Effective estimates of Lyapunov exponents for random products of positive matrices

To cite this article: M Pollicott 2021 Nonlinearity 34 6705

View the article online for updates and enhancements.
Effective estimates of Lyapunov exponents for random products of positive matrices

M Pollicott

Department of Mathematics, Warwick University, Coventry CV4 7AL, United Kingdom

E-mail: masdbl@warwick.ac.uk

Received 4 December 2020, revised 11 June 2021
Accepted for publication 19 July 2021
Published 16 August 2021

Abstract
In this note we describe estimates on the error when calculating the Lyapunov exponent for random products of positive matrices using dynamical determinants. This extends the results in (Jurga N and Morris I 2019 Nonlinearity 32 4117–46; Pollicott M 2010 Invent. Math. 181 209–26) by drawing upon a new approach introduced in (Jenkinson O and Pollicott M 2018 Adv. Math. 325 87–115; Jenkinson O, Pollicott M and Vytnova P 2018 J. Stat. Phys. 170 221–53).

Keywords: numerical estimates, dynamical determinants, transfer operators, thermodynamic formalism, Lyapunov exponents
Mathematics Subject Classification numbers: 37M25, 37C30.

(Some figures may appear in colour only in the online journal)

1. Introduction
The importance of Lyapunov exponents for random matrix products has been illustrated by the renowned work of Kesten [4], Furstenberg [3], Osceledec [9], and others, and its influence in other areas of mathematics and physics. Lyapunov exponents for positive matrices, in particular, play a useful role in such topics as entropy rates in information theory [5] and in the theory of hidden Markov processes.

Let us consider for definiteness the case of $2 \times 2$ matrices, the general case being similar. Assume that $A^{(0)} = (a_{ij}^{(0)})$ and $A^{(1)} = (a_{ij}^{(1)})$ are two such matrices with strictly positive entries.
Moreover, for simplicity, we will assume that both matrices have determinant 1, otherwise the contribution from the determinant can be easily compensated for. Let $0 < p_1, p_2 < 1$ with $p_1 + p_2 = 1$. We can define the Lyapunov exponent

$$\lambda = \lambda(A_1, A_2; p_1, p_2)$$

$$= \lim_{n \to +\infty} \frac{1}{n} \left( \sum_{i_1, \ldots, i_n \in \{1, 2\}} (p_{i_1} \cdots p_{i_n}) \log \|A_{i_1} \cdots A_{i_n}\| \right),$$

the existence of the limit following by subadditivity.

Typically there is no simple closed form for $\lambda$ and so we often want to estimate the associated Lyapunov exponent numerically, as accurately as possible, given the data we compute from finite products of matrices. Working from the definition is not very efficient in this regard, but there are a variety of different methods available. One empirically efficient approach was proposed in [10] and was further developed in [8]. The method involves introducing a convergent power series

$$d(z, s) := 1 + \sum_{n=1}^{\infty} a_n(s) z^n \quad \text{(for } s \in \mathbb{R}),$$

where the coefficient $a_n(s)$ is expressed in terms of the $2^{n+1}$ possible finite products of up to $n$ matrices, and writing

$$\lambda = -\frac{1}{2} \frac{\partial d(1, s)}{\partial t} \bigg|_{t=0} \left( \frac{\partial d(z, 0)}{\partial z} \bigg|_{z=1} \right)^{-1}.$$

This will be described in more detail in section 2.

Ultimately, the value of any computed numerical quantity depends on its known accuracy. This in turn depends on the bounds on the size of terms $a_n(t)$. A useful indication of the size of the expected error comes from the size of the coefficient $a_n(t)$. This is formalized in section 3, but has already been studied in [8, 10].

In this note we want to concentrate on the validated accuracy of these approximations to these coefficients, and thus to the Lyapunov exponent. There are what we shall call basic \textit{Euler bounds} coming directly from the method used in [10], perhaps suitably refined using different functional spaces as in [6, 8]. But the updated approach in [6] using what we shall call \textit{computed bounds} usually leads to better bounds or, consequently, to more accurate estimates on the Lyapunov exponents. We will elaborate on these different bounds in section 3.

In this brief note we compare these two approaches for estimating the error and illustrate the improvement in the error through a number of examples. This is especially pronounced when the matrices are less hyperbolic (i.e., their eigenvalues have absolute values close to 1). This is illustrated by the following example.

\textbf{Example 1.1.} Let us consider a parameterised family of a pair of positive matrices

$$A_1(t) = \begin{pmatrix} 1 + t & t \\ t & 1 \end{pmatrix} \quad \text{and} \quad A_2(t) = \begin{pmatrix} 1 & t \\ 1 & 1 + t \end{pmatrix}, \quad \text{for } t > 0.$$
2. An approach using determinants

To begin, we first describe the basic approach from [10] which is based on fixed points for contractions of the interval.

