Analysis Math., **45 (4)** (2019), 705–726
DOI: 10.1007/s10476-019-0003-4

BOUNDS ON AUTOCORRELATION COEFFICIENTS OF RUDIN–SHAPIRO POLYNOMIALS

J.-P. ALLOUCHE[1], S. CHOI[2], A. DENISE[3], T. ERDÉLYI[4,*] and B. SAFFARI[5]

[1]CNRS, IMJ-PRG, Sorbonne Université, 4 Place Jussieu, Case 247, F-75252 Paris Cedex 05, France
e-mail: jean-paul.allouche@imj-prg.fr

[2]Department of Mathematics, Simon Fraser University, 8888 University Drive, Burnaby, British Columbia V5A 1S6, Canada
e-mail: schoia@sfu.ca

[3]LRI et I2BC Université Paris-Sud, CNRS,Université Paris-Saclay, F-91405 Orsay Cedex, France
e-mail: alain.denise@lri.fr

[4]Department of Mathematics, Texas A&M University, College Station, TX 77843-3368, USA
e-mail: terdelyi@math.tamu.edu

[5]Départment de Mathématiques, Faculté des Sciences d'Orsay, Université Paris-Sud, F-91405 Orsay Cedex, France
e-mail: Bahman.Saffari@math.u-psud.fr

(Received September 12, 2018; revised March 25, 2019; accepted April 10, 2019)

Dedicated to the memory of Professor Jean-Pierre Kahane

Abstract. We study the autocorrelation coefficients of the Rudin–Shapiro polynomials, proving in particular that their maximum on the interval $[1, 2^n)$ is bounded from below by $C_1 2^{\alpha n}$ and is bounded from above by $C_2 2^{\alpha' n}$ where $\alpha = 0.7302852\ldots$ and $\alpha' = 0.7302867\ldots$.

1. Introduction

The Rudin–Shapiro polynomials were discovered by H. S. Shapiro in 1951 [9] (also see the paper of M. J. E. Golay the same year, and the footnote on the first page of [2]). They were studied later by W. Rudin in a 1959 paper [8] as recalled, e.g., in [1]. These polynomials, also called the Shapiro-Rudin polynomials, are constructed as follows.

* Corresponding author.

Key words and phrases: Golay–Rudin–Shapiro polynomial, autocorrelation coefficient, trigonometric polynomial.

Mathematics Subject Classification: 42A05, 11L03.

Let $P_0(x) = 1$ and $Q_0 = 1$. For any integer $n \geq 1$, define

$$\begin{cases} P_n(z) := P_{n-1}(z) + z^{2^{n-1}} Q_{n-1}(z), \\ Q_n(z) := P_{n-1}(z) - z^{2^{n-1}} Q_{n-1}(z). \end{cases} \tag{1.1}$$

Polynomials $P_n(z)$ and $Q_n(z)$ are called Rudin–Shapiro polynomials. Note that $P_n(z)$ and $Q_n(z)$ are polynomials with ± 1 coefficients of degree $L_n - 1$ where $L_n = 2^n$. It is well known that

$$|P_n(e^{it})|^2 + |Q_n(e^{it})|^2 = L_{n+1} = 2^{n+1}, \quad \text{for any } t \in [0, 2\pi). \tag{1.2}$$

By the Fourier coefficients of f at k, we mean the coefficient for the term z^k, or simply

$$\widehat{f}(k) = (f)^\wedge(k) := \frac{1}{2\pi} \int_0^{2\pi} f(e^{it}) e^{-ikt}\, dt.$$

By the definition we have

$$\begin{aligned} |P_n|^2(z) &= z^{2^{n-1}} \overline{P_{n-1}(z)} Q_{n-1}(z) + \overline{z}^{2^{n-1}} P_{n-1}(z) \overline{Q_{n-1}(z)} + 2^n, \\ \left(\overline{P_n} Q_n\right)(z) &= 2|P_{n-1}(z)|^2 - z^{2^{n-1}} \overline{P_{n-1}(z)} Q_{n-1}(z) \\ &\quad + \overline{z}^{2^{n-1}} P_{n-1}(z) \overline{Q_{n-1}(z)} - 2^n, \\ \left(P_n \overline{Q_n}\right)(z) &= 2|P_{n-1}(z)|^2 + z^{2^{n-1}} \overline{P_{n-1}(z)} Q_{n-1}(z) \\ &\quad - \overline{z}^{2^{n-1}} P_{n-1}(z) \overline{Q_{n-1}(z)} - 2^n. \end{aligned} \tag{1.3}$$

We are interested in estimating $\max_k |(|P_n|^2)^\wedge(k)|$. If we write

$$|P_n(e^{it})|^2 = \sum_{k=-L_n+1}^{L_n-1} a_k z^k,$$

then

$$(|P_n|^2)^\wedge(k) = \begin{cases} a_k, & \text{when } -L_n + 1 \leq k \leq L_n - 1, \\ 0, & \text{when } |k| > L_n. \end{cases}$$

We will prove

THEOREM 1. *If P_n and Q_n are the n-th Rudin–Shapiro polynomials and*

$$|P_n(z)|^2 = \sum_{k=-L_n+1}^{L_n-1} a_k z^k, \quad |z| = 1$$

($a_0 = L_n,\ a_k = a_{-k},\ k \ge 1$), *then*

$$C_1 \lambda^{n/2} \le \max_{1 \le k \le L_n - 1} |a_k| \le C_2((1.00000100000025)\lambda)^{n/2}$$

with absolute constants $C_1, C_2 > 0$ *where*

$$\lambda := \frac{(71 + 6\sqrt{177})^{1/3} + (71 - 6\sqrt{177})^{1/3} + 5}{3} = 2.75217177\cdots$$

is the real root of $x^3 - 5x^2 + 12x - 16 = 0$.

Also, if we let

$$\left(\overline{P_n} Q_n\right)(z) = \left(P_n \overline{Q_n}\right)(1/z) = \sum_{k=-L_n+1}^{L_n-1} b_k z^k, \quad |z| = 1$$

($b_0 = 2 - L_n,\ b_k = \overline{b_{-k}},\ k \ge 1$), *then*

$$C_3 \lambda^{n/2} \le \max_{1 \le k \le L_n - 1} |b_k| \le C_2((1.00000100000025)\lambda)^{n/2}$$

with an absolute constant $C_3 > 0$.

Corollary 1. *Under the same assumptions as in Theorem 1, we have*

$$C_1 2^{(\frac{\log \lambda}{2 \log 2})n} < \max_{1 \le k \le L_n - 1} |a_k| < C_2 2^{(0.7302867\cdots)n}$$

and

$$C_3 2^{(\frac{\log \lambda}{2 \log 2})n} < \max_{1 \le k \le L_n - 1} |b_k| < C_2 2^{(0.7302867\cdots)n}.$$

Note that $\frac{\log \lambda}{2 \log 2} = 0.7302852\cdots$.

We remark that the constants C_1, C_2 and C_3 are absolute and effective and the numerical values of these constants can be obtained by determining the implicit constants in the proofs.

As an application of Theorem 1 we employ an old result of Littlewood, see Lemma 5 in Section 4. The following theorem tells much about the oscillation of the modulus of the Rudin–Shapiro polynomials around their normalized L_2 norm over the unit circle. For further investigations in this direction see [3] and [4].

Theorem 2. *If* P_n *and* Q_n *are the* n*-th Rudin–Shapiro polynomials,*

$$R(t) := |P_n(e^{it})|^2 = \sum_{k=-L_n+1}^{L_n-1} a_k e^{ikt}$$

$(a_0 = L_n,\ a_k = a_{-k},\ k \geq 1)$, *and*

$$\max_{1\leq k\leq L_n-1} |a_k| \leq C2^{(\lambda_0-\varepsilon)n}$$

with an absolute constants $C > 0$, $\lambda_0 \geq 1/2$ and $\varepsilon > 0$, then there are absolute constants $A > 0$ and $B > 0$ such that the equation $R(t) = (1+\eta)2^n$ has at least $A2^{(2-2\lambda_0)n}$ distinct solutions in $(-\pi, \pi)$ whenever $\eta \in [-B, B]$ and n is sufficiently large.

In view of Theorem 1, we have

COROLLARY 2. *There is absolute constant $A > 0$ such that the equation $R(t) = (1+\eta)2^n$ has at least $A2^{0.5394282n}$ distinct solutions in $(-\pi, \pi)$ whenever $|\eta| \leq 2^{-8}$.*

2. Upper bound for the autocorrelation coefficients

M. Taghavi in [10] claimed

$$\max_{1\leq k\leq L_n-1} |a_k| \leq (3.2134)2^{(0.7303)n}\,.$$

However, as Allouche and Saffari observed, in his proof Taghavi used an incorrect statement saying that the spectral radius of the product of some matrices is independent of the order of the factors (see the review by the first named author in Zentralblatt 0921.11042). So what he ended up with cannot be viewed as a correctly proved result. We obtain a better upper bound here.

