PROCEEDINGS OF THE ROYAL SOCIETY A

Proc. R. Soc. A (2007) **463**, 2813–2832
doi:10.1098/rspa.2007.0014
Published online 7 August 2007

Padé approximants of random Stieltjes series

By Jens Marklof, Yves Tourigny* and Lech Wołowski

School of Mathematics, University of Bristol, Bristol BS8 1TW, UK

We consider the random continued fraction

$$S(t) := \frac{1}{s_1 + \frac{t}{s_2 + (t/(s_3 + \cdots))}}, \quad t \in \mathbb{C} \backslash \mathbb{R}_-,$$

where s_n are independent random variables with the same gamma distribution. Every realization of the sequence defines a Stieltjes function that can be expressed as

$$S(t) = \int_0^\infty \frac{\sigma(\mathrm{d}x)}{1 + xt}, \quad t \in \mathbb{C} \backslash \mathbb{R}_-,$$

for some measure σ on the positive half-line. We study the convergence of the finite truncations of the continued fraction or, equivalently, of the diagonal Padé approximants of the function S. Using the Dyson–Schmidt method for an equivalent one-dimensional disordered system and the results of Marklof *et al.*, we obtain explicit formulae (in terms of modified Bessel functions) for the almost sure rate of convergence of these approximants, and for the almost sure distribution of their poles.

Keywords: Padé approximation; disordered system; random continued fraction

1. Introduction

Let $\boldsymbol{s} = (s_1, s_2, \ldots)$ be a sequence of positive real numbers and consider the analytic continued fraction

$$S(t) := \frac{1}{s_1 + \frac{t}{s_2 + (t/(s_3 + \cdots))}}, \quad t \in \mathbb{C} \backslash \mathbb{R}_-. \tag{1.1}$$

This continued fraction defines a *Stieltjes function*; it can be represented in integral form as

$$S(t) = \int_0^\infty \frac{\sigma(\mathrm{d}x)}{1 + tx}, \tag{1.2}$$

* Author for correspondence (y.tourigny@bris.ac.uk).

Electronic supplementary material is available at http://dx.doi.org/10.1098/rspa.2007.0014 or via http://www.journals.royalsoc.ac.uk.

Received 2 May 2007
Accepted 5 July 2007

for some measure σ supported on the non-negative half-line such that the *moments*

$$m_n := \int_0^\infty x^n \sigma(\mathrm{d}x), \quad n \in \mathbb{N} \tag{1.3}$$

exist. By an obvious use of the geometric series, every Stieltjes function can be expanded formally in powers of t,

$$S(t) \sim \sum_{j=0}^{\infty} m_j (-t)^j \quad \text{as } t \to 0+. \tag{1.4}$$

Hence, S is the *moment-generating function* of the measure σ. It is a well-known fact of great practical importance that, given the first n of the moments, one may construct the rational function

$$S_n(t) := \frac{P_n(t)}{Q_n(t)} = \frac{1}{s_1 + \frac{t}{s_2 + \cdots + (t/s_n)}}, \quad t \in \mathbb{C} \backslash \mathbb{R}_-, \tag{1.5}$$

where

$$\deg P_n = \begin{cases} (n/2) - 1 & \text{if } n \text{ is even} \\ (n-1)/2 & \text{if } n \text{ is odd} \end{cases} \quad \text{and} \quad \deg Q_n = \begin{cases} n/2 & \text{if } n \text{ is even} \\ (n-1)/2 & \text{if } n \text{ is odd.} \end{cases}$$

This truncation of the continued fraction (1.1) has a MacLaurin expansion whose nth partial sum agrees with that of the series (1.4). Hence, S_n is a diagonal (if n is odd) or near-diagonal (if n is even) *Padé approximant* of S.

Now suppose that S_n are independent positive random variables with the same distribution, say μ. We shall consider the following questions.

(i) What are the almost sure analytic properties of these Stieltjes functions?
(ii) What is the almost sure leading asymptotic behaviour of the error $S(t) - S_n(t)$ as $n \to \infty$?

These questions are of interest because Padé approximation is widely used in applied mathematics as a practical means of accelerating the convergence of the partial sums of series obtained by perturbation methods. As pointed out by Bender & Orszag (1978), the consideration of many particular cases where the S_n are *deterministic* reveals a wide range of large-n behaviours. Our motivation for studying the random case is to gain some insight into the asymptotic behaviour of Padé approximation in the 'generic' case.

Our study of diagonal Padé approximation reduces to aspects of the large-n behaviour of the denominators Q_n. For this reason, as in the deterministic case, the cornerstone of the analysis is the three-term recurrence relation

$$Q_{n+1} = tQ_{n-1} + s_{n+1} Q_n, \quad Q_{-1} = 0, \quad Q_0 = 1. \tag{1.6}$$

(The P_n satisfy the same recurrence relation, albeit with different initial conditions.) This recurrence relation makes a link between Padé approximation and a rich set of other mathematical entities, such as orthogonal polynomials, products of random matrices and discrete Schrödinger-like operators. By exploiting results that are well known in these related fields, one may obtain—

for a very large class of distributions μ of the coefficients S_n—some partial answers to the questions stated earlier. Our contribution in the present paper is to elaborate the particular case where μ is the gamma distribution. More precisely, we obtain explicit formulae (in terms of Bessel functions) for the leading term in the asymptotic behaviour of the error of Padé approximation and for the asymptotic distribution of the poles, as well as the location of the essential spectrum of the measure σ.

In the remainder of this introductory section, we describe briefly the key ideas underlying the analysis. Then we summarize our main results in the form of a theorem.

(*a*) *The moment problem*

The Stieltjes moment problem is, given a sequence $\{m_n\}_{n\in\mathbb{N}}$, to determine whether or not there exists a measure σ such that equation (1.3) holds for every n. Historically, mathematical objects such as the analytic continued fraction (1.1), orthogonal polynomials and Padé approximants were introduced as tools in the study of this moment problem (Akhiezer 1961; Nikishin & Sorokin 1988; Simon 1998). Stieltjes (1894) showed that a necessary and sufficient condition for the *existence* of a measure σ with the prescribed moments is

$$\forall n \in \mathbb{Z}_+, \quad s_n > 0. \tag{1.7}$$

He also showed that

$$\sum_{n=1}^{\infty} s_n = \infty \tag{1.8}$$

is a necessary and sufficient condition for the *uniqueness* of the measure.

Let us assume that condition (1.7) holds and describe in very broad terms one way of 'reconstructing' σ from its moments (see Akhiezer (1961) and Nikishin & Sorokin (1988) for a detailed treatment).

Recall that x' is a *point of increase* of the measure σ if

$$\forall \varepsilon > 0, \quad \int_{\max\{x'-\varepsilon,0\}}^{x'+\varepsilon} \sigma(\mathrm{d}x) > 0.$$

The *spectrum* of σ is the set of its points of increase and will be denoted spec(σ). For $x>0$, we shall denote by δ_x the probability measure on $\mathbb{R}_+$ whose only point of increase is x.

