Growth and Decay of Random Fibonacci Sequences
Author(s): Mark Embree and Lloyd N. Trefethen
Source: *Proceedings: Mathematical, Physical and Engineering Sciences*, Vol. 455, No. 1987 (Jul. 8, 1999), pp. 2471-2485
Published by: The Royal Society
Stable URL: http://www.jstor.org/stable/53482
Accessed: 21/06/2011 04:10

THE ROYAL SOCIETY

Growth and decay of random Fibonacci sequences

BY MARK EMBREE AND LLOYD N. TREFETHEN

Oxford University Computing Laboratory, Wolfson Building, Parks Road, Oxford OX1 3QD, UK

Received 18 September 1998; accepted 7 December 1998

For $0 < \beta < \beta^* \approx 0.702\,58$, solutions to the random recurrence $x_{n+1} = x_n \pm \beta x_{n-1}$ decay exponentially as $n \to \infty$ with probability one, whereas for $\beta > \beta^*$, they grow exponentially. By formulating the problem as a Markov chain involving random matrix products and computing its invariant measure—a fractal—the Lyapunov constant $\sigma(\beta) = \lim_{n\to\infty} |x_n|^{1/n}$ is determined numerically for a wide range of values β, and its dependence on β is observed to be non-smooth. (The limit is defined in the almost sure sense.) This generalizes recent work of Viswanath, who proved $\sigma(1) = 1.131\,988\,24\ldots$. By a simple rescaling, these results also apply to the more general random recurrence $x_{n+1} = \alpha x_n \pm \beta x_{n-1}$ for fixed α and β. These random recurrence relations have links with many fields, including ergodic theory, dynamical systems, heavy-tailed statistics, spectral theory, continued fractions, and condensed matter physics.

Keywords: Fibonacci sequence; Lyapunov constant; Markov chain; random matrix products; random recurrences; Viswanath's constant

1. Introduction

In a remarkable paper, Viswanath (1999) considers the large n behaviour of solutions to the 'random Fibonacci recurrence',

$$x_{n+1} = \pm x_n \pm x_{n-1}, \tag{1.1}$$

where the signs are chosen independently and with equal probabilities, and $x_0 = x_1 = 1$. Computer experiments, as in figure 1, show exponential growth with n. The problem of large n behaviour of (1.1) has been mentioned at least since 1963, when Furstenberg (1963) established exponential growth with probability 1, but Viswanath's contribution represents an intriguing new development. By an ingenious application of a Stern–Brocot tree (Graham *et al.* 1994), he proves that solutions to (1.1) satisfy

$$\lim_{n\to\infty} |x_n|^{1/n} = C \quad \text{a.s.} \tag{1.2}$$

with $C = 1.131\,988\,24\ldots$, a number that we propose to call *Viswanath's constant.* (The abbreviation 'a.s.' stands for *almost surely*, which means that any individual sequence satisfies the property with probability 1.) This growth constant is 0.4% greater than the fourth root of the familiar constant $(\sqrt{5}+1)/2 = 1.618\,033\,98\ldots$ for the standard Fibonacci sequence.

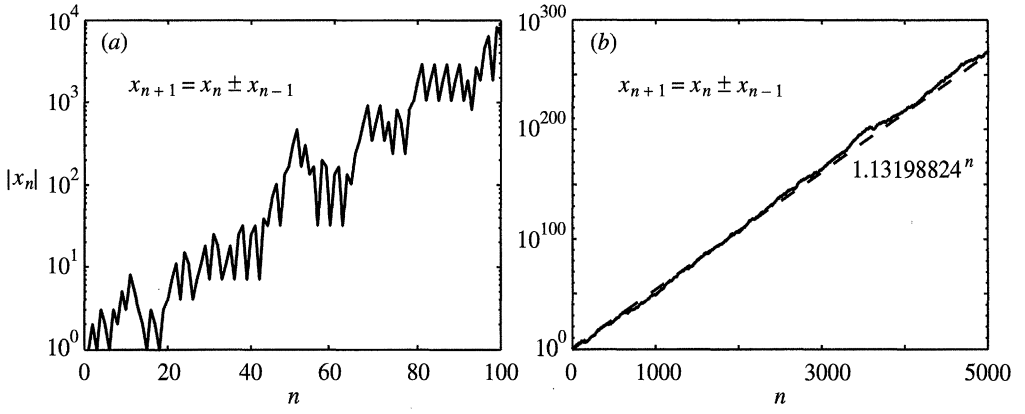

Figure 1. Random Fibonacci sequences, 100 (a) and 5000 (b) iterations. The dashed line represents exponential growth at the rate given by Viswanath's constant, $\sigma(1) = 1.131\,988\,24\ldots$.

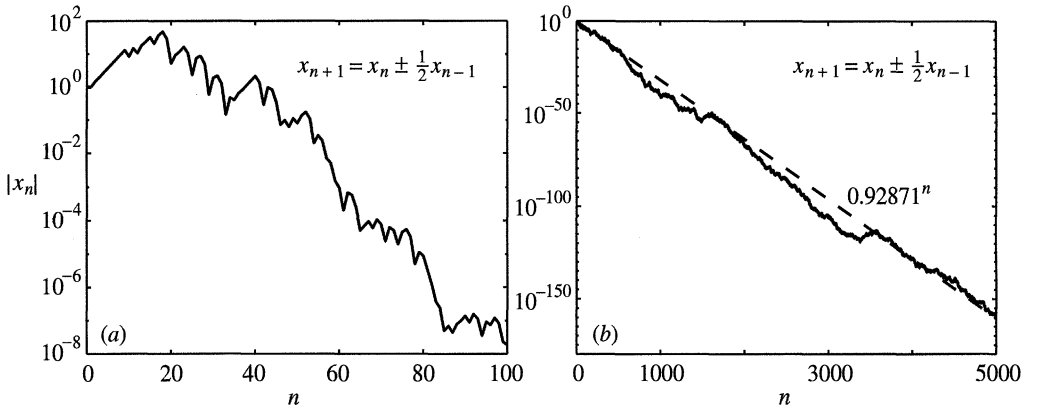

Figure 2. Solutions to the modified recurrence $x_{n+1} = x_n \pm \frac{1}{2}x_{n-1}$, 100 iterations and 5000 iterations. The dashed line represents exponential decay at the rate $\sigma(1/2) \approx 0.929$.

