Journal of Universal Computer Science, vol. 2, no. 5 (1996), 380-395
submitted: 13/5/96, accepted: 13/5/96, appeared: 28/5/96

IS FINITE PRECISION ARITHMETIC USEFUL FOR PHYSICS ? [1]

Françoise Chaitin-Chatelin
CERFACS and CEREMADE (University Paris IX Dauphine), 42 av. G. Coriolis, 31057 Toulouse cedex, France. Email: `chatelin@cerfacs.fri`.

Abstract: Both empirical sciences and computations are fundamentally restricted to measurements/computations involving a finite amount of information. These activities deal with the FINITE – some finite precision numbers, coming out from measurements, or from calculations run for some finite amount of time. By way of contrast, as Leibniz expressed it, mathematics is the science of the INFINITE, which contains the concept of continuum. The related concepts of limit points, derivatives and Cantor sets also belong to mathematics, the realm of the infinite, and not to the world of the finite. One is then lead to wonder about the basis for the "unreasonable effectiveness of mathematics in the natural sciences" (Wigner (1960)). This puzzling situation gave birth, over the centuries, to a very lively philosophical discussion between mathematicians and physicists. We intend to throw into the debate a few simple examples drawn from practice in numerical analysis as well as in finite precision computations. By means of these examples, we illustrate some aspects of the subtle interplay between the discrete and the continuous, which takes place in Scientific Computing, when solving some equations of Physics.

Is Nature better described by discrete or continuous models at its most intimate level, that is below the atomic level? With the theory of quantum physics, it seems that the question has received a significant push towards a discrete space. However one can argue equally that the time variable in Schrödinger's equation is continuous. We will not get involved in the scholarly dispute between the continuous and the discrete. Instead, we will show on simple examples taken from Scientific Computing, the subtlety of the interplay between the continuous and the discrete, which can take place in computations, be it with finite precision or exact arithmetic.

1 Inexact versus exact arithmetic

Almost all the mathematical real numbers require, to be exactly represented in a given basis, an *infinite* amount of digits. Therefore exact computations require an infinite amount of information. On the contrary, with the finite precision arithmetic of the computer, each machine number is represented by a *finite* number of digits, say p (usually p bits in base 2). Consequently computations in finite precision deal with a finite amount of information. It is reasonable to expect that the result of a calculation done with p digits would tend to the exact result if p would tend to infinity. This is indeed very often the case. For a presentation of

[1] C. Calude (ed.). *The Finite, the Unbounded and the Infinite, Proceedings of the Summer School "Chaitin Complexity and Applications"*, Mangalia, Romania, 27 June – 6 July, 1995.

the theory of computability in finite precision, the reader is referred to Chaitin-Chatelin and Frayssé (1996).

Does this mean that one can consider the set of numbers produced by computers as just a practical approximation of the set $\mathbb{R}$ of the real numbers in mathematics? The *paradox of Newcomb-Borel* (Chatelin (1991)), to be expounded now, casts some doubts on this naive view. In 1909, Emile Borel proved that almost all real numbers are normal: in any base, the first digit and all subsequent digits are uniformly distributed (Borel (1909)). However, almost 30 years before, the astronomer Simon Newcomb (1881) had proved that the law of probability of the occurrence of natural numbers is such that all mantissae of their logarithms are equally probable. For Newcomb, *natural* numbers are numbers which occur in Nature, numbers which have been processed, either by computation or by natural phenomena. He proved his theorem after noticing how much faster the first pages of logarithmic tables wear out than the last ones. This result about natural numbers remained mostly unnoticed during one century, and has been frequently rediscovered by engineers and scientists (Benford (1938), Knuth (1969), Stewart (1993)). It follows from Newcomb's theorem that the law of probability of the first digit is logarithmic and not uniform (Feldstein and Goodman (1976)). In base 10, the number 1 is at least 6 times more frequent than 9 as first digit (see Figure 1). This obviously contradicts the well-known theorem of Borel, thus creating an apparent paradox. The way out of this paradox is as follows:

The set $\mathbb{R}$ of mathematical real numbers differs significantly from the set of natural numbers produced by Nature (measurements or computations).

Why is it so? This seems to be a signature of the *nonlinearity of the laws of Nature*. A theorem in Turner (1982) establishes that if uniformly distributed numbers are multiplied together, the resulting distribution converges to the logarithmic one, when the number of the factors tends to infinity. However in practice, a product of *three* numbers is enough to obtain a probability which differs from the theoretical limit by no more than 2.4% (see Table 1). By comparison, a uniform density probability for the first digit would yield the value $1/9 = 0.111$.

However, as k increases, the k^{th} digit tends to become evenly distributed (Feldstein and Goodman (1976)). Table 2 gives, as a function of the base β, the value k for which the limit value $1/\beta$ is achieved for the probability distribution with an accuracy of at least 4 decimal digits.

2 Finite versus infinite in exact arithmetic

Singularities in a model correspond to cases where the solution has no continuous derivatives in some parameter. The singularities of a general differentiable map can be very complex, but they are a set of measure zero Sard (1942). However their computational influence can be devastating:

l.s.d.	1	2	3	4	5	6	7	8	9
p_2	0.301	0.188	0.132	0.099	0.077	0.063	0.053	0.046	0.042
p_3	0.308	0.178	0.124	0.095	0.077	0.065	0.057	0.051	0.046
p_∞	0.301	0.176	0.125	0.097	0.079	0.067	0.058	0.051	0.046

Table 1: Probability distribution p_k for the leading significant digit (l.s.d.) of a product of k numbers, $k = 2, 3, \infty$.