The scheme of the proof of Theorem 1 is as follows. Starting from the recursive relations of Rudin–Shapiro polynomials, we develop recursive relations of the autocorrelation coefficients of $|P_n|^2$, $\overline{P_n}Q_n$ and $P_n\overline{Q_n}$ (see ω_n below) in Lemma 1. We find that the multiplying matrix in each inductive step comes from four 3×3 matrices A, B, C and D (defined below). However, after applying n inductive steps, the resulting matrices M (a product of n matrices) have 4^n different possible matrices which is difficult too handle. By studying the products of A, B, C and D carefully, we show in Lemma 2 that these 4^n matrices M are basically generated multiplicatively by two matrices M_1 and B only and B is of order 4. This nice factorization of M makes us successfully estimating the spectral norm of M induced by L_2 norm in Theorem 3 and hence gives an estimates for the autocorrelation coefficients in Theorem 1.

By induction on (1.3), we have

$$(|P_n|^2)^\wedge(2k) = (\overline{P_n}Q_n)^\wedge(2k) = (P_n\overline{Q_n})^\wedge(2k) = 0.$$

For $n \geq 1$, let S_n be the set of all odd integers k with $-L_n < k < L_n$ and let

$$S_n^\tau := \left\{ k \in S_n : (\tau - 3)2^{n-1} < k \leq (\tau - 2)2^{n-1} \right\}$$

so that S_n is the disjoint union of S_n^1, S_n^2, S_n^3 and S_n^4. (The definition of S_n^τ is slightly different in [10] in order to make $S_1 = \{-1, +1\}$ is a union of S_1^1, S_1^2, S_1^3 and S_1^4.)

For $n \geq 2$, let k_n be an odd integer in S_n. We define k_{n-1} and k_n' from k_n as follows

$$k_n' := \begin{cases} k_n + 2^n & \text{if } k_n \in S_n^1 \cup S_n^2, \\ k_n - 2^n & \text{if } k_n \in S_n^3 \cup S_n^4, \end{cases} \tag{2.1}$$

and

$$k_{n-1} := \begin{cases} k_n & \text{if } k_n \in S_n^2 \cup S_n^3, \\ k_n' & \text{if } k_n \in S_n^1 \cup S_n^4. \end{cases} \tag{2.2}$$

It is easy to see that if $k_n \in S_n$, then $k_{n-1} \in S_{n-1}$ and $k_n' \in S_n$.

Let

$$\omega_n(k_n) := \begin{pmatrix} (|P_n|^2)^\wedge(k_n) \\ (\overline{P_n} Q_n)^\wedge(k_n') \\ (P_n \overline{Q_n})^\wedge(k_n') \end{pmatrix}.$$

Lemma 1 below gives a recursive relation of $\omega_n(k_n)$.

LEMMA 1. *For $n \geq 2$, we have*

$$\omega_n(k_n) = \mathcal{M}\omega_{n-1}(k_{n-1}), \tag{2.3}$$

where $\mathcal{M}$ is one of the following four 3×3 matrices, whenever k_n is in S_n^1, S_n^2, S_n^3 and S_n^4 respectively,

$$A = \begin{pmatrix} 0 & 0 & 1 \\ 2 & -1 & 0 \\ 2 & 1 & 0 \end{pmatrix}, \; B = \begin{pmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \; C = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & -1 \end{pmatrix}, \; D = \begin{pmatrix} 0 & 1 & 0 \\ 2 & 0 & 1 \\ 2 & 0 & -1 \end{pmatrix}.$$

PROOF. This is Lemma 1 in [10] but there are some typos in the original paper. Let record a correct version here. First note that

$$\begin{aligned} k_n \in S_n^1 \cup S_n^3 &\Rightarrow k_{n-1} \in S_{n-1}^3 \cup S_{n-1}^4, \\ k_n \in S_n^2 \cup S_n^4 &\Rightarrow k_{n-1} \in S_{n-1}^1 \cup S_{n-1}^2. \end{aligned} \tag{2.4}$$

Let $k_n \in S_n^1$. By (2.1) and (2.2), $k_{n-1} = k_n' = k_n + 2^n$, so (2.4), together with (2.1) and (2.2) again, imply that $k_{n-1}' = k_{n-1} - 2^{n-1} = k_n + 2^{n-1}$.

Now using (1.2), we have

$$\begin{aligned}(|P_n|^2)^\wedge(k_n) &= \big(z^{2^{n-1}}\overline{P_{n-1}}Q_{n-1}\big)^\wedge(k_n) + \big(\overline{z}^{2^{n-1}}P_{n-1}\overline{Q_{n-1}}\big)^\wedge(k_n)\\ &= \big(\overline{P_{n-1}}Q_{n-1}\big)^\wedge(k_n-2^{n-1}) + \big(P_{n-1}\overline{Q_{n-1}}\big)^\wedge(k_n+2^{n-1})\\ &= \big(\overline{P_{n-1}}Q_{n-1}\big)^\wedge(k_n-2^{n-1}) + \big(P_{n-1}\overline{Q_{n-1}}\big)^\wedge(k'_{n-1}).\end{aligned}$$

As $k_n-2^{n-1}<1-2^{n-1}$, by (1.3), we have $(\overline{P_{n-1}}Q_{n-1})^\wedge(k_n-2^{n-1})=0$ which implies that

$$(|P_n|^2)^\wedge(k_n) = \big(P_{n-1}\overline{Q_{n-1}}\big)^\wedge(k'_{n-1}). \tag{2.5}$$

Next using (1.2), we have

$$\begin{aligned}&(\overline{P_n}Q_n)^\wedge(k'_n)\\ &= 2\big(|P_{n-1}|^2\big)^\wedge(k'_n) - \big(z^{2^{n-1}}\overline{P_{n-1}}Q_{n-1}\big)^\wedge(k'_n) - \big(\overline{z}^{2^{n-1}}P_{n-1}\overline{Q_{n-1}}\big)^\wedge(k'_n)\\ &= 2\big(|P_{n-1}|^2\big)^\wedge(k_{n-1}) - \big(\overline{P_{n-1}}Q_{n-1}\big)^\wedge(k'_n-2^{n-1}) - \big(P_{n-1}\overline{Q_{n-1}}\big)^\wedge(k'_n+2^{n-1}).\end{aligned}$$

(Note that there are some typos for the above formulas in [10].)

Since $k'_n+2^{n-1}>2^{n-1}-1$, by (1.3), $(P_{n-1}\overline{Q_{n-1}})^\wedge(k'_n+2^{n-1})=0$ and as $k'_n-2^{n-1}=k_n+2^{n-1}=k'_{n-1}$, we have

$$(\overline{P_n}Q_n)^\wedge(k'_n) = 2(|P_{n-1}|^2)^\wedge(k_{n-1}) - (\overline{P_{n-1}}Q_{n-1})^\wedge(k'_{n-1}). \tag{2.6}$$

A similar argument also gives

$$(P_n\overline{Q_n})^\wedge(k'_n) = 2(|P_{n-1}|^2)^\wedge(k_{n-1}) + (\overline{P_{n-1}}Q_{n-1})^\wedge(k'_{n-1}). \tag{2.7}$$

Now (2.5), (2.6) and (2.7) imply that in (2.3), $\mathcal{M}=A$, whenever $k_n\in S_n^1$. Similar calculations yield (2.3) for the other three cases, which completes the proof of the lemma. $\square$

Applying Lemma 1 inductively on n, we get

$$\omega_n(k_n) = \mathcal{M}_{n-1}\cdots\mathcal{M}_2\mathcal{M}_1\omega_1(k_1) \tag{2.8}$$

with $\omega_1(k_1)=\pm\begin{pmatrix}1\\1\\-1\end{pmatrix}k_1=\pm1$, and

$$\mathcal{M}_i\in T_1 := \{A,B,C,D\} \quad i=1,\dots,n-1 \tag{2.9}$$

where T_n is the set of all matrices which are a product of n matrices in $\{A,B,C,D\}$.