Exponential growth of solutions to (1.1) may seem unsurprising, but a seemingly similar problem gives exponential decay. Figure 2 is the analogue of figure 1 for the random recurrence,

$$x_{n+1} = \pm x_n \pm \tfrac{1}{2}x_{n-1}. \tag{1.3}$$

Now, we observe decay at a rate approximately 0.929^n! This decay occurs even though the variances of the random variables x_n increase to ∞ with n: the random signs ensure that the terms $\pm x_n$ and $\pm\frac{1}{2}x_{n-1}$ of (1.3) are uncorrelated, even though they are not independent, and thus the expected values of the squares x_n^2 grow exponentially according to $\langle x_{n+1}^2\rangle = \langle x_n^2\rangle + \frac{1}{4}\langle x_{n-1}^2\rangle$. In other words, we have here an example of what is known in statistics as a *heavy-tailed distribution.*

The decay of solutions to (1.3) can be explained heuristically. Suppose a solution $\{x_n\}$ were of constant size, $x_{n-1} = x_n = 1$. Then the new value x_{n+1} would be $3/2$ or $1/2$ with equal probability. The geometric mean of these numbers, $\sqrt{3/2} = 0.866\ldots$, is less than 1, and this suggests that exponential decay is to be expected after all.

Such an argument is not rigorous, of course, and it does not yield the actual constant $C \approx 0.929$.

Having realized that solutions to (1.1) grow while solutions to (1.3) decay, we asked, what about the general random recurrence

$$x_{n+1} = \pm x_n \pm \beta x_{n-1}? \tag{1.4}$$

How does the growth rate of solutions $|x_n|$, as defined by the limit (1.2), depend on β? We call this almost sure limit $\sigma(\beta)$, the *Lyapunov constant* for (1.5):

$$\sigma(\beta) = \lim_{n\to\infty} |x_n|^{1/n} \quad \text{a.s.}$$

(The existence of $\sigma(\beta)$ follows from results of Furstenberg & Kesten (1960); we shall not give details.) What value $\beta_{\min} \in (0,1)$ maximizes the decay rate? At what value $\beta^* \in (\frac{1}{2},1)$ is the recurrence neutrally stable, so that there is neither exponential growth nor decay?

To answer these questions, we actually study the recurrence,

$$x_{n+1} = x_n \pm \beta x_{n-1}, \tag{1.5}$$

which is more general than (1.4) in the sense that the sequences generated by (1.4) are obtained by randomizing the signs of the sequences generated by (1.5). Thus the symmetric probability distribution for (1.4) can be obtained from the non-symmetric distribution associated with (1.5).

Suppose we modify the recurrence (1.5) to include two constant parameters rather than one,

$$x_{n+1} = \alpha x_n \pm \beta x_{n-1}. \tag{1.6}$$

Despite appearances, (1.6) is no more general than (1.5), for if $\{x_n\}$ is a solution to (1.6), then the sequence defined by $y_n = \alpha^{-n} x_n$ is a solution to (1.5) with β replaced by β/α^2. (We return to this matter in figure 6 below.)

The subject of random recurrence relations has links with many fields, including ergodic theory, dynamical systems, spectral theory, continued fractions, and condensed matter physics. It is also a special case of the larger subject of iterated random functions, recently surveyed by Diaconis & Freedman (1999). Our interest in the recurrence (1.1), and then in its generalizations (1.5) and (1.6), was prompted originally by questions of spectra and pseudo-spectra for random matrices and operators (Trefethen 1997), which we shall address in a separate publication.

2. Fractal dependence on β

We computed answers to the questions raised in the introduction by two methods. One approach, the elementary one, is to do careful Monte Carlo experiments with long runs of the recurrence so as to measure growth rates directly (§ 3). The other, following Viswanath and an extensive literature on the theory of random products of matrices, is to view the recurrence as a Markov chain on a certain one-dimensional state space and then approximate the associated invariant measure by discretization (§ 4). Viswanath found that for $\beta = 1$, an exact solution can be obtained. For general β, we are unaware of an exact solution, so our results are numerical. We do not have a proof of their accuracy, but careful varying of parameters and comparison of our

Table 1. *Lyapunov constants* $\lim_{n\to\infty} |x_n|^{1/n}$ *(a.s.) for solutions to* $x_{n+1} = x_n \pm \beta x_{n-1}$ *for various* β

(The numbers are numerically computed but are believed to be accurate to all the digits listed.)

β	Lyapunov constant $\sigma(\beta)$	β	Lyapunov constant $\sigma(\beta)$
0	1	1	1.131 988 24 ...
1/128	0.999 97	1.1	1.177 10
1/64	0.999 88	1.2	1.218 96
1/32	0.999 51	1.3	1.258 52
1/16	0.998 02	1.4	1.296 45
1/8	0.991 75	1.5	1.333 38
1/4	0.957 97	2	1.504 76
1/2	0.928 71	4	2.060 82
0.6	0.970 17	8	2.870 07
0.7	0.999 08	16	4.028 98
0.702 58 ...	1	32	5.677 16
0.8	1.041 88	64	8.014 33
0.9	1.087 75	128	11.323 82

independent methods of calculation gives us high confidence that the results listed in table 1 are correct to the digits listed.

Our best estimate (numerical, not mathematically proved) of the value β^* at which no exponential growth or decay occurs (Bougerol 1986) is

$$\beta^* \in (0.702\,582, 0.702\,585).$$

Figure 3 repeats the last two figures for $x_{n+1} = x_n \pm \beta^* x_{n-1}$. Here, in contrast to the other cases, performing the calculation a second time with new random numbers would yield a curve with an entirely different appearance. (Try it!)

Our best estimate of the value at which maximal decay occurs is

$$\beta_{\min} \in (0.367\,470, 0.367\,475),$$

with corresponding Lyapunov constant

$$\sigma(\beta_{\min}) \in (0.895\,168, 0.895\,174).$$

A notable result of our computations is that the dependence of $\sigma(\beta)$ on β is not smooth. Figure 4 summarizes this dependence, suggesting that the curve $\sigma(\beta)$ is in fact a fractal. Figure 5 confirms this hypothesis by zooming in on the curve to successively finer scales. It is evident that to the limits of our computational resolution, the irregularity of the curve does not go away.

Figure 6 portrays the results of figure 4 in the form of a stability region in parameter space for the two-parameter recurrence (1.6). The boundary of the stability region is not smooth.

In the past, very few numerical calculations for problems of this kind seem to have been performed. We are unaware of any numerical estimates of the asymptotic behaviour of (1.1), let alone (1.5) or (1.6), before Viswanath (1999). An example of Halperin (1967), however, shows that non-smooth dependence on parameters is sometimes to be expected. That $\sigma(\beta)$ is at least Hölder continuous is a consequence of theorem 1 in a paper by Le Page (1989).