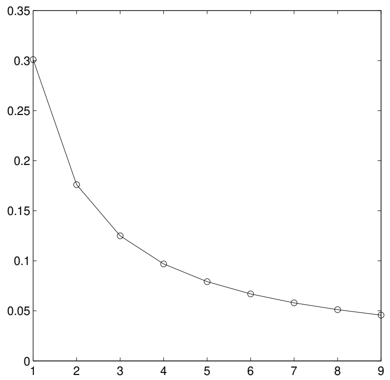

Figure 1: Probability distribution of the l.s.d. in base 10

> When a computation takes place in the neighbourhood of a singularity, results obtained with a numerical approximation can vastly differ from their exact counterparts, even if these approximate results are computed in exact arithmetic.

Because singularities are rare, they are usually not preserved under general perturbations: a double root becomes two close simple roots whenever the equation is slightly perturbed. Therefore, there has been a tendency to underestimate the role of singularities in approximation methods.
This is unfortunate, all the more because singularities play an essential role in the area of theoretical physics concerned with the reduction of theories (Berry (1991)). Often enough, a general theory based on elementary principles is reduced to a particular theory when some dimensionless parameter (say δ) tends to zero, for example, or equivalently, $1/\delta \to \infty$. Then the key question is the nature of the

β	2	3	4	8	10	16
k	13	10	8	5	5	4

Table 2: In base β, digits of rank $\geq k$ are practically uniformly distributed

limit as δ tends to zero, and very often it is highly singular.

The type of singularity is important and the existence of singularities is directly connected to the emergence of new phenomena. For example, statistical mechanics reduces to thermodynamics when $\delta = \frac{1}{N}$ tends to zero, N being the number of particles. The pressure $P(v,t)$ as a function of volume v and temperature t can be derived from the principles of statistical mechanics by large N asymptotics. But the reduction runs into difficulty near a critical point(v_c, t_c) where the compressibility $k = [-v(\partial P/\partial v)_t]^{-1}$ is infinite. The nature of the divergence of k as $t \to t_c$ is a power law, which determines the critical exponents. Thermodynamics is a continuous theory, except at critical points, which correspond to a new state of matter.

Other areas in physics today where singularities underlie some of the most difficult and intensively-studied problems include the limit: wave optics $\to$ geometrical optics, as the wave length tends to 0, or the limit: quantum mechanics $\to$ classical mechanics, as the ratio of the Planck constant to the classical action tends to 0.

As an illustration, we consider the one dimensional convection-diffusion equation

$$u'' + u' = f, \qquad u(0) = u(d) = 0.$$

The associated linear operator is T_d defined by $u \longmapsto T_d u = u'' + u'$. For any finite d, the spectrum of T_d consists of the real negative eigenvalues $\lambda_n = -\frac{1}{4} - \frac{\pi^2}{d^2} n^2$, $n = 1, 2, \ldots$. The limiting case $d = \infty$ corresponds to the single boundary condition $u(0) = 0$ and defines the limit operator T_∞. It can be shown that the spectrum of T_∞ is now the region Π of $\mathbb{C}$ enclosed by the parabola $P = \{z \in \mathbb{C};\ z = -\alpha^2 + i\alpha,\ \alpha \in \mathbb{R}\}$. This shows that the spectrum $\sigma(T_d)$ of T_d is highly discontinuous in the limit when $d \to \infty$. As a result, for z inside the parabola P, $\|(T_d - zI)^{-1}\|$ can be large even for z far away from the eigenvalues of T_d (Reddy and Trefethen (1994)). The same analysis carries over to the operator with explicit convection and diffusion parameters, that is

$$S(\delta, \nu, c)u = \nu u'' + c u', \quad u(0) = u(\delta),$$

provided that the previous parameter d is replaced by the Péclet number $\mathrm{Pe} = \delta c/\nu$. Note that Pe is large, for a fixed δ, when c is large and/or ν is small.
The consequence of the discontinuity of the spectrum when $\mathrm{Pe} \to \infty$ is that, even on a finite interval $[0, \delta]$, when *convection is dominant* (c/ν large), predictions of numerical behaviour, such as $\|(S(\delta, \nu, c) - zI)^{-1}\|$ small for z different from an eigenvalue, which are based on the exact spectrum which lies on the real negative axis, are likely to be misleading for z inside the parabola. Pseudoresonance for z far from the exact eigenvalues is discussed in Reddy and Trefethen (1994).

3 Computer simulations

The idea that computations and empirical sciences should be intimately related has been strongly stated by Fredkin (1982) in the form of a postulate: "There is a one-to-one mapping between what is possible in the real world and what is theoretically possible in the digital simulation world." Along the same line, we have also argued in Chaitin-Chatelin and Frayssé (1996) that

because no equation is exact in the real world, computer simulations can be closer to the physical reality of unstable processes than exact computation.

