GROWTH IN PRODUCTS OF MATRICES: FASTEST, AVERAGE, AND GENERIC

VLADIMIR SHPILRAIN

ABSTRACT. The problems that we consider in this paper are as follows. Let A and B be 2×2 matrices (over reals). Let $w(A, B)$ be a word of length n. After evaluating $w(A, B)$ as a product of matrices, we get a 2×2 matrix, call it W. What is the largest (by the absolute value) possible entry of W, over all $w(A, B)$ of length n, as a function of n? What is the expected absolute value of the largest (by the absolute value) entry in a random product of n matrices, where each matrix is A or B with probability 0.5? What is the Lyapunov exponent for a random matrix product like that? We give partial answer to the first of these questions and an essentially complete answer to the second question. For the third question (the most difficult of the three), we offer a very simple method to produce an upper bound on the Lyapunov exponent in the case where all entries of the matrices A and B are nonnegative.

1. INTRODUCTION

The main problem that we consider in this paper is as follows. Let A and B be 2×2 matrices (over reals). Let $w(A, B)$ be a word of length n. After evaluating $w(A, B)$ as a product of matrices, we get a 2×2 matrix, call it W. What is the largest (by the absolute value) possible entry of W, over all $w(A, B)$ of length n, as a function of n? The latter function is usually (although not always) exponential, so we are looking for an answer in the form $O(s^n)$, i.e., we are looking for the base s of the exponent; this is what we call the *growth rate*, and this is sometimes also called the *joint spectral radius* of the pair (A, B) of matrices, see e.g. [13].

For us, the original motivation to address this question came from the problem of estimating the girth of the Cayley graph of 2-generator (semi)groups of matrices over $\mathbb{Z}_p$. If A and B generate a free sub(semi)group of $SL_2(\mathbb{Z})$, then there cannot be any relations of the form $u(A, B) = v(A, B)$ in $SL_2(\mathbb{Z}_p)$ unless at least one of the entries of the matrix $u(A, B)$ or $v(A, B)$ is at least p. Thus, if the largest entry in a product of n matrices is of the size $O(s^n)$, then the girth of the Cayley graph of the sub(semi)group of $SL_2(\mathbb{Z}_p)$ generated by A and B is $O(\log_s p)$. We note that the problem of estimating the girth of the Cayley graph of 2-generator (semi)groups of matrices in $SL_2(\mathbb{Z}_p)$ got considerable attention from number theorists, see e.g. [2], [7] [9], [14].

The problem of bounding the girth of the Cayley graph of a 2-generator (semi)group is also relevant to security properties of *Cayley hash functions*, see e.g. [1], [3], [15], [25]. A hash function takes, as input, a bit string of an arbitrary length and outputs a bit string of a fixed length (say, 512 or 1024 bits). Cayley hashing is based on a simple idea of using a pair of elements, A and B, of a semigroup S, to hash the "0" and the "1" bit, respectively. Then a bit string is hashed to a product of elements in the natural way. For example, the bit string 1001011 will be hashed to the element $BAABABB$ of a semigroup S. We refer to

[23] for a more detailed discussion on relations between security properties of a Cayley hash function and the girth of the corresponding Cayley graph.

In Section 3, we show how one can compute the *average* growth of entries in a product of n matrices. This turns out to be the easiest of the three problems (fastest, average, generic growth) and is reduced to solving a system of linear recurrence relations with constant coefficients.

A more difficult problem, of interest in the theory of stochastic processes, is to estimate the growth rate of entries in a random product of n matrices over reals, each of which is either A or B with probability $\frac{1}{2}$ (this is what we call *generic* growth rate). For strictly positive matrices, Pollicott [20] reported an algorithm to estimate the Lyapunov exponent $\lambda = \log s$ with any desired precision. Here s is the same as in the first paragraph of the Introduction, and log denotes the natural logarithm. We touch upon this problem in Section 4, although our theoretical contribution in this direction is rather modest: for matrices with nonnegative entries, our method in Section 3 provides an upper bound for the Lyapunov exponent, see Section 4.1. We note that altogether different (analytical) methods for bounding the Lyapunov exponent were used in [10], [21] and [24].

Finally, we mention that studying growth of entries in a product of matrices from a fixed finite set is important for understanding algorithmic complexity (worst-case, average-case, and generic-case) of the word problem in finitely generated groups of matrices, see [18].

2. Fastest Growth

Here are the main two open problems that we address in this section. In what follows, we call a word $w(A, B)$ *positive* if it includes A and B with positive exponents only. In what follows, matrices can be considered over any subring of reals (e.g. $\mathbb{Z}$ or $\mathbb{Q}$).

