

## On the absolute stability regions corresponding to partial sums of the exponential function

DAVID KETCHESON AND LAJOS LÓCZI\*

4700 King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900,  
Saudi Arabia

\*Corresponding author: lajos.loczi@kaust.edu.sa david.ketcheson@kaust.edu.sa

AND

TIHAMÉR A. KOCSIS

Széchenyi István Egyetem, Egyetem tér 1, Győr 9026, Hungary  
katih@sze.hu

[Received on 13 March 2014; revised on 27 July 2014]

Certain numerical methods for initial value problems have as stability function the  $n$ th partial sum of the exponential function. We study the stability region, that is, the set in the complex plane over which the  $n$ th partial sum has at most unit modulus. It is known that the asymptotic shape of the part of the stability region in the left half-plane is a semidisc. We quantify this by providing discs that enclose or are enclosed by the stability region or its left half-plane part. The radius of the smallest disc centred at the origin that contains the stability region (or its portion in the left half-plane) is determined for  $1 \leq n \leq 20$ . Bounds on such radii are proved for  $n \geq 2$ ; these bounds are shown to be optimal in the limit  $n \rightarrow +\infty$ . We prove that the stability region and its complement, restricted to the imaginary axis, consist of alternating intervals of length tending to  $\pi$ , as  $n \rightarrow \infty$ . Finally, we prove that a semidisc in the left half-plane centred at the origin and with vertical boundary lying on the imaginary axis is included in the stability region if and only if  $n \equiv 0 \pmod{4}$  or  $n \equiv 3 \pmod{4}$ . The maximal radii of such semidisks are exactly determined for  $1 \leq n \leq 20$ .

**Keywords:** Taylor polynomials of the exponential function; absolute stability region; disc and semidisc inclusions; boundary behaviour near the imaginary axis.

### 1. Introduction

In this work we study the shape of the region of absolute stability of certain numerical methods for initial value ordinary differential equations (initial value ODEs). Any one-step method for which the number of stages or derivatives used is equal to the order of accuracy must have as stability function the corresponding Taylor polynomial of the exponential. Hence, the region of absolute stability of any such method is the set

$$\mathcal{U}_n := \left\{ z \in \mathbb{C} : \left| \sum_{k=0}^n \frac{z^k}{k!} \right| \leq 1 \right\},$$

that is, the set in the complex plane over which the degree  $n$  Taylor polynomial of the exponential has at most unit modulus. The classes of methods with such stability regions include extrapolation methods, Taylor methods and certain Runge–Kutta methods.

We arrived at the present study while investigating the propagation of roundoff errors in extrapolation methods (Ketcheson *et al.*, 2015). More generally, knowledge about these regions is important in understanding the ODE systems to which these methods may be applied, and the time-step sizes that may be used in such applications. We are especially interested in the integration of high-order dissipative semidiscretizations of hyperbolic partial differential equations. The spectra of such discretizations lie in the left half-plane, with components very close to the imaginary axis. Herein, we investigate the shape of the region  $\mathcal{U}_n$  for finite  $n$ , with emphasis on the left half-plane component and the region near the imaginary axis.

It is convenient to introduce some preliminary notation. Let

$$\mathcal{P}_n(z) := \sum_{k=0}^n \frac{(nz)^k}{k!} \quad (n \in \mathbb{N}^+)$$

denote the scaled  $n$ th partial sum of the exponential function, and let

$$\mathcal{S}_n := \{z \in \mathbb{C} : |\mathcal{P}_n(z)| \leq 1\}.$$

Many results can be stated more compactly by using  $\mathcal{S}_n$  instead of  $\mathcal{U}_n$ . We refer to  $\mathcal{U}_n$  as the *unscaled region* and  $\mathcal{S}_n$  as the *scaled region*. Figures 1 and 2 show the boundaries of the first few unscaled and scaled stability regions, respectively. By  $D_\varrho(z_0)$  we mean the closed disc of radius  $\varrho > 0$  centred at  $z_0 \in \mathbb{C}$ , and let  $\Sigma_1 \subset \mathbb{C}$  denote the Szegő region, which is depicted in Fig. 3 and defined in Section 1.1. The boundary of the Szegő region,  $\partial \Sigma_1$ , is known as the Szegő curve. The closed left half of the complex plane is denoted by  $\{\Re \leq 0\}$ , the imaginary axis is  $\{\Re = 0\}$  and instead of  $D_\varrho(0)$  we simply write  $D_\varrho$ .

We now recall some results from Jeltsch & Nevanlinna (1982) regarding the shape of the scaled stability region for large enough  $n$ . Some sets from Theorems 1.1 and 1.3 are depicted in Fig. 4.

**THEOREM 1.1** (Jeltsch & Nevanlinna, 1982, Lemma 5.1) For any  $\varrho > 0$  there exists a positive index  $n_0(\varrho)$  such that

$$\forall n \geq n_0(\varrho) : D_\varrho(-\varrho) \cap \Sigma_1 \subset \mathcal{S}_n.$$

**THEOREM 1.2** (Jeltsch & Nevanlinna, 1982, Theorem 5.3) For any  $0 < \varrho < 1/(2e)$  there exists a positive index  $n_0(\varrho)$  such that

$$\forall n \geq n_0(\varrho) : D_\varrho(-\varrho) \subset \mathcal{S}_n.$$

**THEOREM 1.3** (Jeltsch & Nevanlinna, 1982, Theorem 5.4) Let

$$\mathcal{S}_\infty := \left\{ z \in \mathbb{C} : \exists \text{ a strictly increasing sequence } n_k \in \mathbb{N}^+ \ (k = 1, 2, \dots) \text{ and } \right. \\ \left. \exists z_{n_k} \in \mathcal{S}_{n_k} \text{ such that } \lim_{k \rightarrow +\infty} z_{n_k} = z \right\}.$$

Then

$$\mathcal{S}_\infty = (D_{1/e} \cap \{\Re \leq 0\}) \cup \partial \Sigma_1;$$

see Fig. 4(b).

In practical numerical analysis, usually small values of  $n$  are relevant. Hence, the above theorems cannot be applied, since they do not specify the value of  $n_0$ . When numerically solving an initial value ODE, the time step is often chosen to ensure *absolute stability*; that is, so that the eigenvalues of the

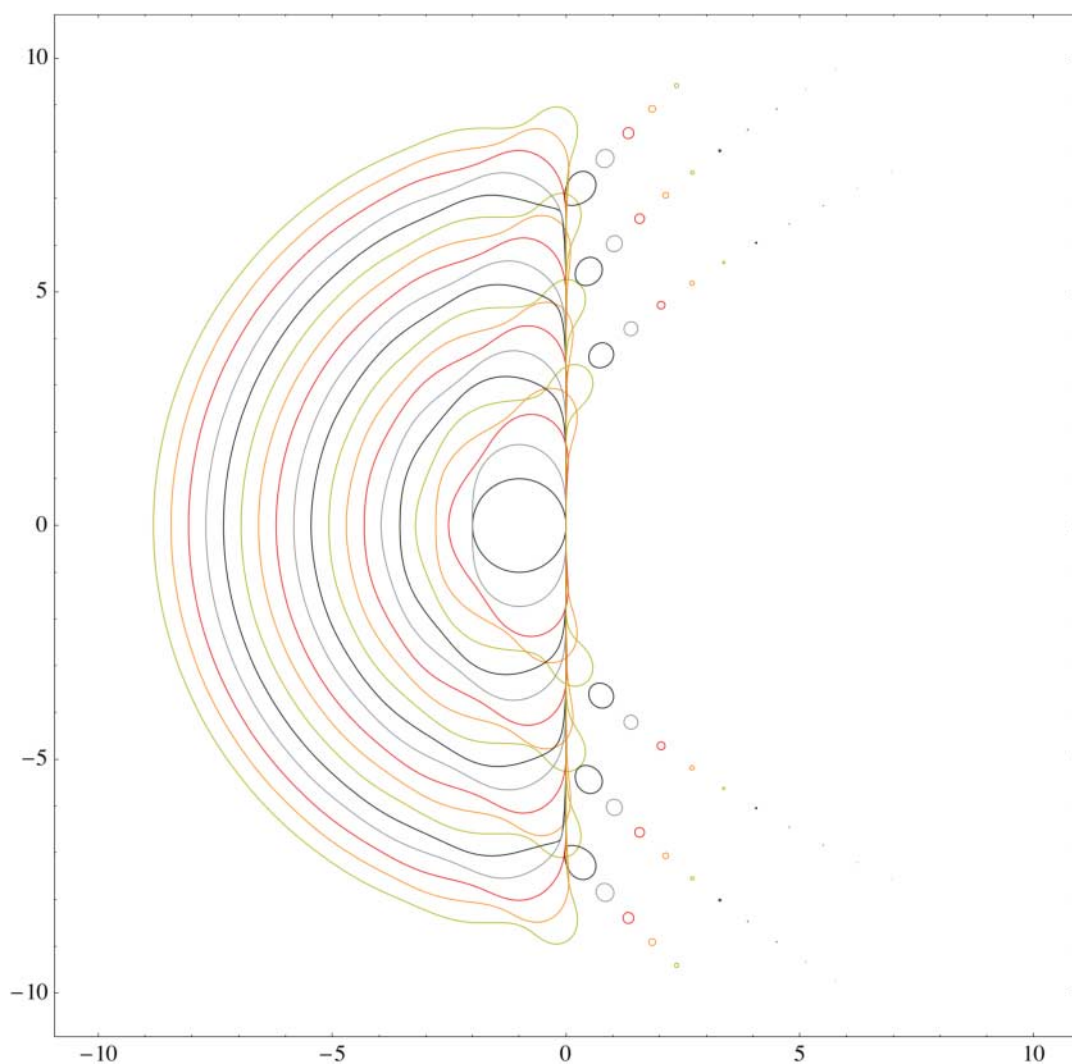


FIG. 1. The boundary curve(s) of the unscaled stability regions  $\mathcal{U}_n$  for  $1 \leq n \leq 20$  in the square  $-10 \leq \Re(z) \leq 10$ ,  $-10 \leq \Im(z) \leq 10$ . In the online version of this article, five colours are used cyclically for different  $n$  values. Compare this figure with Fig. 7 (where 5 and 6 appear as certain block lengths) and Lemma 4.1 (which has ‘period 4’).

solution derivative operator lie inside the stability region of the numerical integrator. Inclusions of the form

$$\mathcal{U}_n \subset D_{\varrho_n} \quad (1.1)$$

or

$$\mathcal{U}_n \cap \{\Re \leq 0\} \subset D_{\varrho_n} \cap \{\Re \leq 0\}, \quad (1.2)$$