We consider any product of two consecutive terms in the matrix product (2.8). Let $j \in \{1, \ldots, n-2\}$, then

$$\mathcal{M}_j \mathcal{M}_{j+1} \in T_2 := \{AC, AD, B^2, BA, C^2, CD, DA, DB\}. \tag{2.10}$$

To see this, suppose that $\mathcal{M}_j = A$, then $k_{n-j} \in S^1_{n-j}$ and so, by (2.4),

$$k_{n-(j+1)} \in S^3_{n-(j+1)} \cup S^4_{n-(j+1)}.$$

Therefore, $\mathcal{M}_{j+1} \in \{C, D\}$, that is, $\mathcal{M}_j \mathcal{M}_{j+1}$ is either AC or AD. Similar arguments yield (2.10) for the cases that $\mathcal{M}_j$ is B, C or D. In other words,

(i) if $\mathcal{M}_j = A, C$, then $\mathcal{M}_{j+1} = C, D$ so that it produces AC, AD, CC, CD,

(ii) if $\mathcal{M}_j = B, D$, then $\mathcal{M}_{j+1} = A, B$ so that it produces BA, BB, DA, DB.

Due to (2.4), T_2 does not contain all 16 possible matrices but only 8 matrices.

We next will show that every element in T_n can be generated by only two matrices B and M_1 in a very simple fashion (see Lemma 2).

We let

$$T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Then we have

$$T^2 = I, \quad C = TBT \quad \text{and} \quad D = TAT.$$

Let

$$M_1 = AT, \quad M_2 = TA, \quad M_3 = BT, \quad M_4 = TB. \tag{2.11}$$

Then we can write T_2 in terms of M_j:

$$\begin{gathered} T_2 = \{AC, AD, B^2, BA, C^2, CD, DA, DB\} \\ = \{(AT)(BT), (AT)(AT), (BT)(TB), (BT)(TA), (TB)(BT), \\ (TB)(AT), (TA)(TA), (TA)(TB)\} \\ = \{M_1M_3, M_1M_1, M_3M_4, M_3M_2, M_4M_3, M_4M_1, M_2M_2, M_2M_4\}. \end{gathered}$$

In view of (i) and (ii), we have the following rules when we multiply M_j together:

(a) after M_1, it is either M_1 or M_3 only;

(b) after M_2, it is either M_2 or M_4 only;

(c) after M_3, it is either M_2 or M_4 only;
(d) after M_4, it is either M_1 or M_3 only.

We now assume n is even and consider the possible factorization of M in T_n.

For any $M \in T_n$, then M is a product of $n/2$ matrices, $M_iM_j \in T_2$. Then

(1) If M is starting from M_1, then in view of (a)–(d), M is one of these forms:

(i) $M = M_1^{\ell_1}$ for $\ell_1 \geq 1$;
(ii) $M = M_1^{\ell_1} M_3$ for $\ell_1 \geq 1$;
(iii) $M = M_1^{\ell_1} M_3 M_4, \ell_1 \geq 1$;
(iv) $M = M_1^{\ell_1} M_3 M_2^{\ell_2}$ for $\ell_1 \geq 1, \ell_2 \geq 1$;
(v) $M = M_1^{\ell_1} M_3 M_2^{\ell_2} M_4$ for $\ell_1 \geq 1, \ell_2 \geq 1$;
(vi) $M = (M_1^{\ell_1} M_3 M_2^{\ell_2} M_4)\, M^* \cdots$, for $\ell_1 \geq 1, \ell_2 \geq 0$, $M^* = M_1$ or M_3.

(2) If M is starting from M_2, then in view of (a)–(d), M is one of these forms:

(i) $M = M_2^{\ell_2}$, for $\ell_2 \geq 1$;
(ii) $M = M_2^{\ell_2} M_4$, for $\ell_2 \geq 1$;
(iii) $M = (M_2^{\ell_2} M_4)\, M^* \cdots$, for $\ell_2 \geq 1$, $M^* = M_1$ or M_3.

(3) if M is starting from M_3, then in view of (a)–(d), M is one of these forms:

(i) $M = M_3 M_2^{\ell_2}$, for $\ell_2 \geq 1$;
(ii) $M = M_3 M_2^{\ell_2} M_4$, for $\ell_2 \geq 1$;
(iii) $M = M_3 M_4$;
(iv) $M = (M_3 M_2^{\ell_2} M_4)\, M^* \cdots$, for $\ell_2 \geq 0$, $M^* = M_1$ or M_3.

(4) if M is starting from M_4, then in view of (a)–(d), M is one of these forms:

(i) $M = M_4$;
(ii) $M = (M_4)\, M^* \ldots$ for $M^* = M_1$ or M_3.

We then rewrite M as a string of block: $M = B_1 B_2 \cdots B_r$ by the appearance of M_4 so that the first $r-1$ blocks are all ending with M_4

$$B_k = (M_{i_1} M_{i_2} \cdots M_{i_t} M_4)$$

for $1 \leq k \leq r-1$ and

$$B_r = (M_{i_1} M_{i_2} \cdots M_{i_t}) M_4 \quad \text{or} \quad B_r = (M_{i_1} M_{i_2} \cdots M_{i_t})$$

where $M_{i_j} \in \{M_1, M_2, M_3\}$.

For the first block B_1, B_1 must be ending with M_4. According to the first matrix of B_1, we have the following 4 cases.

• If B_1 is starting from M_1, then by (iii) and (v) of (1), B_1 is either $M_1^{\ell_1} M_3 M_4$ or $M_1^{\ell_1} M_3 M_2^{\ell_2} M_4$.

• If B_1 is starting from M_2, then by (ii) of (2), B_1 must be $M_2^{\ell_2} M_4$.

• If B_1 is starting from M_3, then by (ii) and (iii) of (3), B_1 is either $M_3M_2^{\ell_2}M_4$ or M_3M_4.

• If B_1 is starting from M_4, then by (i) of (4), B_1 must be M_4.

Hence

$$B_1 \in \{M_1^{\ell_1}M_3M_2^{\ell_2}M_4,\ M_1^{\ell_1}M_3M_4,\ M_2^{\ell_2}M_4,\ M_3M_2^{\ell_2}M_4,\ M_3M_4,\ M_4 : \ell_1,\ell_2 \ge 1\}. \tag{2.12}$$

For $B_2, B_3, \dots, B_{r-1}$, since the last matrix is M_4 in the previous block, then the first matrix in these blocks must be either M_1 or M_3 by requirement (d) above and the last matrix is M_4. Hence

$$B_k \in \{M_1^{\ell_1}M_3M_2^{\ell_2}M_4,\ M_1^{\ell_1}M_3M_4,\ M_3M_2^{l_2}M_4,\ M_3M_4 : \ell_1,\ell_2 \ge 1\} \tag{2.13}$$

for $2 \le k \le r-1$.

For the last block B_r, it is still starting from M_1 or M_3 but it is not necessary ending by M_4, so

$$B_r \in \{M_1^{\ell_1},\ M_1^{\ell_1}M_3,\ M_1^{\ell_1}M_3M_4,\ M_1^{\ell_1}M_3M_2^{\ell_2},\ M_1^{\ell_1}M_3M_2^{\ell_2}M_4,\ M_3M_2^{\ell_2},\ M_3M_2^{\ell_2}M_4,\ M_3M_4 : \ell_1,\ell_2 \ge 1\}. \tag{2.14}$$

If there is only one block, i.e., $r=1$, then either there is no M_4 in M or there is only one M_4 which is the last matrix in M. Then

$$M \in \{M_1^{\ell_1},\ M_1^{\ell_1}M_3,\ M_1^{\ell_1}M_3M_4,\ M_1^{\ell_1}M_3M_2^{\ell_2},\ M_1^{\ell_1}M_3M_2^{\ell_2}M_4,\ M_2^{\ell_2},\ M_2^{\ell_2}M_4,\ M_3M_2^{\ell_2},\ M_3M_2^{\ell_2}M_4,\ M_3M_4,\ M_4 : \ell_1,\ell_2 \ge 1\}. \tag{2.15}$$

As a summary, we have just proved that for even n, if $M \in T_n$, then $M = B_1B_2\dots B_r$ with B_1 in (2.12), B_k in (2.13) for $2 \le k \le r-1$ and B_r in (2.14) or M is either one of (2.15).

For $L \ge 0$, we define the set

$$\mathfrak{L} := \{B^{k_1}M_1^{\ell_1}B^{k_2}M_1^{\ell_2}\cdots B^{k_L}M_1^{\ell_L}B^{k_{L+1}} : k_j \ge 1\,(2 \le j \le L),\ \ell_j \ge 1\,(1 \le j \le L),\ k_1,k_{L+1} \ge 0\}.$$

For $L=0$, we understand that $\prod_{j=1}^{L} M_1^{\ell_j}B^{k_{j+1}} = 1$ so that $\mathfrak{L}$ contains B^{k_1} for $k_1 \ge 0$. The following lemma gives a nice factorization of M in T_n.