2.1. A dynamical formulation

Let $A_1, A_2$ be two $2 \times 2$ matrices with strictly positive entries. We associate to the matrices $A_1 = (a_{ij}^{(1)})_{i,j=1}^2$ and $A_2 = (a_{ij}^{(2)})_{i,j=1}^2$ two Möbius maps $T_1 : [0, 1] \to [0, 1]$ and $T_2 : [0, 1] \to [0, 1]$ given by

$$T_k(x) = \frac{(a_{11}^{(k)} - a_{12}^{(k)})x + a_{12}^{(k)}}{(a_{11}^{(k)} + a_{22}^{(k)}) - (a_{12}^{(k)} - a_{21}^{(k)})x + a_{12}^{(k)} + a_{22}^{(k)}} \quad \text{for } k = 1, 2.$$ 

These correspond to the projective maps associated to the linear actions of the matrices, and the positivity of the matrices ensures that $T_1$ and $T_2$ are both contractions of the interval. We next formally want to associate a (bi-complex) function of two variables $z$ and $s$ formally defined by

$$d(z, s) := \exp \left( -\sum_{n=1}^{\infty} \frac{z^n}{n} \sum_{|\underline{l}|=n} p_\underline{l} \frac{(T_{l_1} \cdots T_{l_n})(x)}{1 - (T_{l_1} \cdots T_{l_n})(x)} \right)$$

where we use the notation:

(a) $\underline{i} = (i_1, \ldots, i_n) \in \{0, 1\}^n$ and $|\underline{i}| = n$;
(b) $p_\underline{i} = p_{i_1} \cdots p_{i_n}$;
(c) $T_{\underline{i}} = T_{i_1} \circ \cdots \circ T_{i_n} : [0, 1] \to [0, 1]$; and

$n = 1, \ldots, 9$. We also show the different bounds for $a_{n=9}(0)$ we get using the traditional Euler bounds and the new computational bounds.

The improvement in the error term becomes particularly significant when $t > 0$ is small. This illustrates that the computed bounds are somewhat better when the hyperbolicity is weaker.

The new approach we describe for estimating the error terms in the Lyapunov exponent in terms of computational bounds is based on similar approaches for estimating other quantities in [6, 7]. In particular, the present manifestation helps to further illustrate the principle of using computed bounds to estimate the error.
Figure 1. \(z(s)\) is the first positive zero for \(z \mapsto d(z, s)\), i.e., \(d(z(s), s) = 0\).

It is easy to see from the definition that the function \(d(z, s)\) converges for \(|z|\) and \(\text{Re}(s)\) sufficiently large. The following result is essentially due to Ruelle [11] (see also corollary 1 in [6, 10]).

**Lemma 2.1 (after Ruelle).** The function \(d(z, s)\) extends as a bi-analytic function to all \((z, s) \in \mathbb{C} \times \mathbb{C}\). Moreover,

(a) For any \(s\) we have that the value at \(z = 0\) is \(d(0, s) = 1\); and

(b) For \(s \in \mathbb{R}^+\) the real valued function \(\mathbb{R}^+ \ni z \mapsto d(z, s) \in \mathbb{R}\) has a smallest real zero \(z(s)\) (i.e., \(d(z(s), s) = 0\)) and this is a simple zero (figure 1).

We now see by the implicit function theorem that \(s \mapsto z(s)\) is differentiable at \(s = 0\). We can express the Lyapunov exponent in terms of the function \(z(t)\). Rather than the original formula from [10] we use a simple alternative formulation due to Morris and Jurga [8]:

**Lemma 2.2.** Since matrices are invertible the Lyapunov exponent is given by \(\lambda = \frac{1}{2}z'(0)\).

We can use lemmas 2.1 and 2.2 and the implicit function theorem to write \(\lambda\) in terms of \(d(z, t)\):

\[
\lambda = -\frac{1}{2} \frac{\partial d(1, s)}{\partial s} \bigg|_{s=0} \left( \frac{\partial d(z, 0)}{\partial z} \right)_{z=1}^{-1},
\]

since \(z(0) = 1\).

**Remark 2.3.** The way in which we have associated to the matrices \(A_1\) and \(A_2\) the Mobius maps \(T_1\) and \(T_2\) is far from unique. Different choices of representations of projective space may lead to different maps, with slightly different properties, but will necessarily lead to the same value of the Lyapunov exponent.

### 2.2. Approximating determinants

Now that (2) gives a formal expression for the Lyapunov exponent in terms of the determinant, we need to use an approximation in order to obtain a numerical estimate.