Problem 1. *Given two 2×2 matrices A and B, find the largest (by the absolute value) possible entry of a matrix $w(A, B)$, as a function of the word length $n = |w|$, over all positive words w of length n.*

For matrices over $\mathbb{Z}$, theoretical as well as experimental evidence suggest that words w that provide fastest growth of matrix entries in this context are *periodic*, i.e., are proper powers. This prompts the following

Conjecture 1. *Given any two 2×2 matrices A and B over $\mathbb{Z}$ (or, more generally, over $\mathbb{Q}$), the fastest growth of the largest (by the absolute value) entry of a matrix $w(A, B)$ over all positive words w of length n is provided by periodic words w. More accurately, for any two 2×2 matrices A and B, there is an integer $r \geq 2$ and a positive word $v(A, B)$ of length r such that the fastest growth of the largest (by the absolute value) entry of a matrix $w(A, B)$ over all positive words $w(A, B)$ of length n divisible by r is provided by words $w(A, B) = (v(A, B))^{\frac{n}{r}}$.*

One gets a somewhat weaker conjecture upon replacing "periodic words" by "eventually periodic words", i.e., words of the form $u(A, B)(v(A, B))^{\frac{n}{r}}$ for fixed $u(A, B)$ and $v(A, B)$.

Conjecture 1 is actually a very special case of a more general Lagarias-Wang finiteness conjecture [13] formulated for arbitrary sets of $d \times d$ matrices over real numbers. A counterexample to the Lagarias-Wang conjecture was constructed in [8]; it is a pair of matrices $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, $B = \begin{pmatrix} \alpha & 0 \\ \alpha & \alpha \end{pmatrix}$, where α is an irrational number whose definition is too complex to reproduce here. After rounding to 3 decimal places, α becomes 0.749.

There is still a possibility that the finiteness conjecture holds for matrices over $\mathbb{Q}$ or at least over $\mathbb{Z}$. In fact, in [19] it was proved that the finiteness conjecture holds for matrices from $SL_2(\mathbb{Z})$ with nonnegative entries. Thus, our Conjecture 1 holds for matrices from $SL_2(\mathbb{Z})$ with nonnegative entries. In fact, the result of [19] is stronger as it establishes that, given two matrices $A, B \in SL_2(\mathbb{Z})$ with nonnegative entries, the fastest growth of entries in their products is provided by powers of A, or B, or AB, or A^2B, or AB^2. (Note that entries of powers of, say, AB and BA have the same growth rate since these matrices are conjugate.) While it is easy to find out, based on the maximum eigenvalue, powers of which particular matrix from a finite collection give the fastest growth, the authors of [19] give an interesting alternative simple method based on computing the trace of a couple of relevant matrices.

We note, in passing, that the latter result (about a particular collection of matrices that could possibly provide the fastest growth) does no longer hold if the matrices are in $GL_2(\mathbb{Z})$ but not in $SL_2(\mathbb{Z})$, see our Section 2.4.

We note though that if at least one of the matrices A or B has at least one negative entry, the situation becomes more mysterious, see Section 2.3. There we have yet another example of a pair of matrices where the fastest growth of the entries is *not* provided by powers of A, or B, or AB, or A^2B, or AB^2.

Below we summarize, in more detail, what is known about the growth of entries in 2-generator semigroups of matrices, i.e., in matrices of the form $w(A, B)$, for various popular instances of A and B.

2.1. **Matrices $A(k)$ and $B(m)$.** Let k and m be real numbers. Denote $A(k) = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$, $B(m) = \begin{pmatrix} 1 & 0 \\ m & 1 \end{pmatrix}$.

First we note that if $k \cdot m \geq 4$, then the semigroup generated by $A(k)$ and $B(m)$ is free. Otherwise, it is a big open problem for what values of k and m this (semi)group is free; see e.g. [12], [22] and references therein. We just mention here that there is no pair (k, m) of rational numbers $0 < k, m < 2$ is known such that the *group* generated by $A(k)$ and $B(m)$ is free. The relevance of freeness of the *semigroup* to questions about products of matrices will be mentioned in Section 4.

1. In [3], it was proved that the maximum growth in products of n matrices of the form $w(A(k), B(k))$ for integers $k \geq 2$ is achieved by the words $w = (A(k)B(k))^{\frac{n}{2}}$ (assuming that n is even). The proof in [3] actually goes through for $k = 1$ as well. The reason why the case $k = 1$ was ignored in [3] is that the *group* generated by $A(1)$ and $B(1)$ is not free. However, what matters for us in the present paper is that the *semigroup* generated by $A(1)$ and $B(1)$ is free.

With these facts established, it is not hard to precisely evaluate the maximum growth rate of the matrix entries in $w(A(k), B(k))$ for various values of $k \geq 1$. This can be done either by solving a system of recurrence relations for the entries of $(A(k)B(k))^{\frac{n}{2}}$ (this is how it was done in [3]) or by computing the eigenvalues of the matrix $A(k)B(k)$.

We use the latter method here. Compute $A(k)B(k) = \begin{pmatrix} 1+k^2 & k \\ k & 1 \end{pmatrix}$. The largest eigenvalue of this matrix is $s = \frac{1}{2}(k^2 + 2 + k\sqrt{k^2+4})$.

In particular, for $k = 2$ the largest eigenvalue of $A(2)B(2)$ is $3 + \sqrt{8}$. Then the largest entry of $(A(2)B(2))^{\frac{n}{2}}$ is $O((\sqrt{3+\sqrt{8}})^{\frac{n}{2}}) = O((1+\sqrt{2})^n)$. Note that $1+\sqrt{2} \approx 2.414$.