Given the m_n, we may define an inner product, say $(\cdot,\cdot)_{\boldsymbol{m}}$, on the space of polynomials as follows: if p and q are two polynomials with coefficients p_i and q_i, respectively, then

$$(p,q)_{\boldsymbol{m}} := \sum_{i,j} m_{i+j} p_i q_j.$$

Knowing the moments, we may compute s_n and S_n. We remark that P_{2n} and Q_{2n} are polynomials of degree $n-1$ and n, respectively. Set

$$\psi_n(\lambda) := \sqrt{s_{2n+1}}\lambda^n Q_{2n}(-1/\lambda). \tag{1.9}$$

Then the fact that S_{2n} matches the moment-generating series to $O(t^{2n})$ implies that ψ_n is orthogonal to every polynomial of degree less than n, in the sense of the inner product $(\cdot,\cdot)_{\boldsymbol{m}}$. It follows (see Akhiezer 1961, ch. 1) that the roots of ψ_n are simple and lie in $\mathbb{R}_+$; denote them by

$$0 \le \lambda_{n,1} < \lambda_{n,2} < \cdots < \lambda_{n,n} < \infty.$$

Gaussian quadrature then defines a discrete measure

$$\sigma_n := \sum_{j=1}^{n} \sigma_{n,j}\delta_{\lambda_{n,j}}, \quad \text{where} \quad \sigma_{n,j} = \left(\sum_{\ell=0}^{n-1} \psi_\ell^2(\lambda_{n,j})\right)^{-1} \tag{1.10}$$

that converges weakly to a measure σ that solves the moment problem. In particular, the inner product $(\cdot,\cdot)_{\boldsymbol{m}}$ coincides with the inner product in $L^2_\sigma(\mathbb{R}_+)$.

The moment problem can also be approached from the point of view of operator theory. The recurrence relation (1.6) for Q_n implies the following recurrence relation for ψ_n:

$$v_0\psi_0 + h_0\psi_1 = \lambda\psi_0 \quad \text{if } n = 0,$$
$$h_{n-1}\psi_{n-1} + v_n\psi_n + h_n\psi_{n+1} = \lambda\psi_n \quad \text{if } n \in \mathbb{Z}_+,$$

where

$$v_n = \begin{cases} \dfrac{1}{s_1 s_2} & \text{if} \quad n = 0, \\ \dfrac{1}{s_{2n+1}}\left(\dfrac{1}{s_{2n}} + \dfrac{1}{s_{2n+2}}\right) & \text{if } n \in \mathbb{Z}_+, \end{cases} \qquad h_n = \frac{1}{s_{2n+2}\sqrt{s_{2n+1}s_{2n+3}}}, \quad n \in \mathbb{N}.$$

These numbers may be used to define a certain Jacobi operator, say $\mathcal{J}$, with a domain contained in the Hilbert space $\ell^2(\mathbb{N})$, as follows: first, we consider sequences $\boldsymbol{\xi} = (\xi_0, \xi_1, \xi_2, \ldots)$ with only finitely many terms and set

$$(\mathcal{J}\boldsymbol{\xi})_n := \begin{cases} v_0\xi_0 + h_0\xi_1 & \text{if} \quad \mathrm{n} = 0, \\ h_{n-1}\xi_{n-1} + v_n\xi_n + h_n\xi_{n+1} & \text{otherwise.} \end{cases} \tag{1.11}$$

Given condition (1.8), it is then possible to extend this definition uniquely to obtain an essentially self-adjoint operator; we use the same symbol $\mathcal{J}$ to refer to this extension. It may then be proved that the moments of the spectral measure of the operator $\mathcal{J}$ are precisely the m_n, and so this spectral measure coincides with σ (see Nikishin & Sorokin 1988).

(*b*) *The density of states*

Consider the finite-dimensional truncation

$$\mathcal{J}_n := \begin{pmatrix} v_0 & h_0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ h_0 & v_1 & h_1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & h_1 & v_2 & h_2 & \cdots & 0 & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & h_{n-3} & v_{n-2} & h_{n-2} \\ 0 & 0 & 0 & 0 & \cdots & 0 & h_{n-2} & v_{n-1} \end{pmatrix},$$

of the operator $\mathcal{J}$. The spectrum of $\mathcal{J}_n$ is the set of zeros of the polynomial ψ_n defined by equation (1.9). By the Chebyshev–Markov–Stieltjes theorem (see Nikishin & Sorokin 1988, §2.8), between any two zeros of ψ_n, there is a point of increase of σ; so we are led to studying the distribution of $\lambda_{n,j}$.

Define a measure κ_n on $\mathbb{R}_+$ by

$$\kappa_n := \frac{1}{n} \sum_{j=1}^{n} \delta_{\lambda_{n,j}}. \tag{1.12}$$

This measure is the normalized eigenvalue counting measure of the matrix $\mathcal{J}_n$. Indeed, we have

$$N_n(\lambda) := \frac{\#\{j : \lambda_{n,j} < \lambda\}}{n} = \int_0^{\lambda} \kappa_n(\mathrm{d}\lambda').$$

The normalized counting measure κ_n has a weak limit, say κ, as $n \to \infty$, and so there is a function N, called the *integrated density of states* of $\mathcal{J}$, defined by

$$N(\lambda) := \int_0^{\lambda} \kappa(\mathrm{d}\lambda) = \lim_{n\to\infty} N_n(\lambda).$$

If κ is absolutely continuous, one can also speak of the *density of states*, say ϱ, defined by

$$\kappa(\mathrm{d}\lambda) = \varrho(\lambda)\mathrm{d}\lambda. \tag{1.13}$$

Although the measures κ and σ may be very different, their essential spectra are the same. In the context of Padé approximation, the integrated density of states describes the distribution of the poles of the approximants.

(*c*) *Krein's string*

There is an interpretation, due to Krein, of the spectrum of the operator $\mathcal{J}$ in terms of the characteristic frequencies of a vibrating string (Akhiezer 1961). Consider a weightless, infinite, perfectly elastic string, tied at one endpoint $x=0$, along which some beads are distributed. Let s_{2n} be the mass of the nth bead, and denote by (x_n, y_n) its position in the xy-plane. We assume that the x_n are fixed and given by the recurrence relation

$$x_{n+1} = x_n + s_{2n+1}, \quad x_0 = 0.$$

For a string of uniform unit tension, the small vertical motion is then described by the discrete wave equation

$$s_{2n}\ddot{y}_n = \frac{y_{n+1} - y_n}{s_{2n+1}} - \frac{y_n - y_{n-1}}{s_{2n-1}}, \quad n \in \mathbb{Z}_+. \tag{1.14}$$

To study the characteristic frequencies of the string, we set

$$y_n = \eta_n \sin(\omega t) \quad \text{and} \quad \xi_n = \frac{\eta_{n+1} - \eta_n}{s_{2n+1}}.$$

Then equation (1.14) reduces to

$$\frac{1}{s_1 s_2}\xi_1 - \frac{1}{s_1 s_2}\xi_0 = -\omega^2 \xi_0,$$

and

$$\frac{1}{s_{2n+1}s_{2n+2}}\xi_{n+1} - \frac{1}{s_{2n+1}}\left(\frac{1}{s_{2n}} + \frac{1}{s_{2n+2}}\right)\xi_n + \frac{1}{s_{2n}s_{2n+1}}\xi_{n-1} = -\omega^2 \xi_n,$$

where $n \geq 1$. Comparing this with the definitions of $\mathcal{J}$ and ψ_n given earlier, it is readily seen that

$$\xi_n = \frac{(-1)^n \xi_0}{\sqrt{s_{2n+1}}}\psi_n(\omega^2).$$

(*d*) *The complex Lyapunov exponent*

Dyson (1953) developed a method for studying the characteristic frequencies of the one-dimensional disordered chain

$$\ddot{y}_n = c_{2n-1}(y_{n+1} - y_n) - c_{2n-2}(y_n - y_{n-1}), \quad n \in \mathbb{Z}_+. \tag{1.15}$$

Here c_{2n-1} and c_{2n-2} are the ratios of the elastic modulus of the nth spring and of the mass of the two particles attached to it. Disorder may be modelled in many ways; for instance, by assuming that the c_n are independent and identically distributed. The approach was later simplified by Schmidt (1957) and applied to the tight-binding Anderson model for a one-dimensional crystal with impurities.