By lemma 2.1, \(d(z, s)\) is biaanalytic and thus has a Taylor expansion in \(z\) at zero of the form

\[
d(z, s) = 1 + \sum_{n=1}^{\infty} a_n(s)z^n,
\]
where \( a_n \to 0 \) super exponentially as \( n \to +\infty \). For any given \( N \geq 1 \) we can consider the polynomial approximations

\[
d^{(N)}(z, s) = 1 + \sum_{n=1}^{N} a_n(s)z^n.
\]

From the definition of \( d(z, t) \) we see that the coefficients of \( d^{(N)}(z, s) \) can be computed using only the \( 2^{N+1} - 1 \) fixed points \( x_i \) where \( |i| \leq N \). Let \( z_N(s) > 0 \) be the smallest positive real zero of \( d^{(N)}(z, s) \) guaranteed by lemma 2.1. We can approximate \( \lambda \) by values \( \lambda_N \) defined by the following analogue of (2):

\[
\lambda_N = -\frac{1}{2} \frac{\partial d^{(N)}(1, s)}{\partial s} \bigg|_{s=0} \left( \frac{\partial d^{(N)}(z, 0)}{\partial z} \bigg|_{z=1} \right)^{-1}.
\]

In practice we would choose \( N \) as large as practicable to get that \( \lambda_N \) is a good approximation to \( \lambda \). We can write this explicitly as

\[
-\frac{1}{2} \sum_{m=1}^{N} a'_n(0) \sum_{m=1}^{N} a_n(0)
\]

(4)

Remark 2.4. To obtain the Taylor coefficients \( a_n(s) \) for \( d(z, s) \) as a function of \( z \) it is convenient to denote

\[
Z_n(s) := \frac{1}{n} \sum_{|i|=n} p_1(T'_i(x_i))^{y_i} / (T'_i(x_i))^{y_i}
\]

and apply the expansion \( e^{-x} = \sum_{k=0}^{\infty} (-x)^k / k! \) to \( d(z, s) = \exp \left( -\sum_{m=1}^{\infty} z^{m} Z_m(s) \right) \). Collecting together the powers of \( z \) gives an explicit formula:

\[
a_n(s) = \sum_{m_1 + \cdots + m_k = n} \frac{(-1)^k}{k!} \prod_{j=1}^{k} Z_{m_j}(s),
\]

where the summation is over the finite set of \( k \)-tuples \( (m_1, \ldots, m_k) \in \mathbb{N}^k \), for any \( 1 \leq k \leq n \). This expression can then be differentiated in \( s \) to get a formula for \( a'_n(s) \).

We can illustrate this method with two simple examples.

Example 2.5. Let \( A_1 = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \) and \( A_2 = \begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix} \). We can then consider the associated two contractions

\[
T_1(x) = \frac{x+1}{x+2} \quad \text{and} \quad T_2(x) = \frac{2x+1}{3x+2}.
\]

Let \( p_1 = p_2 = \frac{1}{2} \). Letting \( N = 10 \), say, we can estimate\(^1\)

\(^1\) We note that Bandtlow and Slipantschuk confirmed the same numerical value using a very different method involving Lagrange–Chebyshev expansions [2].
\[ \lambda \approx \lambda_N \]
\[ = 1.143\,311\,035\,041\,828\,694\,244\,499\,460\,846\,799\,991\,969\,390\,25 \ldots \]

We will return to the accuracy of this estimate in the next section.

**Example 2.6.** Let \( A_1 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \) and \( A_2 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \). We can then consider the associated two contractions
\[ T_1(x) = \frac{1}{10}x + 1 \quad \text{and} \quad T_2(x) = \frac{1}{10}x + 1. \]

Let \( p_1 = p_2 = \frac{1}{2} \). Letting \( N = 10 \), say, we can estimate
\[ \lambda \approx \lambda_N = 0.4660 \ldots \]

We will return to the accuracy of this estimate in the next section.

**Example 1.1 revisited.** We can associate to the matrices \( A_1(t) \) and \( A_2(t) \) the Möbius maps
\[ T_1(x) = \frac{tx + 1}{2(t-1)x + 2} \quad \text{and} \quad T_2(x) = \frac{(1-t)x + t}{(1-2t)x + 1 + 2t}. \]

Now that we have described the method of approximating the Lyapunov exponent using determinant functions we need to ask what confidence we should have in this estimate. The purpose of this note is to give the best possible error bounds. We address this in the next section.

### 3. The error estimates

We now concentrate on efficiently bounding the coefficients \( a_n(s) \) (for \( n \geq N \)) which is the key to estimating the difference between \( \lambda_N \) and \( \lambda \).

#### 3.1. Contraction rates

To formulate a basic bound on the coefficients \( a_n(s) \) (for \( n \geq N \)) we consider a disk \( D = D(c, r) \) in the complex plane \( \mathbb{C} \) centred at \( c \in \mathbb{R} \) and of radius \( r > 0 \) and the image disks \( T_1(D), T_2(D) \subset \mathbb{C} \). Assume these are contained in a concentric disk \( D' = D(c, r') > 0 \) with \( r' < r \):
\[ r' = \max \{|R - c|, |L - c|\} \]
where
\[ L = \min\{T_1(c \pm r), T_1(c \pm r')\} \]
and
\[ R = \max\{T_2(c \pm r), T_1(c \pm r')\}. \]

We can set \( \theta = \frac{r'}{r} \leq (0, 1) \). Typically, we would like to choose \( c, r \) so as to minimize the value of \( \theta \) (figure 2).

\[ \text{If this is not the case then we can modify the construction to use multiple disks to cover the interval [0, 1], although the resulting error bounds are often significantly worse.} \]
Example 2.5 revisited. In example 2.5 we can take the simplest choices $c = \frac{1}{2}$ and $r = \frac{1}{2}$ and then $\theta = \frac{1}{3}$. However, we can hope to improve this with different choices. For example, if we take $c = \frac{11}{20}$ and $r = \frac{4}{5}$ and then $\theta = 0.1893 \ldots$

Example 2.6 revisited. In example 2.6 we can take $c = \frac{1}{2}$ and $r = \frac{1}{2}$ and then $\theta = \frac{\sqrt{2}}{6} = 0.8333 \ldots$ However, increasing to $r = 0.618$ we have a modest improvement to $\theta = 0.8181 \ldots$

Example 1.1 revisited. In table 1 we took the same values $c = \frac{1}{2}$ and $r = \frac{1}{2}$ for each of the values of $t$.

