Physica D 46 (1990) 241–253
North-Holland

CHAOS AND CONTINUED FRACTIONS

R.M. CORLESS, G.W. FRANK and J. Graham MONROE
Department of Applied Mathematics, University of Western Ontario, London, Ontario, Canada

Received 30 October 1989
Accepted 25 May 1990
Communicated by A. Dragt

This paper reports the use of the Gauss map from the theory of simple continued fractions as an example of a chaotic discrete dynamical system. Because of the simplicity of the map and the wealth of classical mathematical results, we are able to gain insight into the interaction between exact dynamical systems and their floating-point simulations. We calculate the correlation dimension and the capacity dimension of the Gauss map, and use these to examine current reconstruction techniques.

1. Introduction

The theory of continued fractions goes back at least to c. 500 A.D. to the work of Āryabhata, and possibly as far back as c. 300 B.C. to Euclid. The theory of chaotic dynamical systems is relatively recent, going back only to the work of Poincaré [23] and Birkhoff [2]. The foundations of the theory of continued fractions, as we know it now, are well established due to the work of Euler, Lagrange, Gauss, and others, while the foundations of chaotic dynamical systems are still evolving. This paper will use the well-established theory of simple continued fractions to critically examine some current methods used in the theory of chaotic dynamical systems.

Ref. [21] gives a good introduction to the classical theory of simple continued fractions, by which we mean continued fractions of the form

$$n_0 + \cfrac{1}{n_1 + \cfrac{1}{n_2 + \cfrac{1}{n_3 + \dots}}},$$

where the n_i are all positive integers, except n_0, which may be zero or negative. We will denote this as $n_0 + [n_1, n_2, n_3, \dots]$, and in what follows n_0 will usually be zero. Simple continued fractions have found applications in Fabry–Perot interferometry [15], and the concept of "noble" numbers used in orbital stability and quasi-amorphous states of matter [26]. For other uses of simple continued fractions in chaos, see ref. [8]. Other types of continued fraction exist; for example, Gautschi [10], Henrici [14], Jones and Thron [16] and others, use functional or analytic continued fractions in approximation theory, since analytic continued fractions can be very effective for computation. We will not be concerned with such continued fractions. We will summarize in the next section all the classical results that we need, without proof. Proofs can be found in refs. [1, 13, 17, 19–21].

2. Summary of classical results

2.1. The Gauss map

We begin with the classical method for finding the continued fraction representation of a number γ. We put n_0 equal to the integer part of γ, the greatest integer less than or equal to γ. If the fractional part of γ is not zero, we put γ_0 equal

to the fractional part of γ. We then invert γ_0, and repeat the process, this time getting n_1 and γ_1, and so on. Note that n_0 may be positive, negative, or zero, but that all the subsequent n_i will be positive, and that each γ_i is in the interval $[0, 1)$. This process gives a unique continued fraction for each starting point γ, and the process terminates if and only if γ is rational. (For any rational γ there is one other simple continued fraction which is only trivially different from the one generated by this algorithm.) This algorithm is related to the Euclidean algorithm for finding the greatest common divisor (gcd) of two integers k and m, in that if we use this method to find the continued fraction of k/m then the last non-zero integer part n_{final} is the gcd of k and m, and further this operation terminates in $\mathcal{O}(\log(\min(k, m)))$ operations [21]. Classically, most attention has been paid to the integers generated by this algorithm, which make up the continued fraction itself. However, Gauss was apparently the first to study the other part of this algorithm, which we present as the following map, called the Gauss map [19] (see fig. 1):

$$G(x) = \begin{cases} 0 & \text{if } x = 0, \\ \dfrac{1}{x} \bmod 1 & \text{otherwise.} \end{cases}$$

We use the notation "mod 1" to mean taking the fractional part. In terms of the Gauss map G, our algorithm then becomes

$$\gamma_{k+1} = \text{fractional part of } 1/\gamma_k = G(\gamma_k)$$
$$n_{k+1} = \text{integer part of } 1/\gamma_k,$$
$$\text{for } k = 0, 1, 2, 3, \ldots$$

and we see that the continued fraction is generated as a byproduct of the iteration of the Gauss map. Thus we expect that any classical results on continued fractions will have implications for the dynamics of the Gauss map.

Note that the jump discontinuities occurring at $x = 1/n$ (for each integer n) may all be removed by mapping onto the circle with the transformation $e^{2\pi i x}$. After this is done, we see that the Gauss map is a map of the circle onto the circle, and may be pictured on a torus, as in fig. 2. The singularity at the origin is not removed by this transformation. For convenience, the singularity is dealt with by artificially making zero a fixed point of the map (this makes our difficulties no worse). Most theorems on the dynamics of dis-

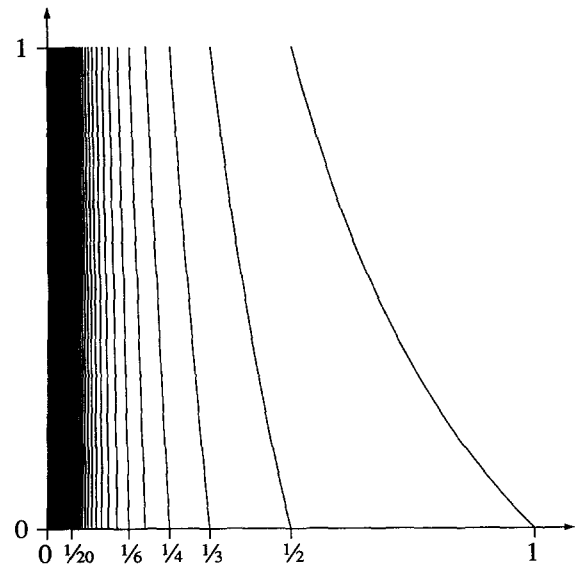

Fig. 1. The graph of the Gauss map $G(x)$. Note that there are an infinite number of jump discontinuities at values of $x = 1/n$, for integers n. In addition, there is a pole at the origin. The darkening of the curve towards the origin is suggestive of the fractional nature of the capacity dimension.