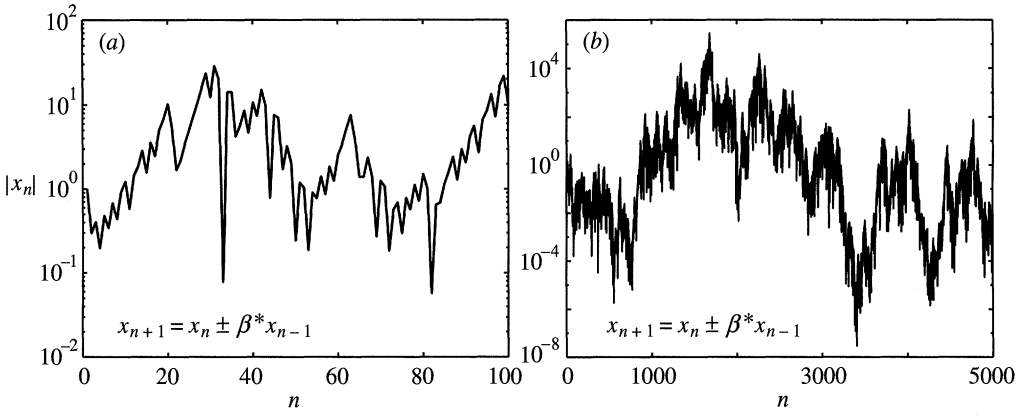

Figure 3. Solutions to the neutrally stable recurrence $x_{n+1} = x_n \pm \beta^* x_{n-1}$ ($\beta^* \approx 0.702\,58$), 100 iterations and 5000 iterations. For this special value of β, no exponential growth or decay occurs.

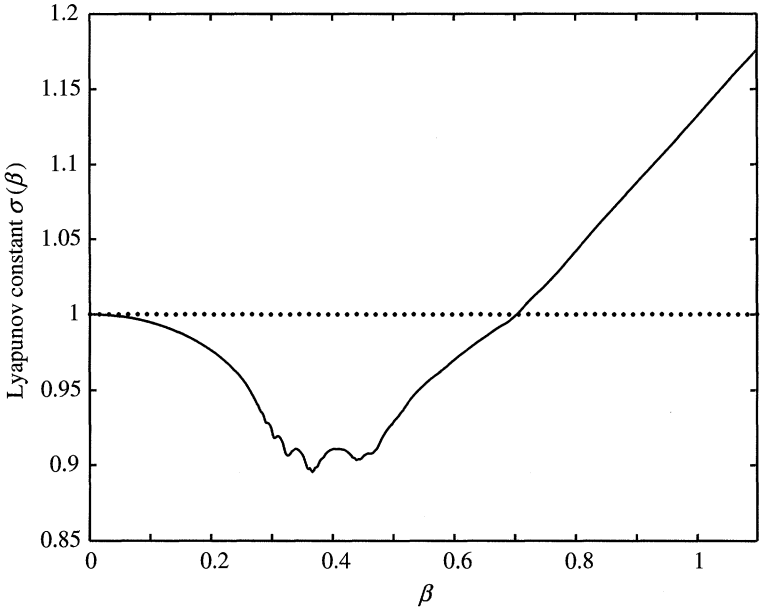

Figure 4. Lyapunov constants for random Fibonacci sequences generated by $x_{n+1} = x_n \pm \beta x_{n-1}$. The curve is believed to be correct to plotting accuracy. See figure 5 for more detail.

3. Monte Carlo calculations

Of course, we began our investigations with Monte Carlo experimentation. The recurrence (1.5) can be expressed in a few lines of any programming language. Indeed, we think the problem of estimating $\sigma(1)$ or β^* on a computer might be appealing for secondary school students.

To calculate $\sigma(\beta)$ by this means to several significant figures, measures must be taken to avoid overflow and underflow. We begin the recurrence in the natural fashion. When the iterates pass an overflow or underflow threshold (we used $|x_n| > 10^{200}$

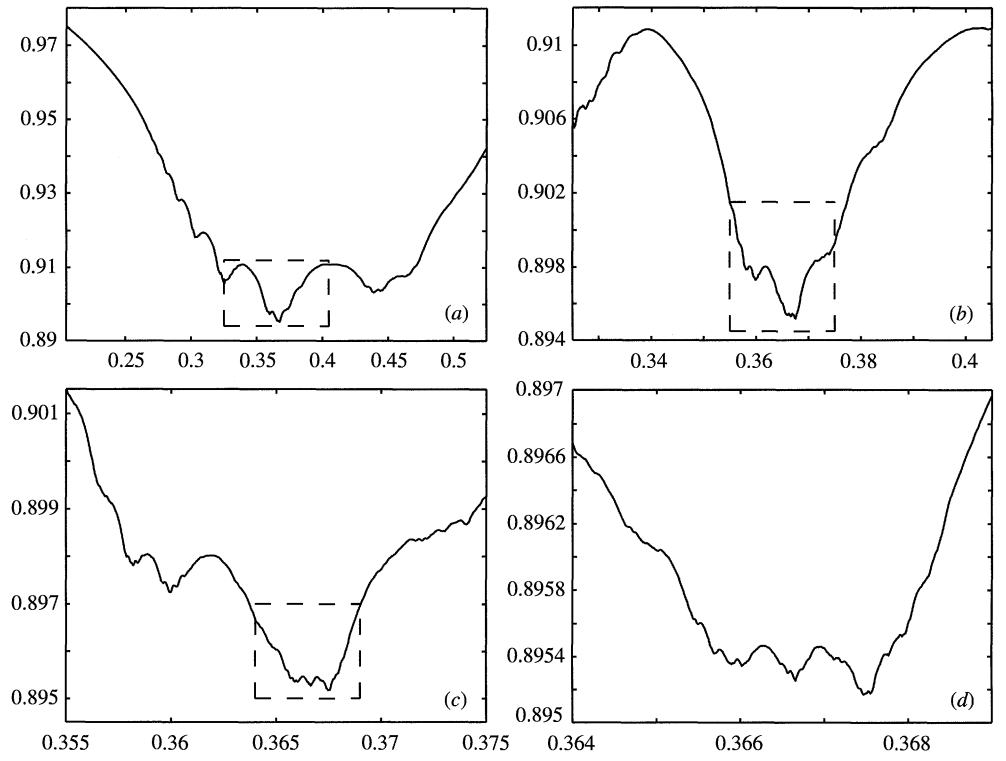

Figure 5. Repetition of figure 4 with β sampled at increasingly finer scales. Each successive plot represents one-quarter of the range of β of the previous one, as indicated by the dashed boxes. For each plot, growth rates for 500 values of β were calculated, and again, the data are believed to be correct to plotting accuracy. Approximately 10^{11} floating point operations went into the preparation of these four plots.

and $|x_n| < 10^{-200}$), we take note of it and restart the recurrence with rescaled values. The total growth at the end of a calculation can then be determined from the current value, x_n, and the number of times the overflow/underflow condition has occurred. This is only one of any number of methods that could be used to scale this computation. The iterates could be rescaled at every step, for example, at a slight extra cost, or at every tenth or hundredth step.