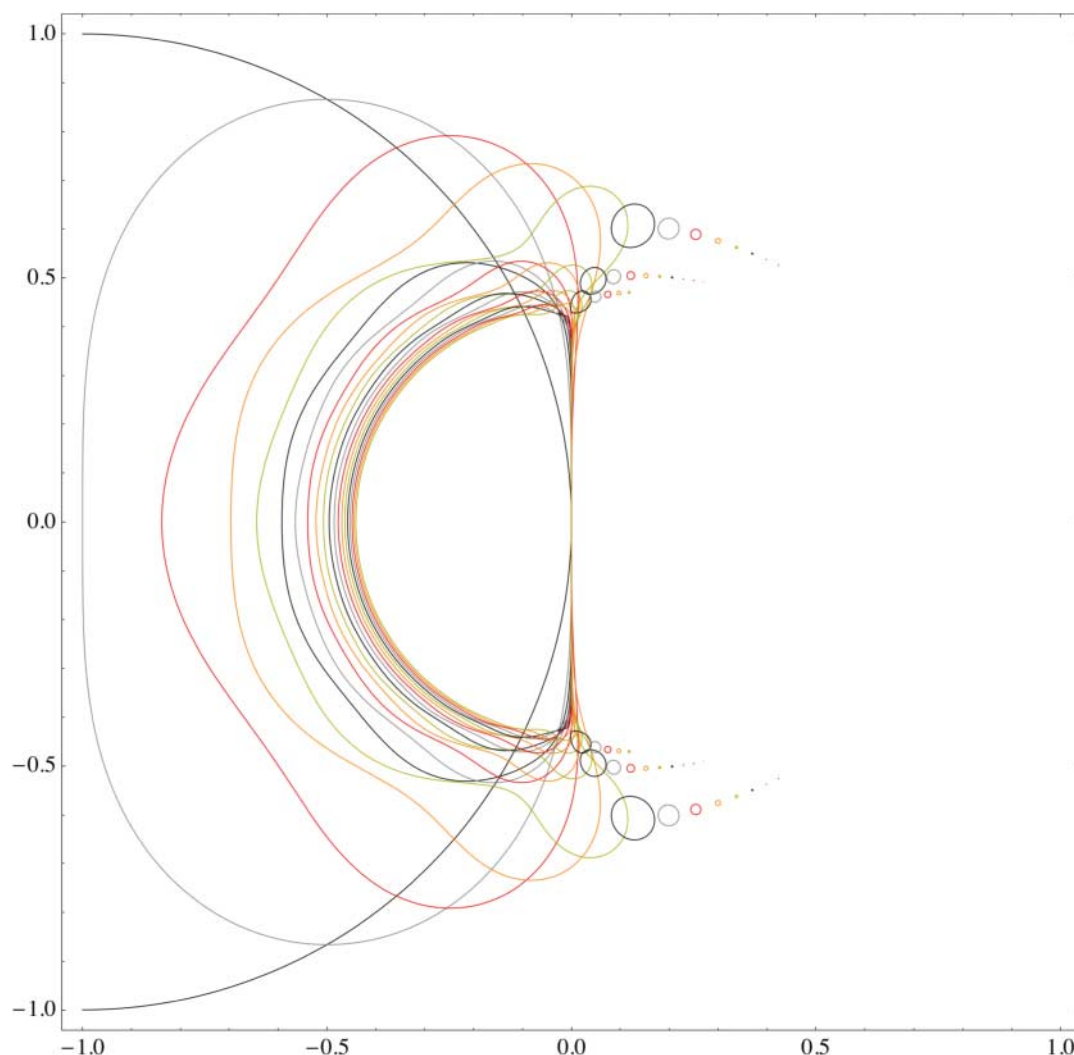


FIG. 2. The boundary curve(s) of the scaled stability regions  $\mathcal{S}_n \subset \mathbb{C}$  for  $1 \leq n \leq 20$  in the square  $-1 \leq \Re(z) \leq 1$ ,  $-1 \leq \Im(z) \leq 1$ . The colour scheme (used in the online version of the article) is the same as that of Fig. 1. Although the general number of connected components is not known, numerical investigations suggest that the set  $\mathcal{S}_n$  is connected for  $1 \leq n \leq 5$ , has three connected components for  $6 \leq n \leq 10$ , and has five connected components for  $n = 11$ , and so on; see also Jeltsch & Nevanlinna (1981, pp. 73–74).

with  $\varrho_n > 0$ , are useful for determining the largest absolutely stable step size. Furthermore, in the above-mentioned work of ours (Ketcheson *et al.*, 2015), we studied the propagation of roundoff errors in one-step methods. We derived bounds on the amplification of such errors based on computing the maxima of certain polynomials over an appropriate region in the complex plane. Ideally, one would like to ensure that roundoff error amplification is bounded under the same criterion of absolute stability. That led us to define, in Ketcheson *et al.* (2015), the *maximum internal amplification factor*, which bounds the

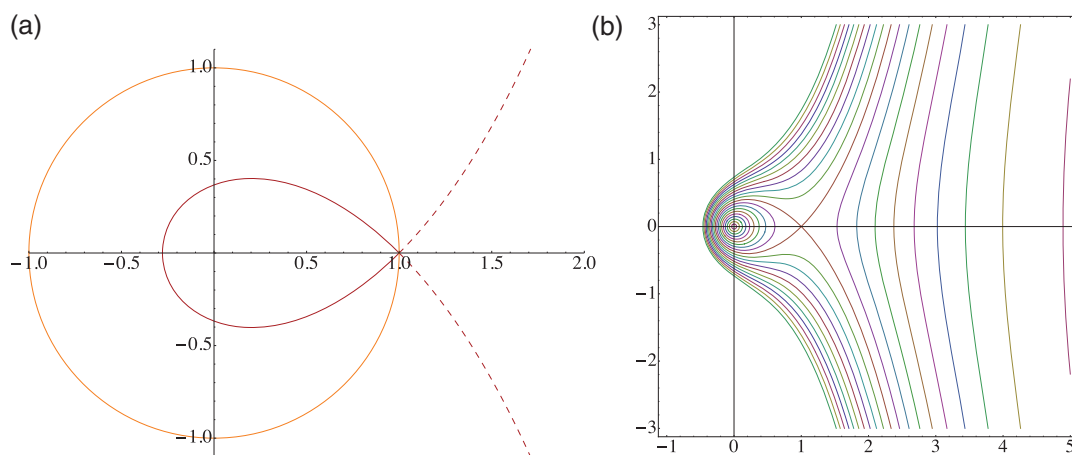


FIG. 3. (a) The Szegő curve  $\partial\Sigma_1$  as the boundary of the Szegő region in the unit disc; (b) some contour lines of  $z \mapsto |ze^{1-z}|$ .

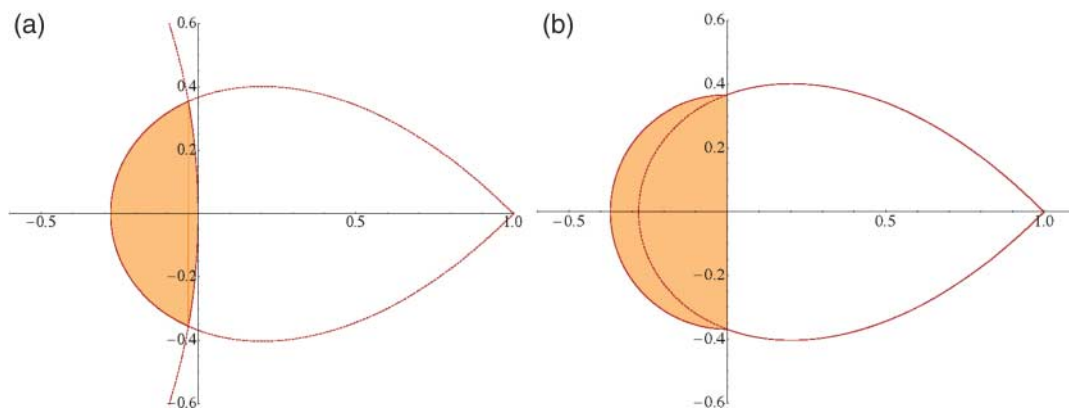


FIG. 4. (a) The set  $D_2(-2) \cap \Sigma_1$  from Theorem 1.1; (b)  $S_\infty$  from Theorem 1.3.

amplification of roundoff errors over one step under the condition of absolute stability. Bounds of type (1.1) or (1.2) are useful for upper-bounding the maximum internal amplification factor.

The structure and main results of this paper are as follows. In Section 1.1 we establish and summarize notation, then in Section 1.2 we review some results relevant to our study. In Section 2 we prove the inclusion  $\mathcal{S}_n \subset D_{1.6}$  for all  $n \geq 2$ , and, for any  $0 < \varepsilon < 0.6$ , the asymptotically optimal bound  $\mathcal{S}_n \subset D_{1+\varepsilon}$  for  $n \geq n_0(\varepsilon)$  large enough. In Section 3 we give similar bounds for the part of  $\mathcal{S}_n$  that lies in the closed left half-plane: the constant 1.6 can be replaced by 0.95 in  $\{\Re \leq 0\}$ , and  $1 + \varepsilon$  by  $1/e + \varepsilon$  ( $\varepsilon > 0$ ). Sections 2 and 3 also contain exact values for the smallest  $D_{\varrho_n}$  containing  $\mathcal{S}_n$  and  $\mathcal{S}_n \cap \{\Re \leq 0\}$ , respectively, for  $1 \leq n \leq 20$ , based on direct computation. Section 4 contains some results regarding the boundary of  $\mathcal{U}_n$  near the imaginary axis (for aesthetic reasons, we consider the boundary of  $\mathcal{U}_n$  and not

that of  $\mathcal{S}_n$ ). In particular, we show that the slices  $\mathcal{U}_n \cap \{\Re = 0\}$  consist of alternating intervals and gaps, with the length of each interval and each gap converging to  $\pi$  as  $n \rightarrow +\infty$  and with offset depending on  $n \bmod 4$ . In Section 5 we prove that  $D_{\varrho_n} \cap \{\Re \leq 0\} \subset \mathcal{S}_n$  for some  $\varrho_n > 0$  if and only if  $n \equiv 0 \bmod 4$  or  $n \equiv 3 \bmod 4$ . We compute the largest such constants  $\varrho_n$  for  $1 \leq n \leq 20$ . Finally, Section 6 contains the proofs of some technical lemmas required for the main results. All computations and figures in this work were performed/created by using *Mathematica* version 9.

**REMARK 1.4** After the completion of the present paper, we discovered that the scenario described in our Section 4 is analogous to the one reported in [Siemieniuch \(1976\)](#). The ‘Nørsett approximations’ discussed in [Siemieniuch \(1976\)](#) are rational approximations to  $e^{-z}$ : the numerator of a Nørsett rational function involves certain sums of Laguerre polynomials, while its denominator has a single multiple root. We wonder what common properties of these different approximations (the Taylor polynomials versus the Nørsett functions) are responsible for the striking similarities between, for example, [Siemieniuch \(1976, Theorem 4.1\)](#) and our Corollary 4.2, or between the ‘asymptotic figure’ [Siemieniuch \(1976, Figure 4.3\)](#) and our Fig. 11. See the last section of [Sanz-Serna \(1980\)](#) as well.

### 1.1 Notation

The notation introduced in earlier paragraphs as well as some new notation needed later are summarized in this section.

In the definition of  $\mathcal{U}_n$  and  $\mathcal{P}_n$ , we use the usual convention that  $0^0 = 1$ . For any  $n \in \mathbb{N}^+$ , the zeros of  $\mathcal{P}_n$  are denoted by  $\{\zeta_k(n)\}_{k=1}^n$ . The real and imaginary parts of a complex number  $z \in \mathbb{C}$  are denoted by  $\Re(z)$  and  $\Im(z)$ , respectively. For  $\delta \in \mathbb{R}$ , by  $\{\Re \leq \delta\}$  we mean  $\{z \in \mathbb{C} : \Re(z) \leq \delta\}$ ; the definition of other similar symbols, such as  $\{\Re > 0\}$ , is analogous. The closed disc centred at  $z_0 \in \mathbb{C}$  with radius  $\varrho > 0$  is  $D_\varrho(z_0)$ ; when  $z_0 = 0$ , we simply write  $D_\varrho$ . For any  $\varrho > 0$ , the half circular arc in the left half-plane is denoted by  $C_\varrho := \{z \in \mathbb{C} : |z| = \varrho\} \cap \{\Re \leq 0\}$ . The boundary of a bounded set  $A \subset \mathbb{C}$  is denoted by  $\partial A$ . We refer to the compact set

$$\Sigma_1 := \{z \in \mathbb{C} : |ze^{1-z}| \leq 1\} \cap D_1$$

as the Szegő region, with boundary,  $\partial \Sigma_1$ , known as the Szegő curve. We use  $|z, w| := |z - w|$  to denote the distance between points  $z, w \in \mathbb{C}$ . Similarly, the distance between two sets  $A, B \subset \mathbb{C}$  is  $|A, B| := \inf_{a \in A, b \in B} |a, b|$ . When, for example,  $A = \{z\}$  is a singleton, we simply write  $|z, B|$  instead of  $|\{z\}, B|$ .

### 1.2 Preliminaries

Szegő showed in his original paper [Szegő \(1924\)](#) that  $\zeta_k(n)$  cluster along the simple closed curve  $\partial \Sigma_1$  as  $n \rightarrow +\infty$ . Many later works refined and extended this result; for example, [Dieudonné \(1935\)](#), [Buckholtz \(1966\)](#), [Newman & Rivlin \(1972, 1976\)](#), [Saff & Varga \(1975, 1976\)](#), [Varga & Carpenter \(1990, 2000, 2001, 2010\)](#), [Carpenter et al. \(1991\)](#), [Yıldırım \(1994\)](#), [Merkle \(1995\)](#), [Kappert \(1996\)](#), [Pritsker & Varga \(1997\)](#), [Walker \(2003\)](#), [Zemyan \(2005\)](#), [Bleher & Mallison \(2006\)](#), [Varga et al. \(2008\)](#) and [Vargas \(2012\)](#).