Roundoff errors are often considered negatively, as a severe limitation on the purity of mathematics, leading in the extreme cases to arithmetic artefacts. The following example puts finite precision arithmetic in a more positive perspective. Consider the family of matrices $A_n = QJ_nQ^*$, where J_n is a Jordan block of order n defined by

$$J_n = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ \vdots & 0 & 1 & \ddots & 0 \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & \dots & 0 & 1 \\ 0 & 0 & \dots & 0 & 0 \end{pmatrix},$$

and Q is a unitary $n \times n$ matrix. J_n is the Jordan form of A_n which has the unique eigenvalue 0 with ascent n. Therefore, the sensitivity of this eigenvalue to perturbations in A_n of size ε is $\varepsilon^{\frac{1}{n}}$: it increases exponentially with n. We illustrate the high sensitivity of the zero eigenvalue by computing the eigenvalues of A_n by the classical QR algorithm (Chatelin (1993)), for $n = 10$, 50, 200 and 500. It is known that the Toeplitz operator J which is the limit in l^2, of J_n as $n \to \infty$, has a spectrum which consists of the closed unit disk (Chatelin (1983)).

The computed spectra are plotted in Figure 2. The difference that we see between the exact eigenvalue 0 and the n computed ones is only the result of the spectral instability, because the eigensolver QR is a reliable algorithm. The role of the computer arithmetic is to make this spectral instability visible, it does not create it.

It is clear that, as n increases, most of the computed eigenvalues tend to cluster first on a circle centered at 0, with radius converging to 1 as n increases, and then to gradually fill the interior for much larger values of n. The exponential spectral instability exhibited above by A_n may seem overwhelming for large n. However, any situation is two-sided, and a more optimistic view can be derived from a look at the alternate perspective: use the computed eigenvalues of A_n to approximate the spectrum of the Toeplitz operator J. In exact arithmetic, the task is hopeless for any finite n: 0 is always at distance 1 from the unit circle. But in finite precision, the spectral information delivered by a reliable software converges towards the border of the limit spectrum as $n \to \infty$. The spectral information computed from A_n is already qualitatively very good for n as low as 500 (which is small when compared to infinity !).

Eigenvalues of the matrix A of order n are singularities of the map $z \longmapsto (A - zI)^{-1}$. They consist of at most n points in the complex plane. The topological dimension of each eigenvalue is therefore $d = 0$. Consider now a multi-

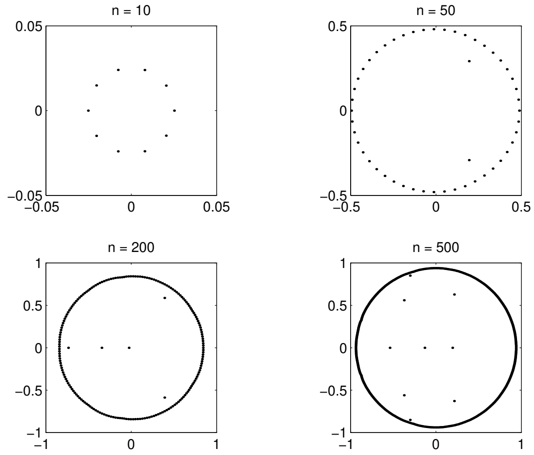

Figure 2: Eigenvalues of A_n computed with QR, $n = 10,\ 50,\ 200,\ 500$

ple eigenvalue such as zero for the Jordan block J_n. One can define (Chaitin-Chatelin (1996)) the *fractal dimension* D of such an eigenvalue. For the eigenvalue 0 of A_n, one gets $D_n = 1 - 1/n$, showing that $D_n \to 1$ as $n \to \infty$. This may explain why the computed eigenvalues tend to cluster, in their vast majority, on a circle which is a line of topological dimension 1. Figure 3 gives the computed spectra for $n = 1000$ and $n = 2000$.

None of the digits of the computed eigenvalues are correct. However, we have seen that they contain meaningful spectral information. Wrong computer results should not be systematically discarded. If they are produced by methods and software which are proven reliable, they contain valuable information on the problem, which can be put to good use.
An arithmetic artefact occurs when computer results differ vastly from their exact counterparts which should be obtained if the algorithm was run in exact arithmetic.

Is an arithmetic artefact always bad?

From a mathematical point of view, the answer is obviously YES. However, this may not be the right way to analyze physical reality: reliable software reflects the instability which is in the model, it does not create it. In extreme situations, the mathematical or numerical instability is so high that we cannot hope for a conventional control of the computational error, as was possible in less ill-conditioned situations. Classical computation seems meaningless at first because none of the digits in the results are correct. In the novel theory called **Qualitative Computing** (Chaitin-Chatelin and Frayssé (1996)), the aim is not anymore to control the computing error (which does not tend to zero) but rather to extract *meaning* from wrong results. When the computation cannot deliver