Lemma 2. *For even n, every $M \in T_n$ can be written in the form of*

$$M = T^{\delta_1}B^{k_1}M_1^{\ell_1}B^{k_2}M_1^{\ell_2}\cdots B^{k_L}M_1^{\ell_L}B^{k_{L+1}}T^{\delta_2}, \tag{2.16}$$

for $\ell_1, \ell_2, \dots, \ell_L \ge 1$, $k_2, k_3, \dots, k_L \ge 1$, $k_1, k_{L+1} \ge 0$ *and* $\delta_1, \delta_2 \in \{0,1\}$. *Also, every* $M \in T_n$ *can be written in the form of*

$$M = T^{\delta_1} B^{k_1} M_1^{\ell_1} B^{k_2} M_1^{\ell_2} \cdots B^{k_L} M_1^{\ell_L} B^{k_{L+1}} T^{\delta_2} \tag{2.17}$$

for $\ell_1, \ell_2, \dots, \ell_L \ge 1$, $k_2, k_3, \dots, k_L \in \{1,2,3\}$, $k_1, k_{L+1} \in \{0,1,2,3\}$ *and* $\delta_1, \delta_2 \in \{0,1\}$.

Proof. Recall that if $M \in T_n$ for even n, then $M = B_1 B_2 \cdots B_r$ with B_1 in (2.12), B_k in (2.13) for $2 \le k \le r-1$ and B_r in (2.14) or M is either one of (2.15).

We now consider the case $M = B_1 B_2 \cdots B_r$. We study the terms in B_1 in (2.12) first. For example, the first the term in the list of (2.12) is

$$\begin{aligned} M_1^{\ell_1} M_3 M_2^{\ell_2} M_4 &= M_1^{\ell_1} BT \underbrace{(TA)(TA)\cdots(TA)}_{\ell_2 \text{ times}} TB \\ &= M_1^{\ell_1} B \underbrace{(AT)(AT)\cdots(AT)}_{\ell_2 \text{ times}} B = M_1^{\ell_1} B M_1^{\ell_2} B \end{aligned} \tag{2.18}$$

by (2.11). Similarly for the other terms in (2.12), we have

$$M_1^{\ell_1} M_3 M_4 = M_1^{\ell_1} B^2, \quad M_2^{\ell_2} M_4 = T M_1^{\ell_2} B,$$
$$M_3 M_2^{\ell_2} M_4 = B M_1^{\ell_2} B, \quad M_3 M_4 = B^2, \quad M_4 = TB$$

and they all belong to $T^{\delta_1} \mathfrak{L} T^{\delta_2}$. Therefore

$$B_1 \in \left\{ M_1^{\ell_1} B M_1^{\ell_2} B, M_1^{\ell_1} B^2, T M_1^{\ell_2} B, B M_1^{\ell_2} B, B^2, TB \right\} \subseteq T^{\delta_1} \mathfrak{L} T^{\delta_2},$$

and, for $2 \le k \le r-1$

$$B_k \in \left\{ M_1^{\ell_1} B M_1^{\ell_2} B, M_1^{\ell_1} B^2, B M_1^{\ell_2} B, B^2 \right\} \subseteq T^{\delta_1} \mathfrak{L} T^{\delta_2}$$

because the matrices in the list of (2.13) are also in (2.12) where $\ell_1, \ell_2 \ge 1$ and $\delta_1, \delta_2 \in \{0,1\}$. We now consider the last block B_r. The matrices in (2.14) not in (2.12) are

$$M_1^{\ell_1}, \quad M_1^{\ell_1} M_3 = M_1^{\ell_1} BT, \quad M_1^{\ell_1} M_3 M_2^{\ell_2} = M_1^{\ell_1} B M_1^{\ell_2} T, \quad M_3 M_2^{\ell_2} = B M_1^{\ell_2} T$$

and therefore

$$\begin{aligned} B_r \in \Big\{ & M_1^{\ell_1}, M_1^{\ell_1} BT, M_1^{\ell_1} B^2, M_1^{\ell_1} B M_1^{\ell_2} T, M_1^{\ell_1} B M_1^{\ell_2} B, \\ & B M_1^{\ell_2} T, B M_1^{\ell_2} B, B^2 \Big\} \subseteq T^{\delta_1} \mathfrak{L} T^{\delta_2} \end{aligned}$$

where $\ell_1, \ell_2 \ge 1$ and $\delta_1, \delta_2 \in \{0,1\}$.

We now consider the remaining case when M is one of (2.15). Then the only matrix in (2.15) but not in (2.12), (2.13) and (2.14) is $M = M_2^{\ell_2} = TM_1^{\ell_2}T \in T^{\delta_1}\mathfrak{L}T^{\delta_2}$ where $\ell_1, \ell_2 \geq 1$ and $\delta_1, \delta_2 \in \{0, 1\}$.

Lastly, it is clearly that if $P_1, P_2 \in \mathfrak{L}$, then $P_1P_2 \in \mathfrak{L}$. Hence, We have proved that every $M \in T_n$ can be written in the form of

$$M = T^{\delta_1}B^{k_1}M_1^{\ell_1}B^{k_2}M_1^{\ell_2}\cdots B^{k_L}M_1^{\ell_L}B^{k_{L+1}}T^{\delta_2}$$

for $k_j \geq 1$ $(2 \leq k \leq L)$, $\ell_j \geq 1$ $(1 \leq j \leq L)$, $k_1, k_{L+1} \geq 0$, and $\delta_1, \delta_2 = 0, 1$. Also $\sum k_j + \sum \ell_j = n$. This proves the first assertion of the lemma.

For the second assertion, we observe that $B^4 = B^2$ and $B^5 = B^3$. This completes the proof of the lemma. $\square$

Note that in (2.16) we have

$$\sum_{j=1}^{L}(k_j + \ell_j) + k_{L+1} = n$$

and in (2.17), we have

$$\sum_{j=1}^{L}(k_j + \ell_j) + k_{L+1} \leq n$$

because $k_j \in \{1, 2, 3\}$ for $2 \leq j \leq L$.

The characteristic polynomial of $M_1^2 = AD$ is $g(x) = x^3 - 5x^2 + 12x - 16$. Then $g(x) = 0$ has one real root λ and two complex roots λ' and $\overline{\lambda'}$ with $|\lambda| > |\lambda'|$ and

$$\lambda = \frac{(71 + 6\sqrt{177})^{1/3} - (-71 + 6\sqrt{177})^{1/3} + 5}{3} = 2.75217177\ldots$$

and

$$\begin{aligned}\lambda' = &-\frac{(71 + 6\sqrt{177})^{1/3} - (-71 + 6\sqrt{177})^{1/3} - 10}{6} \\ &+ i\sqrt{3}\Big(\frac{(71 + 6\sqrt{177})^{1/3} + (-71 + 6\sqrt{177})^{1/3}}{6}\Big) \\ &= 1.12391411\ldots + 2.13316845\ldots i.\end{aligned}$$

Since these eigenvalues are distinct, there is a nonsingular matrix S such that $S^{-1}M_1^2S = \Lambda$ with $\Lambda = \mathrm{diag}[\lambda, \lambda', \overline{\lambda'}]$, the diagonal matrix with diagonal elements λ, λ' and $\overline{\lambda'}$ respectively. Since

$$M_1^{2k} = S\Lambda^k S^{-1} \tag{2.19}$$

where

$$S := \begin{pmatrix} s_1 & s_3 & \overline{s_3} \\ s_2 & s_4 & \overline{s_4} \\ 1 & 1 & 1 \end{pmatrix},$$

$$s_1 := \frac{1}{66}\Big(\big(10-\sqrt{177}\big)\big(71+6\sqrt{177}\big)^{1/3} - \big(10+\sqrt{177}\big)\big(-71+6\sqrt{177}\big)^{1/3} - 22\Big),$$

$$s_2 := \frac{1}{66}\Big(\big(1+\sqrt{177}\big)\big(71+6\sqrt{177}\big)^{1/3} - \big(1-\sqrt{177}\big)\big(-71+6\sqrt{177}\big)^{1/3} + 44\Big),$$

$$s_3 := -\frac{1}{132}\Big(\big(1-i\sqrt{3}\big)\big(10-\sqrt{177}\big)\big(71+6\sqrt{177}\big)^{1/3} - \big(1+i\sqrt{3}\big)\big(10+\sqrt{177}\big)\big(-71+6\sqrt{177}\big)^{1/3} + 44\Big)$$

and

$$s_4 := \frac{1}{132}\Big(-\big(1-i\sqrt{3}\big)\big(1+\sqrt{177}\big)\big(71+6\sqrt{177}\big)^{1/3} + \big(1+i\sqrt{3}\big)\big(1-\sqrt{177}\big)\big(-71+6\sqrt{177}\big)^{1/3} + 88\Big).$$

For any matrix M, let $\|M\|$ be the spectral norm of M induced by the L_2 norm. That is if M is an $n \times n$ matrix, M^* is its conjugate transpose and $\lambda_1, \lambda_2, \dots, \lambda_n$ are the eigenvalues of M^*M (the eigenvalues are all real), then $\|M\| = \sqrt{\max\{\lambda_1, \lambda_2, \dots, \lambda_n\}}$. It is known that the spectral radius of M is less than $\|M\|$. We should remark that spectral norm satisfies the submultiplicativity, i.e, $\|AB\| \le \|A\|\,\|B\|$ but spectral radius does not.