Our rescaling procedure introduces no additional errors into the computation beyond the inevitable rounding errors introduced at every step. Since the machine precision is of the order of 10^{-16}, however, whereas we never take more than 10^9 steps in a run of the recurrence, it is safe to regard the effects of rounding errors as negligible, far smaller than the effect of changing a single sign in the random sequence. One might more justifiably worry about the possibility of biases in the random number generator, and this is one of the reasons why we consider a different computational approach in the next section.

As is typical of such calculations, our Monte Carlo procedure works excellently for obtaining results to low accuracy, but is unsuitable for high accuracy. Experiments confirm that the accuracy of the constants $\sigma(\beta)$ computed by this method is $O(N^{-1/2})$, where N is the length of the sequence. Thus the computational work

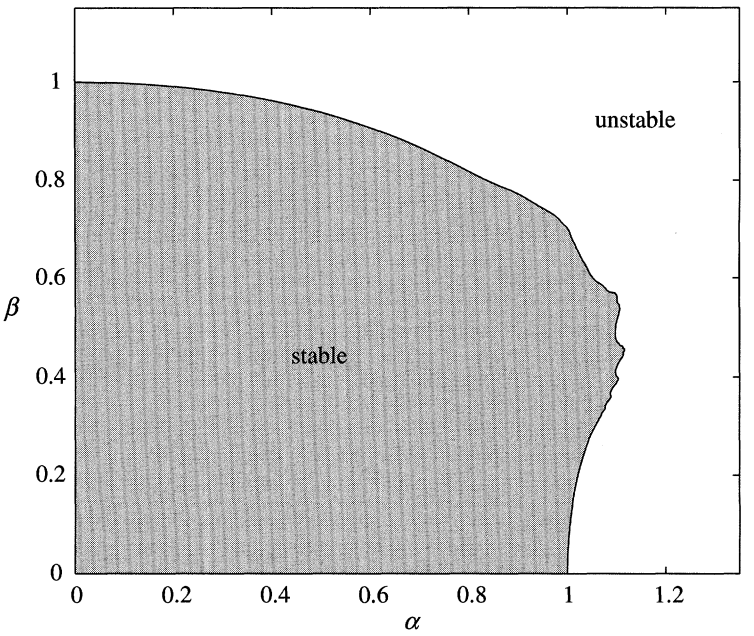

Figure 6. Stability region for the random recurrence $x_{n+1} = \alpha x_n \pm \beta x_{n-1}$. For values of α and β in the shaded region, the Lyapunov constant σ is less than 1, i.e. $|x_n| \to 0$ a.s. as $n \to \infty$.

scales as the square of the number of digits obtained. Some data to this effect are summarized in figure 8, at the end of the next section. On a workstation, we get about four digits of accuracy in a minute of computing.

4. Markov chain calculations and random matrix products

To achieve higher accuracy and to confirm the validity of our Monte Carlo calculations, we turned to an entirely different and mathematically more revealing method modelled on the work of Viswanath (1999). This is to view the random recurrence (1.5) as a Markov chain whose state space consists of normalized vectors $(x_n, x_{n+1})^{\mathrm{T}}$, and to estimate the invariant measure for this Markov chain by discretization. For general information on Markov chains, see Feller (1968), Meyn & Tweedie (1993) and Norris (1997). Existence and uniqueness of an invariant measure in this case is guaranteed by theorems presented by Bougerol (Bougerol & Lacroix 1984).

The non-stochastic recurrence $x_{n+1} = x_n + \beta x_{n-1}$ can be written in the vector form

$$\begin{pmatrix} x_n \\ x_{n+1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ \beta & 1 \end{pmatrix} \begin{pmatrix} x_{n-1} \\ x_n \end{pmatrix},$$

with solutions growing asymptotically at the rate given by the dominant eigenvalue of the matrix. For the stochastic recurrence (1.5), we can work with the same vectors, but the matrix applied at each step is now one of the two possibilities

$$A = \begin{pmatrix} 0 & 1 \\ \beta & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} 0 & 1 \\ -\beta & 1 \end{pmatrix}$$

chosen at random. The recurrence becomes

$$\begin{pmatrix} x_n \\ x_{n+1} \end{pmatrix} = M_n M_{n-1} \cdots M_1 \begin{pmatrix} x_0 \\ x_1 \end{pmatrix},$$

where each M_i is either A or B with equal probability. The growth rates are not governed by the eigenvalues of A and B, but rather by the asymptotic behaviour of the product $M_n M_{n-1} \cdots M_1$, a quantity which cannot be computed in an elementary fashion. Thus the random Fibonacci sequence is a problem in *random matrix products*, a subject whose study was initiated by Bellman (1954). Other notable papers in the field of random matrix products include those by Furstenberg & Kesten (1960), Furstenberg & Kifer (1983), Guivarc'h & Raugi (1985), Le Page (1989), and an excellent introduction to this subject is the lengthy review by Bougerol that constitutes the first half of Bougerol & Lacroix (1984).

To analyse this process as a Markov chain, we first reduce the two-dimensional state space to one dimension by normalizing, so that we are effectively dealing with the projective line. One way to normalize $(x_n, x_{n+1})^{\mathrm{T}}$ is to consider the angle between successive iterates ($x = (\cos\theta, \sin\theta)^{\mathrm{T}}$, $\theta \in (-\pi/2, \pi/2)$). Another is to consider the slope ($x = (1, m)^{\mathrm{T}}$, $m = \tan\theta$). The angles are the simplest conceptually, and the slopes are the simplest algebraically, so we find it convenient to keep both formulations in mind and pass between them as necessary, making no distinction between the two in notation. The essential question becomes, what fraction of the time does the random recurrence spend in the state corresponding to each angle θ or slope m?

This formulation suggests that we look for a probability density function $p(\theta)$ or $p(m)$ that is invariant with respect to the steps of the random recurrence. This is essentially the right idea, but we must be a little more careful, for such a $p(\theta)$ or $p(m)$, defined pointwise, does not exist. What does exist is an invariant measure ν—not absolutely continuous—that would correspond to the integral of p. For any interval $I = [\theta_1, \theta_2]$ or $[m_1, m_2]$, $\nu(I)$ is equal to the proportion of the time, in the limit $n \to \infty$, that the Markov process spends in the corresponding interval. More generally, ν is a Borel measure defined on Borel subsets of $[-\pi/2, \pi/2]$ or $\mathbb{R}$.