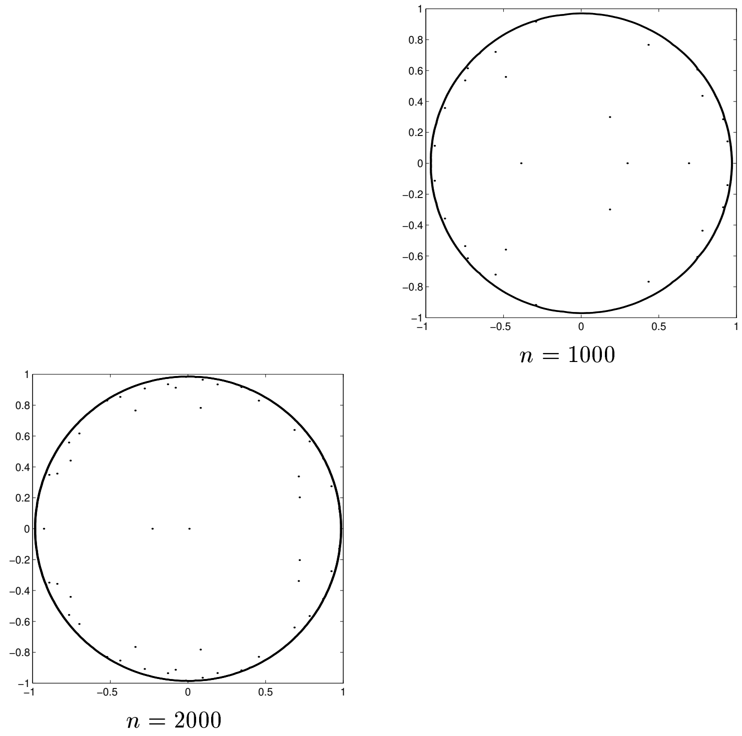

Figure 3: Computed spectra for A_n

full information on the exact solution, because the problem is singular, then the computation –by means of a reliable method– can still deliver relevant *partial information* about the singularity. There can exist a broad spectrum of intermediate situations between right and wrong on the computer.
Methods proved convergent in exact arithmetic may fail to converge in finite precision. There can exist an intermediate state between convergence and divergence, where computer results remain bounded (Chaitin-Chatelin and Frayssé (1996), Chaitin-Chatelin and Gratton (1996). Simple computer simulations commonly found in Linear Algebra are described in Braconnier, Chaitin-Chatelin, and Gratton (1996a, 1996b) which exhibit a chaotic behavior because of the computer arithmetic.

4 When the discrete is better explained by the continuous: the lattice logistic

One area which illustrates well the tension between the finite and the infinite is that of nonlinear dynamics. As an example, we choose the logistic map $f(x) = rx(1-x)$ (Schuster (1989)) which defines the fixed-point equation $x = rx(1-x)$, $0 \leq x \leq 1$, $0 \leq r \leq 4$, where the variable x and the parameter r are two real numbers. The associated iteration is

$$x_0, \; x_{k+1} = f(x_k), \; k \geq 0. \tag{1}$$

We recall at this point the biological origin of the logistic map, in which x_k models the evolution of the normalized population of a species in succeeding generations, $k = 1, 2, \ldots$ (May (1976)) which can take only rational values.
The fixed-point equation has two solutions $x^* = 0$ and $x^* = 1 - 1/r$, $r \neq 0$, which are distinct for all $r \neq 1$.

Figure 4 displays the computed logistic on the stability interval $[0, 4]$. And Figure 5 displays an enlargement of the interval $[3.5, 4]$. The starting point is $x_0 = 0.5$, and for values of r discretized in $[0, 4]$ or $[3.5, 4]$, the computed iterates $\tilde{x}_k$, for $k = 201$ to 250 are plotted. Only the stable solutions for the iterated maps f, $f^2, \ldots, f^p$ appear on the plot. The computed logistic displays, with each

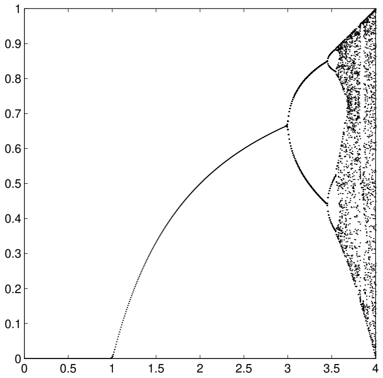

Figure 4: The logistic on $[0, 4]$

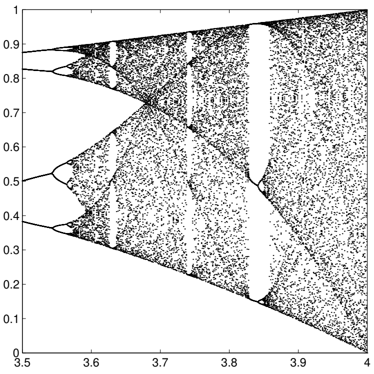

Figure 5: Zoom on $[3.5, 4]$

window of periodicity p, the stability interval for the map $x \longmapsto f^p(x)$. The finite precision on such a simple calculation reveals the intricate stability pattern of the iteration (1) without adding to its inherent mathematical instability. See below. The parameter r unfolds the complete variety of stabilities for the computational scheme (1), ranging from full stability on $[0,3[$, $r \neq 1$ (convergence $x_k \to x^*$ as $k \to \infty$), to no stability for $r > 4$ (divergence $x_k \to -\infty$ as $k \to \infty$, for almost all x_0). In the intermediate zone $[3,4]$, patterns of convergence are observed for subsequences extracted from $\{x_k\}_{\mathbb{N}}$. They can be interpreted as convergence towards the stable solutions of the equation $x = f^p(x)$: they correspond to the convergence of the p subsequences of $\{x_k\}$ starting from the p values $x_0,\ x_1 = f(x_0),\ \ldots, x_{p-1} = f^{p-1}(x_0)$.