3.2. Bounds on $a_n(s)$ terms

In this subsection we specify the bounds on the terms $a_n(s)$ in a fairly formal way. In the appendix A we sketch the ideas. We will also illustrate them with examples in the next section.

We begin by introducing the notation needed to describe the basic Euler error bounds on the coefficients $a_n(s)$.

Notation (basic Euler bounds). Let us denote

$$K_s = \sum_{j=1}^{2^n} \sup_{|z| = r} |T_j(z)|^{s \Re(z)}$$

and

$$E_n = \frac{\theta^{\mu(n+1)/2}}{\prod_{i=1}^{n}(1 - \theta^i)}$$

for $n \geq 1$ and $s \in \mathbb{R}$.

Thus defined, we can recall the original bounds on $a_n(s)$.

Lemma 3.1 (after Ruelle). We can bound $|a_n(s)| \leq K_s E_n$, for $n \geq 0$.

Proof. This essentially dates back to the original work of Ruelle and Grothendeick (cf [10]), although by working in the context of Hardy–Hilbert spaces one can streamline the estimates a little (see [1, 6, 8]), and is sufficient for large values of $n$. \qed
In those examples for which $0 < \theta < 1$ is sufficiently small the bounds given in lemma 3.1 will be reasonably effective for realistic values of $N$. The size of $N$ is determined by the practical considerations of computing $2^{M+1} - 1$ fixed points and the associated weights by derivatives$^3$.

We next introduce the notation needed to describe the new error bounds. This allows us to take advantage of estimates of other ingredients in the bounds which can be numerically computed to high numerical precision.

**Notation (computed constants).** Fix $M \geq N$. We define:

(a) $b_k(s) := r^{-2k} \int_0^1 \left( \sum_{j=1}^{2} p_j T_j(c + r e^{2\pi i t}) T_j(c + r e^{2\pi i t}) - c \right)^2 dt$, for $0 \leq k \leq M$;

(b) $\alpha_m(s) := \sqrt{\sum_{k=m-1}^{M} b_k + \left\| (T_1)' \right\|_\infty + \left\| (T_2)' \right\|_\infty} \frac{2}{\theta (M+1)}$, for $1 \leq m \leq M$; and

(c) $\beta_l(s) := \sum_{m_1 < \cdots < m_l \leq M} \prod_{j=1}^{M} \alpha_{m_j}(s)$ for each $0 \leq l \leq M$,

where $\left\| \cdot \right\|_\infty$ is the supremum norm over $D$.

The choice of $M$ is again dictated by practical considerations. However, the numerical evaluation of the integrals in the definition of $b_k(s)$ can generally be done to suitable precision for $M$ significantly larger than the previous choice of $N$.

We can now associate the values

$$\rho_n(s) := \sum_{r=1}^{n} \beta_r(s) \theta^{M(n-r)} K_4^{n-r} E_{n-r} + \theta^{Mn} K_4^n E_n$$

for $N + 1 \leq n \leq M$.

The expression for $\rho_n(s)$ may appear complicated but it can be readily accurately computed and leads to the following (compare with lemma 3.1).

**Theorem 3.2 (improved bounds on the $|a_n(s)|$).** We can bound the terms

$$|a_n(s)| \leq \begin{cases} \rho_n(s) & \text{for } N + 1 \leq n \leq M \\ K_4^n E_n & \text{for } M < n. \end{cases}$$

**Proof.** The bound for $n > M$ comes from lemma 3.1. The bound for $N + 1 \leq n \leq M$ is explained in the appendix A. $\square$

The improvement is that the bound for $N + 1 \leq n \leq M$ follows by adapting the analysis in [6] (and is better for smaller values of $n$).

**Example 2.5 revisited.** The first 20 values of $\rho_n(0)$ and $K_4^n E_n$ are given in table 2(i). For this particular example the matrices are relatively hyperbolic (i.e., with eigenvalues not close to 1) and the improvement in the bounds coming from theorem 3.2 is not particularly impressive.

**Example 2.6 revisited.** The first 20 values of $\rho_n(0)$ and $K_4^n E_n$ are given in table 2(ii). For this particular example the matrices are less hyperbolic than in the previous example and in this case the improvement in the bounds by using theorem 3.2 is more significant. For example, when $n = 12$ the computational error bounds are reasonably effective, in as much as the error term is significantly smaller than the principal term, whereas the basic bounds are not.