At any point in the Markov process, the current state x was reached by taking one step either from $A^{-1}x$ or from $B^{-1}x$. The invariant measure ν must accordingly have the property

$$\nu(S) = \tfrac{1}{2}\nu(A^{-1}S) + \tfrac{1}{2}\nu(B^{-1}S).$$

In terms of slopes, the two inverse maps are

$$A^{-1}m = \frac{\beta}{m-1}, \qquad B^{-1}m = \frac{\beta}{1-m}.$$

It follows that for slope intervals $S = [m_1, m_2]$, the invariance condition is

$$\nu([m_1, m_2]) = \tfrac{1}{2}\,\nu\!\left(\frac{\beta}{[m_1, m_2] - 1}\right) + \tfrac{1}{2}\,\nu\!\left(\frac{\beta}{1 - [m_1, m_2]}\right). \tag{4.1}$$

(This notation means that the measure of the interval $[m_1, m_2]$ is the mean of the measures of the two sets $\{\beta/(m-1) : m \in [m_1, m_2]\}$ and $\{\beta/(1-m) : m \in [m_1, m_2]\}$.) This simple equation, distantly analogous to what Strang calls a *dilation equation* in the study of wavelets (Strang 1989), is the heart of the analysis of our random recurrence. Our numerical problem is to compute approximations to the measure ν that satisfies it.

Viswanath, working from the recurrence $x_{n+1} = \pm x_n + x_{n-1}$, makes the observation that the inverse maps take intervals in a Stern–Brocot tree to other intervals in the same tree, from which he deduces an exact expression for $\nu([m_1, m_2])$. The presence of $\beta \neq 1$ in the numerators of (4.1) apparently precludes such a clean solution in the general case. Instead, we numerically approximate $\nu([m_1, m_2])$ using piecewise constant functions. Specifically, we subdivide $[-\pi/2, \pi/2]$ into N equally spaced angular intervals, $I_1, \ldots, I_N$. Converting the angular intervals into slope intervals, we apply the inverse maps, which are then converted back into angles. In terms of angular intervals, equation (4.1) takes the form

$$\nu(I_j) = \tfrac{1}{2}\,\nu\left(\tan^{-1}\left(\frac{\beta}{\tan(I_j) - 1}\right)\right) + \tfrac{1}{2}\,\nu\left(\tan^{-1}\left(\frac{\beta}{1 - \tan(I_j)}\right)\right). \tag{4.2}$$

We approximate this equation on the discrete set of intervals $I_1, \ldots, I_N$ by the equation

$$\nu_j = \frac{1}{2}\sum_{k=1}^{N} c_k \nu_k,$$

where ν_j is our discrete approximation to $\nu(I_j)$. Here each coefficient c_k, a number between 0 and 2, represents the amount that I_k overlaps with each of the intervals on the right-hand side of (4.2). These N equations together form a linear system of rank $N - 1$, which can be made consistent by replacing the Nth equation by the conservation condition $\sum_{j=1}^{N} \nu_j = 1$.

The system of equations just described is sparse and non-symmetric. Direct sparse matrix techniques can be used to solve it, but our experiments with the built-in sparse solvers provided in MATLAB® indicate that such methods require $O(N^3)$ operations and $O(N^2)$ storage. This is a prohibitive amount of work, an order of magnitude greater than for the Monte Carlo method, and the memory requirements would preclude treatment of large values of N. On the other hand, we found that we could keep the storage under control and beat the Monte Carlo method in speed by applying a non-symmetric iterative solver, Bi-CGSTAB (Barrett *et al.* 1994; Greenbaum 1997; Van der Vorst 1992). As is typical of Krylov subspace methods, Bi-CGSTAB approximately solves the linear system using only matrix-vector multiplications, inner products, and vector additions. The matrix-vector multiplication can be implemented using a sparse matrix data structure with a fixed number of non-zeros per row. Though our coefficient matrix is not normal, plots of pseudospectra (Trefethen 1997) for low-dimensional instances suggest that the degree of non-normality is limited. The eigenvalues are well clustered, and the condition number grows linearly, $\kappa_2 = O(N)$. Bi-CGSTAB solves the system quickly, even without preconditioning (though perhaps the convergence could be accelerated by a coarse-grid preconditioner), and the number of iterations does not grow very much as we refine the mesh, being 40 or 50 for typical calculations to six or seven digits of accuracy. In the end, our Markov chain algorithm produces accurate estimates of the parameter $\sigma(\beta)$ in an amount of time that grows approximately linearly with N.

Figure 7 depicts invariant measures computed in this manner for four values of β. The curves look like plots of $p(\theta)$, the probability density function that does not exist. In actuality, they are histograms of values $\{\nu_k\}$ based on a subdivision of $[-\pi/2, \pi/2]$ into $N = 256$ equal subintervals. These histograms are well defined, but they do not converge to continuous curves as the discretization is refined. For smaller values of