In fact, [3] gives a precise formula for the largest entry in $(A(2)B(2))^{\frac{n}{2}}$. It is $(\frac{1}{2}+\frac{1}{\sqrt{8}})(1+\sqrt{2})^n + (\frac{1}{2}-\frac{1}{\sqrt{8}})(1-\sqrt{2})^n$.

2. As one would expect, the maximum growth of the entries in matrices $w(A(1), B(1))$ is the slowest among all $w(A(k), B(k))$ for integers $k \geq 1$. The largest entries in the corresponding matrices $(A(1)B(1))^{\frac{n}{2}}$ are $O((\frac{3}{2}+\frac{\sqrt{5}}{2})^{\frac{n}{2}})$. Note that $\sqrt{\frac{3}{2}+\frac{\sqrt{5}}{2}} \approx 1.618$.

3. In [7], the authors have determined the maximum growth of the entries in matrices $w(A(k), B(m))$ for all *integer* $k, m \geq 1$.

Now we establish the following generalization of the above mentioned results from [3] and [7]:

Theorem 1. *The maximum growth of the largest entry in a product of n matrices of the form $w(A(k), B(m))$ for any real numbers $k, m \geq 2$ is achieved by the words $(A(k)B(m))^{\frac{n}{2}}$ and $(B(m)A(k))^{\frac{n}{2}}$ (assuming that n is even). The same holds for $k = m = 1$.*

Proof. First we consider the case $k, m \geq 2$. We are going to use induction by the length of $w = w(A(k), B(m))$, as follows. Let w have length $n + 2$, $n \geq 2$, and let u be the prefix of w of an even length n. Without loss of generality, we can assume that u ends with $A(k)$.

Let the matrix $M = u(A(k), B(m))$ be $\begin{pmatrix} x & y \\ z & t \end{pmatrix}$, where x, y, z, t obviously are nonnegative numbers. Moreover, we can assume that M is not a power of $A(k)$ or $B(m)$ since in those cases, the growth of the largest entry is the slowest (linear). Under this assumption, all the entries of M are going to be strictly positive.

Since we assumed that u ends with $A(k)$, we have $y \geq kx, t \geq kz$, so in particular, the largest entry of the matrix M is in the second column. If u begins with $B(m)$, then the largest entry of the matrix M is in the second row, i.e., it is t. Otherwise, it is y.

Now we are going to list all 4 possible ways to get to w from u by multiplying u by two matrices, $A(k)$ or $B(m)$.

1. $MA(k)A(k) = \begin{pmatrix} x & y+2kx \\ z & t+2kz \end{pmatrix}$.

2. $MB(m)B(m) = \begin{pmatrix} x+2my & y \\ z+2mt & t \end{pmatrix}$.

3. $MB(m)A(k) = \begin{pmatrix} x+my & kx+(km+1)y \\ z+mt & kz+(km+1)t \end{pmatrix}$.

4. $MA(k)B(m) = \begin{pmatrix} (km+1)x+my & y+kx \\ (km+1)z+mt & kz+t \end{pmatrix}$.

We claim that the matrix $MB(m)A(k)$ has the largest entries of the 4 matrices above. To begin, compare the entry $kx+(km+1)y$ of the matrix 3 to the entry $y+2kx$ of the matrix 1. (This is the only non-obvious comparison between the entries of the matrices 3 and 1.) Since we are under the assumption $kx \leq y$, we have $y + 2kx \leq 3y$, whereas $kx + (km + 1)y > 3y$ since $km + 1 > 3$.

Comparison between the entries of the matrices 3 and 2 is similar. Here it is sufficient to notice that $km + 1 > 2m$.

Now to the most nontrivial comparison, between the entries of the matrices 3 and 4. First let us compare $x+my$ to $y+kx$. From the assumption $kx \le y$, we get $y+kx \le 2y$, whereas $x + my > 2y$. Thus, $x + my > y + kx$. In a similar way, we get $z + mt > kz + t$.

Then, let us compare $kx + (km + 1)y$ to $(km + 1)x + my$. Since $kmx \le my$, we have $(km + 1)x + my \le x + 2my$, whereas $kx + (km + 1)y > x + 2my$ because $k > 1$ and $km + 1 > 2$. Therefore, $kx + (km + 1)y > (km + 1)x + my$. In a similar way, we get $kz + (km + 1)t > (km + 1)z + mt$.

Thus, not only is the norm of the matrix 3 greater than that of the matrix 4, but the entries of the second column of the matrix 3 are strictly larger than the entries of the first column of the matrix 4, and the entries of the first column of the matrix 3 are strictly larger than the entries of the second column of the matrix 4.

The bottom line is: when comparing the matrix 3 to any other matrix K on the list, there is a permutation on the columns of K such that after this permutation, the entries of 3 are not less than the corresponding entries of K, and there is at least one entry of 3 that is strictly larger than the corresponding entry of K.

Now we get to a special case $k = m = 1$. In this case, our 4 matrices look as follows.

1. $MA(1)A(1) = \begin{pmatrix} x & y+2x \\ z & t+2z \end{pmatrix}$.

2. $MB(1)B(1) = \begin{pmatrix} x+2y & y \\ z+2t & t \end{pmatrix}$.

3. $MB(1)A(1) = \begin{pmatrix} x+y & x+2y \\ z+t & z+2t \end{pmatrix}$.