Buckholtz gave some quantitative bounds on  $\zeta_k(n)$  by showing that all zeros of  $\mathcal{P}_n$  lie outside the Szegő region and that the zeros asymptotically approach the boundary of the Szegő region ([Buckholtz, 1966](#)). Specifically, for each  $n$  and  $1 \leq k \leq n$ ,

$$|\zeta_k(n), \Sigma_1| \leq \frac{2e}{\sqrt{n}}. \quad (1.3)$$

The set  $\mathcal{S}_n$  and the zeros of  $\mathcal{P}_n$  are closely related; indeed, Theorems 1.1–1.3 rely on the results of Szegő and Buckholtz regarding  $\zeta_k(n)$ .

Let us recall the main idea from Buckholtz (1963, 1966) about the proof of (1.3), since we are going to apply the same reasoning in Sections 2 and 3, but in a more quantitative way. Following Buckholtz (1963), we define the function

$$T_n(z) := \frac{n!}{(nz)^n} \mathcal{P}_n(z) \quad (z \in \mathbb{C} \setminus \{0\}), \quad (1.4)$$

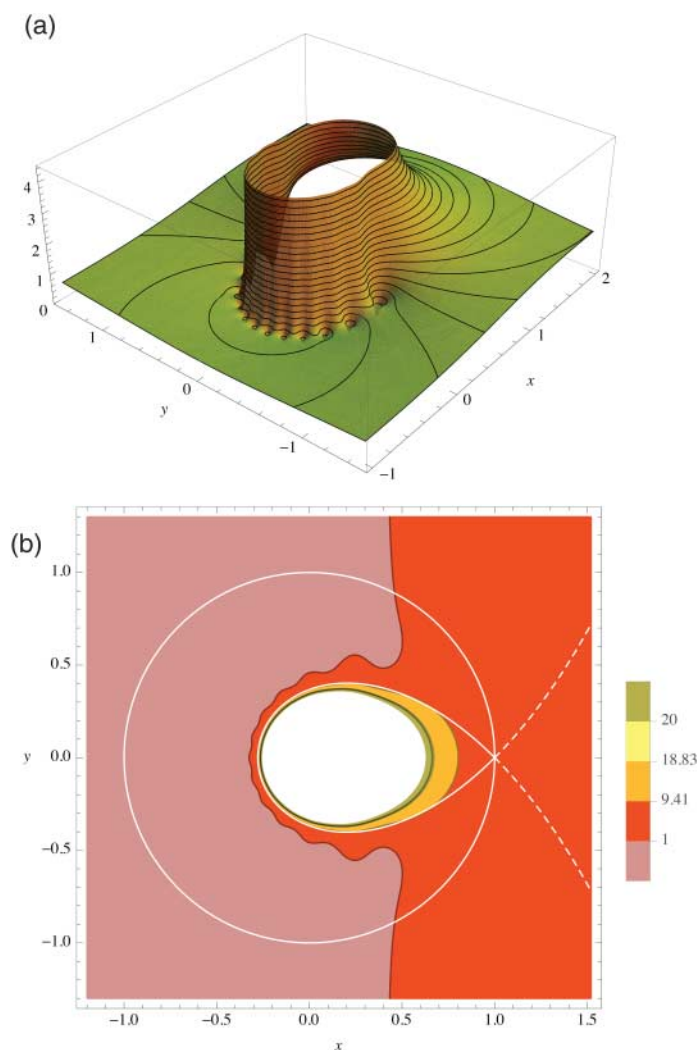


FIG. 5. (a) Some level curves of  $z \mapsto |T_{12}(z)|$  (defined by (1.4), and with  $z = x + iy$ ); (b)  $D_1$  and  $\Sigma_1$  are displayed together with a contour plot of  $|T_{12}|$  showing two specific contours at heights  $e\sqrt{n} = e\sqrt{12} \approx 9.4$  and  $2e\sqrt{n} = 2e\sqrt{12} \approx 18.8$ .



which satisfies the differential equation

$$T_n(z) = \frac{z}{z-1} \left( 1 + \frac{T'_n(z)}{n} \right) \quad (0 \neq z \neq 1), \quad (1.5)$$

and the bounds

$$\forall z \in (\mathbb{C} \setminus \Sigma_1) \cup \partial \Sigma_1 : |T_n(z)| \leq 2e\sqrt{n}, \quad (1.6)$$

$$\forall z \in \mathbb{C}, |z| \geq 1 : |T_n(z)| \leq e\sqrt{n}; \quad (1.7)$$

see Fig. 5. Recall the standard Cauchy inequality for the derivative:

let  $f$  be analytic for all  $z \in D_\delta(z_0)$  and let  $M = \max_{z \in D_\delta(z_0)} |f(z)|$ ; then  $|f'(z_0)| \leq M/\delta$ .

In [Buckholtz \(1966\)](#), this inequality is used with (1.6) to estimate  $|T'_n|$  from above on  $\{z \in \mathbb{C} : |z, \Sigma_1| > 2e/\sqrt{n}\}$ , and then with (1.5) to show  $|T_n| > 0$ , and hence  $|\mathcal{P}_n| > 0$ , on  $\{z \in \mathbb{C} : |z, \Sigma_1| > 2e/\sqrt{n}\}$ .

Finally, we also need the following explicit estimate; see, for example, [Carpenter et al. \(1991\)](#).

LEMMA 1.5 For any  $n \geq 1$ , all roots of  $\mathcal{P}_n$  are located in  $D_1$ .

*Proof.* Since for  $0 \leq m \leq n-1$  we have  $0 \leq n^m/m! \leq n^{m+1}/(m+1)!$ , the Eneström–Kakeya theorem ([Gardner & Govil, 1997](#)) guarantees  $|\zeta_k(n)| \leq 1$  ( $1 \leq k \leq n$ ).  $\square$

## 2. Bounding $\mathcal{S}_n$ by a disc from outside

In this section, we give an asymptotically optimal bound on the smallest disc containing  $\mathcal{S}_n$ , as well as uniform bounds valid for all  $n$ . First we prove the following inclusion for all positive  $n$ .

LEMMA 2.1 We have  $\mathcal{S}_n \subset D_2$  for any  $n \geq 1$ .

*Proof.* From the product representation  $\mathcal{P}_n(z) = (n^n/n!) \prod_{k=1}^n (z - \zeta_k(n))$  at  $z=0$  we obtain  $n^n/n! = \prod_{k=1}^n (-1/\zeta_k(n))$ . Hence, for any  $z \in \mathbb{C}$  we have  $|\mathcal{P}_n(z)| = (n^n/n!) \prod_{k=1}^n |z - \zeta_k(n)| = \prod_{k=1}^n (|z - \zeta_k(n)|/|\zeta_k(n)|) \geq \prod_{k=1}^n |z - \zeta_k(n)|$ , by Lemma 1.5. Therefore, if  $z \in \mathbb{C}$  is chosen such that  $|z| > 2$ , then, due to  $|\zeta_k(n)| \leq 1$  again,  $|z - \zeta_k(n)| > 1$ , so  $|\mathcal{P}_n(z)| > \prod_{k=1}^n 1$ , and thus  $z \notin \mathcal{S}_n$ .  $\square$

Next we refine the above result for  $n$  large enough.

THEOREM 2.2 For each  $\varepsilon \in (0, 1)$  and  $n \geq n_0(\varepsilon) := (1.0085 e/\varepsilon)^2$ , we have  $\mathcal{S}_n \subset D_{1+\varepsilon}$ .

*Proof.* Let us fix  $\varepsilon \in (0, 1)$ . Then for any  $n \geq 1$  and  $|z| \geq 1$  we have  $|T_n(z)| \leq e\sqrt{n}$  by (1.7). So, due to the Cauchy inequality for the derivative,  $|T'_n(z)| \leq e\sqrt{n}/\varepsilon$  for  $|z| \geq 1 + \varepsilon$ . Now for any  $n > (e/\varepsilon)^2$  and  $|z| \geq 1 + \varepsilon$ , we get  $|T'_n(z)|/n \leq e/(\varepsilon\sqrt{n}) < 1$ , so by (1.5) and by (6.2) in Lemma 6.7 with  $\sigma := 1$  we have for such  $n$  and  $z$  values that

$$\begin{aligned} |T_n(z)| &\geq \left( \inf \left\{ \left| \frac{w}{w-1} \right| : w \in \mathbb{C}, |w| \geq 1 + \varepsilon \right\} \right) \left( 1 - \frac{|T'_n(z)|}{n} \right) \\ &\geq \left( \inf \left\{ \left| \frac{w}{w-1} \right| : w \in \mathbb{C}, |w| \geq 1 \right\} \right) \left( 1 - \frac{e}{\varepsilon\sqrt{n}} \right) = \frac{1}{2} \left( 1 - \frac{e}{\varepsilon\sqrt{n}} \right). \end{aligned} \quad (2.1)$$



But  $\frac{1}{2}(1 - e/(\varepsilon\sqrt{n})) \geq \frac{17}{4034}$  for  $n \geq n_0(\varepsilon)$ , hence for any such  $n$  and  $|z| \geq 1 + \varepsilon$  we obtain  $|\mathcal{P}_n(z)| = (n^n |z|^n / n!) |T_n(z)| \geq \frac{17}{4034} n^n / n!$ . Now, by Lemma 6.1,  $n^n / n! \geq e^n / (e\sqrt{n})$ , and the function  $[1, +\infty) \ni n \mapsto e^n / (e\sqrt{n})$  is increasing. So for  $\varepsilon \in (0, 1)$ ,  $n \geq n_0(\varepsilon)$  and  $|z| \geq 1 + \varepsilon$  we obtain

$$|\mathcal{P}_n(z)| \geq \frac{17}{4034} \cdot \frac{e^{n_0(\varepsilon)}}{e\sqrt{n_0(\varepsilon)}} = \frac{17000}{4068289} \cdot \varepsilon \cdot \frac{e^{n_0(\varepsilon)}}{e^2} \geq \frac{17000}{4068289} \cdot 1 \cdot \frac{e^{(1.0085e/1)^2}}{e^2} > 1.03 > 1.$$

□

Finally, we use Theorem 2.2 to improve the bound in Lemma 2.1 for  $n \geq 2$ . In order to do so, we have determined the quantities  $\max_{z \in \mathcal{S}_n} |z|$  as exact algebraic numbers for the first few  $n$  values by transforming the problem from  $\mathbb{C}$  to  $\mathbb{R}^2$  and applying *Mathematica*'s `Maximize` command with objective function  $\sqrt{x^2 + y^2}$ . The resulting values are presented in Table 1. For the sake of brevity, instead of listing any exact algebraic numbers (in the  $n = 15$  case, for example, the algebraic degree of the exact maximum is 420, while in the  $n = 20$  case, the degree is 760), their values are rounded up, so Table 1 provides strict upper bounds on  $\max_{z \in \mathcal{S}_n} |z|$ . These values can be computed quickly: even the  $n = 20$  case was obtained in just 40 seconds (on a laptop, as of 2013). We remark that the maximum is attained within  $\{\Re < 0\}$  for  $1 \leq n \leq 4$ , and within  $\{\Re > 0\}$  for  $5 \leq n \leq 20$ .

**COROLLARY 2.3** For each  $n \geq 2$ , we have  $\mathcal{S}_n \subset D_{1.6}$ .

*Proof.* For  $2 \leq n \leq 20$ , the result follows from the computations given in Table 1. For  $n \geq 21$ , we can apply Theorem 2.2 with  $\varepsilon := 0.6$  because  $21 > n_0(\varepsilon) > 20.87$ . □

**REMARK 2.4** Theorem 2.2 is asymptotically optimal in the sense that the constant 1 in  $D_{1+\varepsilon}$  cannot be replaced by a smaller positive number. This can be seen by noting that  $1 \in \partial \Sigma_1 \subset \mathcal{S}_\infty$  and using Theorem 1.3. On the other hand, at least for the  $3 \leq n \leq 20$  values, Table 1 shows that we have  $\mathcal{S}_n \subset D_1$ . It would be interesting to investigate whether  $\mathcal{S}_n \subset D_1$  holds for all  $n \geq 21$  as well.