Luck (1992) gave a very readable, well-motivated account of the Dyson–Schmidt approach; in brief, it builds on the intimate connection between second-order difference equations, continued fractions and Markov chains. For our purpose, it will be convenient to work with the random difference equation

$$u_{n+1} - u_{n-1} = \frac{s_{n+1}}{\sqrt{t}}u_n, \quad n = 0, 1, 2, \ldots, \tag{1.16}$$

where t is a parameter in $\mathbb{C}\backslash\mathbb{R}_-$, and $\sqrt{\cdot}$ is the branch of the square-root function defined on $\mathbb{C}\backslash\mathbb{R}_-$ that returns a number with a non-negative real part. The following lemma is obvious.

Lemma 1.1. *For every* $t \in \mathbb{C} \backslash \mathbb{R}_-$,

$$Q_n(t) = \left(\sqrt{t}\right)^n u_n,$$

where u_n *solves the difference equation (1.16) with* $u_{-1}=0$ *and* $u_0=1$.

The relevant continued fraction is

$$Z := \sqrt{t}S(t) = \frac{1}{\frac{s_1}{\sqrt{t}} + \frac{1}{(s_2/\sqrt{t})+(1/((s_3/\sqrt{t})+\cdots))}} \tag{1.17}$$

and we write

$$Z_n = \sqrt{t}S_n \tag{1.18}$$

for its truncation. Let u_{-1} and u_0 be complex random variables. Then equation (1.16) defines a sequence of general terms u_n by recurrence. The distribution ν_μ of the random variable Z is a stationary distribution for the Markov chain

$$\hat{\boldsymbol{Z}} = (\hat{Z}_0, \hat{Z}_1, \hat{Z}_2, \ldots), \quad \text{where } \hat{Z}_n := \frac{u_{n-1}}{u_n}. \tag{1.19}$$

(In the terminology of iterated random maps, the random variables Z_n and $\hat{Z}_n$ are, respectively, the *backward* and *forward iterates* associated with the continued fraction; when $u_{-1}=0$ and $u_0=1$, they have the same distribution, but their asymptotic behaviours are very different; see Diaconis & Freedman 1999.)

It follows that the growth of u_n may be quantified by means of the *complex Lyapunov exponent* defined by

$$\Lambda_\mu(t) := -\int_{\mathbb{C}} \ln z \nu_\mu(\mathrm{d}z), \tag{1.20}$$

where ln denotes the principal branch of the logarithm. Indeed, standard results from the theory of Markov chains imply that if $\hat{\boldsymbol{Z}}$ has a unique stationary distribution, then

$$\frac{\ln u_n}{n} = \frac{\ln u_0}{n} - \frac{1}{n}\sum_{j=1}^{n} \ln \hat{Z}_j \xrightarrow[n\to\infty]{} \Lambda_\mu(t) \tag{1.21}$$

for almost every realization of $\boldsymbol{s}$, independently of the choice of $\hat{Z}_0$ (Breiman 1960; Furstenberg 1963; Meyn & Tweedie 1993). In particular, we have the formula

$$\lim_{n\to\infty} \frac{\ln|u_n|}{n} = \mathrm{Re}[\Lambda_\mu(t)].$$

Equation (1.21) is also central to the study of the integrated density of states of the operator $\mathcal{J}$ introduced earlier (Dyson 1953; Schmidt 1957; Luck 1992).

We shall see that, under a very mild assumption on the distribution μ of s_n,

$$N(\lambda) = -\frac{2}{\pi}\,\mathrm{i}[\Lambda_\mu(-1/\lambda + \imath 0+)].$$

(e) Furstenberg's theorem

In order to carry out this programme, we shall also make use of the connection between the Markov chain $\hat{\boldsymbol{Z}}$ and the product of random matrices

$$\mathcal{U}_n := \mathcal{A}_n\mathcal{A}_{n-1}\ldots\mathcal{A}_1, \quad n = 0, 1, 2, \ldots, \tag{1.22}$$

where

$$\mathcal{A}_n := \begin{pmatrix} 0 & 1 \\ 1 & \dfrac{s_n}{\sqrt{t}} \end{pmatrix}. \tag{1.23}$$

The distribution μ from which the s_n are drawn induces, via equation (1.23), a distribution $\tilde{\mu}$ on the group of unimodular 2×2 matrices. The fundamental results of Furstenberg & Kesten (1960) and Furstenberg (1963), which are commonly referred to as 'Furstenberg's theorem', imply in particular that, under very mild assumptions on μ, there is a unique measure $\tilde{\nu}_{\tilde{\mu}}$ on the group of unimodular matrices that is invariant under the action of the random matrix (1.23). Furthermore, the number

$$\gamma_{\tilde{\mu}} := \frac{1}{n}\mathbb{E}(\ln|\mathcal{U}_n|), \tag{1.24}$$

which quantifies the growth of the product $\mathcal{U}_n$ and is independent of the choice of matrix norm $|\cdot|$, may be shown to be *strictly positive*. These results are of great relevance to our problem for two reasons: firstly, any invariant measure $\tilde{\nu}_{\tilde{\mu}}$ yields a measure ν_μ that is stationary for the Markov chain $\hat{\boldsymbol{Z}}$, and vice versa; secondly,

$$\gamma_{\tilde{\mu}} = \mathrm{Re}[\Lambda_\mu(t)].$$

Hence, we deduce at once the uniqueness of the measure ν_μ, as well as the exponential growth of u_n.

Proposition 1.2. *Let the s_n be the independent random variables in $\mathbb{R}_+$ with a common distribution μ that has at least two points of increase. Let $\{u_n\}_{n\in\mathbb{N}}$ be the sequence defined by the recurrence (1.16). Suppose that*

$$\int_{\mathbb{R}_+} s^{\varepsilon}\mu(\mathrm{d}s) < \infty,$$

for some $\varepsilon > 0$. Then, for almost every realization of the sequence $\boldsymbol{s}$, the following holds independently of the starting values $u_0 \neq 0$ and u_{-1}: for Lebesgue almost every $t \in \mathbb{C}\backslash\mathbb{R}_-$,

$$\lim_{n\to\infty}\frac{\ln u_n}{n} = \Lambda_\mu(t) \quad \textit{and} \quad \mathrm{Re}[\Lambda_\mu(t)] > 0$$

Furthermore, if we set $t = -x + \mathrm{i}0\pm$, *then for Lebesgue almost every* $x > 0$,

$$\lim_{n\to\infty} \frac{\ln u_n}{n} = \Lambda_\mu(-x + \mathrm{i}0\pm) \quad \textit{and} \quad \mathrm{Re}[\Lambda_\mu(-x + \mathrm{i}0\pm)] > 0.$$

The proof of this proposition is provided in the electronic supplementary material, appendix A.