$^3$ For example, when $N = 10$ there are over 2000 values.
## Table 2. The values $\rho_n(0)$ and $K_0^n E_n$ for: (i) example 2.5; and (ii) example 2.6.

| $n$ | $\rho_n(0)$          | $K_0^n E_n$          |
|-----|----------------------|----------------------|
| (i) |                      |                      |
| 1   | 2.349 106            | 2.512 430            |
| 2   | 0.680 373            | 1.004 725            |
| 3   | 0.029 410            | 0.073 834            |
| 4   | 0.000 205            | 0.001 021            |
| 5   | $2.449 775 \times 10^{-7}$ | $2.672 299 \times 10^{-6}$ |
| 6   | $5.134 955 \times 10^{-11}$ | $1.323 159 \times 10^{-11}$ |
| 7   | $1.935 018 \times 10^{-15}$ | $1.240 149 \times 10^{-15}$ |
| 8   | $1.328 091 \times 10^{-20}$ | $2.200 306 \times 10^{-20}$ |
| 9   | $1.673 571 \times 10^{-26}$ | $7.389 965 \times 10^{-26}$ |
| 10  | $3.891 633 \times 10^{-33}$ | $4.698 423 \times 10^{-30}$ |
| 11  | $1.675 536 \times 10^{-40}$ | $5.654 738 \times 10^{-37}$ |
| 12  | $1.338 893 \times 10^{-48}$ | $1.288 319 \times 10^{-44}$ |
| 13  | $1.989 254 \times 10^{-57}$ | $5.556 291 \times 10^{-53}$ |
| 14  | $5.503 096 \times 10^{-67}$ | $4.536 252 \times 10^{-62}$ |
| 15  | $2.837 980 \times 10^{-77}$ | $7.010 678 \times 10^{-72}$ |
| 16  | $2.731 066 \times 10^{-88}$ | $2.051 036 \times 10^{-82}$ |
| 17  | $4.908 542 \times 10^{100}$ | $1.135 893 \times 10^{-93}$ |
| 18  | $1.648 925 \times 10^{112}$ | $1.190 836 \times 10^{-105}$ |
| 19  | $1.036 026 \times 10^{125}$ | $2.363 291 \times 10^{-118}$ |
| 20  | $1.218 215 \times 10^{139}$ | $8.878 369 \times 10^{-132}$ |
| (ii)|                      |                      |
| 1   | 6.452 893            | 19.119 303           |
| 2   | 17.138 542           | 82.243 72            |
| 3   | 25.441 420           | 211.550 542          |
| 4   | 23.983 698           | 364.860 167          |
| 5   | 15.400 215           | 448.596 484          |
| 6   | 7.037 361            | 408.266 147          |
| 7   | 2.357 457            | 282.010 308          |
| 8   | 0.591 300            | 150.470 209          |
| 9   | 0.112 795            | 62.813 263           |
| 10  | 0.016 5601           | 20.711 634           |
| 11  | 0.001 888            | 5.433 736            |
| 12  | 0.000 168            | 1.140 633            |
| 13  | 0.000 011            | 0.192 427            |
| 14  | $6.576 476 \times 10^{-7}$ | 0.026 180            |
| 15  | $2.900 652 \times 10^{-8}$ | 0.002 880            |
| 16  | $1.019 080 \times 10^{-9}$ | 0.000 256            |
| 17  | $2.859 584 \times 10^{-11}$ | 0.000 0185           |
| 18  | $6.423 360 \times 10^{-13}$ | 1.094 563 $\times 10^{-6}$ |
| 19  | $1.157 216 \times 10^{-14}$ | 5.244 598 $\times 10^{-8}$ |
| 20  | $1.674 791 \times 10^{-16}$ | 2.047 446 $\times 10^{-9}$ |
3.3. Bounds on $|a'_n(0)|$ and $\sum_{n=1}^{N} n a_n(0)$

Comparing (2) and (4) we see that in order to estimate $\lambda$ we first require the bounds

$$
\sum_{n=N+1}^{\infty} n a_n(0) \leq \sum_{n=N+1}^{M} n \rho_n(s) + \sum_{n=M+1}^{\infty} K_0^n n E_n \quad \text{and}
$$

$$
\sum_{n=n+1}^{\infty} a'_n(0) \leq \frac{1}{\delta} \sum_{n=n+1}^{M} \rho_n(s) + \frac{1}{\delta} \sum_{n=M+1}^{\infty} K_0^n n E_n
$$

using Cauchy’s theorem. As one easily observes, choosing $0 < \theta < 1$ as small as possible helps in making the bound better. This, in turn, depends on the choice of $c$ and $r$.