The variability of stability is generated by the computational bifurcation at the unique singular point (a double root) of the equation $x = f(x)$ that occurs for $r = 1$ and is reflected for $r > 3$ on the computation of the solution $x^* = 1 - 1/r$. The parameter r unfolds the variation of stability for the *iteration with f* as a computational process to solve the equation $x = f(x)$. The unfolding shows how the algebraic singularity of order 2 at $r = 1$ is reflected as a succession of changes in the stability of the computation process.

The computer plots shown in Figures 4 and 5 are meaningful if the computational process (1) can be shown to be **reliable** in some sense (Chaitin-Chatelin and Frayssé (1996)). Each iteration step is very simple (one subtraction and two multiplications). However, for values of r corresponding to the chaotic regime $(3.7 \leq r \leq 4)$, a computed iterate $\tilde{x}_k$ may have lost all significant figures before $k = 100$. It is therefore important to show that this inaccuracy is not created by finite precision, but reflects the *high sensitivity of the computational scheme* (1) *to arbitrary perturbations*, including finite precision as a special case.

This has been done by Hammel, Yorke, and Grebogi (1987). They introduce the ε-pseudotrajectories by inserting a perturbation bounded by ε at each iteration k for N steps at most:

$$|f(z_{k-1}) - z_k| < \varepsilon, \quad 1 \leq k \leq N.$$

Then they show that, given an ε-pseudotrajectory $\{z_k\}$, for a small enough ε, there exist $\delta(\varepsilon)$ and an *exact* trajectory $\{y_k\}_{k=0,\ldots,N}$ of length $N(\varepsilon)$, which approximates, or *shadows* the pseudotrajectory within the precision $\delta(\varepsilon)$, i.e. with $y_k = f(y_{k-1})$, $|y_k - z_k| < \delta$, $0 \leq k \leq N(\varepsilon)$.
Finally they show that a computed trajectory $\{\tilde{x}_k\}$ can be considered as an ε-pseudotrajectory, with ε of the order of machine precision. For example, on a CRAY X-MP with 14 decimal digits, the machine precision is of order 10^{-14}. For $x_0 = 0.4$ and $r = 3.8$, they get $\varepsilon < 3\ 10^{-14}$, $\delta(\varepsilon) \sim 10^{-8}$ and $N(\varepsilon) \sim 10^7$. Similar results are given for other values of x_0 and r. They conjecture that for any map producing the same behaviour as $f = rx(1-x)$, the computed trajectory with machine precision Ψ would correspond to a precision $\delta \sim \Psi^{\frac{1}{2}}$ for the computation of $N \sim \Psi^{-\frac{1}{2}}$ steps.

Despite this definitive roundoff analysis, doubts are continuously expressed about the ability of computer orbits to capture the exact dynamics because of the discreteness of the computer arithmetic. Because all bounded orbits are necessarily periodic on a computer, true chaotic behaviour seems out of reach.

In order to examine the basis for this pessimism, we introduce the lattice logistic considered in Jackson (1991, pp. 216–221). The variable x in $[0, 1]$ is discretized by $\xi^{(j)} = \frac{j}{N+1}, \quad j = 1, \cdots, N$ and $f(x)$ is approximated by

$$\Phi(\xi) = \frac{1}{N+1} \lfloor (N+1) r \xi (1-\xi) \rfloor , \tag{2}$$

where $\lfloor a \rfloor$ denotes the integer part of a and $\xi \in \left\{\xi^{(1)} \cdots \xi^{(N)}\right\}$. The lattice map $\xi \longmapsto \Phi(\xi)$ defines an $\frac{1}{N+1}$-pseudotrajectory of arbitrary length: if we set $\xi_{k+1} = \Phi(\xi_k)$ for an arbitrary $k \geq 0$, then

$$|f(\xi_k) - \xi_{k+1}| = \left| f(\xi_k) - \frac{1}{N+1} \lfloor (N+1) f(\xi_k) \rfloor \right| .$$

Because $|(N+1)f(\xi) - \lfloor (N+1)f(\xi) \rfloor| \leq 1$ for all ξ in the lattice, we conclude that $|f(\xi_k) - \xi_{k+1}| \leq \frac{1}{N+1}$ for all k. In Figures 6 to 9, we display the iterates ξ_k for $k = 100$ to 200, for r discretized at 500 equally spaced points in $[0, 4]$. The values of N are taken to be $N = 10$, 50, 100 and 500. The starting points are the N points $\xi_0^{(j)} = \frac{j}{N+1}$, $j = 1, \cdots, N$.

The unfolding by the parameter r shows how well the periodic behaviour approximates the continuous limit. The discrete is explained by the continuous for N as low as 10 or 50. By way of comparison, there are roughly $2^{52} \sim 10^{15}$ machine numbers in $]0, 1]$, when IEEE double precision is used, as in Figure 4 and 5, where machine numbers have a mantissa of $p = 53$ bits.