For any positive integer k, if ℓ is even, then we have

$$\|M_1^\ell\| = \|S\Lambda^{\ell/2}S^{-1}\| \le \|S\|\,\|\Lambda^{\ell/2}\|\,\|S^{-1}\| \ll \lambda^{\ell/2}$$

because $\|\Lambda^{\ell/2}\| = \lambda^{\ell/2}$. If ℓ is odd, then we have

$$\|M_1^\ell\| = \|M_1M_1^{\ell-1}\| = \|M_1S\Lambda^{(\ell-1)/2}S^{-1}\|$$
$$\le \|M_1\|\,\|S\|\,\|\Lambda^{(\ell-1)/2}\|\,\|S^{-1}\| \ll \lambda^{(\ell-1)/2} \ll \lambda^{\ell/2}.$$

Therefore, for any $\ell \ge 1$, we have

$$\|M_1^\ell\| \ll \lambda^{\ell/2}. \tag{2.20}$$

Since the implicit constant in (2.20) may not be less than 1, so we cannot apply this estimate directly to (2.17) to prove that $\|M\| \ll \lambda^{n/2}$ for any $M \in T_n$ because the implicit constants may be multiplied together to be large when

L is in the order of n. As a result, we need to make use of the factorization in (2.17) when we estimate $\|M\|$. Among all the terms $M_1^\ell B^k$, the term M_1B is the worst because $\|M_1B\| > \lambda$ and needs special consideration (see Lemma 4 below).

The following lemma deals with the product of $2L$ matrices without M_1B factor in (2.16) or (2.17).

Lemma 3. *We have*

$$\big\|(M_1^{\ell_1}B^{k_1})(M_1^{\ell_2}B^{k_2})\cdots(M_1^{\ell_{L-1}}B^{k_{L-1}}M_1^{\ell_L}B^{k_L})\big\| \le ((1.0000005)^2\lambda)^{\frac12\sum_{j=1}^L(k_j+\ell_j)} \tag{2.21}$$

for any (k_j,ℓ_j) *with* $k_j\in\{1,2,3\}$, $\ell_j\ge 1$ *and* $(\ell_j,k_j)\ne(1,1)$ *for* $1\le j\le L$.

Proof. We will prove (3.2) by induction on L.

We first claim that

$$\|M_1^\ell B^k\| \le ((1.0000005)^2\lambda)^{\frac12(\ell+k)} \tag{2.22}$$

for all $k\in\{1,2,3\},\ell\ge 1$ and $(\ell,k)\ne(1,1)$.

If ℓ is even, then

$$\|M_1^\ell B^k\| = \|(M_1^2)^{\ell/2}B^k\| = \|S\Lambda^{\ell/2}S^{-1}B^k\|$$

$$\le \|S\|\cdot\|S^{-1}B^k\|\lambda^{\ell/2} = \frac{\|S\|\|S^{-1}B^k\|}{\lambda^{k/2}}\lambda^{\frac12(\ell+k)} = c_k\lambda^{\frac12(\ell+k)},$$

where $c_k = \frac{\|S\|\|S^{-1}B^k\|}{\lambda^{k/2}}$. If $\ell\ge\frac{\log c_k}{\log 1.0000005}-k$, then $c_k\le 1.0000005^{\ell+k}$. It then follows that

$$\|M_1^\ell B^k\| \le c_k\lambda^{\frac12(\ell+k)} \le ((1.0000005)^2\lambda)^{\frac12(\ell+k)}. \tag{2.23}$$

It remains to check that (2.22) holds for $\ell\le\frac{\log c_k}{\log 1.0000005}-k$. Using `Maple`, we can compute $\|S\|$ and $\|S^{-1}B^k\|$ and find out that

$$\frac{\log c_k}{\log 1.0000005}-k = \begin{cases} 1318902.018 & \text{when } k=1,\\ 991945.7928 & \text{when } k=2,\\ -20445.79861 & \text{when } k=3.\end{cases}$$

We then use `Maple` to check (2.22) holds for $\ell\le 1318902$ when $k=1$ and $\ell\le 991945$ when $k=2$ and do not need to check for $k=3$.

Similarly, if ℓ is odd, then

$$\|M_1^\ell B^k\| = \big\|(M_1^2)^{(\ell-1)/2}M_1B^k\big\| = \big\|S\Lambda^{(\ell-1)/2}S^{-1}M_1B^k\big\|$$

$$\leq \|S\| \left\|S^{-1}M_1B^k\right\| \lambda^{\frac{1}{2}(\ell-1)} = \frac{\|S\|\|S^{-1}M_1B^k\|}{\lambda^{(k+1)/2}} \lambda^{\frac{1}{2}(\ell+k)} = c_k' \lambda^{\frac{1}{2}(\ell+k)},$$

where $c_k' = \frac{\|S\|\|S^{-1}M_1B^k\|}{\lambda^{(k+1)/2}}$. If $\ell \geq \frac{\log c_k'}{\log 1.0000005} - k$, then

$$\|M_1^\ell B^k\| \leq c_k' \lambda^{\frac{1}{2}(\ell+k)} \leq ((1.0000005)^2\lambda)^{\frac{1}{2}(\ell+k)}.$$

It remains to check that (2.22) holds for $\ell \leq \frac{\log c_k'}{\log 1.0000005} - k$. Using `Maple` again, we have

$$\frac{\log c_k'}{\log 1.0000005} - k = \begin{cases} 1187950.952 & \text{when } k = 1, \\ 862238.8518 & \text{when } k = 2, \\ -150152.7391 & \text{when } k = 3. \end{cases}$$

We do not need to check this again because these bounds are smaller than the bounds in the case where ℓ is even. This proves our claim. Hence (3.2) is true for $L = 1$.

We now assume that (3.2) is true for $L - 1$. Then we have

$$\begin{aligned} \left\|M_1^{\ell_1}B^{k_1}M_1^{\ell_2}B^{k_2}\cdots M_1^{\ell_L}B^{k_L}\right\| &\leq \left\|M_1^{\ell_1}B^{k_1}\right\| \cdot \left\|M_1^{\ell_2}B^{k_2}\cdots M_1^{\ell_L}B^{k_L}\right\| \\ &\leq ((1.0000005)^2\lambda)^{\frac{1}{2}(\ell_1+k_1)}((1.0000005)^2\lambda)^{\frac{1}{2}\sum_{j=2}^{L}(\ell_j+k_j)} \\ &= ((1.0000005)^2\lambda)^{\frac{1}{2}\sum_{j=1}^{L}(\ell_j+k_j)}. \end{aligned}$$

This proves our lemma. $\square$

The calculations in the proof of Lemma 3 are simple in using `Maple` and they have been verified independently a couple of times.

The term M_1B is bad because $\|M_1B\| > (1.0000005)^2\lambda$ which does not satisfy (3.2). For handling this term, we have the following lemma.

LEMMA 4. *We have*

$$\|(M_1B)^\ell\| \leq ((1.0000005)^2\lambda)^\ell \tag{2.24}$$

for any $\ell \geq 2$ *and*

$$\|M_1BM_1^\ell B^k\| \leq ((1.0000005)^2\lambda)^{(\ell+k+2)/2} \tag{2.25}$$

for $\ell \geq 1$ *and* $0 \leq k \leq 3$. *Note that* $\|M_1B\| = 2\sqrt{2} > (1.0000005)^2\lambda$ *so that* (2.24) *does not hold for* $\ell = 1$.

PROOF. First we write $M_1B = S_1\Lambda_1S_1^{-1}$, where

$$\Lambda_1 = \operatorname{diag}\left[\frac{1+\sqrt{17}}{2}, \frac{1-\sqrt{17}}{2}, 0\right] \quad \text{and} \quad S_1 = \begin{pmatrix} \frac{-5+\sqrt{17}}{4} & \frac{-5-\sqrt{17}}{4} & 1 \\ \frac{-3+\sqrt{17}}{2} & \frac{-3-\sqrt{17}}{2} & 0 \\ 1 & 1 & 0 \end{pmatrix}.$$

Hence, we have

$$\begin{aligned}\left\|(M_1B)^\ell\right\| &= \left\|S_1\Lambda_1^\ell S_1^{-1}\right\| \le \|S_1\|\cdot\|S_1^{-1}\|\,\|\Lambda_1^\ell\|\\ &= (5.61541131\cdots)\Big(\frac{1+\sqrt{17}}{2}\Big)^\ell \le ((1.0000005)^2\lambda)^\ell\end{aligned}$$

for $\ell \ge 25$ because $\|S_1\| = 4.38008933\ldots$, $\|S_1^{-1}\| = 1.282031231$ and $\|\Lambda_1^\ell\| = \big(\frac{1+\sqrt{17}}{2}\big)^\ell$. For $2 \le \ell \le 24$, we use `Maple` to check directly that (2.24) is true. This proves (2.24).