(*f*) *The gamma distribution: statement of the main result*

Using such machinery, we are able, for a very wide class of distributions μ, to deduce the almost sure exponential nature of the convergence of diagonal Padé approximation and also to deduce the almost sure singularity of the measure σ. A more quantitative study requires the calculation of the complex Lyapunov exponent, but there are very few known instances where it can be expressed in terms of familiar functions.

In his seminal paper on the disordered chain (1.15), Dyson studied in some detail the particular case where the c_n are independent and gamma distributed. Dyson found the invariant distribution of the continued fraction

$$\frac{c_0 t}{1 + \frac{c_1 t}{1 + (c_2 t/(1 + \cdots))}}$$

in the particular case where $t > 0$; he then used analytic continuation to obtain an expansion for the complex Lyapunov exponent at $t < 0$, and hence for the distribution of the characteristic frequencies. Dyson's continued fraction is not equivalent to ours (compare equations (1.15) and (1.14)), and so his analytical results do not transfer to our problem. However, the continued fraction (1.17) with independent gamma-distributed s_n, i.e. where

$$\mu(\mathrm{d}x) := \frac{1}{b^a \Gamma(a)} x^{a-1} \exp(-x/b)\mathrm{d}x, \quad a, b > 0 \tag{1.25}$$

was examined also by Letac & Seshadri (1983); they obtained the probability distribution ν_μ and found an explicit formula for the corresponding Lyapunov exponent for $t > 0$. In a recent paper, we generalized this result by finding ν_μ and the real part of the complex Lyapunov exponent for *every* complex t (see Marklof *et al.* in press); a straightforward extension of these calculations leads to the remarkably simple formula

$$\Lambda_\mu(t) = \partial_a \ln\left[K_a\left(\frac{2\sqrt{t}}{b}\right)\right]. \tag{1.26}$$

In this expression, ∂_a denotes differentiation with respect to a, and K_a is the modified Bessel function of the second kind (see electronic supplementary material, appendix B). The following summarizes the key results of this paper.

Theorem 1.3. *Suppose that the* s_n *are independent draws from the gamma distribution with parameters* $a > 0$ *and* $b > 0$. *Then, for almost every realization of the sequence* $\boldsymbol{s}$, *the following holds*:

(i) *The density of states is given explicitly by the formula*

$$\varrho(\lambda) = -\frac{2}{\pi^2 \lambda} \partial_a \left[\frac{1}{J_a^2\left(\frac{2}{b\sqrt{\lambda}}\right) + Y_a^2\left(\frac{2}{b\sqrt{\lambda}}\right)} \right].$$

(ii) $\mathrm{spec}(\sigma) = [0, \infty)$ *and its absolutely continuous part is empty.*

(iii) *For Lebesgue almost every* $t \in \mathbb{C} \backslash \mathbb{R}_-$,

$$\lim_{n\to\infty} \frac{\ln|S(t) - S_n(t)|}{n} = -2\partial_a \ln\left| K_a\left(\frac{2\sqrt{t}}{b}\right)\right|.$$

(*g*) *Relation to other work*

As stated earlier, our focus in the present paper is on the performance of diagonal Padé approximation, viewed as a method of summing some random series. Related questions have been considered in the past, in different contexts. Foster & Pitcher (1974) studied the convergence of random T-fractions; these are continued fraction expansions which are in a one-to-one correspondence with the space of formal power series, but whose convergents are not Padé approximants. Foster & Pitcher (1974) showed that under very general conditions on the distribution of the coefficients, the difference between two successive convergents tends to zero exponentially fast, and that the exponent is twice the Lyapunov exponent associated with an infinite product of random matrices. Geronimo (1993) studied the random measure (on the unit circle) generated by a three-term recurrence relation with random identically distributed coefficients; he showed the positivity of the corresponding Lyapunov exponent and deduced that the random measure is singular with respect to the Lebesgue measure. This list is not exhaustive (see also Csordas *et al.* (1973) and Mannion (1993)).

The question of the nature of the measure σ has a counterpart in the theory of disordered systems which has been studied extensively in the context of Anderson localization. For example, the tight-binding Anderson model uses a discretized version of the Schrödinger equation with a potential that takes random identically distributed values at every point in a doubly infinite lattice. The resulting operator has a second-order finite-difference form like that of the operator $\mathcal{J}$—in which, more precisely, $h_n = 1$ and v_n are independent and identically distributed—but it acts on sequences in $\ell^2(\mathbb{Z})$. For a very wide choice of the distribution of the potential values v_n, the Lyapunov exponent of the discretized Schrödinger operator is strictly positive, so that, by Ishii's formula, the absolutely continuous spectrum is empty. A more refined study (see, for instance, Carmona & Lacroix (1990) and Pastur & Figotin (1992)) reveals that these operators have a *pure point* spectrum, i.e. the spectrum is the closure of the discrete spectrum. Such operators are said to exhibit the *localization property* because, for this type of spectrum, the generalized eigenfunctions decay exponentially fast as $|n| \to \infty$. The rigorous extension of such detailed results to the semi-infinite case would involve technicalities which are beyond the scope of the present paper.

Our analysis exploits a number of ideas and techniques found in these earlier studies. *We view our main contribution as that of exhibiting an interesting example of a class of random Stieltjes functions for which the leading behaviour of the error of diagonal Padé approximation and the density of states of the corresponding Jacobi operator are given explicitly in terms of special functions.*

The remainder of the paper is devoted to a detailed proof of theorem 1.3: the first statement follows immediately from Dyson's formula for the density of states, which is derived in §2. In §3, we deduce the singularity of the measure σ from the positivity of the real part of the Lyapunov exponent. In §4 we show that the error of diagonal Padé approximation is inversely proportional to the square of u_n; this yields the third statement in the theorem. Finally, in §5, we provide a numerical illustration of our results.

2. The formula for the density of states

So that we can use proposition 1.2, we shall henceforth suppose that the s_n are independent draws from a distribution μ on $\mathbb{R}_+$ such that

(i) μ has at least two points of increase.
(ii) There exists $\varepsilon>0$ such that

$$\int_{\mathbb{R}_+} s^{\varepsilon}\mu(\mathrm{d}s) < \infty.$$

To avoid needless repetitions, we shall not mention these particular assumptions explicitly again in the statement of the intermediate results leading to theorem 2.4. We begin by relating the growth of ψ_n to the complex Lyapunov exponent.

Lemma 2.1. *For almost every realization of the sequence* $\boldsymbol{s}$, *we have*, *for Lebesgue almost every* $\lambda \in \mathbb{C}\backslash\mathbb{R}_+$,

$$\lim_{n\to\infty} \frac{\ln \psi_n(\lambda)}{n} = \mathrm{i}\pi + 2\Lambda_\mu(-1/\lambda),$$

and, *for Lebesgue almost every* $\lambda \in \mathbb{R}_+$,

$$\lim_{n\to\infty} \frac{\ln \psi_n(\lambda + \mathrm{i}0\pm)}{n} = \mathrm{i}\pi + 2\Lambda_\mu(-1/\lambda + \mathrm{i}0\pm).$$

Proof. Let $\lambda \in \mathbb{C}\backslash\mathbb{R}_+$. By definition,

$$\psi_n(\lambda) = \sqrt{s_{2n+1}}\lambda^n Q_{2n}(-1/\lambda).$$

Hence, by lemma 1.1,

$$\psi_n(\lambda) = \sqrt{s_{2n+1}}\lambda^n \left(\sqrt{-1/\lambda}\right)^{2n} u_{2n} = (-1)^n \sqrt{s_{2n+1}} u_{2n},$$

where u_n solves the difference equation (1.16) with $t=-1/\lambda$. This yields

$$\frac{\ln \psi_n(\lambda)}{n} = \frac{\ln s_{2n+1}}{2n} + \mathrm{i}\pi + 2\frac{\ln u_{2n}}{2n}.$$