In Figure 10 (resp., Figure 11) we again display ξ_k for $k = 100$ to 200, for r discretized at 1000 equally spaced points in $[0, 4]$ (resp., $[3.4, 4]$). The unique initial point $\xi_0 = 1/2$ is chosen. Despite the coarseness of the lattice ($N = 10^4$ points in $]0, 1[$ compared to $N \sim 10^{15}$), Figures 10 and 11 display computer plots which are strikingly similar to those of Figure 4 and 5.

In addition, the lattice logistic displays aesthetic patterns at the bifurcation around $r = 3$. Figures 12 to 15 display ξ_k for $k = 300$ to 400, for r discretized at 300 equally spaced points in $[2.96, 3.04]$, for $N = 500$, 1000, 3000 and 10^4. All possible initial points are considered.

5 When a new phenomenon can emerge from the discrete: the fully discrete wave equation

The continuous wave equation

$$\frac{\partial^2 u}{\partial t^2} - v^2 \frac{\partial^2 u}{\partial x^2} = 0$$

is of fundamental importance almost everywhere in Physics. The semidiscretized model (discretization in space) has been used successfully to describe the propagation of waves in homogeneous monodimensional crystals. In a recent work, Iavernaro and Trigiante (1992) have studied the fully discrete wave equation in

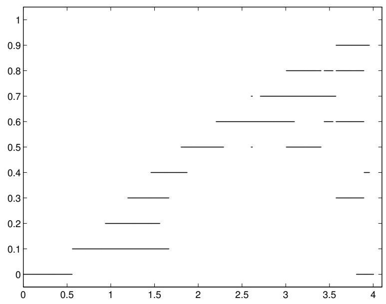

Figure 6: The lattice logistic for $N = 10$

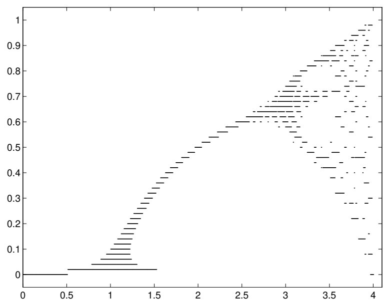

Figure 7: The lattice logistic for $N = 50$

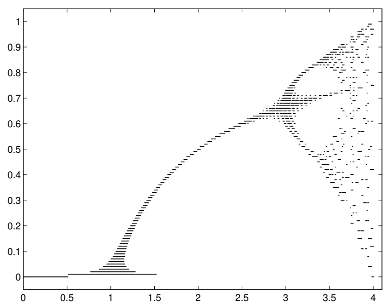

Figure 8: The lattice logistic for $N = 100$

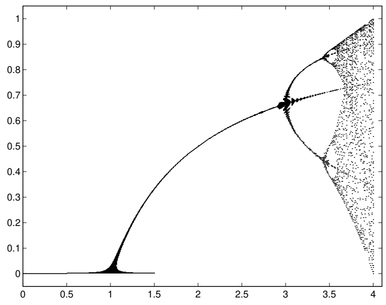

Figure 9: The lattice logistic for $N = 500$

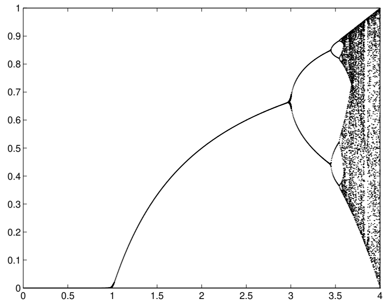

Figure 10: The lattice logistic for $N = 10^4$

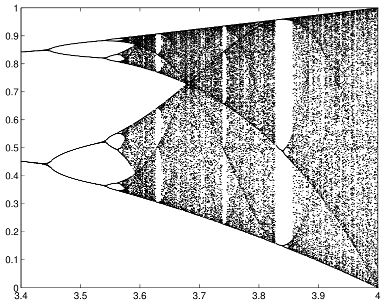

Figure 11: The lattice logistic for $N = 10^4$ (zoom)

one space dimension (Δx is the space mesh, and Δt is the time mesh). The dispersion relations show that, when time is also discretized, the phase and group velocities take the same maximum value $v_{\max}$ in the limit of an infinite wavelength, as in the semidiscrete approximation. But in addition, this can happen for any finite wavelength: dispersion tends to disappear as $v \to v_{\max} = \Delta x/\Delta t$.

The existence of a maximum velocity $v_{\max}$ where dispersion disappears is a consequence of the discretization in time.

Iavernaro and Trigiante (1992) apply this result to the case of a one-dimensional crystal consisting of $(N+1)$ atoms of mass m. For an appropriate Δt, they show the existence of arbitrary periodic solutions traveling at maximum speed without dispersion. They are soliton-like solutions which can occur in a *linear* equation as a result of the discretization in time.

For the light velocity $c = 3\ 10^8 \mathrm{m/s}$, and for the interatomic distance $\Delta x \sim 3\ 10^{-10}\mathrm{m}$, one obtains from $c = \Delta x/\Delta t$, the estimation $\Delta t = 10^{-18}\mathrm{s}$. It is interesting to remark that an estimate of the same order for the time unit Δt has been proposed (Wolf (1989)), based on completely different considerations.