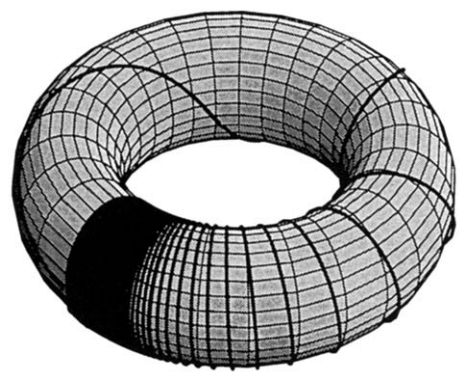

Fig. 2. The graph of the Gauss map $G(x)$ on the torus. Note that all the jump discontinuities have been removed, but that the pole at the origin remains. The darkening of the curve towards the singularity again gives an idea of the fractional nature of the capacity dimension.

crete maps assume continuity, which is thus violated here.

We make the following observation: if we represent a point in the interval $[0,1)$ by its continued fraction, $\gamma_0 = [n_1, n_2, n_3, \ldots]$, then a simple induction shows that $G(\gamma_0) = \gamma_1 = [n_2, n_3, n_4, \ldots]$, $G(\gamma_1) = \gamma_2 = [n_3, n_4, n_5, \ldots]$, $G(\gamma_2) = \gamma_3 = [n_4, n_5, n_6, \ldots]$, and so on. This makes clear the connection between the Gauss map and the shift map of symbolic dynamics. The only difference between them is that here the n_i may be arbitrarily large, whereas in the shift map the elements are taken to be bounded. This difference is a consequence of the singularity at the origin of the Gauss map, and means that the maximum winding number of the Gauss map is infinite.

An analogy is illuminating: if we think of our space as a circular hoop with the origin at one point O on the hoop, our initial point as a dimensionless bead on the hoop, and the Gauss map as taking the bead from its current position clockwise past the origin at least once to its next position on the hoop, then the integers n_i are the number of times the bead passes the origin on the ith iteration (the maximum such number is the "winding number" of the map, and here this is obviously infinite), and the γ_i are the coordinates of the bead on the hoop once it comes to rest. If the bead comes to rest close to the origin on one side, with a small γ_i, then on the next iteration it will be pushed many times around the hoop. If it comes to rest close to the origin on the other side, with a γ_i close to 1, then it will only go past the origin once on its next iteration. We may think of the bead as being pushed around the circle, with the strength of the push being inversely proportional to the distance measured counterclockwise from the point O.

2.2. *Periodic and fixed points of the Gauss map*

The following classical theorem, interpreted in a modern dynamical sense, identifies the fixed and periodic points of the Gauss map.

Theorem (Galois). The number γ has a purely periodic continued fraction, including the first integer n_0, if and only if γ is a "reduced quadratic irrational", which means that γ is a root of a quadratic equation with integer coefficients and, further, that its other root lies in the interval $(-1, 0)$.

Corollary. The periodic points of the Gauss map are the reciprocals of the reduced quadratic irrationals. These numbers are dense in $[0,1)$.

For a proof of the theorem, see ref. [21], or ref. [13]. To prove the corollary, we note that $\gamma = [n_1, n_2, n_3, \ldots]$ is periodic under the Gauss map if and only if its continued fraction is periodic, starting at n_1, by the shift property mentioned in the previous section. Numbers with periodic continued fraction expansions are dense in $[0,1)$ because their continued fractions may start off like any given continued fraction, and thus be arbitrarily close to any given number.

An example of particular interest is τ, the golden ratio, which satisfies $\tau^2 - \tau - 1 = 0$. The other root of this quadratic is $-1/\tau$, which is in the desired interval. The continued fraction of τ is $\tau = 1 + [1, 1, 1, 1, \ldots]$, so $1/\tau$ has the continued fraction $[1, 1, 1, 1, \ldots]$, which shows that $1/\tau$ is a point of period 1 of the Gauss map. We will return to this example later.

There are general results in the theory of chaotic dynamical systems, with which we could hope to establish the character of the set of periodic points of the Gauss map [18, 25, 28]. However, these results deal with the characterisation of the behaviour of *continuous* maps of the interval, extended by Block to maps of the circle [3], and the Gauss map has a singularity at the origin. Thus the hypotheses of these theorems are not weak enough to apply. However, the results of these theorems hold, as will be seen by direct methods. In fact, the methods used in Block's proof may also be used to prove the same results, though these methods are not used here.

We note here that there are infinitely many points of each period. For example, $[n_1, n_2, \ldots, n_k, n_1, n_2, \ldots, n_k, \ldots]$ has period k, for any choice of integers $n_1, n_2, \ldots, n_k$. Having points of arbitrary period is one characteristic of a chaotic map [18]. However, we would like to see if the map is sensitive to initial conditions (SIC) in that nearby initial points have orbits that separate at an exponential rate. This again can be established in an elementary fashion by using a classical result.

Theorem (Lagrange). γ has an ultimately periodic continued fraction, which means that $\gamma = [a_1, a_2, a_3, \ldots, a_i, n_1, n_2, \ldots, n_k, n_1, n_2, \ldots, n_k, \ldots]$ with transients $a_1, a_2, a_3, \ldots, a_i$ at the start of a periodic continued fraction, if and only if γ is a quadratic irrational (γ is a root of a quadratic with integer coefficients).

Corollary. The Gauss map is SIC.

For a proof of Lagrange's theorem, see ref. [13]. To prove the corollary, we note that every rational initial point is "attracted" to the artificial fixed point at 0, while every quadratic irrational is ultimately "attracted" to a periodic orbit. Both sets are dense in the interval $[0, 1)$. The rate of separation may be checked by considering all points in a small interval I, of width ε. By the pigeonhole principle, this interval must contain a rational number of the form p/n, where n is the smallest integer larger than $1/\varepsilon$. The number of iterations of the Gauss map required to reach zero for this initial point is, by the speed of the Euclidean algorithm, $\mathcal{O}(\log(n))$, and thus $\mathcal{O}(\log(\varepsilon))$. But the interval I also contains a purely periodic point since the reduced quadratic irrationals are dense in $[0, 1)$, and thus the rate of separation is exponential.