The first statement in the proposition then follows from the electronic supplementary material, appendix A, corollary 5.3. The proof of the second statement is identical. ∎

Next, we examine the implications of the lemma for the distribution of $\lambda_{n,j}$. By virtue of the recurrence relation satisfied by ψ_n, we can write

$$\psi_n(\lambda) = \sqrt{s_{2n+1}}\left(\prod_{j=1}^{2n} s_j\right)(\lambda - \lambda_{n,1})\ldots(\lambda - \lambda_{n,n}).$$

Let

$$E_n := \{\lambda_{n,j} : 1 \le j \le n\},$$

and let $\lambda \notin E_n$. Then

$$\frac{\ln \psi_n(\lambda)}{n} = \frac{\ln s_{2n+1}}{2n} + \frac{1}{n}\sum_{j=1}^{2n} \ln s_j + \frac{1}{n}\sum_{j=1}^{n} \ln(\lambda - \lambda_{n,j}). \tag{2.1}$$

Proposition 2.2. *Suppose that*

$$\int_{\mathbb{R}_+} |\ln s| \mu(\mathrm{d}s) < \infty.$$

Then for almost every realization of the sequence $\boldsymbol{s}$, *for Lebesgue almost every* $\lambda \in \mathbb{C}\backslash\mathbb{R}_+$,

$$\int_0^\infty \ln|\lambda - \lambda'| \kappa_n(\mathrm{d}\lambda') \xrightarrow[n\to\infty]{} 2\,\mathrm{Re}[\Lambda_\mu(-1/\lambda)] - 2\int_{\mathbb{R}_+} \ln s \mu(\mathrm{d}s).$$

Proof. Consider the real part in equation (2.1). Then

$$\int_0^\infty \ln|\lambda - \lambda'| \kappa_n(\mathrm{d}\lambda') = \frac{1}{n}\sum_{j=1}^{n} \ln|\lambda - \lambda_{n,j}| = \frac{\ln|\psi_n(\lambda)|}{n} - \frac{\ln s_{2n+1}}{2n} - \frac{1}{n}\sum_{j=1}^{2n} \ln s_j.$$

By lemma 2.1, the first term on the right tends to

$$2\,\mathrm{Re}[\Lambda_\mu(-1/\lambda)],$$

the second term tends to zero and, by the ergodic theorem, the third term tends to

$$2\int_{\mathbb{R}_+} \ln s \mu(\mathrm{d}s).$$

∎

Corollary 2.3. *Under the same assumption, for almost every realization of the sequence* $\boldsymbol{s}$, *the sequence* $\{\kappa_n\}_{n\in\mathbb{N}}$ *has a weak limit, say* κ, *which is a probability measure on* $\mathbb{R}_+$. *In particular,*

$$\lim_{n\to\infty} N_n(\lambda) = N(\lambda) := \int_0^\lambda \kappa(\mathrm{d}\lambda').$$

Proof. The proof is a specialization of that given by Goldsheid & Khoruzhenko (2005) in the more general case of a non-Hermitian Jacobi matrix. See electronic supplementary material, appendix C for details. ∎

Theorem 2.4. *Let the s_n be independent random variables in $\mathbb{R}_+$ with a common distribution μ that has at least two points of increase. Suppose also that*

$$\int_{\mathbb{R}_+} |\ln s|\mu(\mathrm{d}s) < \infty \quad \textit{and} \quad \int_{\mathbb{R}_+} s^{\varepsilon}\mu(\mathrm{d}s) < \infty,$$

for some $\varepsilon > 0$. Then, for almost every realization of the sequence $\boldsymbol{s}$, for Lebesgue almost every $\lambda \in \mathbb{R}_+$,

$$N(\lambda) = -\frac{2}{\pi}\mathrm{Im}[\Lambda_\mu(-1/\lambda + \mathrm{i}0+)].$$

Proof. Let $\lambda \in \mathbb{R}_+ \backslash E_n$. Then

$$\ln \psi_n(\lambda + \mathrm{i}0\pm) - \frac{\ln s_{2n+1}}{2n} - \frac{1}{n}\sum_{j=1}^{2n} \ln s_j = \sum_{j=1}^{n} \ln(\lambda - \lambda_{n,j} + \mathrm{i}0\pm)$$

$$= \sum_{\lambda_{n,j}<\lambda} \ln(\lambda - \lambda_{n,j} + \mathrm{i}0\pm) + \sum_{\lambda_{n,j}>\lambda} \ln(\lambda - \lambda_{n,j} + \mathrm{i}0\pm)$$

$$= \sum_{\lambda_{n,j}<\lambda} \ln|\lambda - \lambda_{n,j}| + \sum_{\lambda_{n,j}>\lambda} (\pm \mathrm{i}\pi + \ln|\lambda - \lambda_{n,j}|) = \sum_{j=1}^{n} \ln|\lambda - \lambda_{n,j}| \pm \sum_{\lambda_{n,j}>\lambda} \mathrm{i}\pi.$$

Hence, we have the identity

$$\frac{\ln \psi_n(\lambda + \mathrm{i}0\pm)}{n} = \frac{\ln s_{2n+1}}{2n} + \frac{1}{n}\sum_{j=1}^{2n} \ln s_j + \int_0^\infty \ln|\lambda - \lambda'|\kappa_n(\mathrm{d}\lambda') \pm \mathrm{i}\pi[1 - N_n(\lambda)]. \tag{2.2}$$

Consider the imaginary part; the result then follows from lemma 2.1 and corollary 2.3. ■

Corollary 2.5. *Suppose that the s_n are independent and gamma distributed with parameters a and b. Then, for $\lambda > 0$, we have the following formula for the density of states*:

$$\varrho(\lambda) := N'(\lambda) = -\frac{2}{\pi^2\lambda}\partial_a\left[\frac{1}{J_a^2\left(\frac{2}{b\sqrt{\lambda}}\right) + Y_a^2\left(\frac{2}{b\sqrt{\lambda}}\right)}\right].$$

Proof. Let $\lambda > 0$ and set

$$z = \frac{2}{b\sqrt{\lambda}}.$$

For $t = -1/\lambda + \mathrm{i}0+$, we have

$$K_a\left(\frac{2\sqrt{t}}{b}\right) = K_a(\mathrm{i}z) = -\frac{\pi}{2}e^{-\mathrm{i}a(\pi/2)}[Y_a(z) + \mathrm{i}J_a(z)].$$

Hence,

$$\partial_a K_a\left(\frac{2\sqrt{t}}{b}\right) = -\mathrm{i}\frac{\pi}{2}K_a\left(\frac{2\sqrt{t}}{b}\right) - \frac{\pi}{2}\mathrm{e}^{-\mathrm{i}a(\pi/2)}[\partial_a Y_a(z) + \mathrm{i}\partial_a J_a(z)],$$

and so

$$\Lambda_\mu(t) = \frac{\partial_a K_a\left(\frac{2\sqrt{t}}{b}\right)}{K_a\left(\frac{2\sqrt{t}}{b}\right)} = \overline{K_a\left(\frac{2\sqrt{t}}{b}\right)}\frac{\partial_a K_a\left(\frac{2\sqrt{t}}{b}\right)}{\left|K_a t\left(\frac{2\sqrt{t}}{b}\right)\right|^2}$$

$$= -\mathrm{i}\frac{\pi}{2} + \mathrm{i}\frac{Y_a(z)\partial_a J_a(z) - J_a(z)\partial_a Y_a(z)}{J_a(z)^2 + Y_a(z)^2} + \frac{J_a(z)\partial_a J_a(z) + Y_a(z)\partial_a Y_a(z)}{J_a(z)^2 + Y_a(z)^2}.$$