More precisely, we can write

$$
\left| \frac{\sum_{n=1}^{N} a'_n(0)}{\sum_{n=1}^{N} n a_n(0)} \right| \leq \frac{1}{\sum_{n=1}^{N} n a_n(0)} \left( \sum_{n=1}^{N} |a'_n(0)| \right)
$$

$$
+ \frac{\sum_{n=1}^{N} |a'_n(0)|}{\sum_{n=N+1}^{\infty} n |a_n(0)|} \sum_{n=N+1}^{\infty} n |a_n(0)|.
$$

Obtaining upper bounds on the values $|a'_n(0)|$ for the numerator of the expression in (6) comes from a standard application of Cauchy’s theorem, i.e.,

$$
|a'_n(0)| = \left| \frac{1}{2\pi i} \int_{|z|=\delta} \frac{a_n(s)}{s^2} \, ds \right| \leq \frac{1}{\delta} \sup_{|z|=\delta} |a_n(s)|
$$

for any suitably small $\delta > 0$.

Since $|\sum_{n=1}^{\infty} n a_n(0)| = |\frac{d\lambda(0)}{d\lambda} \sum_{n=1}^{\infty} n a_n(0)| > 0$ (we will give a more explicit lower bound at the end of the appendix A) and, for sufficiently large $N$, we can write

$$
\left| \sum_{n=1}^{N} n a_n(0) \right| \geq \left| \frac{\partial d(z_0)}{\partial z} \right|_{z=1} - \sum_{n=N+1}^{\infty} n |a_n(0)|
$$

$$
\geq \left| \frac{\partial d(z_0)}{\partial z} \right|_{z=1} - \sum_{n=N+1}^{\infty} n K_0^n E_n > 0
$$

we see that the terms in the denominator can be made non-zero.

**Remark 3.3.** For the inequality in (6) to be useful one requires, in particular, an effective lower bound on $|\sum_{n=1}^{\infty} n a_n(0)|$ for the examples under consideration. However, in practise the triangle inequality gives a lower bound

$$
\sum_{n=1}^{\infty} n a_n(0) \geq \sum_{n=1}^{N} n a_n(0) - \sum_{n=N+1}^{\infty} n |a_n(0)|
$$

$$
\geq \sum_{n=1}^{N} n a_n(0) - \sum_{n=M+1}^{\infty} n \rho_n(0) - \sum_{n=M+1}^{\infty} n K_0^n E_n,
$$

where we have used the estimates in theorem 3.2 to bound the tail of the series.
The main term $|\sum_{n=1}^{N} na_n(0)|$ on the last line of (8), which also appears on the right-hand side of (6), can be numerically computed in specific examples. The remaining two terms are effectively bounded, and in the case of the examples we consider they are considerably smaller than the first term and, in particular, this was used in the bounds for the Lyapunov exponents for the examples in the next subsection.

3.4. Bounds on the Lyapunov exponents

Finally, we can use (6) and lemma 2.2 to write

$$|\lambda - \lambda_N| \leq \frac{1}{2} \sum_{n=1}^{\infty} |\sum_{n=1}^{N} |a_n(0)| \sum_{n=N+1}^{\infty} n|a_n(0)|$$

for large $N$.

**Example 2.5 revisited again.** In this example we can take $N = 9$, $M = 100$ and $\delta = \frac{1}{20}$, say, and estimate

$$\lambda = 1.143311035102949245843251853655882994025 \ldots$$

$$\pm 3 \times 10^{-21}.$$

**Example 2.6 revisited again.** In this example we can take $N = 12$, $M = 100$ and $\delta = \frac{1}{20}$ and estimate

$$\lambda = 0.4660 \pm 0.003.$$

Although the error bound in this case is significantly worse than in the previous example it replaces an ineffective Euler bound.

**Acknowledgment**

The author gratefully acknowledges support from “ERC-Advanced Grant 833802-Resonances and EPSRC-EP/T001674/1”.

**Appendix A. Proof of theorem 3.2**

We present the proof of theorem 3.2, which is based on estimates in [6].

**A.1. The Hilbert space and the operators**

We begin by defining the Hilbert space of functions we will consider. We can consider the Hardy–Hilbert space $\mathcal{H}^2 = \mathcal{H}^2(D)$ of analytic functions $f : D \rightarrow \mathbb{C}$ with the norm

$$\|f\|^2 = \int_0^1 |f(c + re^{2\pi i t})|^2 \, dt.$$

Equivalently, writing $f(z) = \sum_{n=0}^{\infty} c_n m_n(z) \in \mathcal{H}^2$, where $m_n(z) := (z - c)^n r^{-n}$, we have that $\|f\|^2 = \sum_{n=0}^{\infty} |c_n|^2$.

We next want to define the linear operators on $\mathcal{H}$ that are at the heart of the proof. We can associate to the maps $T_1$ and $T_2$ linear operators on $\mathcal{H}^2 = \mathcal{H}^2(D)$ defined by

$$L_2 w(z) = p_1 T_1'(z) w(T_1 z) + p_2 T_2'(z) w(T_2 z).$$
where \( s \in \mathbb{R} \). The operators \( L_s \) are trace class and we can write \( d(z, s) = \det(I - zL_s) \) [6]. Moreover, we can bound the coefficients by

\[
|a_n(s)| \leq \sum_{i_1 < \cdots < i_n} A_{i_1}(L_s) \ldots A_{i_n}(L_s), \tag{A1}
\]

for \( n \geq 1 \), where the summation is over indices \( i_1, \ldots, i_n \in \mathbb{N} \) satisfying \( i_1 < \cdots < i_n \) and

\[
A_i(L_s) = \inf \{ \| L_s - \kappa \| : \text{rank}(\kappa) \leq i - 1 \},
\]

for \( i \in \mathbb{N} \), denotes the \( i \text{th approximation number} \). Here the infimum is taken over all linear operators of rank at most \( k - 1 \).