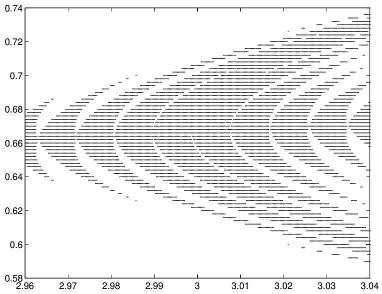

Figure 12: First bifurcation for $N = 500$

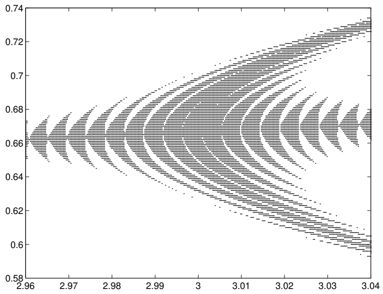

Figure 13: First bifurcation for $N = 1000$

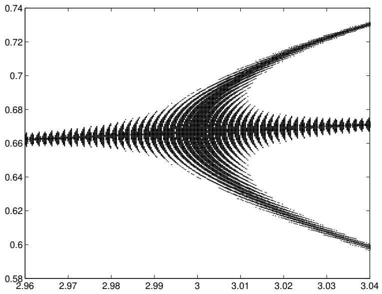

Figure 14: First bifurcation for $N = 3000$

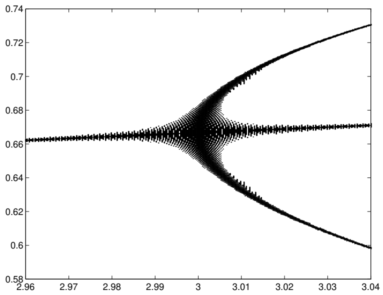

Figure 15: First bifurcation for $N = 10^4$

6 Is the Mandelbrot set really that complex?

The Mandelbrot set (Devaney (1989)) is associated with the logistic map in the complex plane $\mathbb{C}$. We consider the iteration:

$$z_0 = 1/2, \quad z_{k+1} = \lambda z_k(1 - z_k), \quad k \geq 0 \tag{3}$$

for z_k and λ complex. Depending on λ, (3) may either converge, remain bounded or diverge. The Mandelbrot set $\mathcal{M}$ is defined as the set of λ in $\mathbb{C}$ such that $|z_k| \not\to \infty$ as $k \to \infty$. Because of its fractal structure, the Mandelbrot set has been described as the most complex mathematical object (Penrose (1989)). And the question arises to decide whether a given complex number belongs to $\mathcal{M}$ or not. Blum, Shub, and Smale (1989) have proposed a theoretical framework according to which this question cannot be answered: it remains undecidable.

However, from the point of view of *algorithmic complexity* (Chaitin (1987, 1996)), the algorithmic content of (3) is extremely low. With this algorithmic viewpoint, $\mathcal{M}$ is not at all complex. This is in full accord with the fact that only a few steps of (3) are required to get a good picture of $\mathcal{M}$.

To approximate $\mathcal{M}$, we can plot for any λ in $\mathbb{C}$ the map: $\lambda \longmapsto k_\lambda$ defined as the index of the first iteration such that $|z_k(z_k - 2)| \geq 8$, where $\{z_k\}$ are computed by (3). Such a map is called the *divergence portrait* of the complex iteration (3). We have used a maximum number of iterations equal to 32. The result is displayed on Figure 16. The 32 regions corresponding to $k_\lambda = 1$ up to 32 are displayed according to a scale of 9 levels of gray. Black corresponds to $k_\lambda >$ 32. One recognizes in black a very convincing image of the Mandelbrot set. This set defined as $\{\lambda; |z_k| \not\to \infty\}$ corresponds to the largest values k_λ. The left (resp. right) plot corresponds to the region $[-2, 4] \times [-3, 3]$ (resp. $[-0.13, 0.13] \times [1, 1.26]$) in $\mathbb{C}$.

7 Conclusion

From the variety of computing examples which has been presented, no clear conclusion can be drawn in favour of either of the two aspects of Nature, the discrete or the continuous. "Subtle is the Lord" as Einstein used to say. In the absence of a definitive clue in either direction, it seems wise to keep in mind the two aspects at the same time, and to make them interfere constructively. On the road to this goal, scientific computing can be a tool of choice because it enables one to bring face to face the finite and the infinite, by comparing computer simulations to exact computations, whenever possible.

Acknowledgment

The author wishes to thank her co-workers of the Qualitative Computing Group at CERFACS for their help in conducting the numerical experiments.