We deduce the formulae

$$\mathrm{Re}\left[\Lambda_\mu\left(-\frac{1}{\lambda} + \mathrm{i}0+\right)\right] = \frac{J_a(z)\partial_a J_a(z) + Y_a(z)\partial_a Y_a(z)}{J_a(z)^2 + Y_a(z)^2}, \tag{2.3}$$

and

$$\mathrm{Im}\left[\Lambda_\mu\left(-\frac{1}{\lambda} + \mathrm{i}0+\right)\right] = -\frac{\pi}{2} + \frac{Y_a(z)\partial_a J_a(z) - J_a(z)\partial_a Y_a(z)}{J_a(z)^2 + Y_a(z)^2}. \tag{2.4}$$

The proposition, together with the last equation, then yields

$$1 - N(\lambda) = \frac{2}{\pi}\frac{Y_a(z)\partial_a J_a(z) - J_a(z)\partial_a Y_a(z)}{J_a(z)^2 + Y_a(z)^2} = -\frac{2}{\pi}\partial_a\left[\arctan\frac{Y_a(z)}{J_a(z)}\right].$$

Differentiating both sides with respect to λ, we find

$$-\varrho(\lambda) = -\frac{2}{\pi}\frac{\mathrm{d}z}{\mathrm{d}\lambda}\frac{\mathrm{d}}{\mathrm{d}z}\partial_a\left[\arctan\frac{Y_a(z)}{J_a(z)}\right] = \frac{2}{\pi}\frac{1}{b\lambda^{3/2}}\frac{\mathrm{d}}{\mathrm{d}z}\partial_a\left[\arctan\frac{Y_a(z)}{J_a(z)}\right].$$

We obtain the desired result by changing the order of differentiation on the right-hand side, and making use of the identity

$$J_a(z)Y_a'(z) - Y_a(z)J_a'(z) = \frac{2}{\pi z}.$$

■

Corollary 2.6. *Under the same assumption, for almost every realization of the sequence* $\boldsymbol{s}$,

$$\mathrm{spec}(\sigma) = [0, \infty).$$

Proof. By differentiating the identity (see Watson 1966, §13.73)

$$J_a^2(z) + Y_a^2(z) = \frac{8}{\pi^2}\int_0^\infty K_0(2z\sinh t)\cosh(2at)\mathrm{d}t$$

with respect to a, we deduce that ϱ is strictly positive. ■

3. Singularity of the spectrum

Corollary 2.6 implies in particular that the radius of convergence of the generating series of the moments is zero almost surely; in other words, the random Stieltjes functions that we have constructed are not analytic at the origin.

The problem of determining the nature of the spectrum is more delicate. Every measure may be decomposed into three disjoint parts: its absolutely continuous, singular continuous and discrete parts, denoted by σ_{ac}, σ_{sc} and σ_{d}, respectively. Ishii (1973) and Yoshioka (1973) showed that the spectrum of the absolutely continuous part is given by the formula

$$\mathrm{spec}(\sigma_{\mathrm{ac}}) = \overline{\left\{\lambda > 0 : \lim_{t \to -\frac{1}{\lambda} + \mathrm{i}0+} \mathrm{Re}\Lambda_{\mu}(t) = 0\right\}}. \tag{3.1}$$

This result may be established by examining the resolvent of the Jacobi operator $\mathcal{J}$. The following result is essentially equivalent:

Proposition 3.1. *Let the s_n be independent random variables in $\mathbb{R}_+$ with a common distribution μ that has at least two points of increase. Assume that*

$$\int_{\mathbb{R}_+} |\ln s| \mu(\mathrm{d}s) < \infty \quad \textit{and} \quad \int_{\mathbb{R}_+} s^{\varepsilon} \mu(\mathrm{d}s) < \infty,$$

for some $\varepsilon > 0$. Then, for almost every realization of the sequence $\boldsymbol{s}$,

$$\mathrm{spec}(\sigma_{ac}) = \emptyset.$$

Proof. By lemma 2.1 and electronic supplementary material, appendix A, corollary 5.5, for almost every realization of $\boldsymbol{s}$, for Lebesgue almost every $\lambda \in \mathbb{R}_+$, we have

$$\lim_{n \to \infty} \frac{\ln|\psi_n(\lambda)|}{n} > 0. \tag{3.2}$$

Now, let $\eta > 0$ and consider the set

$$\mathbb{S}_{\eta} := \left\{\lambda \in \mathbb{R}_+ : \psi_n(\lambda) = o\left(\sqrt{n}[\ln n]^{1+\eta}\right) \quad \text{as } n \to \infty\right\}.$$

By equation (3.2), this set has Lebesgue measure zero almost surely. On the other hand, it follows from the Men'shov–Rademacher theorem (see Nikishin & Sorokin 1988, proposition 8.3) that for σ almost every $\lambda \in \mathbb{R}_+$,

$$\psi_n(\lambda) = o\left(\sqrt{n}[\ln n]^{1+\eta}\right) \quad \text{as } n \to \infty.$$

So, almost surely, for every σ *measurable set A,*

$$\int_A \sigma(\mathrm{d}\lambda) = \int_{A \cap \mathbb{S}_{\eta}} \sigma(\mathrm{d}\lambda) = \int_{A \cap \mathbb{S}_{\eta}} \sigma_{\mathrm{ac}}(\mathrm{d}\lambda) + \int_{A \cap \mathbb{S}_{\eta}} \sigma_{\mathrm{sc}}(\mathrm{d}\lambda) + \int_{A \cap \mathbb{S}_{\eta}} \sigma_{\mathrm{d}}(\mathrm{d}\lambda)$$

$$= \int_{A \cap \mathbb{S}_{\eta}} \sigma_{\mathrm{sc}}(\mathrm{d}\lambda) + \int_{A \cap \mathbb{S}_{\eta}} \sigma_{\mathrm{d}}(\mathrm{d}\lambda) = \int_A \sigma_{\mathrm{sc}}(\mathrm{d}\lambda) + \int_A \sigma_{\mathrm{d}}(\mathrm{d}\lambda).$$

■

4. The rate of convergence

Let $\{u_n\}_{n\in\mathbb{N}}$ be the sequence defined by the recurrence relation (1.16) with the starting values $u_{-1}=0$ and $u_0=1$. Also, denote by $\mathcal{T}$ the shift operator on the space of complex sequences, i.e.

$$\mathcal{T}(s_1, s_2, \dots) = (s_2, s_3, \dots).$$

In order to emphasize the dependence of the continued fraction (1.17) on the sequence $\boldsymbol{s}$, we shall sometimes write $Z(\boldsymbol{s})$ and $Z_n(\boldsymbol{s})$ instead of Z and Z_n. We have the following convenient representation of the error

Lemma 4.1.

$$Z(\boldsymbol{s}) - Z_n(\boldsymbol{s}) = \frac{(-1)^n}{u_n}\left(\prod_{j=0}^{n} Z(\mathcal{T}^j \boldsymbol{s})\right).$$