2.3. *Aperiodic points*

Non-quadratic irrationals have continued fraction expansions, as well. By Lagrange's theorem, these continued fractions are aperiodic, and hence the orbits of these initial points under the Gauss map are aperiodic. Note that most numbers in $[0, 1)$ are thus aperiodic. We examine some beautiful examples, beginning with one due to Euler:

(1) e (the base of the natural logarithms) has an aperiodic continued fraction expansion $e = 2 + [1, 2, 1, 1, 4, 1, 1, 6, \ldots]$. The elements of the orbit of this initial point are always of the form $[1, 2N, 1, 1, \ldots]$, $[2N, 1, 1, \ldots]$, or $[1, 1, 2N, \ldots]$, which tend to 1, 0, and $\frac{1}{2}$ respectively. Thus the ω-limit set of this orbit is the set $\{1, 0, \frac{1}{2}\}$, which, unlike the ω-limit sets of continuous maps, is *not* invariant under the Gauss map since $G(1) = G(\frac{1}{2}) = 0$, so G applied to this set simply gives 0. In other words, we have an asymptotically periodic orbit which is not asymptotic to a real orbit of the map. This cannot happen for a discrete dynamical system with a continuous map.

(2) (Stark [27]) If x is the positive root of $x^3 - 3600x^2 + 120x - 2 = 0$, then $x = 3599 + [1, 28, 1, 7198, 1, 29, 388787400, 23, 1, 8998, 1, 13, 1, 10284, 1, 2, 25400776804, 1, 1, \ldots]$ which has very large entries placed irregularly throughout. This intermittency is also a typical feature of a chaotic system [12].

(3) (Lambert, 1770 – cf. ref. [21]) The continued fraction for π is not known, in the sense that no pattern has been identified. It begins $\pi = 3 + [7, 15, 1, 292, 1, 1, 1, 2, 1, 3, 1, 14, 2, \ldots]$ and some 17 000 000 elements of this continued fraction have been computed by Gosper [4]. There are many open questions about this continued fraction – for example, it is not known if the elements of the continued fraction are bounded. We will return to this example in a later section.

2.4. *Lyapunov exponents*

We showed earlier that the separation of orbits initially close to each other occurred at an exponential rate. We would like to examine the Lyapunov exponents of the Gauss map, to see if we can explicitly measure the rate of separation. The Lyapunov exponents of orbits of the Gauss

map are defined as [8]

$$\lambda(\gamma) = \lim_{n\to\infty} \frac{1}{n} \ln\left(\prod_{i=0}^{n} |G'(\gamma_i)| \right)$$

whenever this limit exists. Nearby orbits will separate from the orbit of γ at an average rate of $e^{\lambda k}$, after k iterations of G. Khintchin [17] derived a remarkable theorem with which we could show the Lyapunov exponent of almost all (in the sense of Lebesgue measure) orbits to be $\pi^2/(6\ln 2)$. Easier ways have since been found to establish this result, using ergodic theory. We summarize the ergodic results in the next section. In this section we simply note that for any rational initial point, the above limit does *not* exist. Further, for any periodic orbit the calculation can be made explicitly, to give Lyapunov exponents that *differ* from the almost-everywhere value. For example, the fixed points $\alpha_N = [N, N, N, N, \ldots]$ have Lyapunov exponents

$$\lambda(\alpha_N) = 2\ln(1/\alpha_N) \sim \ln(N) + N^{-2} - \tfrac{3}{2}N^{-4} + \mathcal{O}(N^{-6})$$

so that there are orbits with arbitrarily large Lyapunov exponents. Note also that for the orbit of e, the limit defining the Lyapunov exponent is infinite.

The special case $N = 1$ gives τ, the golden ratio. Thus $\lambda(1/\tau) = 2\ln\tau = 0.96\ldots$, which is smaller than the almost-everywhere Lyapunov exponent. In fact, we have:

Theorem. No orbit of the Gauss map has a Lyapunov exponent smaller than $\lambda(1/\tau) = 2\ln\tau$.

Proof. Let $\gamma = [n_1, n_2, n_3, \ldots,]$ be any initial point in $(0, 1)$ such that $\lambda(\gamma)$ exists. We will show that the product $\prod_{i=0}^{N}(1/\gamma_i^2)$ which appears in the definition of $\lambda(\gamma)$ must be at least τ^{2N} (N sufficiently large), which will prove the theorem. We consider two subsequent elements γ_k and γ_{k+1} of the orbit of γ. If $k = N$, enlarge the product by one term. Note γ_k and γ_{k+1} are related by $\gamma_k = 1/(n_{k+1} + \gamma_{k+1})$. If $\gamma_k \le 1/\tau$ then the contribution of γ_k^{-2} to the product is at least τ^2. If instead $\gamma_k > 1/\tau$ then $\gamma_k\gamma_{k+1} = \gamma_{k+1}/(n_{k+1} + \gamma_{k+1}) = 1 - n_{k+1}\gamma_k \le 1 - \gamma_k < 1 - 1/\tau = 1/\tau^2$ so the contribution of $1/\gamma_k^2\gamma_{k+1}^2$ to the product is at least τ^4. This proves the theorem.

Remark. There are infinitely many initial points γ in $(0, 1)$ with this Lyapunov exponent. For example, all the numbers $\gamma = [n_1, n_2, n_3, \ldots, n_k, 1, 1, 1, 1, \ldots]$, that is, all the numbers whose continued fractions ultimately end in 1's, have Lyapunov exponent $2\ln\tau$. These are the so-called noble numbers [26], noticed for their resistance to chaos, and we see here that they all share the (still positive) minimum possible Lyapunov exponent under the Gauss map.