4. $MA(1)B(1) = \begin{pmatrix} 2x+y & y+x \\ 2z+t & z+t \end{pmatrix}$.

We are still under the assumptions $kx \le y$ and $kz \le t$, i.e., $x \le y$ and $z \le t$ in this case. Note that we cannot have $x = y$ and $z = t$ at the same time since this would make the matrix M singular, which is impossible. Without loss of generality, we can assume that $x < y$. Also, recall that under our assumptions x, y, z, t are all strictly positive.

Comparison between the matrices 3 and 2 is obvious. To compare the matrices 3 and 1, we note that $x + 2y > y + 2x$. Indeed, taking everything to the left gives $y - x > 0$, which holds under our assumptions. Similarly, $z + 2t \ge t + 2z$. Then, $x + y > x$ and $z + t > z$ because $y, t > 0$.

When comparing the matrices 3 and 4, we see that the first column of the matrix 3 is the same as the second column of the matrix 4. Comparing entries in the second column of 3 to those in the first column of 4, we have $x + 2y > y + 2x$ and $z + 2t \ge 2z + t$.

Again, the bottom line is: when comparing the matrix 3 to any other matrix K on the list, there is a permutation on the columns of K such that after this permutation, the entries of 3 are not less than the corresponding entries of K, and there is at least one entry of 3 that is strictly larger than the corresponding entry of K. This completes the proof.

□

2.2. **Other pairs of matrices.** Our Theorem 1 can help handle some other pairs of matrices as well. For example, let $X = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, $Y = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$. Then, let $U = XY = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$, $V = YX = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$.

Then the fastest growth of the entries in a product of the matrices U and V is given by matrices U^n or V^n because in these products, the matrices X and Y alternate and this gives the fastest growth according to Theorem 1. The largest eigenvalue of either matrix (U or V) is 2, so the largest entries of U^n and V^n are $O(2^n)$.

A similar argument applies, in particular, to the pair of matrices (XY, Y) as well as to pairs of matrices of the form $(A(k)B(m), B(m)A(k))$ with $k, m \geq 2$.

Another example is given by the matrices, say, $U = XYX = \begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix}$, $V = Y = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$. Again by Theorem 1, the fastest growth of the entries in a product of the matrices U and V is given by matrices $(UV)^{\frac{n}{2}}$ since $UV = XYXY$.

Now let us look at the pair (XY, XY^2); this pair was used as an example in Pollicott's paper [20]. Denote $U = XY = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$, $V = XY^2 = \begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix}$. It may seem that the fastest growth should be provided by matrices U^n since that would give an alternating product of X and Y. However, here $V = XY^2$ is just a single matrix, not a product of matrices, so it is obviously more "beneficial" for a faster growth to have powers of V rather than powers of U or any combination of U and V.

The largest eigenvalue of the matrix V is $2+\sqrt{3}$, so the largest entry in V^n is $O((2+\sqrt{3})^n)$.

2.3. **Matrices $A(2)$ and $B(-2)$.** This pair of matrices is rather mysterious in the sense that it is not at all clear what positive words $w(A, B)$ would give the fastest growth of the largest entry. Note that the matrix $A(2)B(-2) = \begin{pmatrix} -3 & 2 \\ -2 & 1 \end{pmatrix}$ has the eigenvalue $\lambda = -1$ of multiplicity 2, so the entries of the matrices $(A(2)B(-2))^{\frac{n}{2}}$ grow linearly in n.

According to computer experiments, the largest (by the absolute value) entries among all $w(A(2), B(-2))$ of length n occur in $(A^2B^2)^{\frac{n}{4}}$, and these entries are $O((\sqrt{2+\sqrt{3}})^n)$ since $(2+\sqrt{3})^2 = 7+\sqrt{48}$ is the largest (by the absolute value) eigenvalue of the matrix $A^2B^2 = \begin{pmatrix} -15 & 4 \\ -4 & 1 \end{pmatrix}$. Note that $\sqrt{2+\sqrt{3}} \approx 1.93$.

2.4. **A pair of binary matrices.** The following pair of binary matrices was considered in the paper by Jungers and Blondel [11]: $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

The matrix B is in $GL_2(\mathbb{Z})$ but not in $GL_2(\mathbb{Z})$. Here the fastest growth of the entries seems to be provided by powers of the matrix $A^3B = \begin{pmatrix} 3 & 1 \\ 1 & 0 \end{pmatrix}$, and therefore the joint spectral radius of the matrices A and B in this case is $s = (\frac{3+\sqrt{13}}{2})^{\frac{1}{4}} \approx 1.348$.

3. Average Growth

Perhaps surprisingly, computing the average growth rate of the entries in a random product of n matrices A and B (where each factor is A or B with probability $\frac{1}{2}$) is easier than computing the maximum growth rate (Section 2) or generic growth rate (Section 4).

Our method is based on solving a system of linear recurrence relations with constant coefficients. To illustrate this method, let us first determine the average growth rate of the entries in a random product of n matrices $A(k)$ and $B(k)$ (see Section 2.1), for some small integer values of k.