Proof. Using the recurrence relations (1.6) and the identity

$$Z(\mathcal{T}^n \boldsymbol{s}) = \frac{1}{\frac{s_{n+1}}{\sqrt{t}} + Z(\mathcal{T}^{n+1}\boldsymbol{s})},$$

it is straightforward to show (by induction on n) that

$$Z(\mathcal{T}^n \boldsymbol{s}) = -\frac{1}{\sqrt{t}}\frac{Q_n S - P_n}{Q_{n-1}S - P_{n-1}}.$$

Then

$$\prod_{j=0}^{n} Z(\mathcal{T}^j \boldsymbol{s}) = (-1/\sqrt{t})^{n+1}\prod_{j=0}^{n}\frac{Q_j S - P_j}{Q_{j-1}S - P_{j-1}} = (-1/\sqrt{t})^{n+1}\frac{Q_n S - P_n}{Q_{-1}S - P_{-1}}$$

$$= (-1/\sqrt{t})^n Q_n \sqrt{t}(S - S_n).$$

Lemma 1.1 then yields the desired result. ∎

Theorem 4.2. *Let the s_n be positive independent random variables with a common distribution μ that has at least two points of increase. Suppose also that there exists $\varepsilon>0$ such that*

$$\int_{\mathbb{R}_+} s^{\varepsilon}\mu(\mathrm{d}s) < \infty.$$

Then, for almost every realization of the sequence $\boldsymbol{s}$, for Lebesgue almost every $t\in\mathbb{C}\backslash\mathbb{R}_-$,

$$\lim_{n\to\infty}\frac{\ln|S(t) - S_n(t)|}{n} = -2\,\mathrm{Re}[\Lambda_\mu(t)].$$

Proof. Let $t\in\mathbb{C}\backslash\mathbb{R}_-$ be fixed. We have

$$Z(\boldsymbol{s}) = \sqrt{t}S(t) \quad \text{and} \quad Z_n(\boldsymbol{s}) = \sqrt{t}S_n(t).$$

Hence, by lemma 4.1,

$$\frac{\ln|S(t) - S_n(t)|}{n} = -\frac{\ln|t|}{2n} - \frac{\ln|u_n|}{n} + \frac{1}{n}\sum_{j=1}^{n} \ln|Z(\mathcal{T}^j \boldsymbol{s})|. \tag{4.1}$$

Consider the almost sure limit of each of the three terms on the right-hand side of this equality as $n \to \infty$. The first term tends to zero. By proposition 1.2, the second term tends to

$$-\mathrm{Re}\big[\Lambda_\mu(t)\big].$$

Finally, it follows easily from the ergodic theorem that the third term tends to the same limit. The result follows from a standard argument involving the use of Fubini's theorem. ■

5. A numerical illustration

Following the example of Dyson (1953), it is instructive to begin with an examination of the (deterministic) case where

$$\mu = \mu^\infty := \delta_1.$$

Set

$$S^\infty(t) := \frac{1}{1 + \frac{t}{1+(t/(1+\cdots))}}, \quad t \in \mathbb{C}\backslash\mathbb{R}_-.$$

Then

$$S^\infty(t) = \frac{1}{1+\sqrt{1+4t}} = \int_0^\infty \frac{\sigma^\infty(\mathrm{d}x)}{1+xt},$$

where

$$\sigma^\infty(\mathrm{d}x) = \begin{cases} \dfrac{1}{2\pi}\sqrt{4/x - 1}\,\mathrm{d}x & \text{if } 0 < x < 4, \\ 0 & \text{if } x > 4. \end{cases}$$

Note that μ^∞ has only one point of increase, and so the hypothesis of proposition 3.1 is not satisfied. Indeed, for this choice of μ, the spectrum of the measure σ is absolutely continuous.

In this case, the continued fraction (1.17) reduces to

$$Z^\infty = \frac{2}{\sqrt{1/t+4} + 1/\sqrt{t}},$$

and so the complex Lyapunov exponent is

$$\Lambda_{\mu^\infty}(t) = -\int_{\mathbb{C}} \ln z\,\nu_{\mu^\infty}(\mathrm{d}z) = -\int_{\mathbb{C}} \ln z\,\delta_{Z^\infty}(\mathrm{d}z) = \ln\frac{\sqrt{1/t+4} + 1/\sqrt{t}}{2}.$$

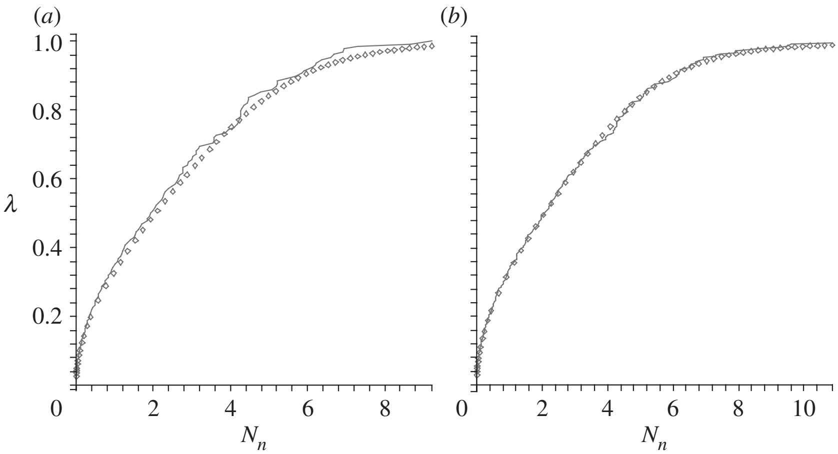

Figure 1. Counting measure N_n (solid line) for a particular realization of the sequence $\boldsymbol{s}$ when $\mu = \mu^a$ with $a=8$: (*a*) $n=128$ and (*b*) $n=256$. For comparison, points corresponding to values of the integrated density of states N are also shown.

In particular, an elementary calculation shows that

$$N^{\infty}(\lambda) = \begin{cases} 1 - \dfrac{2}{\pi} \arccos \dfrac{\sqrt{\lambda}}{2} & \text{if } 0 < \lambda < 4, \\ 1 & \text{if } \lambda > 4. \end{cases}$$

Next, let $a \in \mathbb{Z}_+$ and denote by μ^a the gamma distribution with $b = 1/a$. As Dyson remarked, this distribution has mean 1 and variance $1/a$ and so we may, for large a, view it as a perturbation of μ^{∞}. Indeed, using our explicit formula for Λ_{μ^a} together with the large-order expansions in Abramowitz & Stegun (1964), §9.7, we find

$$\Lambda_{\mu^a}(t) \sim \Lambda_{\mu^{\infty}}(t) + \frac{1}{a} \frac{1+8t}{2(1+4t)} + O\left(\frac{1}{a^2}\right) \quad \text{as } a \to \infty, \quad a \in \mathbb{Z}_+.$$

Likewise setting

$$\beta := \arccos \frac{\sqrt{\lambda}}{2}, \quad 0 < \lambda < 4.$$

And using the large-order expansions in Abramowitz & Stegun (1964), §9.3, we obtain, for $a \in \mathbb{Z}_+$,

$$\varrho(\lambda) \sim \varrho^{\infty}(\lambda) - \frac{1}{a^2} \frac{\cos \beta}{32\pi \sin^3 \beta} (13 + 38 \cot^2 \beta + 25 \cot^4 \beta) + O\left(\frac{1}{a^4}\right) \quad \text{as } a \to \infty.$$

This expansion breaks down at $\lambda = 4$; as a increases, $\varrho(\lambda)$ diverges to infinity there but tends to zero exponentially fast for $\lambda > 4$.