2.5. Ergodic results

The Gauss map is well-known in ergodic theory (see ref. [1] or ref. [19]). The results are summarized here, for contrast with the results of the sections previous and following.

The Gauss map preserves the Gauss measure

$$\mu(\mathrm{A}) = \frac{1}{\ln 2}\int_{\mathrm{A}} \frac{1}{1+x}\,\mathrm{d}\lambda,$$

where λ is the Lebesgue measure. Thus the Gauss map is ergodic, and almost all (in the sense of either the Lebesgue or Gauss measure) initial points have orbits which have the interval $[0, 1]$ as ω-limit set. Thus the *only* attractor whose basin of attraction has nonzero measure is the interval $[0, 1]$.

By the ergodicity of the map, we may explicitly calculate the Lyapunov exponent as follows:

$$\lambda(\gamma) = -2\lim_{n\to\infty}\frac{1}{n}\left(\sum_{i=0}^{n}\ln(\gamma_i)\right) = -\frac{2}{\ln 2}\int_0^1 \frac{\ln(x)}{1+x}\,\mathrm{d}\lambda = \frac{\pi^2}{6\ln 2} = 2.3731\ldots,$$

which holds for almost all initial points γ. This is

of interest, since there are few nontrivial maps for which the Lyapunov exponent can be calculated explicitly.

The metric entropy of the Gauss map is calculated in ref. [19] to be exactly the same as the Lyapunov exponent above. This is for good reason, as in the case of an ergodic map the notions of Lyapunov exponent and metric entropy coincide. It is interesting to contrast this with the *topological* entropy of the Gauss map, which we now show to be infinite, using the methods of Block et al. [3].

Decompose the interval $(0,1)$ into the infinite set of subintervals $(1/(k+1), 1/k)$ for $k = 1,2,3,\ldots$. Then G maps each subinterval onto the unit interval, so each subinterval covers all others (see fig. 1). Thus every entry in the infinite-dimensional covering matrix is unity. The largest eigenvalue of any n by n submatrix is n. Thus the topological entropy of the Gauss map is at least $\log(n)$, and as we may take n arbitrarily large, the topological entropy of the Gauss map is infinite. To succinctly describe this condition, we say the Gauss map is *hyperchaotic*.

3. The floating-point Gauss map

All of the results of the previous sections are valid for the familiar domain of the real numbers. However, when we work in any fixed-precision system, we have two difficulties:

(1) Not all real numbers are even representable in the system, and

(2) arithmetic does not have the properties we are used to.

For example, defining u as the smallest machine representable number which when added to 1 gives a number different from 1 when stored, we see that $G(\delta)$ is computed as 0, whenever δ is any number between 0 and u. This effectively limits the power of the singularity of the Gauss map.

To return to the analogy of the introduction, we consider the domain of machine-representable numbers not as a smooth circle but as a slotted ring, with the number of slots on the ring corresponding to the number of machine-representable numbers in our system. In this analogy, u corresponds to the width of the slots. Now our bead can only occupy one of the slots on the ring, and not just any arbitrary position, and the floating-point Gauss map takes the bead from one slot, winding around the ring as many times as are indicated by the integer part, and finally putting the bead into another slot. We see now that the maximum winding number of the floating-point Gauss map is finite, and the slot next to the origin is the one with this winding number.

A more evident difficulty is that all of the representable points are rational, and we know that the exact Gauss map takes these initial points to zero eventually. So if we define a floating-point Gauss map as

$$\hat{G}(x) = \begin{cases} 0 & \text{if } x = 0, \\ \dfrac{1}{x} \bmod 1 & \text{otherwise.} \end{cases}$$

where now the operations of division and "mod 1" take place over the floating-point domain, with attendant roundoff error, we have to answer some new dynamical questions:

(1) Are there any orbits which do not go to 0?

(2) Is the behaviour of the floating-point Gauss map similar to the exact Gauss map? In particular, is $\hat{G}$ chaotic?

(3) Can we define an appropriate Lyapunov exponent for this map?

(4) Is numerical work with $\hat{G}$ useful at all for study of G?

Not surprisingly, some orbits under $\hat{G}$ do terminate at 0, though often not when we expect them to. However, on some machines, some orbits never hit 0, being periodic. For example on the HP28S the initial point $\gamma_0 = 0.3$ gives $\gamma_1 = 0.333\,333\,333\,3$, $\gamma_2 = 0.000\,000\,000\,3$, and $\gamma_3 = 0.3 = \gamma_0$, with period 3. Note that under the exact Gauss map the second iterate (γ_2) of this initial point is zero. Since the set of machine-represent-

able numbers is finite, *all* orbits are ultimately periodic (perhaps with period 1, as at $x = 0$). Note that the behaviour of $\hat{G}$ depends strongly on the floating-point implementation. For example, with the Apple SANE numerics implementation, the starting point $\gamma_0 = 0.3$ gives an orbit with either a long transient regime or a period in excess of 65 000.

Since all orbits are ultimately periodic, and there are only a finite number of such orbits, the floating-point Gauss map (and indeed any machine simulation of any dynamical system) is *not* chaotic in the usual sense. Arbitrarily small perturbations in the initial conditions are not even allowed, so the sensitivity of the map to such perturbations is moot. The definition of the Lyapunov exponent for the exact Gauss map seems not to be relevant here: the presence of the derivative $G'(x)$ in the definition of Lyapunov exponent measures the effect of such arbitrarily small perturbations. However, if we define an approximate Lyapunov exponent for the first N iterations of the floating point Gauss map as

$$\lambda_N(\gamma) = \frac{1}{N} \ln\left(\prod_{i=0}^{N} |\hat{G}'(\gamma_i)| \right),$$

whenever the elements of the orbit are nonzero, then this in some sense measures the average sensitivity of the first N elements of the orbit under the *exact* Gauss map to arbitrarily small perturbations. This "Lyapunov exponent" is what is calculated in practice for a great many numerical simulations of dynamical systems, and if it is positive this is taken as evidence for chaos in the underlying system [12].