For $\ell \ge 6$, we have

$$\begin{aligned}\left\|M_1BM_1^\ell B^k\right\| &= \left\|M_1BM_1^4M_1^{\ell-4}B^k\right\| \le \|M_1BM_1^4\|\cdot\|M_1^{\ell-4}B^k\|\\ &\le ((1.0000005)^2\lambda)^3((1.0000005)^2\lambda)^{(\ell-4+k)/2} \le ((1.0000005)^2\lambda)^{(\ell+k+2)/2}\end{aligned}$$

by Lemma 3 with $(\ell-4,k)\ne(1,1)$ and

$$\|M_1BM_1^4\| = 19.97828015\ldots < 20.84624870 = ((1.0000005)^2\lambda)^3.$$

For $1 \le \ell \le 5$ and $1 \le k \le 3$, we use `Maple` to check that (2.25) is true for these ℓ and k. Among these 15 terms, we have

$$\frac{\|M_1BM_1^\ell B^k\|}{((1.0000005)^2\lambda)^{(\ell+k+2)/2}} \le 0.92894011 < 1.$$

This proves (2.25). $\square$

THEOREM 3. *We have*

$$\|M\| \ll ((1.0000005)^2\lambda)^{n/2} \tag{2.26}$$

for any $M \in T_n$.

PROOF. If n is even, for any $M\in T_n$, by Lemma 2,

$$M = T^{\delta_1}B^{k_1}M_1^{\ell_1}B^{k_2}M_1^{\ell_2}\cdots B^{k_L}M_1^{\ell_L}B^{k_{L+1}}T^{\delta_2}$$

for $\ell_1,\ell_2,\ldots,\ell_L \ge 1$, $k_2,k_3,\ldots,k_L\in\{1,2,3\}$, $k_1,k_{L+1}\in\{0,1,2,3\}$ and $\delta_1,\delta_2 \in\{0,1\}$.

We consider the following two cases: (1) $k_{L+1}=1,2,3$ and (2) $k_{L+1}=0$.

Case (1): If $k_{L+1}=1,2,3$, then

$$\begin{aligned}\|M\| &\le \big(\|T^{\delta_1}B^{k_1}\|\cdot\|T^{\delta_2}\|\big)\left\|M_1^{\ell_1}B^{k_2}M_1^{\ell_2}\cdots B^{k_L}M_1^{\ell_L}B^{k_{L+1}}\right\|\\ &\ll \left\|M_1^{\ell_1}B^{k_2}M_1^{\ell_2}\cdots B^{k_L}M_1^{\ell_L}B^{k_{L+1}}\right\|.\end{aligned} \tag{2.27}$$

We further split the consideration into 4 cases depending on how M contains the term M_1B as a factor.

(i) If there is no $(\ell_j, k_{j+1}) = (1,1)$, then, by Lemma 3, we have

$$\big\|M_1^{\ell_1}B^{k_2}M_1^{\ell_2}\cdots B^{k_L}M_1^{\ell_L}B^{k_{L+1}}\big\| \le ((1.0000005)^2\lambda)^{n/2}. \tag{2.28}$$

(ii) If there is $(\ell_j, k_{j+1}) = (1,1)$ for $1 \le k \le L-1$, then by grouping as a product of $(M_1^{\ell_j}B^{k_{j+1}}M_1^{\ell_{j+1}}B^{k_{j+2}}) = (M_1BM_1^{\ell_{j+1}}B^{k_{j+2}})$ and possibly $(M_1^{\ell_L}B^{k_{L+1}})$ for L is odd and using (3.2), (2.23) and (2.25), we have

$$\big\|M_1^{\ell_1}B^{k_2}M_1^{\ell_2}\cdots B^{k_L}M_1^{\ell_L}B^{k_{L+1}}\big\| \le ((1.0000005)^2\lambda)^{n/2}. \tag{2.29}$$

(iii) If $(\ell_L, k_{L+1}) = (\ell_{L-1}, k_L) = \cdots = (\ell_{t+1}, k_{t+2}) = (1,1)$ but $(\ell_t, k_{t+1}) \neq (1,1)$, for some $1 \le t \le L-2$, then

$$\begin{aligned}\big\|M_1^{\ell_1}B^{k_2}M_1^{\ell_2}\cdots B^{k_L}M_1^{\ell_L}B^{k_{L+1}}\big\| &\le \big\|M_1^{\ell_1}B^{k_2}\cdots M_1^{\ell_t}B^{k_{t+1}}\big\|\cdot\|(M_1B)^{L-t}\|\\ &\le ((1.0000005)^2\lambda)^{\sum_{j=1}^{t}(\ell_j+k_{j+1})/2}\|(M_1B)^{L-t}\|\end{aligned}$$

by the same argument as in (i) or (ii). In view of (2.24), since $L-t \ge 2$, we have $\|(M_1B)^{L-t}\| \le ((1.0000005)^2\lambda)^{L-t}$ and hence

$$\begin{aligned}&\big\|M_1^{\ell_1}B^{k_2}M_1^{\ell_2}\cdots B^{k_L}M_1^{\ell_L}B^{k_{L+1}}\big\|\\ &\quad\le ((1.0000005)^2\lambda)^{\sum_{j=1}^{t}(\ell_j+k_{j+1})/2+(L-t)}\\ &\quad= ((1.0000005)^2\lambda)^{\sum_{j=1}^{L}(\ell_j+k_{j+1})/2} \le ((1.0000005)^2\lambda)^{n/2}\end{aligned} \tag{2.30}$$

(iv) if $(\ell_L, k_{L+1}) = (1,1)$ but $(\ell_{L-1}, k_L) \neq (1,1)$, then

$$\begin{aligned}\big\|M_1^{\ell_1}B^{k_2}M_1^{\ell_2}\cdots B^{k_L}M_1^{\ell_L}B^{k_{L+1}}\big\| &\le \big\|M_1^{\ell_1}B^{k_2}M_1^{\ell_2}\cdots B^{k_L}\big\|\,\|M_1B\|\\ &\ll ((1.0000005)^2\lambda)^{n/2-1} \ll ((1.0000005)^2\lambda)^{n/2}\end{aligned} \tag{2.31}$$

by (i), (ii) and (iii). Hence in view of (2.27)–(2.31), we have

$$\|M\| \ll ((1.0000005)^2\lambda)^{n/2}.$$

Case (2): If $k_{L+1} = 0$, then

$$\begin{aligned}\|M\| \ll \big\|M_1^{\ell_1}B^{k_2}M_1^{\ell_2}\cdots B^{k_L}M_1^{\ell_L}\big\| &\le \big\|M_1^{\ell_1}B^{k_2}M_1^{\ell_2}\cdots B^{k_L}\big\|\cdot\|M_1^{\ell_L}\|\\ &\ll \lambda^{\ell_L/2}\big\|M_1^{\ell_1}B^{k_2}M_1^{\ell_2}\cdots B^{k_L}\big\| \ll ((1.0000005)^2\lambda)^{n/2}\end{aligned}$$

by (2.20) and the term $\|M_1^{\ell_1}B^{k_2}M_1^{\ell_2}\cdots B^{k_L}\|$ can be treated in the same way as Case (1). Hence we get (2.26) for n is even.

If n is odd, then $M = M'\mathcal{M}$ with $M' \in T_{n-1}$ and $\mathcal{M} \in T_1$. So

$$\|M\| \le \|M'\| \cdot \|\mathcal{M}\| \ll \|M'\| \ll ((1.0000005)^2\lambda)^{(n-1)/2}$$

by applying (2.26) for M' as $n-1$ is even. This proves (2.26) for every odd n.

As a result, we have

$$|a_{k_n}|, |b_{k_n'}| \le \sqrt{|a_{k_n}|^2 + 2|b_{k_n'}|^2} = \|\omega_n(k_n)\|$$
$$\le \|M\| \, \|\omega_1(k_1)\| \ll ((1.0000005)^2\lambda)^{n/2}.$$

This completes the proof of Theorem 3. $\square$

In view of Theorem 3, we prove the upper bound in Theorem 1.

We remark that the upper bound and lower bound agree with 6 decimal digits of the constants in the exponents. In order to improve our upper bound to agree with 7 decimal digits, we need to verify Lemma 3 for $\ell \le 3.297256652 \cdot 10^7$ for $k = 1$ and $\ell \le 2.479868687 \cdot 10^7$ with (1.00000002) in the place of (1.000005).