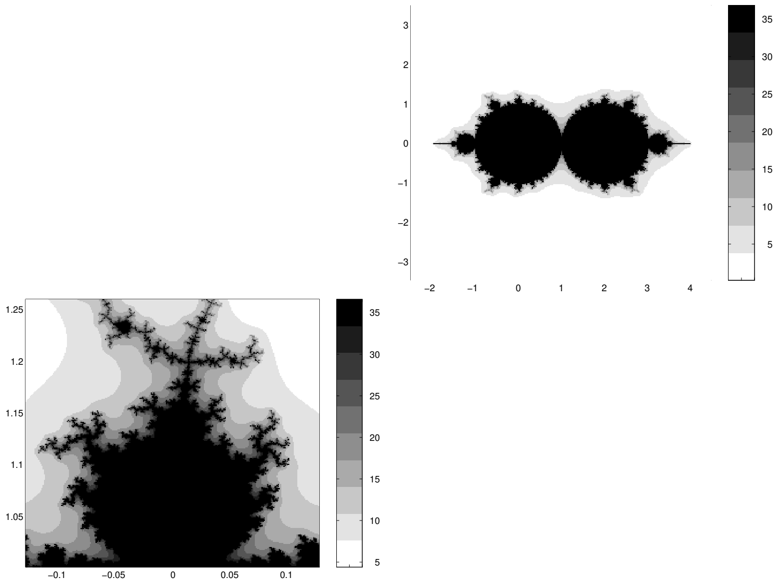

Figure 16: The Mandelbrot set

References

F. Benford, (1938), *On the law of anomalous numbers*, Proc. Amer. Philos. Soc., **78**, 551–572.

M. V. Berry, (1991), *Asymptotics, singularities, and the reduction of theories*, in **Proceedings of the 9th Congress of Logic, Methodology and Philosophy of Science**, D. Prawitz, B. Skyrms, and D. Westerståhl, eds., Studies in Logic and the Foundations of Mathematics, Uppsala, Kluwer Academic Publishers, Dordrecht, The Netherlands.

L. Blum, M. Shub, and S. Smale, (1989), *On a theory of computation and complexity over the real numbers: NP-completeness, recursive functions and universal machines*, Bull. AMS, **21**, 1–46.

E. Borel, (1909), *Les probabilités dénombrables et leurs applications arithmétiques*, Rend. Circ. Mat. Palermo, **27**, 247–271.

T. Braconnier, F. Chaitin-Chatelin, and S. Gratton, (1996a), *Entre convergence et divergence: exemples de chaos arithmétique en Algèbre Linéaire lié à la nonnormalité.* Work in progress.

T. Braconnier, F. Chaitin-Chatelin, and S. Gratton, (1996b), *Chaotic behavior for linear- and eigen- solvers applied on nonnormal matrices.* In preparation.

F. Chaitin-Chatelin and V. Frayssé, (1996), **Lectures on Finite Precision Computations**, SIAM, Philadelphia.

F. Chaitin-Chatelin and S. Gratton, (1996), *Convergence of successive iteration methods in finite precision under high nonnormality*, BIT. To appear.

F. Chaitin-Chatelin, (1996), *Is nonnormality a serious computational difficulty in practice ?* IFIP WoCo 7 **"The Quality of Numerical Software"**, Oxford, 8–12 July.

G. J. Chaitin, (1987), **Information, randomness and incompleteness**, World Scientific, Singapore.

G. J. Chaitin, (1996). URL `http://www.research.ibm.com/people/c/chaitin`.

F. Chatelin, (1983), **Spectral approximation of linear operators**, Academic Press, New York.
F. Chatelin, (1991), *Les ordinateurs calculent faux et c'est tant mieux.* Journées Observatoire de Paris.
F. Chatelin, (1993), **Eigenvalues of matrices**, Wiley, Chichester. Enlarged Translation of the French Edition with Masson.
R. L. Devaney, (1989), **An introduction to chaotic dynamical systems**, Addison-Wesley, Redwood City, California, second ed.
A. Feldstein and R. Goodman, (1976), *Convergence estimates for the distribution of trailing digits*, J. ACM, **23**, 287–297.
E. Fredkin, (1982), *Digital information mechanics.* Talk given in January, Moskito Island, British Virgin Islands.
S. H. Hammel, J. A. Yorke, and C. Grebogi, (1987), *Do numerical orbits of chaotic dynamical processes represent true orbits ?*, J. Complexity, **3**, 136–145.
F. Iavernaro and D. Trigiante, (1992), *Interesting physical properties of the discrete wave equation*, in **Proceedings of the First World Congress of Nonlinear Analysis**, vol. 2, Berlin, Walter de Gruyter, 1193–1202.
E. A. Jackson, (1991), **Perspectives of nonlinear dynamics**, Cambridge University Press, Cambridge. Tome 1.
D. E. Knuth, (1969), **The art of computer programming. Vol. 2 : Seminumerical algorithms**, Addison-Wesley, Reading Mass.
R. May, (1976), *Simple mathematical models with very complicated dynamics*, Nature, **261**, 459–467.
S. Newcomb, (1881), *Note on the frequency of use of different digits in natural numbers*, Amer. J. Math., **4**, 39–40.
R. Penrose, (1989), **The Emperor's new mind**, Oxford University Press, Oxford.
S. C. Reddy and L. Trefethen, (1994), *Pseudospectra of the convection-diffusion operator*, SIAM J. Appl. Math., **54**, 1634–1649.
A. Sard, (1942), *The measure of the critical values of differentiable maps*, Bull. AMS, **48**, 883–896.
H. G. Schuster, (1989), **Deterministic Chaos. An introduction**, VCH, Weinheim.
I. Stewart, (1993), *1 est plus probable que 9*, Pour la science, **190**, 94–98.
P. R. Turner, (1982), *The distribution of leading significant digits*, IMA J. Numer. Anal., **2**, 407–412.
E. Wigner, (1960), *The unreasonable effectiveness of mathematics in the natural sciences*, Comm. pure appl. Math., **13**, 1–14.
C. Wolf, (1989), *Discrete time effects and the spectrum of hydrogen*, Eur. J. Phys., **10**, 197–199.