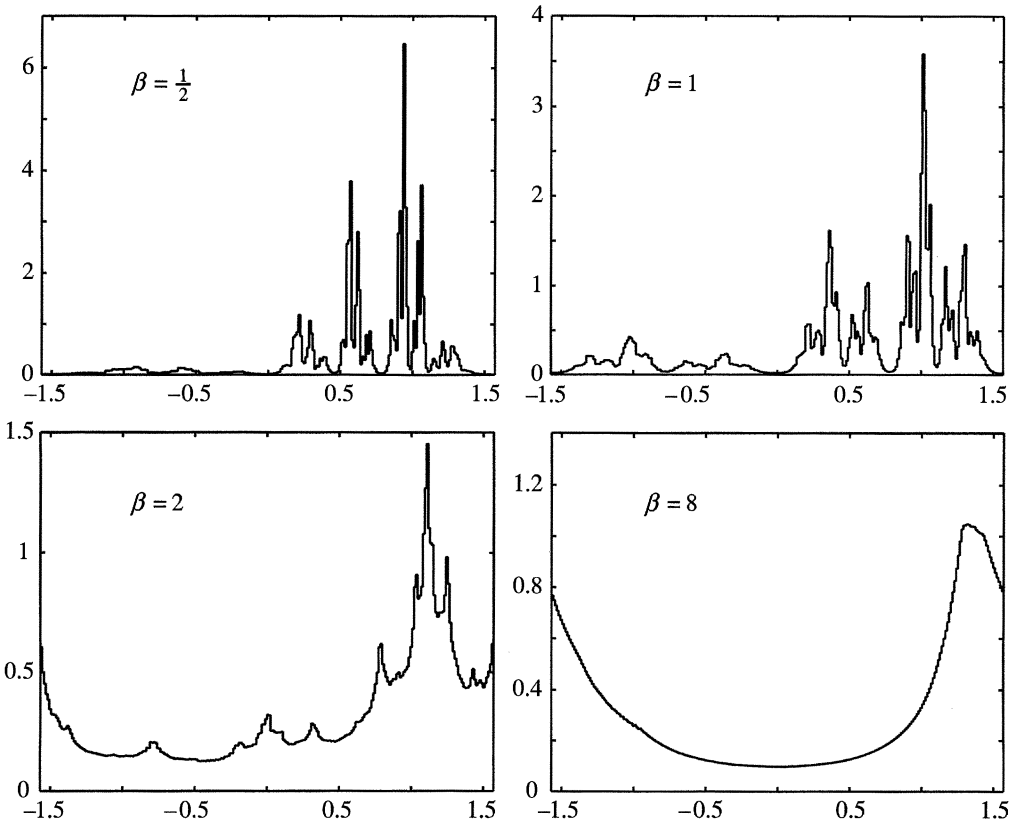

Figure 7. Discrete histogram approximations $\{\nu_k\}$ to the invariant measure ν using $N = 256$ intervals on $[-\pi/2, \pi/2]$, for four values of β. For small β, most of the measure is concentrated in $\theta > 0$: sign changes in the recurrence are exponentially rare (see figure 9).

β, the fractal nature of the invariant measure is strikingly apparent in the figure. For larger β, the measure is smoother.

The histograms of figure 7 were obtained by solving linear systems of size $N = 256$. From such approximations one can obtain Lyapunov constants accurate to about two digits. For the six-digit results reported in table 1, we increased N to 2^{21}. This could be done via a matrix of dimension 2^{21}, but instead, following Viswanath, we used a matrix of dimension 2^{20} by taking advantage of a symmetry property. Since only the behaviour of $|x_n|$ is being measured, one can make use of the recurrence (1.4) with two random signs, thereby attaining a probability distribution that is even with respect to θ, which is then computed only for values $\theta > 0$. In the end, the largest matrix we deal with is of dimension $2^{20} = 1\,048\,576$, with approximately 10^7 non-zero entries, filling 100 megabytes of memory on our Sun Ultra 30 workstation. (The precise figures depend on β.) Each matrix-vector multiplication in the Bi-CGSTAB iteration takes about 1 s.

This completes our explanation of how the invariant measure ν is computed. It remains to describe how we get from ν to the Lyapunov constant $\sigma(\beta)$. The answer is that once ν is known, $\sigma(\beta)$ can be obtained by an integral due originally to Fursten-

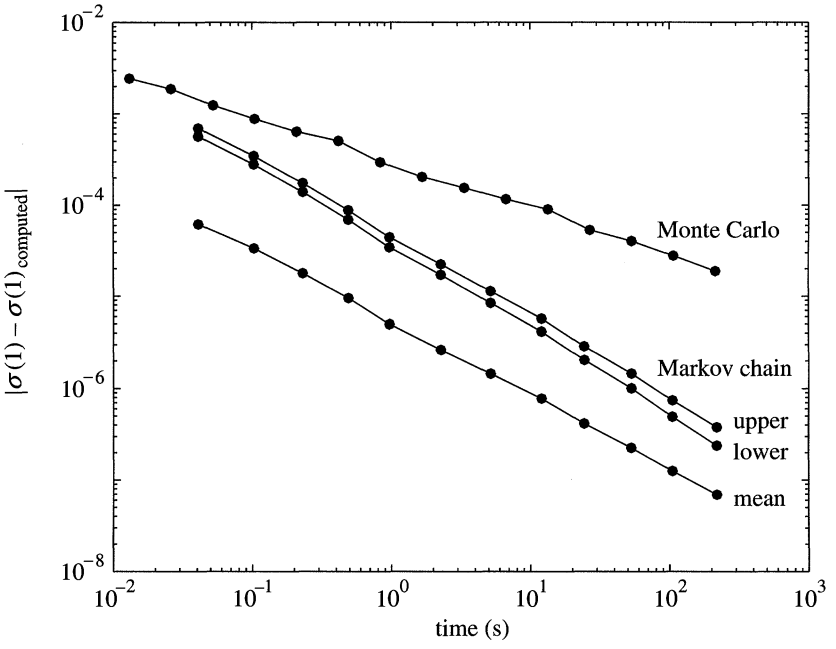

Figure 8. Computation times for the Monte Carlo and Markov chain methods for determining the number $\sigma(1)$. The x-axis represents elapsed time on a Sun Ultra 30 workstation, and the errors represented on the y-axis were computed by taking advantage of Viswanath's eight-digit result $\sigma(1) = 1.131\,988\,24\ldots$. For higher-accuracy computations, the Monte Carlo method cannot compete.

berg and sometimes known as Thouless's formula,

$$\log(\sigma(\beta)) = \int \operatorname{amp}(x)\,\mathrm{d}\nu,$$

where $\operatorname{amp}(x)$ denotes the amplification factor associated with the state x,

$$\operatorname{amp}(x) = \frac{1}{2}\log\frac{\|Ax\|_2}{\|x\|_2} + \frac{1}{2}\log\frac{\|Bx\|_2}{\|x\|_2} = \frac{1}{4}\log\left(\frac{\beta^4 + 4m^4}{(1+m^2)^2}\right). \tag{4.3}$$

By exploiting the monotonicity of this function in most subintervals, we calculate approximate upper and lower bounds for $\sigma(\beta)$. As we refine the size of the mesh used to approximate ν, experiments show that these upper and lower estimates converge linearly to values consistent with our Monte Carlo experiments. Even for relatively low accuracies, the Markov chain solution is faster. Figure 8 illustrates these convergence properties.

5. Discussion and generalizations

The complex behaviour of our simple recurrence suggests numerous problems for further investigation. We shall mention a few of them.

A careful study of the fractal nature of the curve $\sigma(\beta)$ would be interesting. For example, what can be said about the fractal dimension (Falconer 1990)? Does the

irregularity in figure 4 diminish only gradually outside the region near $\beta = 0.4$, or does the curve become genuinely smooth for certain β?

Rigorous analysis of $\sigma(\beta)$ would probably be feasible in the limits $\beta \to 0$ and $\beta \to \infty$. Consideration of the geometric mean of $1-\beta$ and $1+\beta$ suggests the asymptotic behaviour $\sigma(\beta) \sim 1 - \frac{1}{2}\beta^2$ as $\beta \to 0$, convincingly matched by the results of table 1, and the next term in the series appears to be about $-1.75\beta^4$. A similar argument suggests $\sigma(\beta) \sim \beta^{1/2}$ as $\beta \to \infty$, also confirmed by the data, and the next term in that series appears to be about $0.114\beta^{-1/2}$.