3. Lower bound for the autocorrelation coefficients

This is in fact a result in [11] but its proof contains some mistakes and typos. For example, k_n should be $\frac{1}{3}(2L_n - 1)$ instead of $\frac{1}{3}(L_n + 1)$ as stated in [11] when n is odd. Also the coefficients of (6) and (7) in [11] are incorrect (cf. (3.3) and (3.4) below).

In view of (1.1) and (1.3), for $|z| = 1$ we have

$$|P_n(z)|^2 = z^{2^{n-1}}\overline{P_{n-1}(z)}Q_{n-1}(z) + \overline{z}^{2^{n-1}}P_{n-1}(z)\overline{Q_{n-1}(z)} + 2^n.$$

Using (1.3) again, we have

$$|P_n(z)|^2 = 2\big(z^{L_{n-1}} + \overline{z}^{L_{n-1}}\big)|P_{n-2}(z)|^2 - \big(z^{3L_{n-2}} - \overline{z}^{L_{n-2}}\big)\overline{P_{n-2}(z)}Q_{n-2}(z)$$
$$+ \big(z^{L_{n-2}} - \overline{z}^{3L_{n-2}}\big)P_{n-2}(z)\overline{Q_{n-2}(z)} - L_{n-1}\big(z^{L_{n-1}} + \overline{z}^{L_{n-1}}\big) + 2^n.$$

Let

$$k_n = \frac{1}{3}(2L_n + (-1)^n) \quad \text{and} \quad k_n' := k_n - L_n. \tag{3.1}$$

Thus, for $k_n \in S_n$, we have

$$(|P_n(z)|^2)^\wedge(k_n) = 2\Big(\big(z^{L_{n-1}} + \overline{z}^{L_{n-1}}\big)|P_{n-2}(z)|^2\Big)^\wedge(k_n)$$

$$
\begin{aligned}
&- \Big(\big(z^{3L_{n-2}} - \overline{z}^{L_{n-2}}\big)\overline{P_{n-2}(z)}\,Q_{n-2}(z)\Big)^\wedge(k_n)\\
&+ \Big(\big(z^{L_{n-2}} - \overline{z}^{3L_{n-2}}\big)P_{n-2}(z)\overline{Q_{n-2}(z)} - L_{n-1}\big(z^{L_{n-1}} + \overline{z}^{L_{n-1}}\big)\Big)^\wedge(k_n)\\
&= 2\big(|P_{n-2}(z)|^2\big)^\wedge(k_n - L_{n-1}) + 2\big(|P_{n-2}(z)|^2\big)^\wedge(k_n + L_{n-1})\\
&-\big(\overline{P_{n-2}(z)}\,Q_{n-2}(z)\big)^\wedge(k_n - 3L_{n-2}) + \big(\overline{P_{n-2}(z)}\,Q_{n-2}(z)\big)^\wedge(k_n + L_{n-2})\\
&+\big(P_{n-2}(z)\overline{Q_{n-2}(z)}\big)^\wedge(k_n - L_{n-2}) - \big(P_{n-2}(z)\overline{Q_{n-2}(z)}\big)^\wedge(k_n + 2L_{n-2})\\
&-L_{n-1}(1)^\wedge(k_n - L_{n-1}) + L_{n-1}(1)^\wedge(k_n + L_{n-1}).
\end{aligned}
$$

In view of (3.1), we have $k_n+L_{n-1}, k_n+L_{n-2}, k_n-L_{n-2}, k_n+2L_{n-2} \notin S_{n-2}$, $k_n \neq \pm L_{n-1}$ so that their corresponding Fourier coefficients are 0 and

$$k_n - 3L_{n-2} = k_{n-2} - L_{n-2} = k'_{n-2}.$$

It follows that

$$(|P_n(z)|^2)^\wedge(k_n) = 2(|P_{n-2}|^2)^\wedge(k_{n-2}) - \big(\overline{P_{n-2}}\,Q_{n-2}\big)^\wedge(k'_{n-2}). \tag{3.2}$$

Similarly, we have

$$
\begin{aligned}
&\overline{P_n(z)}\,Q_n(z)\\
&= 2\big(\overline{z}^{L_{n-1}} - z^{L_{n-1}}\big)|P_{n-2}(z)|^2 + \big(z^{3L_{n-2}} + 2z^{L_{n-2}} + \overline{z}^{L_{n-2}}\big)\overline{P_{n-2}(z)}\,Q_{n-2}(z)\\
&-\big(\overline{z}^{3L_{n-2}} - 2\overline{z}^{L_{n-2}} + z^{L_{n-2}}\big)P_{n-2}(z)\overline{Q_{n-2}(z)} + L_{n-1}\big(z^{2L_{n-2}} - \overline{z}^{2L_{n-2}}\big)
\end{aligned}
$$

and hence

$$
\begin{aligned}
\big(\overline{P_n(z)}\,Q_n\big)^\wedge(k'_n) &= 2(|P_{n-2}|^2)^\wedge(k_{n-2})\\
&+ \big(\overline{P_{n-2}}\,Q_{n-2}\big)^\wedge(k'_{n-2}) + 2\big(P_{n-2}\overline{Q_{n-2}}\big)^\wedge(k'_{n-2})
\end{aligned} \tag{3.3}
$$

and

$$
\begin{aligned}
P_n(z)\overline{Q_n}(z) &= 2\big(z^{L_{n-1}} - \overline{z}^{L_{n-1}}\big)|P_{n-2}(z)|^2 - \big(z^{3L_{n-2}} - 2z^{L_{n-2}} + \overline{z}^{L_{n-2}}\big)\\
&+ \big(\overline{z}^{3L_{n-2}} + 2\overline{z}^{L_{n-2}} + z^{L_{n-2}}\big)P_{n-2}(z)\overline{Q_{n-2}(z)} - L_{n-1}\big(z^{2L_{n-2}} - \overline{z}^{2L_{n-2}}\big)
\end{aligned}
$$

and hence

$$
\begin{aligned}
\big(P_n(z)\overline{Q_n}\big)^\wedge(k'_n) &= -2\big(|P_{n-2}|^2\big)^\wedge(k_{n-2})\\
&- \big(\overline{P_{n-2}}\,Q_{n-2}\big)^\wedge(k'_{n-2}) + 2\big(P_{n-2}\overline{Q_{n-2}}\big)^\wedge(k'_{n-2}).
\end{aligned} \tag{3.4}
$$

Therefore, we have

$$\omega_n = M_1\omega_{n-2}$$

for $n \ge 2$ where $M_2 = \left(\begin{smallmatrix} 2 & -1 & 0 \\ 2 & 1 & 2 \\ -2 & -1 & 2 \end{smallmatrix}\right)$. Also we have $\Lambda_2 = S'^{-1} M_2 S'$ where Λ is 3×3 diagonal matrix with $\lambda, (\lambda'), (\overline{\lambda'})$ entries. where

$$S' := \begin{pmatrix} s_5 & s_6 & \overline{s_6} \\ s_7 & s_8 & \overline{s_8} \\ 1 & 1 & 1 \end{pmatrix}$$

and

$$s_5 := \frac{-11(1+\sqrt{177})(71+6\sqrt{177})^{1/3} - (103-7\sqrt{177})(71+6\sqrt{177})^{2/3} + 242}{1452},$$

$$s_6 := \frac{1}{2904}\Big(11(1-i\sqrt{3})(\sqrt{177}+1)(71+6\sqrt{177})^{1/3} + (1+i\sqrt{3})(103-7\sqrt{177})(71+6\sqrt{177})^{2/3} + 484\Big),$$

$$s_7 := \frac{1}{726}\Big(11(-21+\sqrt{177})(71+6\sqrt{177})^{1/3} + (-39+5\sqrt{177})(71+6\sqrt{177})^{2/3}\Big),$$

$$s_8 := \Big(11(1-i\sqrt{3})(21-\sqrt{177})(71+6\sqrt{177})^{1/3} + (1+i\sqrt{3})(39-5\sqrt{177})(71+6\sqrt{177})^{2/3}\Big).$$

