}}\right)$.

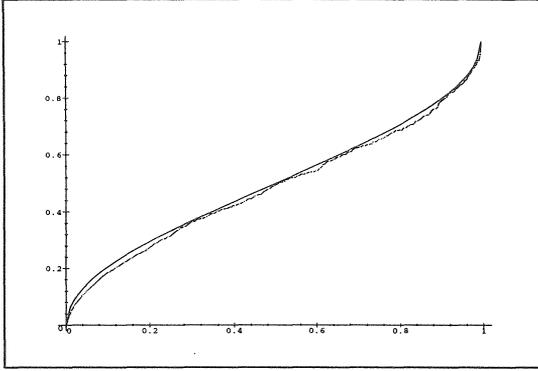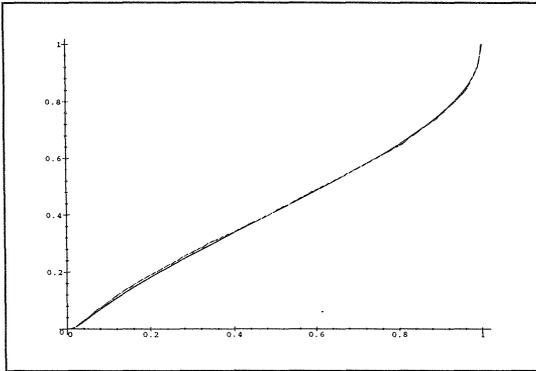Figure 3. $H^{(3/2)}_{0.01,\,10^6}(u)$ against $\frac{2}{\pi}\arcsin\left(\sqrt{u}\right)$.

Figure 4. $H^{(2)}_{0.001,\,10^7}(u)$ against $h_{10^7}(u;\widehat{c})/h_N(1;\widehat{c})$.

MATHEMATICS DEPARTMENT, UNIVERSITY OF QUEENSLAND, BRISBANE, 4072 AUSTRALIA
E-mail address: pmd@maths.uq.oz.au

MATHEMATICS DEPARTMENT, UNIVERSITY OF QUEENSLAND, BRISBANE, 4072 AUSTRALIA. PERMANENT ADDRESS: INSTITUTE OF INFORMATION TRANSMISSION PROBLEMS, RUSSIAN ACADEMY OF SCIENCE, 19, BOL'SHOI KARETNYI ST., MOSCOW, 101447, RUSSIA
E-mail address: ap@maths.uq.oz.au