### A.2. The bounds on the approximation numbers

We have two different bounds on the approximation numbers.

(a) We have the basic bound \( A_n(L_s) \leq K_s \theta^n \), for \( n \geq 1 \) (see lemma 2.1, and corollary 1 in [6]).

(b) We have a better bound on \( A_n(L_s) \), for \( n \geq 1 \). Let \( \kappa_n: \mathcal{H}^2 \to \mathcal{H}^2 \) be the orthogonal projection onto the \( n \)-dimensional space

\[
E_n = \text{span}\{ m_k(z) := (z - c)^j r^k : 0 \leq k \leq n - 1 \}
\]

(where we call that \( c \) and \( r \) are the centre and radius of \( D \), respectively) then \( A_n(L_s) \leq \| L_s(I - \kappa_n) \| \). For any \( f(z) = \sum_{k=0}^{\infty} c_k m_k \in \mathcal{H} \), a simple application of the Cauchy–Schwarz inequality gives that

\[
\| L_s(I - \kappa_n)(f) \| \leq \left\| \sum_{k=n-1}^{\infty} c_k L_s(m_k) \right\| \leq \sum_{k=n-1}^{\infty} |c_k| \| L_s(m_k) \|
\]

\[
\leq \sqrt{\sum_{k=n}^{\infty} \| L_s(m_k) \|^2 \sum_{k=n}^{\infty} |c_k|^2} \leq \| f \|. \tag{A2}
\]

In particular, combining this bound with the definition of \( A_n(L_s) \) gives

\[
A_n(L_s) \leq \sqrt{\sum_{k=n-1}^{\infty} \| L_s(m_k) \|^2}. \tag{A2}
\]

We can write

\[
b_k(s) = \| L_s(m_k) \|^2 = \int_0^1 \left| p_1 T_1'(c + r e^{2\pi s}) m_k(T_1(c + r e^{2\pi s})) + p_1 T_2'(c + r e^{2\pi s}) m_k(T_2(c + r e^{2\pi s})) \right|^2 \, dt
\]

(which was introduced in section 3.2) and then we can bound the tail of the infinite series appearing in (A2) using
\[ b_k(s) \leq \frac{1}{r^n} \left( \max_{j=1,2} \sup_{|z-c| \leq r} |T_j^\prime(z)|^{\Re(s)} \right)^2 \left( \sup_{|z-c| \leq 2r} |z-c|^4 \right)^2 \]
\[ \leq \left( \max_{j=1,2} \sup_{|z-c| \leq r} |T_j^\prime(z)|^{\Re(s)} \right)^2 \left( \frac{\theta r}{2} \right)^{2k} \]
\[ = C \theta^{2k} \]

for \( k \geq M \), where \( C := \left( \sum_{j=1}^2 \sup_{|z-c| \leq r} |T_j^\prime(z)|^{\Re(s)} \right)^2 \). We can then bound the terms \( b_k(s) \) for \( k = M+1, \ldots, \infty \):
\[ \sum_{k=M+1}^{\infty} b_k(s) \leq \sum_{k=M+1}^{\infty} C \theta^{2k} = \frac{C \theta^{2M+1}}{1 - \theta^2} \tag{A3} \]

(compare with lemma 5 in [6]). Combining (A2) and (A3) gives the bound \( A_n(L_s) \leq \alpha_n(s) \).

We can now rewrite the right-hand side of (A1) as:
\[ |a_n(s)| \leq \sum_{i_1 < \cdots < i_n} A_{i_1}(L_{i_1}) \cdots A_{i_n}(L_{i_n}) \]
\[ \leq \min_{\sum_{i=1}^{M+1} M_{i}} \left( \sum_{i_1 < \cdots < i_{M+1}} A_{i_1}(L_{i_1}) \cdots A_{i_{M+1}}(L_{i_{M+1}}) \right) \left( \sum_{M+1 < i_{M+1} < \cdots < i_n} A_{i_{M+1}}(L_{i_{M+1}}) \cdots A_{i_n}(L_{i_n}) \right) \]
\[ + \left( \sum_{M \leq i_1 < \cdots < i_n} A_{i_1}(L_{i_1}) \cdots A_{i_n}(L_{i_n}) \right) \]
\[ \leq \min_{\sum_{i=1}^{M+1} M_{i}} \left( \sum_{i_1 < \cdots < i_{M+1}} \alpha_{i_1}(s) \cdots \alpha_{i_{M+1}}(s) \right) \left( \sum_{M+1 < i_{M+1} < \cdots < i_n} K_i \theta^{i_{M+1}} \cdots \theta^{i_n} \right) \]
\[ + \left( \sum_{M \leq i_1 < \cdots < i_n} K_i \theta^{i_1} \cdots \theta^{i_n} \right) \]
\[ = \min_{\sum_{i=1}^{M+1} M_{i}} \left( \sum_{i_1 < \cdots < i_{M+1}} \beta_{i_1}(s) K_{i_1}^{\theta^{i_1}} \theta^{M_{i_1}} \right) \left( \sum_{M+1 < i_{M+1} < \cdots < i_n} \theta^{i_{M+1}} \cdots \theta^{i_n} \right) \left( \sum_{i_1 < \cdots < i_n} \theta^{i_1} \cdots \theta^{i_n} \right) \]