But what if the calculated orbit has no counterpart in the exact system? If roundoff errors introduced into the calculation produce an orbit that is unlike any in the exact system, this approximate Lyapunov exponent would be spurious. We will give a proof in the following section, which uses the techniques of backward error analysis, that shows orbits under the floating-point Gauss map are "machine close" to corresponding orbits under the exact Gauss map. A general theorem of this nature has been proved for hyperbolic invariant sets, by Bowen [12]. Here a direct proof is more appropriate and informative. This means that the approximate Lyapunov exponent defined above will accurately reflect the Lyapunov exponent of some orbit of the exact Gauss map, provided N is large enough that transient effects have been minimized, and not so large that accumulated roundoff error in the sum degrades the result.

We contrast this behaviour with what happens when continuous maps are made discrete by finite difference schemes. Yamaguti and Ushiki [31] and Ushiki [30] have shown that finite difference formulae applied to non-chaotic continuous systems may produce chaotic numerical solutions, assuming the calculations are carried out exactly. In contrast we have shown here that a chaotic discrete map becomes nonchaotic when implemented in fixed-precision arithmetic.

A further difficulty is that all of the orbits of $\hat{G}$ are ultimately periodic, and periodic orbits of G have Lyapunov exponents that are different from the almost-everywhere value (which is usually the exponent of physical interest). It is not immediately clear that these Lyapunov exponents calculated from $\hat{G}$ will tell us anything useful about the exact map G.

On closer examination, however, we see that if the period of an orbit is long, then the orbit behaves for a long time as if it were aperiodic. Hence we may expect that the computed Lyapunov exponent of a long period orbit will be close to $\pi^2/6 \ln 2 = 2.373\ldots$. This is what happens in practice, since many initial points seem to give long period orbits. For example, if we compute the first 100 000 elements of the orbit of 0.73 under $\hat{G}$ on the HP28S, we get a computed $\lambda = 2.36992$. This is within 0.2% of the expected value of the Lyapunov exponent of the exact Gauss map (though notice that the Lyapunov exponent of the orbit of G starting at 0.73 is not even defined – we *rely* on the roundoff error to give us our results, which is somewhat unusual).

3.1. Orbits under $\hat{G}$ are close to orbits under G

The following theorem justifies the remarks of the previous section. The basic idea of its proof is that given some initial point $\hat{y}$ the floating-point Gauss map also generates an initial point y whose continued fraction is exactly equal to $[a_1, a_2, a_3, \ldots]$, where the a_k are all (machine-representable) integers. This initial point y has a G-orbit that is everywhere within a small multiple of u, the machine epsilon, of the $\hat{G}$-orbit of $\hat{y}$. The technique of the proof is of interest for more than just the Gauss map, because similar techniques can be used to prove that numerical simulations of orbits of some continuous systems are machine close to exact orbits of some nearby initial point (for a descriptive review of work by Yorke, Grebogi, and Hammel establishing similar results for continuous maps see ref. [6]).

Theorem. If $x_0, x_1, x_2, x_3, \ldots$ is the sequence of iterates of $\hat{G}$, and $a_1, a_2, a_3, \ldots$ is the sequence of (machine-representable) integers that arise in the process, then $y = [a_1, a_2, a_3, \ldots]$ has an orbit under G whose elements are close to $x_0, x_1, x_2, x_3, \ldots$ in a sense to be made precise, and, in particular, y is close to x_0.

We will show first that we may approximate an element of the orbit of y by a certain rational number. We then show, using a common model of floating-point arithmetic, that the corresponding x_k is "machine close" to this same rational number. This last will be seen to depend on the fact that if you run the Gauss map backwards, errors are damped instead of amplified.

Proof. Consider $y_k = [a_{k+1}, a_{k+2}, a_{k+3}, \ldots]$. The rational numbers

$$p_n/q_n = [a_{k+1}, a_{k+2}, a_{k+3}, \ldots, a_{k+n}]$$

satisfy

$$|y_k - p_n/q_n| < 1/q_n^2 \text{ and } q_n \geq F_n$$

where F_n is the nth Fibonnacci number, from elementary properties of simple continued fractions (see ref. [21] or ref. [17] for details). This means that given an $\varepsilon > 0$, we can find an n so that $|y_k - p_n/q_n| < \varepsilon$.

To prove the second part, we use the common model of floating-point division that states that if the floating point numbers a, b, and c satisfy $a \div b = c$, where the division takes place over the floating-point numbers, then there is a number δ with $|\delta| < u$ so that $c(1+\delta) = a/b$ exactly. Note that we do not model the addition, since this will be seen to be unnecessary.

If the orbit $x_0, x_1, x_2, x_3, \ldots$ has been produced by a floating-point system satisfying this model, then for each n there is a number δ_{k+n} with $|\delta_{k+n}| < u$ such that

$$(1+\delta_{k+n})x_{k+n} = \frac{1}{a_{k+n+1} + x_{k+n+1}},$$

where we may consider the addition as exact, since a_{k+n+1} is a machine-representable integer, defined by this process, and x_{k+n+1} is a machine-representable floating point number. If we put $\varepsilon_{k+n+1} = x_{k+n+1}/a_{k+n+1}$ then we have

$$(1+\varepsilon_{k+n+1})(1+\delta_{k+n})x_{k+n} = \frac{1}{a_{k+n+1}}.$$

Now put

$$z_{k+m} = [a_{k+m+1}, a_{k+m+2}, a_{k+m+3}, \ldots, a_{k+n+1}]$$

for $m = 1, 2, \ldots, n$, and put $\varepsilon_{k+m} = z_{k+m} - x_{k+m}$ for $m = 0, 1, 2, \ldots, n$.

Note that $\varepsilon_k = z_k - x_k$ is the error we wish to estimate, since by the first part we can estimate the error $z_k - y_k$.