Precise statements could be made about the statistical aspects of the random variables x_n, $|x_n|$, $|x_n|^{1/n}$, or $n^{-1}\log|x_n|$. According to results developed over the years by Furstenberg, Tutubalin, Virtser, Le Page, Guivarc'h, Goldsheid, and others, the distributions of $n^{-1}\log|x_n|$ and $|x_n|^{1/n}$ converge to Gaussians of standard deviation $O(n^{-1/2})$ as $n \to \infty$. (If one distribution converges to a Gaussian, so must the other, since the standard deviations are diminishing to zero.) This is a kind of 'central limit theorem for random matrix products' (see Bougerol & Lacroix (1984), Goldsheid & Guivarc'h (1996), Guivarc'h & Raugi (1986) and the references therein).

Another question suggests itself that might be of physical interest in certain applications. In figure 7 it is apparent that for large β, a considerable amount of measure is associated with values $\theta < 0$, corresponding to sign changes in the random sequence. For small β, however, sign changes become very rare. Let us define the *sign-flip frequency* $f(\beta)$ to be the proportion of values x_n with $x_n x_{n-1} < 1$ in the limit $n \to \infty$. Figure 9 indicates the exponential decrease of the sign-flip frequency as β reduces to 1/4. Don Coppersmith of IBM T. J. Watson Research Centre pointed out the significance of the value 1/4 to us and provided the remarkably accurate estimate

$$f(\beta) \approx 2^{-\pi/(2\sqrt{\beta-1/4})}, \tag{5.1}$$

based on counting how many successive fortuitous sign choices are required to achieve a sign flip. Sign-flip frequencies for other related problems have been considered previously by Simon & Taylor (1985).

The above remarks concern further investigations that might be made of our stochastic recurrence problem as formulated in (1.5). In addition, numerous ideas for modification of the problem suggest themselves.

The coin might be weighted, so that $+$ is chosen with probability p and $-$ with probability $1-p$. In the limit $p \to 1$ we recover the non-stochastic recurrence relation $x_{n+1} = x_n + \beta x_{n-1}$.

The coefficients $\{-1, 1\}$ in the recurrence $x_{n+1} = \pm x_n \pm \beta x_{n-1}$ might be replaced by, for example, standard normally distributed random numbers or uniform random numbers in $[-1, 1]$. (For definiteness, we now choose the variant with two random coefficients rather than one.) In the former case, since the coefficients are what are known as *stable* random variables (Feller 1971), an analytic solution can probably be obtained by using the methods of Viswanath & Trefethen (1998); for $\beta = 1$ and $\beta = 1/2$ our numerical experiments give growth rates of approximately 0.995 and 0.797, respectively. In the latter case, perhaps no closed form solution is possible; numerical experiments suggest results for $\beta = 1$ and $\beta = 1/2$ of approximately 0.7499 and 0.5804. For problems like these involving a continuum of matrices governed by a probability density function rather than just a measure, theorem 2 of Le Page (1989) implies that the dependence of $\sigma(\beta)$ on β is C^∞.

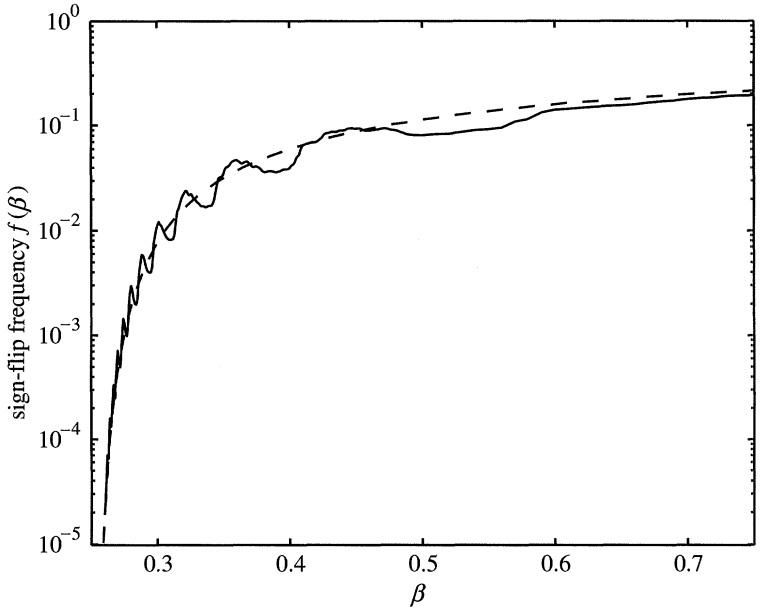

Figure 9. Sign-flip frequency as a function of β. The frequency is 0 for $\beta \leqslant 1/4$. The dashed line indicates the approximation (5.1) derived by Coppersmith.

The length of the recurrence might be increased. In our eyes the most interesting such generalization is the triangular system of recurrences of growing length

$$\left.\begin{aligned} x_1 &= \pm \alpha x_0, \\ x_2 &= \pm \alpha x_1 \pm \alpha^2 x_0, \\ x_3 &= \pm \alpha x_2 \pm \alpha^2 x_1 \pm \alpha^3 x_0, \\ &\vdots \end{aligned}\right\} \tag{5.2}$$

where α is a parameter. For $\alpha = 1$, we obtain the system of recurrences studied in Viswanath & Trefethen (1998), except with $\{-1, 1\}$ rather than normally distributed coefficients; computations show that the growth rate is approximately 1.3272. Other values of α are mathematically equivalent, via a rescaling $y_n = \alpha^{-n} x_n$, as in the remark following (1.6). In particular, for $\alpha = \alpha^* \approx 1/1.3272 \approx 0.7535$, the system (5.2) is neutrally stable. A matrix version of this observation is that if A_n is an $n \times n$ triangular matrix with 1 on the diagonal and random entries $\pm\alpha^j$ on the jth subdiagonal, then $\|A_n^{-1}\|$ increases exponentially as $n \to \infty$ for $\alpha > \alpha^*$, but only increases very slowly for $\alpha < \alpha^*$. For $\alpha < \frac{1}{2}$, there is no growth at all.

One might take the distribution of coefficients to be complex. If $\{-\beta, \beta\}$ in (1.5) is replaced by the complex circle $\{e^{i\theta}\beta\}$ with a uniform probability distribution, for example, then the dependence of σ on β is as shown in figure 10; the Lyapunov constants for $\beta = 1/2$ and $\beta = 1$ are approximately 1.02 and 1.19, respectively. Now there is no decay for any value of β. A heuristic explanation for this is that as a consequence of Cauchy's theorem, the geometric mean of the complex numbers on the circle $1 + \beta e^{i\theta}$ is equal to 1, not smaller. The curve $\sigma(\beta)$ is smooth, a fact that again follows from theorem 2 of Le Page (1989).