When \( n \leq M \) this completes the proof of theorem 3.2. \( \square \)

**Remark 3.4.** It is possible to further improve the bounds, at the cost of introducing more complicated expressions. In particular, we can get an even more refined estimate if we introduce the following notation. Let \( N \leq Q \leq M \)

(a) \( J(s) := K_s \sqrt{1 + \theta^{2M+2-Q}} \)
(b) \( \frac{\alpha_i}{\beta_i}(t) := \beta_i(t) + \sum_{i=0}^{M-1} J^{M-i}(t) \beta_{M-i}(t) \theta^{M-i} E_{n-i} \).

We can then bound the terms
\[ |a_n(s)| \leq \min \left\{ K_s^\alpha E_{n, \frac{\beta}{\alpha}}(s) \right\} . \]
The proof again follows the lines of the arguments in [6] and we omit the details.

A lower bound on $\sum_{n=1}^{N} n \lambda_n(0)$. Finally, we return to the lower bound from section 3.3. If the transfer operator $L_{s}$ has eigenvalues $\{\lambda_n(s)\}_{n=1}^{\infty}$ with $|\lambda_1(s)| \geq |\lambda_2(s)| \geq \cdots$ then we can write $d(z, s) = \prod_{n=1}^{\infty} (1 - z \lambda_n(s))$ and

$$\frac{\partial d(z, 0)}{\partial z} \bigg|_{z=1} = \prod_{n=2}^{\infty} (1 - \lambda_n(0))$$

(cf [11] and lemma 4.7 in [8]). Using Weyl’s inequality and the basic bound on approximation numbers

$$|\lambda_n(0)| \leq \prod_{k=1}^{n} |\lambda_k(0)|^{1/n} \leq \prod_{k=1}^{n} A_{k}(L_{0})^{1/n} \leq K_{0} \left( \prod_{k=1}^{n} \theta^{k} \right)^{1/n}$$

$$= K_{0} \theta^{(n+1)/2} < \frac{1}{2}$$

for $n > n_0 := \left\lfloor \frac{2 \log(2 K_{0})}{\log \theta} \right\rfloor$. Finally, we can bound

$$\left| \frac{\partial d(z, 0)}{\partial z} \bigg|_{z=1} \right| \geq \prod_{k=2}^{n_0} |1 - \lambda_k(0)| \prod_{n=n_0+1}^{\infty} \left( 1 - \frac{1}{2} \theta^{(n-n_0)/2} \right)$$

$$\geq |1 - \lambda_2(0)|^{n_0-1} \prod_{n=n_0+1}^{\infty} \left( 1 - \frac{1}{2} \theta^{(n-n_0)/2} \right) > 0$$

as required.

References

[1] Bandtlow O F and Jenkinson O 2008 On the Ruelle eigenvalue sequence Ergod. Theor. Dynam. Syst. 28 1701–11
[2] Bandtlow O and Slipantschuk J 2020 Lagrange approximation of transfer operators associated with holomorphic data (arXiv:2004.03534)
[3] Furstenberg H 1963 Noncommuting random products Trans. Am. Math. Soc. 108 377
[4] Furstenberg H and Kesten H 1960 Products of random matrices Ann. Math. Stat. 31 457–69
[5] Jacquet P, Serroussi G and Szpankowski W 2008 On the entropy of a hidden Markov process Theor. Comput. Sci. 395 203–19
[6] Jenkinson O and Pollicott M 2018 Rigorous effective bounds on the Hausdorff dimension of continued fraction Cantor sets: a hundred decimal digits for the dimension of $E_2$ Adv. Math. 325 87–115
[7] Jenkinson O, Pollicott M and Vytnova P 2018 Rigorous computation of diffusion coefficients for expanding maps J. Stat. Phys. 170 221–53
[8] Jurga N and Morris I 2019 Effective estimates on the top Lyapunov exponents for random matrix products Nonlinearity 32 4117–46
[9] Oseledets V I 1968 A multiplicative ergodic theorem. Characteristic Lyapunov, exponents of dynamical systems Tr. Mosk. Mat. Obsc. 19 179–210
[10] Pollicott M 2010 Maximal Lyapunov exponents for random matrix products Invent. Math. 181 209–26
[11] Ruelle D 1976 Zeta-functions for expanding maps and Anosov flows Invent. Math. 34 231–42