So now

$$\begin{aligned}(1+\delta_{k+m})x_{k+m} &= \frac{1}{a_{k+m+1} + x_{k+m+1}} \\ &= \frac{1}{a_{k+m+1} + z_{k+m+1} - \varepsilon_{k+m+1}} \\ &= z_{k+m}\frac{1}{1 - \varepsilon_{k+m+1}z_{k+m}}\end{aligned}$$

from whence, on cross-multiplying and expanding, we get the recurrence relation

$$\varepsilon_{k+m} = \delta_{k+m}x_{k+m} - (1+\delta_{k+m})z_{k+m}x_{k+m}\varepsilon_{k+m+1},$$

from which we may derive an upper bound on $\varepsilon_k = z_k - x_k$, and we note at this point that z_k is one of the rationals which approximates y_k. Note that the first term in this recurrence relation is essentially the roundoff error introduced at this particular step, while the second term is the error from one level below in the continued fraction, multiplied by a "shrinkage factor" $z_{k+m}x_{k+m}$.

As in the proof that τ has the minimum Lyapunov exponent, we are unable to say anything useful about z_{k+m} directly, but we are able to bound $z_{k+m}z_{k+m+1}$, which is easily shown to be less than $\frac{1}{2}$. With some simple estimates on the above recurrence this gives

$$\varepsilon_{k+m} \leq \begin{cases} 4u + \dfrac{1-4u}{2^{(n+1-m)/2}} & n-m \text{ is odd,} \\ 4u + \dfrac{1-3u}{2^{(n-m)/2}} & n-m \text{ is even} \end{cases}$$

and since as $n \to \infty$, $z_k \to y_k$, we have at last

$$|x_k - y_k| \leq 4u.$$

Thus there is a nearby initial point y_0 whose orbit under G follows as near as can be expected the computed orbit $x_0, x_1, x_2, x_3, \ldots$ of the floating-point Gauss map.

Our earlier example of $x_0 = 0.3$ gave a periodic orbit on the HP28S, which has $u = 10^{-11}$. The nearby initial point with this orbit under G is

$$\begin{aligned} y &= [3,3,3333333333,3,3,\ldots] \\ &= \tfrac{1}{2}(\sqrt{111111111128888888889} - 33333333333) \\ &\doteq 0.3 + 0.2999999999976 \times 10^{-12}. \end{aligned}$$

As a further curiosity, we note that the machine representation of $1/\tau$ on the HP28S is a fixed point of $\hat{G}$ allowing us to calculate the exact continued fraction of $1/\tau$ from a finite machine.

4. Dimension estimates of the exact Gauss map

In this section we exhibit a numerical estimate of the correlation dimension of the Gauss map. We used Maple [5] to generate, starting with π accurate to 8470 places, the first 8217 elements of the *exact* continued fraction for π, and then used this continued fraction to generate the first 8192 elements of the orbit of π under the exact Gauss map. We note that we can follow the exact orbit easily once we have the continued fraction representation because G is the shift map on the continued fraction representation of the initial point, and to get any particular element of the orbit to relative error less than 6×10^{-6}, we need only use at most the first twenty-five integers in its continued fraction to generate the decimal representation. We used enough precision in π to start with to guarantee $8192 + 25$ integers in the continued fraction, so we could guarantee the accuracy of the first 8192 elements of the orbit. We chose 8192, a pure power of 2, for convenience should we use a fast Fourier transform on the data. We then analyzed this orbit as if it were a set of experimental data, generated by some unknown process.

We believe this to be instructive in light of the recent enthusiasm for chaotic time series analysis of experimental data sets. The past decade has seen the development of a set of useful computational techniques which permit the "reconstruction" of geometric and dynamical attributes of nonlinear dynamical systems from limited experimental information. One such method is the calculation of the *correlation dimension*, a probabilistic measure of the extent to which the trajectory of a dynamical system is space-filling [11].

We accordingly treat the partial orbit $\{G^n(\pi)\}_{n=0}^{k}$ as sequential elements of our time series. Although the iterated Gauss map is a discrete dynamical system, the reconstruction

methods assume that the time series is generated by discretely sampling a *continuous* dynamical system governed by a set of ordinary differential equations [29]. We apply the method of delays [22] to this time series, embedding the univariate time series in $\mathbb{R}^d$ as a set of d-vectors $\{x_i\}_{i=1}^N$ where d is the embedding dimension and the maximum number of vectors N varies with the particular implementation of the method. If the original time series consisted of post-transient sampling of motion on an attractor, the method of delays produces a trajectory which lies on an equivalent attractor (modulo a diffeomorphism) projected to $\mathbb{R}^d$. The correlation dimension of the recovered attractor is then found using an algorithm developed by Grassberger and Procaccia [11].

We begin by computing the correlation function

$$C_d(r) = \frac{1}{N^2} \sum_{i \neq j=1}^{N} \Theta\big(r - |x_i - x_j|\big),$$

where $\Theta(z)$ is the Heaviside unit step function, $\Theta(z) = 1$ if $z > 0$, and 0 otherwise, and $|x_i - x_j|$ is the Euclidean distance between the points. The correlation function $C_d(r)$ provides a weighted count of the number of pairs of points within the trajectory which lie within a distance r of each other. If the correlation function exhibits exponential scaling behaviour, $C_d(r) \sim r^\nu$, for some range of small r-values, then the exponent ν is referred to as the correlation dimension of the recovered attractor. Noninteger values of ν are indicative of fractal scaling structure, typical of the strange attractors associated with chaotic dynamical systems [12].