TABLE 1 For  $n \geq 3$ , the exact values of  $\max_{z \in \mathcal{S}_n} |z|$  are rounded up. In the given range  $1 \leq n \leq 20$ , the minimum of  $\max_{z \in \mathcal{S}_n} |z|$  is attained for  $n = 8$

$n$	$\max_{z \in \mathcal{S}_n}  z $	$n$	$\max_{z \in \mathcal{S}_n}  z $
1	2	11	0.664
2	$\frac{1}{2}\sqrt{2(1+\sqrt{2})} \approx 1.099$	12	0.670
3	0.847	13	0.676
4	0.741	14	0.682
5	0.690	15	0.687
6	0.665	16	0.692
7	0.6546	17	0.697
8	0.6523	18	0.702
9	0.6542	19	0.707
10	0.659	20	0.711

REMARK 2.5 One can ask how the computations leading to the results in Table 1 can be reproduced in a way that is independent of *Mathematica*'s `Maximize` command. One possibility is to apply the method of Lagrange multipliers with Lagrange function

$$\mathbb{R}^3 \ni (x, y, \lambda) \mapsto F(x, y, \lambda) := x^2 + y^2 + \lambda \left( \left| \sum_{k=0}^n \frac{(x + iy)^k}{k!} \right|^2 - 1 \right) \in \mathbb{R}.$$

Then with the help of resultants one can successively eliminate, say,  $y$  and  $\lambda$  from the polynomial system

$$\partial_x F(x, y, \lambda) = 0, \quad \partial_y F(x, y, \lambda) = 0, \quad \partial_\lambda F(x, y, \lambda) = 0,$$

to get a univariate polynomial in  $x$ . This would give finitely many points  $(x, y) \in \mathbb{R}^2 = \mathbb{C}$  among which the farthest point of  $\mathcal{S}_n$  from the origin will occur. A similar approach could be used to obtain the values in Table 2 in Section 3, or Table 3 in Section 5.

### 3. Bounding $\mathcal{S}_n$ in the left half-plane by a semidisc from outside

Figures 2 and 4(b) suggest that the size of the bounding discs in Section 2 is dictated by the locations of zeros of  $\mathcal{P}_n$  in the right half-plane. In numerical analysis, we are often interested in stability behaviour only in the left half-plane. For  $n > 4$ , the portion of  $\mathcal{S}_n$  in the left half-plane is apparently contained in a smaller disc. Theorem 3.1 and Corollary 3.2 give an asymptotic result, while Theorem 3.4 is valid for all  $n \geq 3$ .

THEOREM 3.1 For any  $\varepsilon > 0$  there exists  $n_0(\varepsilon) \in \mathbb{N}^+$  such that  $n \geq n_0(\varepsilon)$  implies

$$\mathcal{S}_n \cap \{\Re \leq 0\} \subset D_{1/e+\varepsilon} \cap \{\Re \leq 0\}.$$

*Proof.* Let us fix any  $\varepsilon > 0$ , and set  $\varrho_\varepsilon := 1/e + \varepsilon$  and  $\delta_\varepsilon := \varepsilon e / \sqrt{e^2 + 1}$ . Then for any  $\varrho \geq \varrho_\varepsilon$  and  $z \in C_\varrho$ , we know by Lemma 6.5 that  $|z, \Sigma_1| \geq \delta_\varepsilon$ . On the other hand, (1.6) implies  $|T_n(z)| \leq 2e\sqrt{n}$  for  $z \notin \Sigma_1$ , then the Cauchy inequality for the derivative gives that  $|T'_n(z)| \leq 2e\sqrt{n}/\delta_\varepsilon$  for  $|z, \Sigma_1| \geq \delta_\varepsilon$ . So we get

$$\exists \tilde{n}(\varepsilon) \in \mathbb{N}, \quad \forall n \geq \tilde{n}(\varepsilon), \quad \forall \varrho \geq \varrho_\varepsilon, \quad \forall z \in C_\varrho: \quad \frac{|T'_n(z)|}{n} \leq \frac{2e}{\delta_\varepsilon \sqrt{n}} < 1.$$

Then for such  $n$  and  $z$  values we proceed as in (2.1): by  $\varrho_\varepsilon > 1/e$  and (6.2) in Lemma 6.7 with  $\sigma := 1/e$  we have

$$\begin{aligned} |T_n(z)| &\geq \left( \inf \left\{ \left| \frac{w}{w-1} \right| : w \in \mathbb{C}, |w| \geq \varrho_\varepsilon, \Re(w) \leq 0 \right\} \right) \left( 1 - \frac{|T'_n(z)|}{n} \right) \\ &\geq \left( \inf \left\{ \left| \frac{w}{w-1} \right| : w \in \mathbb{C}, |w| \geq \frac{1}{e} \right\} \right) \left( 1 - \frac{2e}{\delta_\varepsilon \sqrt{\tilde{n}(\varepsilon)}} \right) \\ &= \frac{1}{1+e} \left( 1 - \frac{2e}{\delta_\varepsilon \sqrt{\tilde{n}(\varepsilon)}} \right) > 0. \end{aligned}$$

TABLE 2 For  $n \geq 3$ , the exact values of  $\max_{z \in \mathcal{S}_n \cap \{\Re \leq 0\}} |z|$  are rounded up

$n$	$\max_{z \in \mathcal{S}_n \cap \{\Re \leq 0\}}  z $	$n$	$\max_{z \in \mathcal{S}_n \cap \{\Re \leq 0\}}  z $
1	2	11	0.496
2	$\frac{1}{2} \sqrt{2(1 + \sqrt{2})} \approx 1.099$	12	0.486
3	0.847	13	0.480
4	0.741	14	0.476
5	0.680	15	0.474
6	0.597	16	0.458
7	0.566	17	0.453
8	0.546	18	0.450
9	0.534	19	0.449
10	0.527	20	0.448

This yields, again by Lemma 6.1, for  $n \geq \tilde{n}(\varepsilon)$ ,  $\varrho \geq \varrho_\varepsilon$  and  $z \in C_\varrho$ , that

$$|\mathcal{P}_n(z)| = \frac{n^n |z|^n}{n!} |T_n(z)| \geq \frac{|z|^n e^n}{(1+e)e\sqrt{n}} \left(1 - \frac{2e}{\delta_\varepsilon \sqrt{\tilde{n}(\varepsilon)}}\right) \geq \frac{(1+\varepsilon e)^n}{(1+e)e\sqrt{n}} \left(1 - \frac{2e}{\delta_\varepsilon \sqrt{\tilde{n}(\varepsilon)}}\right),$$

and the last right-hand side is  $> 1$  for all  $n$  larger than a suitable  $n_0(\varepsilon)$ .  $\square$

By repeating the above proof with natural modifications, we obtain the following (more effective) version.

**COROLLARY 3.2** Let us set  $\varrho_n := 1/e + (2.2\sqrt{e^2 + 1})/\sqrt{n}$ . Then for any  $n \geq 2$  we have

$$\mathcal{S}_n \cap \{\Re \leq 0\} \subset D_{\varrho_n} \cap \{\Re \leq 0\}.$$

**REMARK 3.3** Like Theorem 2.2, Theorem 3.1 is asymptotically optimal: the constant  $1/e$  in  $D_{1/e+\varepsilon}$  cannot be replaced by a smaller positive number, since  $-1/e \in \mathcal{S}_\infty$  (cf. Remark 2.4).

Table 2 contains the values  $\max\{|z| : z \in \mathcal{S}_n \cap \{\Re \leq 0\}\}$  for  $1 \leq n \leq 20$ . It can be seen from Figs 1 and 2 that, for larger  $n$  values, the set  $\mathcal{S}_n \cap \{\Re \leq 0\}$  is close to a semidisc centred at the origin; see also the radial slices at the end of Section 5. For  $1 \leq n \leq 20$ , the set  $\mathcal{S}_n \cap \{\Re \leq 0\}$  can be covered by a disc with radius slightly less than  $1/e + (\ln n)/(2en) + 1.64/n$ ; cf. Theorems 1.3 and 3.1, Remark 3.6, and observation  $O_3$  at the beginning of Section 4. Also note that the asymptotic series of  $\sqrt[n]{n!}/n^n$  (as  $n \rightarrow +\infty$ ) starts with  $1/e + (\ln n)/(2en) + (\ln(2\pi))/(2en)$ .

**THEOREM 3.4** For each  $n \geq 3$  we have

$$\mathcal{S}_n \cap \{\Re \leq 0\} \subset D_{0.95} \cap \{\Re \leq 0\}.$$

*Proof.* The computations in Table 2 prove the statement of the theorem for  $3 \leq n \leq 20$ . So, due to Lemma 2.1 also, it is enough to show that  $|\mathcal{P}_n(z)| > 1$  for any  $z \in C_\varrho$  with  $\varrho \in [0.95, 2]$  and  $n \geq 21$ .

Let us set  $\delta_\varrho := (\varrho e - 1)/(2\sqrt{e^2 + 1}) > 0$ . Then (1.6) implies  $|T_n(z)| \leq 2e\sqrt{n}$  for  $z \notin \Sigma_1$ , and the Cauchy inequality for the derivative gives that  $|T'_n(z)| \leq 2e\sqrt{n}/\delta_\varrho$  for  $|z, \Sigma_1| \geq \delta_\varrho$  and, say,  $\Re(z) \leq \delta_\varrho$ . Now from (1.5) we obtain

$$|T_n(z)| \leq \left( \sup \left\{ \left| \frac{w}{w-1} \right| : w \in \mathbb{C}, |w, \Sigma_1| \geq \delta_\varrho, \Re(w) \leq \delta_\varrho \right\} \right) \left( 1 + \frac{2e}{\delta_\varrho \sqrt{n}} \right), \quad (3.1)$$

for  $z$  with  $|z, \Sigma_1| \geq \delta_\varrho, \Re(z) \leq \delta_\varrho$ . We note that for any  $z$  with  $|z, \Sigma_1| \geq 2\delta_\varrho, \Re(z) \leq 0$  we have

$$D_{\delta_\varrho}(z) \subset \{w \in \mathbb{C} : |w, \Sigma_1| \geq \delta_\varrho, \Re(w) \leq \delta_\varrho\},$$

so (3.1) and the Cauchy inequality for the derivative again yield

$$\frac{|T'_n(z)|}{n} \leq \frac{1}{n\delta_\varrho} \left( \sup \left\{ \left| \frac{w}{w-1} \right| : w \in \mathbb{C}, |w, \Sigma_1| \geq \delta_\varrho, \Re(w) \leq \delta_\varrho \right\} \right) \left( 1 + \frac{2e}{\delta_\varrho \sqrt{n}} \right), \quad (3.2)$$

for any  $z$  with  $|z, \Sigma_1| \geq 2\delta_\varrho, \Re(z) \leq 0$ . We check with the help of Lemma 6.7 that the right-hand side of (3.2) is  $< 1$  for all  $n \geq 21$  and  $\varrho \in [0.95, (1 + \sqrt{1 + e^2})/e] \cup ((1 + \sqrt{1 + e^2})/e, 2]$ .