3.1. Average growth for products of $A(1)$ and $B(1)$.

Let $\begin{pmatrix} a_n & b_n \\ c_n & d_n \end{pmatrix}$ denote the result of multiplying n matrices where each factor is $A(1)$ or $B(1)$ with probability $\frac{1}{2}$. Denote the expectation of a_n by $\bar{a}_n$, etc. Then, using linearity of the expectation, we have the following recurrence relations for the expectations:

$$\bar{a}_n = \tfrac{1}{2}\bar{a}_{n-1} + \tfrac{1}{2}(\bar{a}_{n-1} + \bar{b}_{n-1}) = \bar{a}_{n-1} + \tfrac{1}{2}\bar{b}_{n-1}.$$
$$\bar{b}_n = \tfrac{1}{2}(\bar{a}_{n-1} + \bar{b}_{n-1}) + \tfrac{1}{2}\bar{b}_{n-1} = \tfrac{1}{2}\bar{a}_{n-1} + \bar{b}_{n-1}.$$

From this, $\bar{b}_n = 2\bar{a}_{n-1} - \frac{3}{2}\bar{a}_n$, and then $\bar{a}_n = 2\bar{a}_{n-1} - \frac{3}{4}\bar{a}_{n-2}$. Solving the latter recurrence relation for $\bar{a}_n$, we get that $\bar{a}_n$ is a linear combination of $(\frac{3}{2})^n$ and $(\frac{1}{2})^n$. Therefore, $\bar{a}_n$ is $O((\frac{3}{2})^n)$, and so is $\bar{b}_n$. The recurrence relations for $\bar{c}_n$ and $\bar{d}_n$ are similar.

3.2. Average growth for products of $A(2)$ and $B(2)$.

Using the same notation in this case, we get

$$\bar{a}_n = \tfrac{1}{2}\bar{a}_{n-1} + \tfrac{1}{2}(\bar{a}_{n-1} + 2\bar{b}_{n-1}) = \bar{a}_{n-1} + \bar{b}_{n-1}.$$
$$\bar{b}_n = \tfrac{1}{2}(2\bar{a}_{n-1} + \bar{b}_{n-1}) + \tfrac{1}{2}\bar{b}_{n-1} = \bar{a}_{n-1} + \bar{b}_{n-1}.$$

From this we see that $\bar{b}_n = \bar{a}_n$, whence $\bar{a}_n = 2\bar{a}_{n-1}$. Therefore, $\bar{a}_n = O(2^n)$ and so is $\bar{b}_n$. The recurrence relations for $\bar{c}_n$ and $\bar{d}_n$ are similar.

3.3. Average growth for products of $A(2)$ and $B(-2)$.

Using the same notation in this case, we get

$$\bar{a}_n = \tfrac{1}{2}\bar{a}_{n-1} + \tfrac{1}{2}(\bar{a}_{n-1} - 2\bar{b}_{n-1}) = \bar{a}_{n-1} - \bar{b}_{n-1}.$$
$$\bar{b}_n = \tfrac{1}{2}(2\bar{a}_{n-1} + \bar{b}_{n-1}) + \tfrac{1}{2}\bar{b}_{n-1} = \bar{a}_{n-1} + \bar{b}_{n-1}.$$

From this, $\bar{b}_n = 2\bar{a}_{n-1} - \bar{a}_n$, and then $\bar{a}_n = 2\bar{a}_{n-1} - 2\bar{a}_{n-2}$ and $\bar{b}_n = 2\bar{b}_{n-1} - 2\bar{b}_{n-2}$. Solving the recurrence relation for $\bar{a}_n$, we get that $\bar{a}_n$ is a linear combination of $(1+i)^n$ and $(1-i)^n$. Since $(1+i)^n = -(\sqrt{2})^n$ if n is divisible by 4, we have $|\bar{a}_n| = O((\sqrt{2})^n)$, and so is $|\bar{b}_n|$. The recurrence relations for $|\bar{c}_n|$ and $|\bar{d}_n|$ are similar.

3.4. Average growth for matrices from Pollicott's paper.

In [20], the author has considered the following two matrices: $A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix}$.

Using the same method as in the previous subsections, we have the following system of recurrence relations for the expectations:

$$\bar{a}_n = \tfrac{1}{2}(2\bar{a}_{n-1} + \bar{b}_{n-1}) + \tfrac{1}{2}(3\bar{a}_{n-1} + 2\bar{b}_{n-1}) = \tfrac{5}{2}\bar{a}_{n-1} + \tfrac{3}{2}\bar{b}_{n-1}.$$
$$\bar{b}_n = \tfrac{1}{2}(\bar{a}_{n-1} + \bar{b}_{n-1}) + \tfrac{1}{2}(\bar{a}_{n-1} + \bar{b}_{n-1}) = \bar{a}_{n-1} + \bar{b}_{n-1}.$$

From the latter recurrence relation we have $\frac{3}{2}\bar{b}_n = \frac{3}{2}\bar{a}_{n-1} + \frac{3}{2}\bar{b}_{n-1} = \frac{3}{2}\bar{a}_{n-1} + (\bar{a}_n - \frac{5}{2}\bar{a}_{n-1}) = \bar{a}_n - \bar{a}_{n-1}$. Hence $\frac{3}{2}\bar{b}_{n-1} = \bar{a}_{n-1} - \bar{a}_{n-2}$.