The correlation dimension was computed for time series consisting of 8000 iterates of the exact Gauss map with initial point $\gamma_0 = \pi - 3 = 0.14159\ldots$. The method of delays was used to reconstruct multidimensional trajectories with embedding dimensions ranging from 1 to 4. The results are presented in fig. 3. We see that the

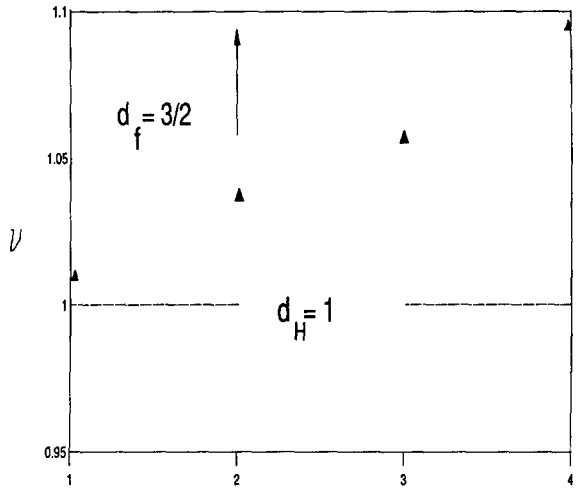

Fig. 3. The correlation dimension (ν) of the first 8000 elements of the orbit of π under the exact Gauss map, embedded in dimensions 1 through 4, calculated by the method of delays. The error reported by the least-squares fit used to generate these dimension estimates is on the order of ± 0.002. The Hausdorff dimension d_H is 1 for all embedding dimensions. The capacity dimension d_f is exactly $\frac{3}{2}$ for an embedding dimension of 2.

trajectory appears to possess a correlation dimension somewhat in excess of unity, with $\nu \approx 1.1$. The reconstruction was also performed for 4000 iterates of the floating-point Gauss map, and the results were indistinguishable. The error in ν due to the least-squares fit used to generate these dimension estimates is on the order of ± 0.002. The error in ν due to the finite sampling of the orbit is *unknown*, though it is probably at least an order of magnitude larger than the least-squares error, and probably increases as the embedding dimension increases. In light of these observations, the fact that ν is observed to be in excess of unity may not be significant.

In general, it is puzzling to obtain a correlation dimension in excess of unity for a map known to be one-dimensional. The explanation serves to illustrate the caution one must take when applying these methods to data sets obtained from systems with unknown governing equations. First, recall our assumption to treat the iterates of the Gauss map as time series output from a set of ordinary differential equations. This was to allow

the invocation of Takens' theorem, which applies only to continuous dynamical systems. While all continuous systems can generate discrete maps (for example by Poincaré sections) not all discrete systems arise from sections of continuous systems [7]. It is simple to show that the iterates of the Gauss map cannot arise from a "nice" one-dimensional continuous dynamical system, as follows.

Consider a continuous-time one-dimensional dynamical system defined on the unit interval by $dx/dt = f(x)$, with initial condition x_0. Assume this to have a unique solution for any initial condition in $[0, 1]$ [7]. Then it follows that a time-series of measurements of $x(t)$ must be (in the absence of fixed points) either monotonically increasing or monotonically decreasing, and cursory examination of the orbit of the Gauss map shows that it is neither. Therefore, a typical orbit of the Gauss map cannot be generated by a well-behaved continuous-time one-dimensional dynamical system.

The method of delays attempts to fit the discrete data by a single continuous trajectory. For this time series, the trajectory must fold back on itself to accommodate non-monotonic behaviour, which can only be supported if the trajectory is *embedded* in a space with dimension greater than unity. This, however, does not mean that the *attractor itself* must have dimension greater than unity.

Next, we examine the capacity dimension in the two-dimensional embedding. The attractor has the shape pictured in fig. 1 (or equivalently in fig. 2), which is precisely the graph of G. This is because all pairs $(x, G(x))$ are on the graph of G, and G is ergodic, so orbits are dense on the interval and hence pairwise dense on the graph of G. We see that the only possible cause of a dimension in excess of unity is the singularity at the origin.

When we examine the capacity dimension d_f, we cover the square $[0, 1] \times [0, 1]$ with equal-sized boxes [9]. We take boxes of size $1/n$ by $1/n$. Consider the columns of boxes in $0 < x < 1/J_1$, where J_1 is the smallest integer greater than $n^{1/2}$. Then if $j > J_1$ the interval $[1/(j+1), 1/j]$ is of width less than $1/n$, so every column of boxes in the interval $(0, 1/J_1)$ contains a line segment of G, and thus every box in the column contains points of G. This gives us $N(1/n) > n^{3/2}$, where $N(1/n)$ is the number of boxes of width $1/n$ that contain points of G. Thus the capacity dimension is at least $\frac{3}{2}$. In fact it is exactly $\frac{3}{2}$, since by counting the number of columns that fit one segment of G in *two* columns, and that fit one segment in *three* columns, and so on to the last interval, which takes $n/2$ columns to cover the portion of G, we see that the number of boxes containing points of G in the square is

$$N(1/n) \sim n\left(n^{1/2} + \frac{1}{2}\frac{n^{1/2}}{\sqrt{2}+1} + \frac{1}{3}\frac{n^{1/2}}{\sqrt{3}+\sqrt{2}} + \frac{1}{4}\frac{n^{1/2}}{\sqrt{4}+\sqrt{3}} + \cdots + \frac{1}{n/2}\frac{n^{1/2}}{\sqrt{n/2}+\sqrt{n/2-1}} \right),$$

which is of smaller order than $n^{3/2} \log n$. Hence $d_f = \frac{3}{2}$ exactly.

The $\log n$ in that last expression is of practical importance. As n gets large, and the box size gets small, this behaviour is unimportant compared to the exponent, that is, the capacity dimension. However, for *finite* box sizes, which any reconstruction technique must use, there is a danger that the $\log n$ may be confused with power-law behaviour. Indeed, for the range of r about 2^{-20} we get $n \log_2 n$ approximately equal to $n^{1.2}$. This spurious exponent can be avoided only by using very small box sizes, which in general requires very long time series, or by explicitly fitting for logarithmic behaviour. It is entirely possible, then, that the correlation dimension calculation as done in the previous section is picking up logarithmic behaviour and not power-law behaviour.

Finally, we remark that the Hausdorff dimension d_H of the graph of G is 1. This is because

the attractor is a countable union of simple curve pieces, and the Hausdorff dimension of a countable union is the supremum of the Hausdorff dimensions of the pieces [24], which here are all unity. This observation can be generalized to the embedding of an orbit of this map in any finite dimension. This map thus provides a nontrivial example of a map where the capacity or "fractal" dimension, d_f, is *not* equal to the Hausdorff dimension d_H.