If n is even, then inductively we have

$$\omega_n = M_2^{n/2}\omega_0$$

where $\omega_0 = \left(\begin{smallmatrix} 0 \\ 1 \\ 1 \end{smallmatrix}\right)$ because $|P_0|^2(z) = (\overline{P_0}Q_0)(z) = (P_0\overline{Q_0})(z) = 1$ and $k_0 = 1$, $k'_0 = 0$. In view of (2.19), we have

$$\omega_n = S'\Lambda^{n/2}(S')^{-1}\omega_0$$

with $\Lambda = \mathrm{diag}[\lambda, \lambda', \overline{\lambda'}]$. Therefore, there are constants a, b, c such that

$$\Big|(|P_n|^2)^\wedge\Big(\frac{1}{3}(2L_n+1)\Big)\Big| = \big|a\lambda^{n/2} + b(\lambda')^{n/2} + \overline{b}(\overline{\lambda'})^{n/2}\big|$$
$$\ge |a|\lambda^{n/2} - 2\,|b|\,|\lambda'|^{n/2} \gg \lambda^{n/2}$$

where $a = -0.38215952\cdots$ and $b = 0.19107976\ldots + i0.88541019\ldots$ and $|\lambda| > |\lambda'|$. This proves

$$\max_{1\le k\le L_n-1} |a_k| \ge \Big|(|P_n|^2)^\wedge\Big(\frac{1}{3}(2L_n+1)\Big)\Big| \gg \lambda^{n/2}.$$

If n is odd, then inductively, we have

$$\omega_n = M_2^{(n-1)/2}\omega_1$$

where $\omega_1 = \binom{P1}{\substack{1\\-1}}$ because

$$|P_1|^2(z) = \overline{z} + 1 + z, \quad (\overline{P_1}Q_1)(z) = \overline{z} - z, \quad (P_1\overline{Q_1})(z) = -\overline{z} + z$$

and $k_1 = 1$, $k_1' = -1$. In view of (2.19), we have

$$\omega_n = S'\Lambda^{(n-1)/2}(S')^{-1}\omega_1$$

with $\Lambda = \mathrm{diag}[\lambda, \lambda', \overline{\lambda'}]$. Therefore, there are constants a', b', c' such that

$$\begin{aligned}\Big|(|P_n|^2)^\wedge\Big(\frac{1}{3}(2L_n-1)\Big)\Big| &= \big|a'\lambda^{(n-1)/2} + b'(\lambda')^{(n-1)/2} + \overline{b'}(\overline{\lambda'})^{(n-1)/2}\big|\\ &= \big|(a'\lambda^{-1/2})\lambda^{n/2} + (b'\lambda'^{-1/2})(\lambda')^{n/2} + (\overline{b'}\lambda'^{-1/2})(\overline{\lambda'})^{n/2}\big|\\ &\geq (|a'|\lambda^{-1/2})\lambda^{n/2} - (2|b'||\lambda'|^{-1/2})|\lambda'|^{n/2} \gg \lambda^{n/2}\end{aligned}$$

where $a' = 0.28744961\ldots$ and $b' = 0.3562751947 - i\,0.3300357859$. This proves for all odd n, we have

$$\max_{1\leq k\leq L_n-1} |a_k| \geq \Big|(|P_n|^2)^\wedge\Big(\frac{1}{3}(2L_n-1)\Big)\Big| \gg \lambda^{n/2}.$$

The lower bounds for $\max_{1\leq k\leq L_n-1}|b_k|$ can be proved in the same way. This proves the lower bounds for Theorem 1.

4. Application of Theorem 1

Let

$$M_q(f) = \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} |f(\theta)|^q\,d\theta\right)^{1/q}, \quad q > 0.$$

Our next lemma is due to Littlewood, see [6, Theorem 1].

Lemma 5. *Let $n \geq 1$. If the trigonometric polynomial f of the form*

$$f(t) = \sum_{m=0}^{L_n-1} b_m \cos(mt + \alpha_m)\,, \qquad b_m, \alpha_m \in \mathbb{R}\,,$$

satisfies

$$M_1(f) \geq c\mu\,, \quad \mu := M_2(f)\,,$$

where $c > 0$ is a constant, $b_0 = 0$,

$$s_{\lfloor (L_n-1)/h \rfloor} = \sum_{m=1}^{\lfloor (L_n-1)/h \rfloor} \frac{b_m^2}{\mu^2} \le 2^{-9} c^6$$

for some constant $h > 0$, and $v \in \mathbb{R}$ satisfies

$$|v| \le V = 2^{-5} c^3 \,,$$

then

$$\mathcal{N}(f, v) > D h^{-1} c^5 2^n \,,$$

where $\mathcal{N}(f, v)$ denotes the number of real zeros of $f - v\mu$ in $(-\pi, \pi)$, and $D > 0$ is an absolute constant.

Proof of Theorem 2. Let the trigonometric polynomial f be defined by $f(t) := R(t) - L_n$. We show that f satisfies the assumptions of the Lemma with $c = 1/4$ and $h := 2^{(2\lambda-1)n}$ if n is sufficiently large. Note that Littlewood in [6] evaluated $M_4(P_n)$ and found that $M_4(P_n) \sim (4^{n+1}/3)^{1/4}$. Hence $\mu := M_2(f) \sim (1/3)^{1/2} 2^n$. Also, it follows from (1.2) that $M_\infty(f) \le 2^n$, hence

$$(1/3) 2^{2n} \sim (M_2(f))^2 \le M_1(f) M_\infty(f) \le 2^n M_1(f)$$

implies that $M_1(f) \ge \mu/2$ if n is sufficiently large. Obviously f is of the form

$$f(t) = \sum_{m=1}^{L_n-1} b_m \cos(mt) \,, \quad b_m \in \mathbb{R} \,,$$

where the assumptions imply that

$$s_{\lfloor (L_n-1)/h \rfloor} = \sum_{m=1}^{\lfloor (L_n-1)/h \rfloor} \frac{b_m^2}{\mu^2} \le \Big(\frac{2^n - 1}{h} \Big) \Big(\frac{4 (C 2^{(\lambda-\varepsilon)n})^2}{\mu^2} \Big)$$

$$\le \Big(\frac{2^n}{2^{(2\lambda-1)n}} \Big) \Big(\frac{4 C^2 2^{(2\lambda - 2\varepsilon)n}}{2^n/4} \Big) \le 16 C^2 2^{-2\varepsilon n} \le 2^{-9} c^6$$

if n is sufficiently large. So f satisfies the assumptions of Lemma 3.6 with $c = 1/2$ and $h := 2^{(2\lambda-1)n}$ if n is sufficiently large, indeed. Thus Lemma 5 implies that

$$\mathcal{N}(f, v) > D h^{-1} c^5 2^n = A 2^{(2-2\lambda)n}$$

whenever

$$|\eta| 2^n = |v| \mu \le V 2^n = 2^{-5} c^3 2^n = 2^{-5} 2^{-3} 2^n = 2^{-8} 2^n$$

and hence we can choose $B = 2^{-8}$ in the theorem if n is sufficiently large.
$\square$

5. Concluding remark

A very first version by a subset of the authors of the present paper gave weaker bounds for the (finite) correlation coefficients using a less precise method. We hope to revisit these questions in the near future, with a detailed discussion about the different correlation coefficients that can be defined in this context.

Acknowledgement. The authors wish to acknowledge the anonymous reviewers for reviewing the paper, including the computational details, thoroughly.

References

[1] J. Brillhart and L. Carlitz, Note on the Shapiro polynomials, *Proc. Amer. Math. Soc.*, **25** (1970) 114–118.

[2] J. Brillhart and P. Morton, A case study in mathematical research: the Golay–Rudin–Shapiro sequence, *Amer. Math. Monthly*, **103** (1996), 854–869.

[3] T. Erdélyi, On the oscillation of the modulus of the Rudin–Shapiro polynomials on the unit circle, *Mathematika* (2019) (accepted).

[4] T. Erdélyi, Improved results on the oscillation of the modulus of the Rudin–Shapiro polynomials on the unit circle, *Proc. Amer. Math. Soc.* (2019) (accepted).

[5] M. J. E. Golay, Statistic multislit spectrometry and its application to the panoramic display of infrared spectra, *J. Optical Soc. America*, **41** (1951) 468–472.

[6] J. E. Littlewood, The real zeros and value distributions of real trigonometrical polynomials, *J. London Math. Soc.*, **41** (1966) 336–342.

[7] J. E. Littlewood, *Some Problems in Real and Complex Analysis*, D. C. Heath and Co. Raytheon Education Co. (Lexington, Mass., 1968).

[8] W. Rudin, Some theorems on Fourier coefficients, *Proc. Amer. Math. Soc.*, **10** (1959) 855–859.

[9] H. S. Shapiro, Extremal Problems for Polynomials and Power Series, M.S. Thesis, Massachusetts Institute of Technology, Department of Mathematics (1951), available at dspace.mit.edu/handle/1721.1/12198.

[10] M. Taghavi, Upper bounds for the autocorrelation coefficients of the Rudin–Shapiro polynomials, *Korean J. Comput. Appl. Math.*, **4** (1997), 39–46.

[11] M. Taghavi, Autocorrelation for a class of polynomial with coefficients defined on T, *Iran. J. Sci. Technol. Trans. A Sci.*, **31** (2007), 391–396.