Therefore, the first recurrence relation for $\bar{a}_n$ now gives us $\bar{a}_n = \frac{5}{2}\bar{a}_{n-1} + (\bar{a}_{n-1} - \bar{a}_{n-2}) = \frac{7}{2}\bar{a}_{n-1} - \bar{a}_{n-2}$. Solving this recurrence relation gives $\bar{a}_n = O((\frac{7+\sqrt{33}}{4})^n)$. Note that $s = \frac{7+\sqrt{33}}{4} \approx 3.186$. Then $\lambda = \log s$ (this is what can be called the "average Lyapunov exponent") is approximately 1.159. This gives a pretty good upper bound for the actual Lyapunov exponent in this case, see Section 4.

3.5. **A "shortcut".** Instead of solving a system of recurrence relations, one can just compute the matrix $\frac{1}{2}(A+B)$, and the largest (by the absolute value) eigenvalue of this matrix will give the average growth rate we are looking for.

In particular, for the pair $(A(k), B(k))$ (with positive k), we have $\frac{1}{2}(A(k)+B(k)) = \begin{pmatrix} 1 & \frac{k}{2} \\ \frac{k}{2} & 1 \end{pmatrix}$. The largest eigenvalue of this matrix is $1+\frac{k}{2}$, so the average growth rate of entries in a random product of matrices is $1+\frac{k}{2}$.

However, justifying this "shortcut" would involve considering random variables whose values are matrices rather than numbers. This would take us too far away from the main theme of the present paper, to the realm of what has become known as "free probability" (i.e., probability theory for non-commuting random variables), so we just refer an interested reader to [16] or [17].

4. Generic Growth and the Lyapunov Exponent

Now we are going to look at the growth of the entries in a random product of n matrices $A(2)$ and $B(2)$ (where each factor is $A(2)$ or $B(2)$ with probability $\frac{1}{2}$). The first question is: how is this different from the situation we considered in Section 3?

To answer this, consider the following simple example. Let $X_1, \ldots, X_n$ be independent random variables, each taking just two values, 0 and 2, with probability $\frac{1}{2}$. Then the expectation of their product $X_1X_2\cdots X_n$ is the product of their expectations and is therefore equal to 1. At the same time, the product $X_1X_2\cdots X_n$ is equal to 0 with probability $1-2^{-n}$.

Similar phenomenon occurs in our situation. We note that, while the average (or expected) growth rate of the largest entry of a product matrix is not too hard to compute in most cases (see our Section 3), computing generic growth rate is a lot harder.

Pollicott [20] considered a random product of n matrices where each matrix is either A or B with probability $\frac{1}{2}$ and studied the growth of the *norm* of such a product as a function of n. To relate what we call the *growth rate* s (see the Introduction) to what is called the *Lyapunov exponent* λ in the theory of stochastic processes, we mention that by a famous result of Kesten and Furstenberg [5],

$$\lambda = \lim_{n\to\infty} \frac{1}{n} \log ||M_1M_2\cdots M_n||, \tag{1}$$

where log denotes the natural logarithm.

There are several well-known ways to define the norm of a matrix, but it is easy to verify that the value of λ does not depend on the chosen matrix norm. If we use the sum of the absolute values of all the entries of a matrix M as its norm $||M||$, then it is particularly obvious that we have $\lambda = \log s$ by the above formula.

Pollicott mentions that "it is a fundamental problem to find both an explicit expression for λ and a useful method of accurate approximation". Whereas an explicit formula for λ remains out of reach even for "nice" pairs of matrices like $A(2)$ and $B(2)$ (see Section 2.1), Pollicott [20] offered an algorithm that allows to determine the growth rate of the largest entry in a random product of n non-singular *strictly positive* matrices A and B with any desired precision.

To illustrate his method, Pollicott used the following two matrices: $A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix}$. First, we note:

Proposition 1. *The matrices A and B generate a free semigroup.*

Proof. Denote $X = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, $Y = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$. It is well known that X and Y generate a free semigroup. Then, $A = XY$, $B = XY^2$.

Since neither of the words XY and XY^2 is a suffix of the other, A and B generate a free semigroup. □

The reason why we care about the semigroup generated by A and B being free is to avoid possible discomfort (for (semi)group theorists) in connection with the equality (1). Specifically, it may happen that if the semigroup is not free, the product of n matrices in (1) is equal to a product of less than n matrices, and then one may have technical issues in defining the limit. If both matrices are strictly positive, then this can happen only in some very rare cases (say, when $AB = BA$), but if not, then the situation changes. For example, if $C = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, $D = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$, then $CDC = DCD$, although this particular relation does not change the length of any product.

Somewhat less obviously, if $A = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$, $B = \begin{pmatrix} 3 & 5 \\ 0 & 5 \end{pmatrix}$, then $AB^{10}A^2BA^2BA^{10} = B^2A^6B^2A^2BABABA^2B^2A^2BAB^2$, see [6].