5. Conclusions

The Gauss map has been shown to be a good example of a chaotic discrete dynamical system, in that it exhibits in an accessible fashion all the common features of such systems. In particular, the dimensions of the attractor embedded in two dimensions are $d_H = 1$, $\nu \approx 1.1$, and $d_f = 3/2$. Note that the last digit on the correlation dimension may not be significant if logarithmic behaviour is being confused with power-law behaviour in the reconstruction. The map is simple enough that the relationship of numerical simulation of the map to the exact map can be explored effectively. We find that the numerical simulation of the map behaves significantly differently, in that the numerical simulation is not chaotic, but is still useful in that the Lyapunov exponent of the exact map can be accurately calculated from the simulation. We have in fact shown that this behaviour of numerical simulation is general. This last result shows that arguing whether a given map is chaotic or merely has orbits with very long period may not be relevant, since the observable behaviour of the map is not significantly different in either case.

Acknowledgements

This work was carried out with the assistance of NSERC and ITRC. The original inspiration for this paper occurred in a course on chaos given by Professor M.A.H. Nerenberg. We are grateful to Professors Nerenberg, G.C. Essex, and T. Lookman for many useful discussions. Professors David Stoutemyer and Patrick Mann provided kind assistance with the plot appearing in fig. 2. The literature search was assisted by Ms. Pauline Seto. Direct PostScript was used to generate fig. 1 Mathematica and PostScript were used to generate fig. 2.

References

[1] P. Billingsley, Ergodic Theory and Information (Wiley, New York, 1965).
[2] G.D. Birkhoff, Sur quelqes courbes fermees remarquables, Bull. Soc. Math. France 60 (1982) 1–26.
[3] L. Block, J. Guckenheimer, M. Misiurewicz and L. Young, Periodic points of one dimensional maps, in: Springer Lecture Notes in Mathematics, Vol. 819 (Springer, Berlin, 1979) pp. 18–34.
[4] J.M. Borwein and P.B. Borwein, Pi and the AGM: a Study in Analytic Number Theory and Computational Complexity (Wiley, New York, 1987).
[5] B.W. Char, K.O. Geddes, G.H. Gonnet, M.B. Monagan and S.M. Watt, The Maple User's Manual, 5th Ed., WATCOM (1988).
[6] B.A. Cipra, Computer-drawn pictures stalk the wild trajectory, Science 241 (1988) 1162–1163.
[7] D.R.J. Chillingworth, Differential Topology with a View to Applications (Pitman, San Francisco, 1976).
[8] R.L. Devaney, An Introduction to Chaotic Dynamical Systems (Benjamin/Cummings, Menlo Park, 1985).
[9] J.D. Farmer, E. Ott and J.A. Yorke, The dimension of chaotic attractors, Physica D 7 (1983) 153–180.
[10] W. Gautschi, Efficient computation of the complex error function, SIAM J. Numer. Analysis 7 (1970) 187–198.
[11] P. Grassberger and I. Procaccia, Characterization of strange attractors, Phys. Rev. Lett. 50 (1985) 346–349.
[12] J. Guckenheimer and P. Holmes, Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields (Springer, New York, 1983).
[13] G.H. Hardy and E.M. Wright, An Introduction to the Theory of Numbers, 5th Ed. (Oxford Univ. Press, Oxford, 1979).
[14] P. Henrici, Applied and Computational Complex Analysis, Vol. 2 (Wiley, New York, 1977).
[15] K. Ikeda and M. Mizuno, Frustrated instabilities in nonlinear optical resonators, Phys. Rev. Lett. 53 (1984) 1340–1343.
[16] W.B. Jones and W.J. Thron, Continued Fractions: Analytic Theory and Applications (Addison–Wesley, Reading, 1980).
[17] A.Y. Khintchin, Continued Fractions (Noordhoff, Groningen, 1963).

[18] T.Y. Li and J.A. Yorke, Period three implies chaos, Amer. Math. Monthly 82 (1975) 985–992.
[19] R. Mañé, Ergodic Theory and Differentiable Dynamics (Springer, Berlin, 1987).
[20] I. Niven, Irrational Numbers, MAA Carus Mathematical Monograph Series, Vol. 11 (New Jersey, 1956).
[21] C.D. Olds, Continued Fractions (Random House, Toronto, 1963).
[22] N.H. Packard, J.P. Crutchfield, J.D. Farmer and R.S. Shaw, Geometry from a time series, Phys. Rev. Lett. 45 (1980) 712.
[23] H. Poincaré, Les Methodes Nouvelles de la Mécanique Celeste, Vols. 1–3 (Gauthier-Villars, Paris, 1899).
[24] D. Ruelle, Chaotic Evolution and Strange Attractors (Cambridge Univ. Press, Cambridge, 1989).
[25] A.N. Šaarkovskii, Coexistence of cycles of a continuous map of a line into itself, Ukr. Math. Z. 16 (1964) 61–71.
[26] M.R. Schroeder, Number Theory in Science and Communication (Springer, Berlin, 1984).
[27] H.M. Stark, An explanation of some exotic continued fractions found by Brillhart, Computers in Number Theory (Proc. Science Research Council Atlas Symposium No. 2, Oxford, 1971, eds, A.O.L. Atkin and B.J. Birch (Academic Press, London, 1971) pp. 21–25.
[28] P. Štefan, A theorem of Šaarkovskii on the existence of periodic orbits of continuous endomorphisms of the real line, Commun. Math. Phys. 54 (1977) 237–248.
[29] F. Takens, Lecture Notes in Mathematics, eds. D.A. Rand and L.S. Young (Springer, Berlin, 1981).
[30] S. Ushiki, Central difference scheme and chaos, Physica D 4 (1982) 407–424.
[31] M. Yamaguti and S. Ushiki, Chaos in numerical analysis of ordinary differential equations, Physica D 3 (1981) 618–626.