The following computations were performed in multiple-precision floating-point arithmetic with the Maple software package; the eigenvalues and eigenvectors of the matrix $\mathcal{J}_n$ were calculated using the `Eigenvals` function, which implements the QR algorithm. Figure 1 corresponds to the case where $\mu = \mu^a$ with $a = 8$; it illustrates the convergence of the counting measure N_n to the integrated density of states N as n increases—thus confirming the validity of our formula for N.

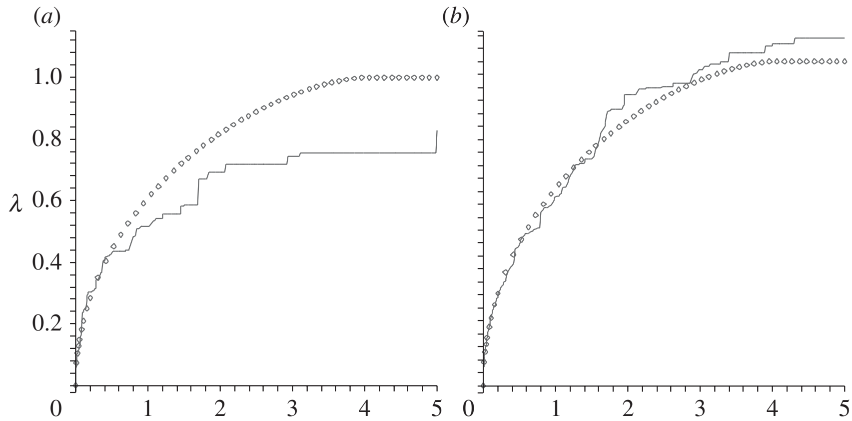

Figure 2. Plot (solid line) of $\int_0^\lambda \sigma_n(\mathrm{d}\lambda')$, with $n=256$, corresponding to particular realizations of $\boldsymbol{s}$ when $\mu=\mu^a$ where (*a*) $a=8$ and (*b*) $a=64$. For comparison, points corresponding to values of the function $\int_0^\lambda \sigma^\infty(\mathrm{d}\lambda')$ are also shown.

While the Lyapunov exponent and the density of states are non-random, the measure $\boldsymbol{\sigma}$ is random. We note that $\boldsymbol{\sigma}^\infty$ is absolutely continuous whereas, for every $a\in\mathbb{Z}_+$, almost every realization of $\boldsymbol{\sigma}$ is singular. Figure 2 shows the approximation

$$\int_0^\lambda \boldsymbol{\sigma}_n(\mathrm{d}\lambda'),$$

of the integrated measure for particular realizations corresponding to $a=8$ and 64. Here $\boldsymbol{\sigma}_n$ is the discrete measure defined by the quadrature formula (1.10), and $n=256$.

The authors gratefully acknowledge the support of the Engineering and Physical Sciences Research Council (United Kingdom) under grant GR/S87461/01 and of the Leverhulme Trust under a Philip Leverhulme Prize (J.M.). The second author would also like to thank Professor Alain Comtet for his hospitality during a visit to the Laboratoire de Physique Théorique et Modèles Statistiques, Université de Paris-Sud (Orsay), and for many useful discussions while some of this work was carried out.

References

Abramowitz, M. & Stegun, I. A. 1964 *Handbook of mathematical functions*. New York, NY: Dover.

Akhiezer, N. I. 1961 *The classical moment problem and related problems in analysis*. Moscow, Russia: Fizmatgiz. [Transl. Oliver & Boyd 1965.]

Bender, C. & Orszag, S. A. 1978 *Advanced mathematical methods for scientists and engineers*. Columbus, OH: McGraw-Hill.

Breiman, L. 1960 The strong law of large numbers for a class of Markov chains. *Ann. Math. Stat.* **31**, 801–803.

Carmona, R. & Lacroix, J. 1990 *Spectral theory of random Schrödinger operators*. Boston, MA: Birkhäuser.

Csordas, G., Hilden, H. & Pitcher, T. 1973 Random functions in H^∞. *Z. Wahrsch. Verw. Gebiete* **26**, 325–334. (doi:10.1007/BF00534896)

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

Dyson, F. J. 1953 The dynamics of a disordered linear chain. *Phys. Rev.* **92**, 1331–1338. (doi:10.1103/PhysRev.92.1331)

Foster, S. & Pitcher, T. 1974 Convergence properties of random T-fractions. *Z. Wahrsch. Verw. Gebiete* **29**, 323–330. (doi:10.1007/BF00532716)

Furstenberg, H. 1963 Non commuting random products. *Trans. Amer. Math. Soc.* **108**, 377–428. (doi:10.2307/1993589)

Furstenberg, H. & Kesten, H. 1960 Products of random matrices. *Ann. Math. Stat.* **31**, 377–428.

Geronimo, J. S. 1993 Polynomials orthogonal on the unit circle with random recurrence coefficients. In *Methods of approximation theory in complex analysis and mathematical physics (Leningrad, 1991)*. Springer lecture notes in mathematics, no. 1550, pp. 43–61.

Goldsheid, I. & Khoruzhenko, B. 2005 The Thouless formula for random non-hermitian Jacobi matrices. *Isr. J. Math.* **148**, 331–346.

Ishii, K. 1973 Localization of eigenstates and transport phenomena in the one-dimensional disordered system. *Suppl. Prog. Theor. Phys.* **53**, 77–138.

Letac, G. & Seshadri, V. 1983 A characterisation of the generalised inverse gaussian distribution by continued fractions. *Z. Wahrsch. Verw. Gebiete* **62**, 485–489. (doi:10.1007/BF00534200)

Luck, J. M. 1992 *Systèmes désordonnés unidimensionnels*. Frederick, MD: ALEA.

Mannion, D. 1993 Products of 2×2 matrices. *Ann. Appl. Probab.* **3**, 1189–1218.

Marklof, J., Tourigny, Y., & Wołowski, L. In press. Explicit invariant measures for products of random matrices. *Trans. Amer. Math. Soc.*

Meyn, P. & Tweedie, R. 1993 *Markov chains and stochastic stability*. Berlin, Germany: Springer.

Nikishin, E. M. & Sorokin, V. N. 1988 *Rational approximation and orthogonality*. Moscow, Russia: Nauk. [Transl. *Am. Math. Soc.*, 1991.]

Pastur, L. & Figotin, A. 1992 *Spectra of random and almost-periodic operators*. Berlin, Germany: Springer.

Schmidt, H. 1957 Disordered one-dimensional crystals. *Phys. Rev.* **105**, 425–441. (doi:10.1103/PhysRev.105.425)

Simon, B. 1998 The classical moment problem as a self-adjoint finite difference operator. *Adv. Math.* **117**, 82–203. (doi:10.1006/aima.1998.1728)

Stieltjes, T. J. 1894 Recherches sur les fractions continues. *Ann. Fac. Sci. Toulouse* **8**, 1–122.

Watson, G. N. 1966 *A treatise on the theory of Bessel functions*. Cambridge, UK: Cambridge University Press.

Yoshioka, Y. 1973 On the singularity of the spectral measures of a semi-infinite random system. *Proc. Jpn Acad.* **49**, 665–668.