4.1. **Upper bounds on the Lyapunov exponent.** Denote the average and generic growth rates of the entries in a product of multiple copies of given matrices A and B by s_{ave} and s_{gen}, respectively. If all entries of A and B are nonnegative, then $s_{gen} \le s_{ave}$ (Cauchy–Schwarz type inequality).

Therefore, for random products of matrices with nonnegative entries, the Lyapunov exponent λ is less than or equal to $\log s_{ave}$. This gives an easily computable upper bound on λ. We note that in [21] and [24], upper and lower bounds on λ were obtained using altogether different techniques, with [24] specifically addressing the case where $A = A(k)$, $B = B(m)$ (in the notation of our Section 2.1).

For the pair of matrices $A(k)$ and $B(m)$, Corollary 1 from [24] gives the following upper bound: $\lambda \le \frac{1}{4}[c + \log(\sqrt{km} + 1/\sqrt{km}) + \frac{1}{2}\log(1+km)]$, where c is a constant approximately equal to 1.0157.

In the case of matrices $A(1)$ and $B(1)$ (see Section 3.1), we have $s_{ave} = 1.5$, so our upper bound for λ in this case is $\log 1.5 \approx 0.405$. To compare, the upper bound provided by Corollary 1 from [24] is 0.514, so our upper bound is better in this case.

For the matrices $A(2)$ and $B(2)$, Corollary 1 from [24] gives $\lambda \le 0.684$. We know from Section 3.2 that $s_{ave} = 2$ in this case. Therefore, $\lambda \le \log 2 \approx 0.693$. Thus, for $A(2)$ and $B(2)$ the upper bound provided by Corollary 1 from [24] is better than ours.

More generally, for the matrices $A(k)$ and $B(k)$, the upper bound from [24, Corollary 1] gives $\lambda \le \frac{1}{4}[c + \log(k + \frac{1}{k}) + \frac{1}{2}\log(1 + k^2)]$, which is asymptotically equal to $\frac{1}{2}\log k$. At the same time, our method (see Section 3.5) gives $\lambda \le \log(1 + \frac{k}{2})$. Thus, our method gives a tighter upper bound on λ for small (positive) k, whereas the method of [24] gives a tighter upper bound for larger k.

For Pollicott's matrices $A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix}$, we have $s_{ave} = \frac{7+\sqrt{33}}{4} \approx 3.186$, so our upper bound for λ in this case is $\log 3.186 \approx 1.159$, whereas Pollicott [20] gives the following approximation for the actual value of λ: 1.1433...

For the matrices $A = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$, $B = \begin{pmatrix} 5 & 2 \\ 2 & 5 \end{pmatrix}$ from the paper [10], we have $s_{ave} = 5.5$, so our upper bound for λ in this case is $\log 5.5 \approx 1.7$. The upper bound in [10] is 1.66.., so it is tighter.

To summarize, our upper bounds are sometimes better, sometimes not, compared to those of [10] and [24]. At the same time, our combinatorial method is much simpler than analytical methods of [10], [21], and [24].

4.2. **Computer experiments.** We have run numerous computer experiments to approximate the generic growth rate (and therefore the Lyapunov exponent) for various pairs of matrices, and the results are summarized below. The results below were obtained by computing random products of 1,000,000 matrices.

1. For the matrices $A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix}$ from Pollicott's paper [20], we have the generic growth rate $s \approx 3.136$, so $\lambda = \log s \approx 1.143$ (when rounded to 3 decimal places), which agrees with the results in [20].

2. For the matrices $A(2)$ and $B(2)$ (see Section 3.2), we have the generic growth rate $s \approx 1.9$, so $\lambda = \log s \approx 0.64$.

3. For the matrices $A(2)$ and $B(-2)$ (see Section 3.3), the generic growth rate is $s \approx 1.68$, so $\lambda = \log s \approx 0.52$.

5. Summary

In this section, we compare maximal, average, and generic growth rates of the entries in a product of matrices for several particular pairs of matrices. We denote these growth rates by s_{max}, s_{ave}, and s_{gen}, respectively. Also recall that the Lyapunov exponent λ is the natural logarithm of s_{gen}, and that $s_{gen} \le s_{ave}$ for matrices with nonnegative entries.

• Among all pairs of matrices from $GL_2(\mathbb{Z})$ with nonnegative entries, the pair of binary matrices from Section 2.4 seems to have the smallest joint spectral radius $s_{max} = (\frac{3+\sqrt{13}}{2})^{\frac{1}{4}} \approx 1.348$.

• For the pair $(A(1), B(1))$, we have: $s_{max} = \frac{1}{2}(1 + \sqrt{5}) \approx 1.618$, $s_{ave} = 1.5$, $s_{gen} \approx 1.49$ (the latter is based on computer experiments), so $\lambda = \log s_{gen} \approx 0.4$.

• For the pair $(A(2), B(2))$, we have: $s_{max} = 1 + \sqrt{2} \approx 2.41$, $s_{ave} = 2$, $s_{gen} \approx 1.9$ (the latter is based on computer experiments), so $\lambda = \log s_{gen} \approx 0.64$.

• For the pair $(A(2), B(-2))$, we have: $s_{max} = \sqrt{2+\sqrt{3}} \approx 1.93$ (based on computer experiments), $s_{ave} = \sqrt{2}$, $s_{gen} \approx 1.68$ (the latter is based on computer experiments), so $\lambda = \log s_{gen} \approx 0.52$.