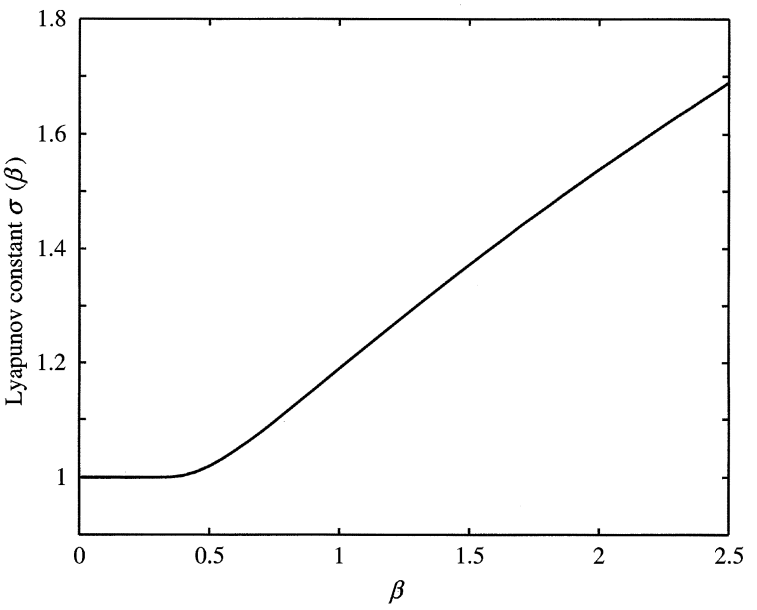

Figure 10. Analogue of figure 4 for the complex recurrence $x_{n+1} = x_n + e^{i\theta}\beta x_{n-1}$, $\theta \in [-\pi, \pi]$, computed via Monte Carlo experiments. Now the curve is smooth, and there is no decay for small β.

Finally, of course, it would be interesting to pursue the implications of the behaviour of our random recurrences in various fields of application, such as those mentioned in the abstract.

We are grateful to Divakar Viswanath for motivating this work, for teaching us about random matrix products, and for detailed remarks on a draft manuscript, including pointers to key items in the literature. For advice and suggestions of various kinds we thank Peter Donnelly, Charles Goldie, Ilya Goldsheid, Des Higham, Sam Howison, Gilbert Strang, Herbert Wilf, and Tong Zhang. Don Coppersmith pointed out to us the cut-off of the sign-flip frequency at $\beta = 1/4$ and provided the estimate (5.1). L.N.T. further acknowledges earlier discussions about the system (5.2) with Persi Diaconis, Harry Kesten, and Andrew Odlyzko.

References

Barrett, R., Berry, M., Chan, T. F., Demmel, J., Donato, J., Dongarra, J., Eijkhout, V., Pozo, R., Romine, C. & Van der Vorst, H. 1994 *Templates for the solution of linear systems: building blocks for iterative methods.* Philadelphia: SIAM.

Bellman, R. 1954 Limit theorems for non-commutative operations. I. *Duke Math. J.* **21**, 491–500.

Bougerol, P. 1986 Oscillation des produits de matrices aléatoires dont l'exposant de Lyapounov est nul. In *Lyapunov exponents* (ed. L. Arnold & V. Wihstutz). Springer Lecture Notes in Mathematics, vol. 1186, pp. 27–36.

Bougerol, P. & Lacroix, J. 1984 *Products of random matrices with applications to infinite-dimensional Schrödinger operators.* Basel: Birkhäuser.

Diaconis, P. & Freedman, D. 1999 Iterated random functions. *SIAM Rev.* **41**, 45–76.

Falconer, K. 1990 *Fractal geometry: mathematical foundations and applications.* Chichester: Wiley.

Feller, W. 1968 *An introduction to probability theory and its applications*, vol. 1, 3rd edn, pp. 372–443. New York: Wiley.

Feller, W. 1971 *An introduction to probability theory and its applications*, vol. 2, 2nd edn, pp. 169–176. New York: Wiley.

Furstenberg, H. 1963 Noncommuting random products. *Trans. Am. Math. Soc.* **108**, 377–428.

Furstenberg, H. & Kesten, H. 1960 Products of random matrices. *Ann. Math. Stat.* **31**, 457–469.

Furstenberg, H. & Kifer, Y. 1983 Random matrix products and measures on projective spaces. *Israel J. Math.* **46**, 12–32.

Goldsheid, I. Y. & Guivarc'h, Y. 1996 Zariski closure and the dimension of the Gaussian law of the product of random matrices, I. *Prob. Theory Relat. Fields* **105**, 109–142.

Graham, R. L., Knuth, D. E., & Patashnik, O. 1994 *Concrete mathematics*, 2nd edn, pp. 116–123. Reading, MA: Addison-Wesley.

Greenbaum, A. 1997 *Iterative methods for solving linear systems.* Philadelphia: SIAM.

Guivarc'h, Y. & Raugi, A. 1985 Frontière de Furstenberg, propriétés de contraction et théoremes de convergence. *Z. Wahrsch. Verw. Gebiete* **69**, 187–242.

Guivarc'h, Y. & Raugi, A. 1986 Products of random matrices: convergence theorems. *Contemporary Math.* **50**, 31–54.

Halperin, B. 1967 Properties of a particle in a one-dimensional random potential. *Adv. Chem. Phys.* **31**, 123–177.

Meyn, S. P. & Tweedie, R. L. 1993 *Markov chains and stochastic stability.* London: Springer.

Le Page, E. 1989 Régularité du plus grand exposant caractéristique des produits de matrices aléatoires indépendantes et applications. *Ann. Inst. Henri Poincaré* **25**, 109–142.

Norris, J. R. 1997 *Markov chains.* Cambridge University Press.

Simon, B. & Taylor, M. 1985 Harmonic analysis on SL(2, R) and smoothness of the density of states in the one-dimensional Anderson model. *Commun. Math. Phys.* **1985**, 1–19.

Strang, G. 1989 Wavelets and dilation equations: a brief introduction. *SIAM Rev.* **31**, 614–627.

Trefethen, L. N. 1997 Pseudospectra of linear operators. *SIAM Rev.* **39**, 383–406.

Van der Vorst, H. 1992 Bi-CGSTAB: a fast and smoothly converging variant of BiCG for the solution of nonsymmetric linear systems. *SIAM J. Sci. Stat. Comput.* **19**, 631–644.

Viswanath, D. 1999 Random Fibonacci sequences and the number 1.131 988 24 .... *Math. Comp.* (In the press.)

Viswanath, D. & Trefethen, L. N. 1998 Condition numbers of random triangular matrices. *SIAM J. Matrix Analysis Applic.* **19**, 564–581.