Now by Lemma 6.4 we see that for any  $w \in \mathbb{C}$  and  $|w, \Sigma_1| \geq 2\delta_\varrho$  we have  $|w| \geq \frac{1}{4}$ . So by (6.2) in Lemma 6.7 with  $\sigma := \frac{1}{4}$  and by (1.5) again we show that for  $n \geq 21, |z, \Sigma_1| \geq 2\delta_\varrho$  and  $\Re(z) \leq 0$  we have

$$\begin{aligned} |T_n(z)| &\geq \left( \inf \left\{ \left| \frac{w}{w-1} \right| : w \in \mathbb{C}, |w, \Sigma_1| \geq 2\delta_\varrho, \Re(w) \leq 0 \right\} \right) \left( 1 - \frac{|T'_n(z)|}{n} \right) \\ &\geq \left( \inf \left\{ \left| \frac{w}{w-1} \right| : w \in \mathbb{C}, |w| \geq 1/4 \right\} \right) \left( 1 - \frac{|T'_n(z)|}{n} \right) = \frac{1}{5} \left( 1 - \frac{|T'_n(z)|}{n} \right) \\ &\geq \frac{1}{5} \left( 1 - \frac{1}{n\delta_\varrho} \left( \sup \left\{ \left| \frac{w}{w-1} \right| : w \in \mathbb{C}, |w, \Sigma_1| \geq \delta_\varrho, \Re(w) \leq \delta_\varrho \right\} \right) \left( 1 + \frac{2e}{\delta_\varrho \sqrt{n}} \right) \right). \end{aligned}$$

We use Lemma 6.7 one more time to verify that this last lower estimate of  $|T_n(z)|$  is  $\geq 4 \times 10^{-4}$  for  $n \geq 21$  and  $\varrho \in [0.95, (1 + \sqrt{1 + e^2})/e] \cup ((1 + \sqrt{1 + e^2})/e, 2]$ .

Finally, from Lemma 6.5 and the definition of  $\delta_\varrho$  we see that  $\varrho \in [0.95, 2], z \in C_\varrho$  and  $n \geq 21$  imply  $|z, \Sigma_1| \geq 2\delta_\varrho$  and  $\Re(z) \leq 0$ , and hence

$$|T_n(z)| \geq 4 \times 10^{-4}.$$

Therefore, for  $\varrho \in [0.95, 2], z \in C_\varrho$  and  $n \geq 21$  (by Lemma 6.1 also) we obtain

$$|\mathcal{P}_n(z)| = \frac{n^n |z|^n}{n!} |T_n(z)| \geq \frac{n^n \cdot 0.95^n}{n!} \cdot 4 \times 10^{-4} \geq \frac{e^n \cdot 0.95^n}{e\sqrt{n}} \cdot 4 \times 10^{-4} > 14422 > 1. \quad \square$$

**REMARK 3.5** One could slightly reduce the constant 0.95 in  $D_{0.95}$  by performing another iteration (or several more iterations) in the proof above (an estimate on  $T \rightsquigarrow$  an estimate on  $T' \rightsquigarrow$  an estimate on  $T$ , and so on), but then the starting index 21 in  $n \geq 21$  would jump to a much higher value.

**REMARK 3.6** The main obstacle to improving the bounds appearing in Theorems 2.2 and 3.4 is that we could only guarantee the positivity of  $|1 + T'_n(z)/n|$  by applying the inequality  $|a - b| \geq ||a| - |b||$  and choosing large  $n$  values. For example, in order for Corollary 3.2 to be stronger than the estimate in

Corollary 2.3 we need  $n \geq 27$ , and to be stronger than the estimate in Theorem 3.4 we need  $n \geq 120$ . The convergence rate  $\mathcal{O}(1/\sqrt{n})$  presented in Corollary 3.2 naturally follows from the estimates on  $|T_n|$  and  $|T'_n|$ ; we do not know whether this rate can be improved. As a related result, [Carpenter et al. \(1991\)](#) prove that the convergence rate of the zeros  $\zeta_k(n)$  in (1.3) to  $\partial \Sigma_1$  is exactly  $\mathcal{O}(1/\sqrt{n})$  as  $n \rightarrow +\infty$ , but the rate improves to  $\mathcal{O}((\ln n)/n)$  if zeros in a small disc  $D_\varepsilon(1)$  ( $\varepsilon \in (0, 1]$  arbitrary but fixed) are ignored.

#### 4. The stability region near the imaginary axis

In this section we are concerned with the imaginary part of the ‘upper vertical slice’ of  $\mathcal{U}_n$  along the imaginary axis,

$$\mathcal{V}_n^+ := \left\{ \Im(z) : z \in \mathbb{C}, \Re(z) = 0, \Im(z) \geq 0, \left| \sum_{k=0}^n \frac{z^k}{k!} \right| \leq 1 \right\} \quad (n \in \mathbb{N}^+).$$

We focus only on the upper half-plane due to symmetry, and use the unscaled regions because they lead to simpler expressions. By also taking into account the explicit representations given in Lemma 4.1, it is easy to determine the set  $\mathcal{V}_n^+$  for a particular  $n$ ; see Fig. 6 for  $1 \leq n \leq 20$ , and Fig. 7 for  $1 \leq n \leq 100$ . Based on these figures and on the exact representation of the end points of the shaded intervals as `Root` objects in *Mathematica*, the following observations are made.

- O<sub>1</sub>. For  $n \equiv 0 \pmod{4}$  or  $n \equiv 3 \pmod{4}$ , the connected component of  $\mathcal{V}_n^+$  containing the origin is an interval of positive length, while for  $n \equiv 1 \pmod{4}$  or  $n \equiv 2 \pmod{4}$  the corresponding interval is the singleton  $\{0\}$ .
- O<sub>2</sub>. For  $n \geq 1$ ,  $\mathcal{V}_n^+$  consists of disjoint compact intervals whose end points tend to the grid  $\{\pi\ell : \ell \in \mathbb{N}\}$  as  $n \rightarrow +\infty$  for even  $n$ , or to the grid  $\{0\} \cup \{\pi/2 + \pi\ell : \ell \in \mathbb{N}\}$  as  $n \rightarrow +\infty$  for odd  $n$ . For a fixed and large enough  $n$ , nondegenerate intervals in  $\mathcal{V}_n^+$  or in the complement of  $\mathcal{V}_n^+$  alternate one after another as we move away from the origin; the pattern starts according to the rule described by O<sub>1</sub>.

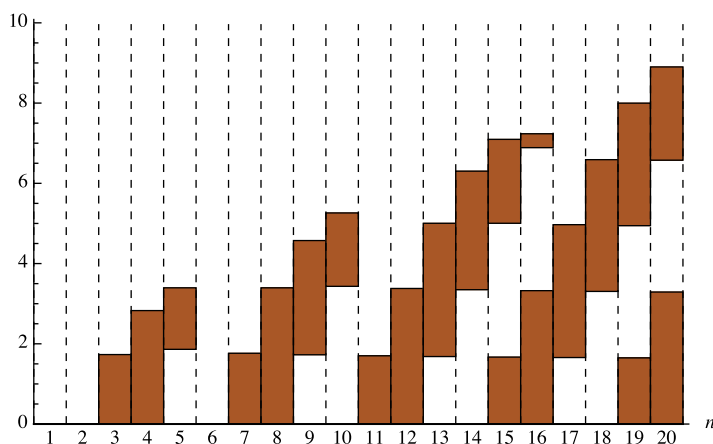


FIG. 6. For each  $1 \leq n \leq 20$ , the intervals constituting  $\mathcal{V}_n^+$  are represented as shaded rectangles.

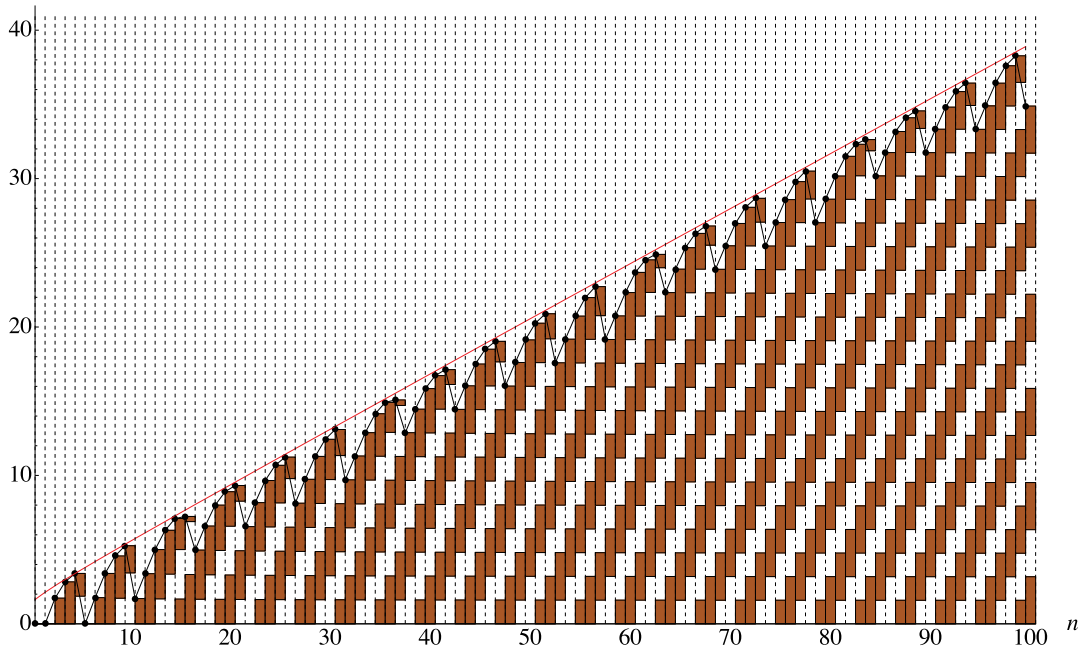


FIG. 7. Extension of Fig. 6 for  $1 \leq n \leq 100$ . Black dots (with piecewise linear interpolation) show the sequence  $\max(\mathcal{V}_n^+)$ . The blocks consisting of monotone nondecreasing adjacent elements have length 5, 5, 6, 5, 5, 5, 6, 5, 5, 5, 6, 5, 5, 5, 6, 5, 5, 5, 6, 5, 5, 5 in the given range (cf. the ‘separation of bubbles’ in Fig. 1). The top red curve, being the graph of  $n \mapsto n/e + (\ln n)/(2e) + 1.2604$ , is an upper estimate in the given range.

O<sub>3</sub>. The farthest point of  $\mathcal{V}_n^+$  from the origin is bounded from above by  $n/e + (\ln n)/(2e) + 1.2604$  for  $1 \leq n \leq 100$  (see the top red curve in Fig. 7).

In the following we explain observations O<sub>1</sub> and O<sub>2</sub>, and indicate how O<sub>3</sub> is related to some earlier results in the literature. At the end of the section we also investigate the boundary curve of  $\mathcal{U}_n$  that oscillates around the imaginary axis: intersection points of the boundary curve and the upper semiaxis define the end points of the shaded intervals in the corresponding column of Fig. 7—we illustrate the amplitude of these oscillations for different  $n$  values.

The first lemma explicitly describes the absolute value of the  $n$ th Taylor polynomial of the exponential function along the imaginary axis, simultaneously providing us with ‘finite truncations’ of the identity  $\cos^2 y + \sin^2 y = 1$ . It is convenient to refer to the  $E$ -polynomial, which was used in Wanner *et al.* (1978) to study the behaviour of rational functions along the imaginary axis. For the polynomials studied here, the  $E$ -polynomial takes the form

$$E_n(y) := \left| \sum_{k=0}^n \frac{(iy)^k}{k!} \right|^2 - 1 \quad (y \in \mathbb{R}, n \in \mathbb{N}^+).$$

LEMMA 4.1 For any  $y \in \mathbb{R}$  and  $n \geq 1$  integer we have

$$E_n(y) = \begin{cases} -\frac{y^{n+1}}{n!} \sum_{k=1}^{n/2} \frac{(-1)^{k+1} y^{2k-1}}{(2k-1)!(k+n/2)} & \text{if } n \equiv 0 \pmod{4}, \\ \frac{y^{n+1}}{n!} \sum_{k=0}^{(n-1)/2} \frac{(-1)^k y^{2k}}{(2k)!(k+(n+1)/2)} & \text{if } n \equiv 1 \pmod{4}, \\ \frac{y^{n+1}}{n!} \sum_{k=1}^{n/2} \frac{(-1)^{k+1} y^{2k-1}}{(2k-1)!(k+n/2)} & \text{if } n \equiv 2 \pmod{4}, \\ -\frac{y^{n+1}}{n!} \sum_{k=0}^{(n-1)/2} \frac{(-1)^k y^{2k}}{(2k)!(k+(n+1)/2)} & \text{if } n \equiv 3 \pmod{4}. \end{cases}$$

The straightforward proof of the above lemma is omitted. The following corollary shows that the sign of the lowest-order term of  $E_n$  explains  $O_1$ .