• For the pair of matrices in Pollicott's paper [20] (see our Section 3.4), we have: $s_{max} = 2 + \sqrt{3} \approx 3.73$, $s_{ave} = \frac{7+\sqrt{33}}{4} \approx 3.186$, $s_{gen} \approx 3.136$ (the latter is based on computer experiments), so $\lambda = \log s_{gen} \approx 1.143$.

We note that the only pair that has $s_{gen} > s_{ave}$ is the pair $(A(2), B(-2))$. This is because the matrix $B(-2)$ has a negative entry.

References

[1] Y. Aikawa, H. Jo, S. Satake, *Left-right Cayley hashing: a new framework for provably secure hash functions*, Math. Cryptology **3** (2023), 53—65.

[2] J. Bourgain, A. Gamburd, *Uniform expansion bounds for Cayley graphs of $SL_2(\mathbb{F}_p)$*. Ann. of Math. (2) **167** (2008), 625–642.

[3] L. Bromberg, V. Shpilrain, A. Vdovina, *Navigating in the Cayley graph of $SL_2(\mathbb{F}_p)$ and applications to hashing*, Semigroup Forum **94** (2017), 314–324.

[4] A. Chorna, K. Geller, V. Shpilrain, *On two-generator subgroups of $SL_2(\mathbb{Z})$, $SL_2(\mathbb{Q})$, and $SL_2(\mathbb{R})$*, J. Algebra **478** (2017), 367–381.

[5] H. Furstenberg, H. Kesten, *Products of random matrices*, Ann. Math. Statist. **31** (1960), 457–469.

[6] P. Gawrychowski, M. Gutan, A. Kisielewicz, *On the problem of freeness of multiplicative matrix semigroups*, Theoretical Computer Science **411** (2010), 1115–1120.

[7] S. Han, A. M. Masuda, S. Singh, J. Thiel, *Maximal entries of elements in certain matrix monoids*, Integers **20** (2020), paper No. A31.

[8] K. G. Hare, I. D. Morris, N. Sidorov, J. Theys, *An explicit counterexample to the Lagarias-Wang finiteness conjecture*, Adv. Math. **226** (2011), 4667–4701.

[9] H. A. Helfgott, *Growth and generation in $SL_2(\mathbb{Z}/p\mathbb{Z})$*, Ann. of Math. (2) **167** (2008), 601–623.

[10] N. Jurga and I. Morris, *Effective estimates of the top Lyapunov exponents for random matrix products*, Nonlinearity **32** (2019), 4117–4146.

[11] R. M. Jungers, V. D. Blondel, *On the finiteness property for rational matrices*, Linear Algebra Appl. **428** (2008), 2283–2295.

[12] S. Kim, T. Koberda, *Non-freeness of groups generated by two parabolic elements with small rational parameters*, Michigan Math. J. **71** (2022), 809–833.

[13] J. C. Lagarias, Y. Wang, *The finiteness conjecture for the generalized spectral radius of a set of matrices*, Linear Algebra Appl. **214** (1995) 17–42.

[14] M. Larsen, *Navigating the Cayley graph of $SL_2(\mathbb{F}_p)$*, Int. Math. Res. Notes **27** (2003), 1465-–1471.

[15] C. Le Coz, C. Battarbee, R. Flores, T. Koberda, D. Kahrobaei, *Post-quantum hash functions using $SL_n(\mathbb{F}_p)$*, https://arxiv.org/abs/2207.03987

[16] J. A. Mingo, R. Speicher, *Free Probability and Random Matrices. Fields Institute Monographs*, Vol. **35**, Springer, New York, 2017.

[17] A. Nica, R. Speicher, *Lectures on the Combinatorics of Free Probability*, Cambridge University Press, 2006.

[18] A. Olshanskii, V. Shpilrain, *Linear average-case complexity of algorithmic problems in groups,* https://arxiv.org/abs/2205.05232

[19] G. Panti, D. Sclosa, *The finiteness conjecture holds in $(SL_2(\mathbb{Z} \geq 0))^2$*, Nonlinearity **34** (2021), 5234–5260.

[20] M. Pollicott, *Maximal Lyapunov exponents for random matrix products*, Invent. Math. **181** (2010), 209–226.

[21] M. Pollicott, *Effective estimates of Lyapunov exponents for random products of positive matrices*, Nonlinearity **34** (2021), 6705–6718.

[22] I. Smilga, *New sequences of non-free rational points*, C. R. Math. Acad. Sci. Paris **359** (2021), 983–989.

[23] V. Shpilrain, B. Sosnovski, *Cayley hashing with cookies*, preprint. https://arxiv.org/abs/2402.04943

[24] R. Sturman, J.-L. Thiffeault, *Lyapunov exponents for the random product of two shears*, J. Nonlinear Sci. **29** (2019), 593–620.

[25] J.-P. Tillich and G. Zémor, *Hashing with SL_2*, in CRYPTO 1994, Lecture Notes Comp. Sci. **839** (1994), 40–49.

Department of Mathematics, The City College of New York, New York, NY 10031
Email address: shpilrain@yahoo.com