COROLLARY 4.2 For any positive integer  $n$  with  $n \equiv 1 \pmod{4}$  or  $n \equiv 2 \pmod{4}$ , there exists  $\varrho_n > 0$  such that

$$\mathcal{V}_n^+ \cap [0, \varrho_n] = \{0\}.$$

On the other hand, for any positive integer  $n$  with  $n \equiv 0 \pmod{4}$  or  $n \equiv 3 \pmod{4}$ , there exists  $\varrho_n > 0$  such that

$$[0, \varrho_n] \subset \mathcal{V}_n^+.$$

*Proof.* By Lemma 4.1 we have for  $n \equiv 0 \pmod{4}$  and  $y \in \mathbb{R}$  that

$$E_n(y) = -\frac{1}{(1+n/2)n!} y^{n+2} + \mathcal{O}(y^{n+4}),$$

as  $y \rightarrow 0$ , so for some  $\varrho_n > 0$  sufficiently small,  $E_n(y) \leq 0$  for all  $y \in [0, \varrho_n]$ . But

$$\mathcal{V}_n^+ = \{y \in \mathbb{R} : y \geq 0, E_n(y) \leq 0\} \quad (n \in \mathbb{N}^+),$$

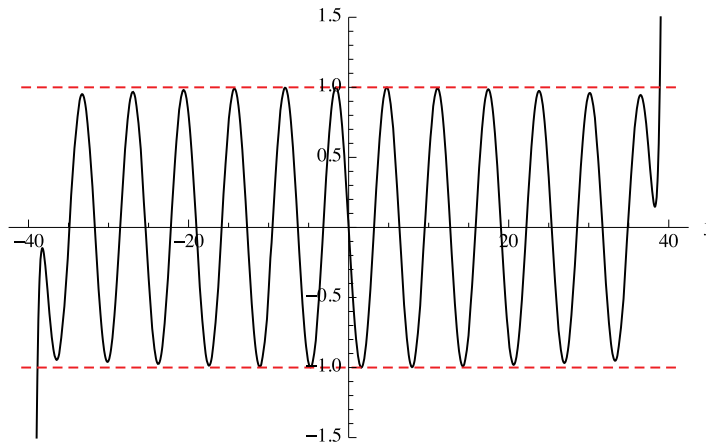
therefore this case is finished. The proof for the other three congruence classes is analogous.  $\square$

We begin explaining  $O_2$  by noting that for each  $n \geq 1$ ,  $\lim_{y \rightarrow +\infty} E_n(y) = +\infty$  and  $E_n(0) = 0$ , so there exist an index  $k_0(n) \in \mathbb{N}^+$  and some mutually disjoint nonempty compact intervals  $I_{n,k}$  ( $k = 1, 2, \dots, k_0(n)$ ) such that

$$\mathcal{V}_n^+ = \{y \in \mathbb{R} : y \geq 0, E_n(y) \leq 0\} = \bigcup_{k=1}^{k_0(n)} I_{n,k}. \quad (4.1)$$

These  $I_{n,k}$  intervals are just the shaded vertical rectangles in Fig. 6 or 7. It of course can happen that  $k_0(n) = 1$ , or some of the  $I_{n,k}$  intervals are singletons. We order the intervals  $I_{n,k}$  in the natural way so that  $\min I_{n,1} = 0$  and  $\max I_{n,k} < \min I_{n,k+1}$  ( $k = 1, 2, \dots, k_0(n) - 1$ ). Let us consider now the scaled



FIG. 8. The graph of  $\tilde{E}_{100}$  over the interval  $[-n/e - 4, n/e + 4]$ .

counterpart of the polynomial  $E_n$ :

$$\tilde{E}_n(y) := \frac{(n+1)!}{2y^{n+1}} E_n(y) \quad (y \in \mathbb{R} \setminus \{0\}).$$

Then  $\mathcal{V}_n^+ = \{0\} \cup \{y \in \mathbb{R} : y > 0, \tilde{E}_n(y) \leq 0\}$ . The advantage of  $\tilde{E}_n$  over  $E_n$  is that for

$$n \equiv 0, n \equiv 1, n \equiv 2 \text{ and } n \equiv 3 \pmod{4},$$

$\tilde{E}_n(y)$  is a perturbation of the truncated series for

$$-\sin y \equiv -\sum_{k=1}^{\infty} \frac{(-1)^{k+1} y^{2k-1}}{(2k-1)!}, \cos y, \sin y \text{ and } -\cos y,$$

respectively; see Fig. 8. This motivates us to use some tools from complex analysis to study the positive real roots of  $\tilde{E}_n$ . In what follows, we again focus only on the  $n \equiv 0 \pmod{4}$  case, because the explanation of  $O_2$  for the other three congruence classes is analogous.

Let  $f_m$  denote the extension of  $\tilde{E}_n$  with  $n = 4m$  to all of  $\mathbb{C}$ :

$$f_m(z) = -\frac{(4m+1)}{2} \sum_{k=1}^{2m} \frac{(-1)^{k+1} z^{2k-1}}{(2k-1)!(k+2m)} \quad (z \in \mathbb{C}, m \in \mathbb{N}^+). \quad (4.2)$$

Clearly,

$$\mathcal{V}_{4m}^+ = \{y \in \mathbb{R} : y \geq 0, f_m(y) \leq 0\}. \quad (4.3)$$

The following lemma makes the word ‘perturbation’ above more precise.

**LEMMA 4.3** The sequence of polynomials  $f_m$  defined by (4.2) converges uniformly to the function  $-\sin$  on compact subsets of  $\mathbb{C}$  as  $m \rightarrow +\infty$ .

*Proof.* Let us fix some  $\varrho > 0$ . We know that the sequence of Maclaurin polynomials  $\tilde{f}_m(z) := -\sum_{k=0}^{2m-1} ((-1)^k z^{2k+1} / (2k+1)!)$  converges uniformly to  $-\sin$  on  $D_\varrho$  as  $m \rightarrow +\infty$ , so, by using the triangle inequality, it is enough to show that

$$\lim_{m \rightarrow +\infty} \sup\{|f_m(z) - \tilde{f}_m(z)| : z \in D_\varrho\} = 0.$$

But

$$\begin{aligned} |f_m(z) - \tilde{f}_m(z)| &= \left| \sum_{k=0}^{2m-1} \frac{(-1)^k z^{2k+1}}{(2k+1)!} \cdot \frac{2k+1}{2k+2+4m} \right| \\ &\leq |z| \sum_{k=0}^{2m-1} \frac{|z|^{2k}}{(2k)!} \cdot \frac{1}{2k+2+4m} \leq \frac{\varrho}{4m} \sum_{k=0}^{\infty} \frac{\varrho^{2k}}{(2k)!} = \frac{\varrho \cosh \varrho}{4m} \rightarrow 0, \end{aligned} \quad (4.4)$$

as  $m \rightarrow +\infty$ . □

**REMARK 4.4** Note that the first inequality in (4.4) holds with equality for purely imaginary  $z$  values, and  $\varrho \cosh \varrho$  is large when  $\varrho$  is large. These already suggest—together with numerical computations—that the uniform convergence of  $f_m$  to  $-\sin$  on  $D_\varrho$  is rather slow, compared with the uniform convergence of  $\tilde{f}_m$  to  $-\sin$ .

The uniform convergence of  $f_m$  to  $-\sin$  on compact subsets of  $\mathbb{C}$  implies that the roots of  $f_m$  converge to the roots of  $-\sin$ . Let us elaborate on this in the following theorem to finish our explanation of  $O_2$  in the  $n = 4m$  case. The theorem expresses the fact that for any fixed  $k^* \in \mathbb{N}^+$  there exists an index  $m_0(k^*)$  such that for each  $m > m_0(k^*)$  we have  $k_0(4m) \geq k^*$  in decomposition (4.1), and the sequence of intervals  $I_{4m,k^*} \subset \mathcal{V}_{4m}^+$  converges to the interval  $[(2k^* - 2)\pi, (2k^* - 1)\pi]$  as  $m \rightarrow +\infty$ . We essentially repeat the proof of Hurwitz's theorem via Rouché's theorem as given in [Titchmarsh \(1939, p. 119\)](#) (cf. [Varga & Carpenter, 2001, pp. 374–375](#)).

**THEOREM 4.5** For any fixed  $k^* \geq 1$ ,

$$\lim_{m \rightarrow +\infty} \min I_{4m,k^*} = (2k^* - 2)\pi \quad \text{and} \quad \lim_{m \rightarrow +\infty} \max I_{4m,k^*} = (2k^* - 1)\pi.$$

*Proof.* We set  $\varrho := (2k^* - 1)\pi + \pi/3$ ,  $\varepsilon := \pi/4$  and let  $\eta \equiv \eta(\varrho, \varepsilon) := \min\{|-\sin(z)| : z \in \partial D_\varrho \cup \partial D_\varepsilon\}$ . Then for any  $\ell \in \mathbb{Z}$ ,  $0 < \eta \leq \min\{|-\sin(z)| : z \in \partial D_\varepsilon(\ell\pi)\}$ . By Lemma 4.3, there is an index  $m_0(\varrho, \varepsilon) > 0$  such that for all  $m > m_0(\varrho, \varepsilon)$  we have

$$|f_m - (-\sin)| < \eta \leq |-\sin| \quad \text{on } \partial D_\varrho \cup \bigcup_{\ell=-2k^*+1}^{2k^*-1} \partial D_\varepsilon(\ell\pi).$$

So let us fix  $m > m_0(\varrho, \varepsilon)$  arbitrarily. We show that the decomposition (4.1) of  $\mathcal{V}_{4m}^+$  consists of at least  $k^*$  intervals—that is,  $k_0(4m) \geq k^*$ —and  $\max I_{4m,k^*} < \varrho$ .

Indeed, Rouché's theorem asserts, on the one hand, that both  $-\sin$  and  $f_m$  have the same number of zeros (counted with multiplicity) in  $D_\varrho$ , that is, the number of zeros is  $4k^* - 1$  for both functions; on the other hand, for each  $\ell \in [-2k^* + 1, 2k^* - 1] \cap \mathbb{Z}$  they have the same number of zeros in  $D_\varepsilon(\ell\pi)$  as well, that is, they both have a unique zero in each such disc. Since  $f_m$  has real coefficients, this means that  $f_m$  has a unique real zero in  $D_\varepsilon(\ell\pi)$  for each  $\ell \in [-2k^* + 1, 2k^* - 1] \cap \mathbb{Z}$ , also implying that all

complex zeros of  $f_m$  in  $D_\varrho$  are real and simple. But  $f_m(0) = 0$ , so the number of zeros of  $f_m$  in the interval  $(0, \varrho)$  is exactly  $2k^* - 1$ . Let  $0 < y_{4m,1} < y_{4m,2} < \dots < y_{4m,2k^*-1} < \varrho$  denote these zeros of  $f_m$ , and set  $y_{4m,0} := 0$ . By the fact that zeros of  $f_m$  are simple, and by taking into account Corollary 4.2 and (4.3), we have  $f_m(y) \leq 0$  for  $y \in [y_{4m,0}, y_{4m,1}]$ , and hence  $f_m(y) \leq 0$  for  $y \in \bigcup_{k=0}^{k^*-1} [y_{4m,2k}, y_{4m,2k+1}]$ , and  $f_m(y) > 0$  for  $y \in (\bigcup_{k=1}^{k^*-1} (y_{4m,2k-1}, y_{4m,2k})) \cup (y_{4m,2k^*-1}, \varrho]$ . In other words,  $\mathcal{V}_{4m}^+ \cap [0, \varrho] = \bigcup_{k=0}^{k^*-1} [y_{4m,2k}, y_{4m,2k+1}]$ .

Finally, by repeating the above argument with some  $\varepsilon_j \rightarrow 0^+$  ( $\varepsilon_j \leq \varepsilon$ ,  $j = 1, 2, \dots$ ) instead of  $\varepsilon$ , the convergence of the end points of the  $I_{4m,k^*}$  intervals to the corresponding multiples of  $\pi$  as  $m \rightarrow +\infty$  is also established.  $\square$

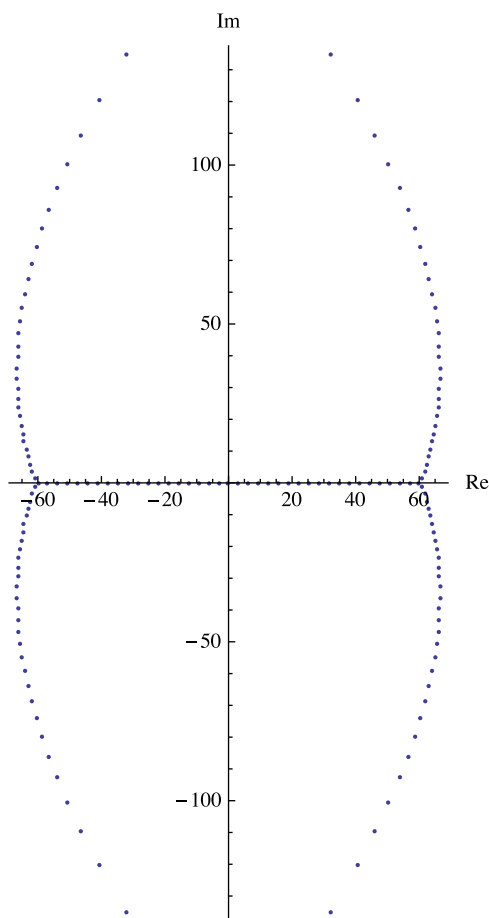
**REMARK 4.6** The convergence of the end points of the intervals  $I_{4m,k^*}$  (as  $m \rightarrow +\infty$  and  $k^*$  is fixed) to the corresponding multiples of  $\pi$  seems to be monotone if the first few ( $I_{4,k^*}, I_{8,k^*}, \dots$ ) intervals are ignored, but we did not prove this. Moreover, Remark 4.4 indicates why this convergence is relatively slow.

**REMARK 4.7** It is interesting to apply the notions of order stars (Wanner *et al.*, 1978) to the current setting. By Wanner *et al.* (1978, Proposition 3), we know that near the origin, the order star for  $\mathcal{P}_n$  and its complement each consist of  $n + 1$  alternating sectors of equal angular size. This leads easily to a proof of Corollary 4.2. Meanwhile, Wanner *et al.* (1978, Propositions 2 and 4) indicate that the order star for  $\mathcal{P}_n$  has  $n$  bounded dual fingers, each containing one zero of  $\mathcal{P}_n$ . The dual fingers correspond (near the origin) to the sectors belonging to the complement of the order star, so approximately half of them start in the right half-plane. But according to Szegő, more than half of the zeros of  $\mathcal{P}_n$  lie in the left half-plane. This means that a certain fraction of the dual fingers must cross the imaginary axis. These crossings correspond to the gaps in  $\mathcal{V}_n^+$ . If one supposes that each crossing and each gap between crossings have equal length, one obtains that the width of each finger where it crosses the imaginary axis must be  $\pi$ .

As for  $O_3$ , we note that the upper bound  $n/e + (\ln n)/(2e) + 1.2604$  (or its possibly modified version for  $n > 100$ ) is an upper bound on the largest positive root of  $\tilde{E}_n$ . Since these polynomials are uniformly close to  $\pm \sin$  or  $\pm \cos$  on compact sets of  $\mathbb{C}$  for large  $n$  values (Lemma 4.3), it is reasonable to expect (but we do not investigate this further here either) that the analysis presented in Varga & Carpenter (2001) is applicable in the current situation as well, at least for  $n$  large enough: compare our Fig. 9 with Varga & Carpenter (2001, Figures 1.1 and 1.2), or our expression  $n/e + (\ln n)/(2e) + 1.2604$  with Szegő's asymptotic result (Varga & Carpenter, 2001, Formula (1.12)) and its improvement (Varga & Carpenter, 2001, Section 5).

Let us close the section by presenting some further observations based on computations. In the numerical integration of oscillatory problems, one is interested in the size of the component of  $\mathcal{V}_n^+$  that is connected to the origin. The largest connected component of  $\mathcal{V}_n^+$  ( $1 \leq n \leq 100$ ) containing 0 occurs for  $n = 8$ , when  $\mathcal{V}_8^+ = [0, y_{8,1}]$  with  $y_{8,1} \approx 3.3951402205749$ ; the largest scaled component (i.e., when the  $n$ th column of Fig. 7 is shrunk by a factor  $n$ ) occurs for  $n = 4$ . For large  $n$ , the connected component of  $\mathcal{V}_n^+$  ( $1 \leq n \leq 100$ ) containing 0 remains small, but the boundary of  $\mathcal{U}_n$  lies very close to the imaginary axis and the magnitude of the amplification factor along the imaginary axis is only slightly greater than unity (indeed, indistinguishable in double precision) over a relatively large interval. We can exactly determine the distance from the boundary of  $\mathcal{U}_n$  to the imaginary axis for a given  $n$  with *Mathematica*: as an illustration, we chose  $n = 6$  (Fig. 10(a)),  $n = 20$  (Fig. 10(b)) and  $n = 100$  (Fig. 11). Below we describe the technique we used to create these figures.

For a given  $y \geq 0$ , the real solutions  $x$  of  $|\sum_{k=0}^n ((x + iy)^k / k!)| \leq 1$  are computed and the solution with the smallest absolute value is denoted by  $x_n^{\min}(y)$ . In *Mathematica*, the function  $y \mapsto x_n^{\min}(y)$

FIG. 9. The complex zeros of the polynomial  $f_{40}$  defined in (4.2).

(where  $y$  runs over some interval) can be represented as a piecewise defined function composed of `Root` objects. However, the  $x_n^{\min}(\cdot)$  function typically spans several orders of magnitude, for example,

$$\begin{aligned} x_{100}^{\min}(0.1) &\approx 10^{-262}, & x_{100}^{\min}(1) &\approx 9 \times 10^{-161}, & x_{100}^{\min}(10) &\approx -5 \times 10^{-60}, \\ x_{100}^{\min}(20) &\approx 2 \times 10^{-29}, & x_{100}^{\min}(30) &\approx -1.5 \times 10^{-11}, & x_{100}^{\min}(38.1) &\approx -0.639. \end{aligned}$$

We add that any approximate real number above can be exactly be represented as a root of an integer polynomial of degree 200, whose coefficients can typically be written altogether by approximately 76000 digits. So to display the graph of  $x_n^{\min}(\cdot)$  in a meaningful way, some scaling has to be applied. For fixed  $n$  values, Figs 10 and 11 actually display the curves

$$y \mapsto \text{sign}(x_n^{\min}(y)) \cdot \frac{-1}{\log_{10} |x_n^{\min}(y)|}, \quad (4.5)$$

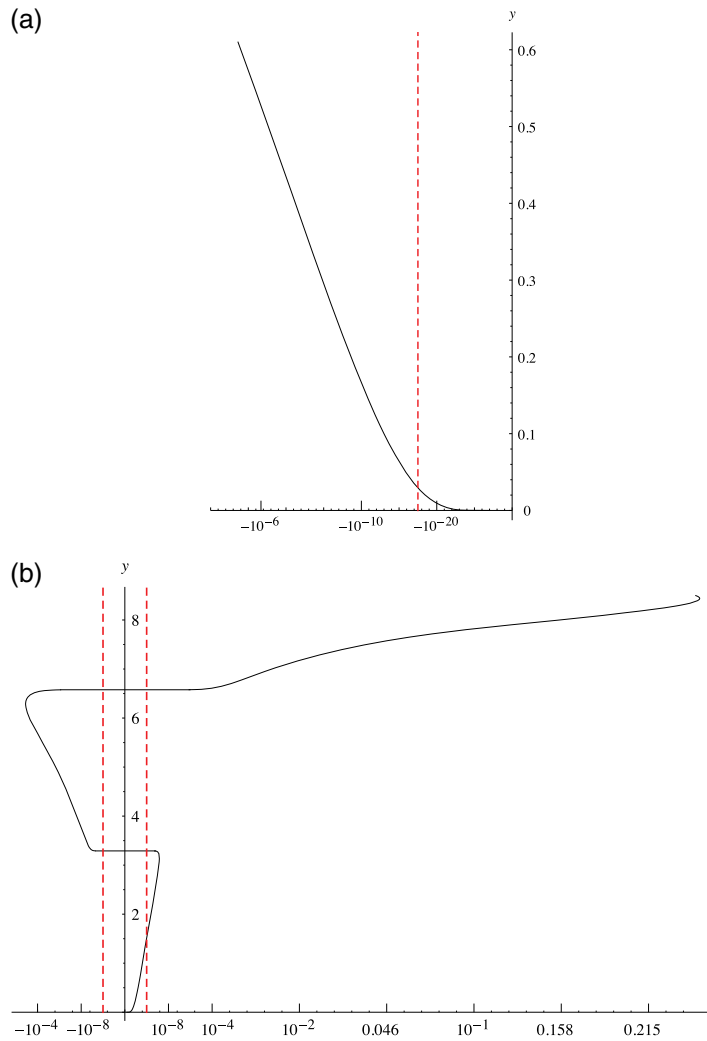


FIG. 10. The boundary curve of  $\mathcal{U}_n$  closest to the imaginary axis, on an 'inverse logarithmic scale', for (a)  $n=6$  and (b)  $n=20$ ; see (4.5).

with  $y$  values measured along the vertical axis (corresponding to the imaginary axis) and function values along the horizontal one (corresponding to the real axis). Since for each  $n$  and  $y$  value we now have  $|x_n^{\min}(y)| < 1$ , the sign correction ensures that a point on the figure is in the open left half-plane (or right half-plane) if and only if  $\text{sign}(x_n^{\min}(y)) = -1$  (or  $1$ ). The intersection points of the graph of  $x_n^{\min}(\cdot)$  and the vertical axis correspond to the end points of the shaded intervals in Fig. 7. The curve segments bounded by the vertical red dashed lines (placed at  $\pm 10^{-16}$ ) correspond to stability region boundaries that are 'invisible' by using machine precision. As for the vertical (black or red) dotted lines in Fig. 11, they measure the amplitude of the oscillations (i.e., the local extrema) of (4.5) in the interval  $[-10^{-16}, 10^{-16}]$ ,

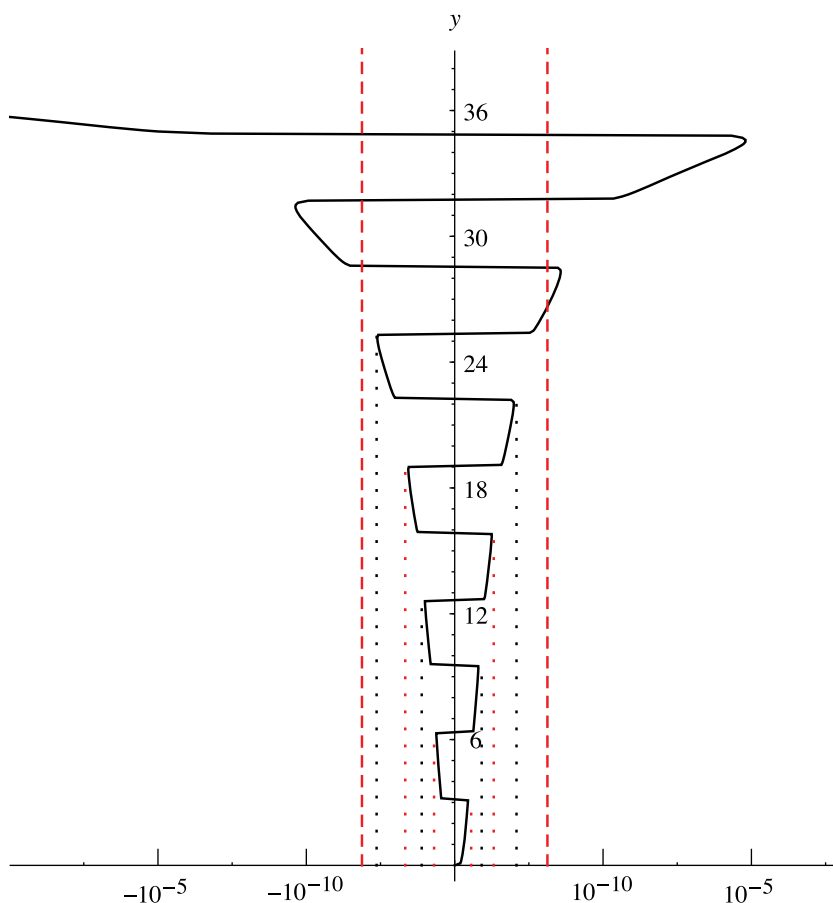


FIG. 11. The boundary curve of  $\mathcal{U}_{100}$  closest to the imaginary axis, on an ‘inverse logarithmic scale’; see (4.5).

and are found approximately at

$$-10^{-19}, -10^{-30}, -10^{-45}, -10^{-72}, 10^{-90}, 10^{-55}, 10^{-38}, 10^{-24}.$$

### 5. Semidisks contained in $\mathcal{S}_n$ in the left half-plane

Results of this section are formulated mostly in terms of the scaled stability region  $\mathcal{S}_n$ .

First we prove a theorem giving a necessary and sufficient condition for a small enough semidisc in the left half-plane and centred at the origin to be contained in  $\mathcal{S}_n$ . Then we present the radius of the maximal such semidisks for  $n \leq 20$ . Finally, we show some figures concerning the radial slices of  $\mathcal{S}_n \cap \{\Re \leq 0\}$ .

**THEOREM 5.1** Let  $n$  be a positive integer. Then

$$\exists \varrho > 0: \quad D_{\varrho} \cap \{\Re \leq 0\} \subset \mathcal{S}_n,$$

if and only if  $n \equiv 0 \pmod{4}$  or  $n \equiv 3 \pmod{4}$ .

*Proof.* Corollary 4.2 establishes that  $D_\varrho \cap \{\Re \leq 0\} \not\subset \mathcal{S}_n$  for any  $\varrho > 0$  and  $n \equiv 1 \pmod{4}$  or  $n \equiv 2 \pmod{4}$ , so to finish the proof we show that a sufficiently small closed left semidisc centred at the origin is contained in  $\mathcal{S}_n$  for  $n \equiv 0 \pmod{4}$  or  $n \equiv 3 \pmod{4}$ .

To this end we fix an  $n \geq 1$  with  $n \equiv 0 \pmod{4}$  or  $n \equiv 3 \pmod{4}$  and note that

$$\left| \sum_{k=0}^n \frac{(x + iy)^k}{k!} \right|^2 = 1 + 2x + P_n(x, y),$$

where  $x, y \in \mathbb{R}$  and the polynomial  $P_n$  has the form

$$P_n(x, y) = \sum_{k \geq 0, \ell \geq 0, k+\ell \geq 2, \ell \text{ is even}} a_{k,\ell} x^k y^\ell,$$

with some real coefficients  $a_{k,\ell}$ . Then with  $x = r \cos \varphi$  and  $y = r \sin \varphi$  ( $r \geq 0$ ,  $\pi/2 \leq \varphi \leq \pi$ ; due to symmetry, only the upper left quadrant is considered) we have

$$\partial_\varphi \left( \left| \sum_{k=0}^n \frac{(r \exp(i\varphi))^k}{k!} \right|^2 \right) = (r \sin \varphi)(-2 + rQ_n(r, \cos \varphi, \sin \varphi)),$$

where  $Q_n$  is a suitable real polynomial in three variables. Consequently, there exists  $r_n^* > 0$  such that for any fixed  $0 \leq r \leq r_n^*$  the function

$$\left[ \frac{\pi}{2}, \pi \right] \ni \varphi \mapsto \left| \sum_{k=0}^n \frac{(r \exp(i\varphi))^k}{k!} \right|^2$$

is nonincreasing, hence its maximal value occurs (for example) at  $\varphi = \pi/2$ . So by the second part of Corollary 4.2, for  $|z| \leq \min(r_n^*, \varrho_n)$  and  $\Re(z) \leq 0$  we have

$$\left| \sum_{k=0}^n \frac{z^k}{k!} \right|^2 \leq \left| \sum_{k=0}^n \frac{(i|z|)^k}{k!} \right|^2 \leq 1.$$

□

For any  $3 \leq n \leq 20$  and  $n \equiv 0 \pmod{4}$  or  $n \equiv 3 \pmod{4}$ , we have determined the maximal  $\varrho_n^* > 0$  radius as an exact algebraic number such that  $D_{\varrho_n^*} \cap \{\Re \leq 0\} \subset \mathcal{S}_n$  (see Table 3) as follows. It is seen from the definition of  $\mathcal{V}_n^+$  (and by taking into account the scaling) that the length of the largest interval in  $\mathcal{V}_n^+$  containing 0 is an upper bound on  $n\varrho_n^*$ . Let  $y_{n,1} > 0$  denote the length of this largest interval. We first exactly determine  $y_{n,1}$  with *Mathematica*'s Reduce (by locating the smallest positive root of the appropriate polynomial in Lemma 4.1), then show (with Reduce again) that no real numbers  $x$  and  $y$



can satisfy the system

$$x^2 + y^2 \leq (y_{n,1})^2, \quad x \leq 0, \quad \left| \sum_{k=0}^n \frac{(x + iy)^k}{k!} \right| > 1,$$

proving that  $\varrho_n^* = y_{n,1}/n$ . Interestingly, the above simple approach breaks down for  $n = 4$ : it turns out that  $\varrho_4^* < y_{4,1}/4$ ; see Fig. 12.

Thus, it seems that the radius of the maximal semidisc included in  $\mathcal{U}_n$  tends to  $\pi/2$  for  $n \equiv 3 \pmod{4}$  and to  $\pi$  for  $n \equiv 0 \pmod{4}$ . But it is clear from material in earlier sections that, excepting a small region near the imaginary axis,  $\mathcal{U}_n$  covers a much larger semidisc (of radius approximately  $n/e$ ). It is interesting to examine by how much the boundary of

$$\mathcal{S}_n \cap \{\Re \leq 0\} \cap \{\Im \geq 0\}$$

deviates from the asymptotic semidisc shape indicated by Theorem 1.3.

For a fixed  $n \geq 1$  and some  $\varphi \in [\pi/2, \pi]$ , let us define the nonempty set

$$\mathcal{R}_n(\varphi) := \{r \in \mathbb{R} : r \geq 0, \quad r e^{i\varphi} \in \mathcal{S}_n\}. \quad (5.1)$$

For each  $2 \leq n \leq 20$  we determined  $\mathcal{R}_n(\varphi)$  for approximately 160 different  $\varphi \in [\pi/2, \pi]$  values. These investigations suggest that if a small wedge near the imaginary axis is ignored, then  $\mathcal{S}_n \cap \{\Re \leq 0\} \cap \{\Im \geq 0\}$  is a starlike set in  $\mathbb{C}$  with respect to the origin. In other words, we conjecture that for each  $n \geq 1$  there exists a  $\varphi_n \in [0, \pi/2)$  with  $\varphi_n \ll 1$  such that for every  $\varphi$  with  $\pi/2 + \varphi_n \leq \varphi \leq \pi$ , the set  $\mathcal{R}_n(\varphi)$  is a compact interval with  $\min(\mathcal{R}_n(\varphi)) = 0$ . For example, the following values of  $\varphi_n$  seem to be appropriate:  $\varphi_4 = 0$ ;  $\varphi_6 = 0$  with  $\mathcal{R}_6(\pi/2) = \{0\}$  but  $\mathcal{R}_6(\pi/2 + \varepsilon)$  (for any  $0 < \varepsilon \leq \pi/2$ ) being a nondegenerate interval; and there is a suitable  $\varphi_{20}$  already in  $(0, 5 \times 10^{-5})$ . Figure 13 shows the graphs of

$$\left[ \frac{\pi}{2} + \varphi_n, \pi \right] \ni \varphi \mapsto \max(\mathcal{R}_n(\varphi)), \quad (5.2)$$

TABLE 3 For  $3 \leq n \leq 20$ , the table contains the maximal  $\varrho_n^*$  values (whenever they are positive) such that  $D_{\varrho_n^*} \cap \{\Re \leq 0\} \subset \mathcal{S}_n$ . Instead of listing exact algebraic numbers (apart from the first row, and with degrees given in the last column), the values of  $\varrho_n^*$  are rounded down. For convenience, the maximal inner radius for each unscaled stability region is also given (as  $n\varrho_n^*$ ). The exceptional  $n = 4$  case is displayed in Fig. 12

$n$	$\varrho_n^*$	$n\varrho_n^*$	The algebraic degree of $\varrho_n^*$
3	$\sqrt{3}/3 \approx 0.577$	$\approx 1.732$	2
7	$\approx 0.252$	$\approx 1.764$	6
11	$\approx 0.154$	$\approx 1.701$	10
15	$\approx 0.111$	$\approx 1.668$	14
19	$\approx 0.086$	$\approx 1.649$	18
4	$\approx 0.653$	$\approx 2.615$	24
8	$\approx 0.424$	$\approx 3.395$	6
12	$\approx 0.281$	$\approx 3.379$	10
16	$\approx 0.207$	$\approx 3.324$	14
20	$\approx 0.164$	$\approx 3.290$	18

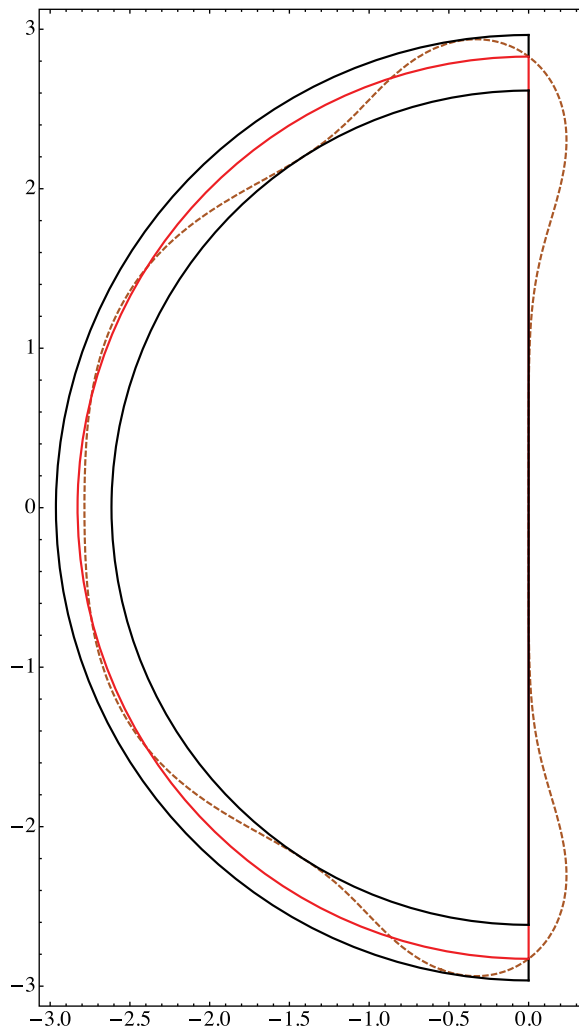


FIG. 12. The figure shows the smallest left semidisc (outer semidisc) that contains / the largest left semidisc (inner semidisc) that is contained in  $\mathcal{U}_4$  (dashed curve). The middle semidisc (red in the online version of the article) has radius  $y_{4,1} = \sqrt{8}$  (note that  $\mathcal{V}_4^+$  is connected, so  $y_{4,1} = \max(\mathcal{V}_4^+)$ ); see the corresponding row of Table 3 as well.

for  $n=4$ ,  $n=6$  and  $n=20$ , with linear interpolation between the approximately 160 different  $\varphi$  values in each case. Note that—for a particular  $n$ —the value  $\max_{\pi/2 \leq \varphi \leq \pi}(\max(\mathcal{R}_n(\varphi)))$  is found in Table 2. Meanwhile,  $\min_{\pi/2 \leq \varphi \leq \pi}(\max(\mathcal{R}_4(\varphi)))$  corresponds to  $\varrho_4^*$  in Table 3; also compare the orange curve in Fig. 13 and the dashed brown curve in Fig. 12. As for the brown curve in Fig. 13,  $\min_{\pi/2 \leq \varphi \leq \pi}(\max(\mathcal{R}_6(\varphi))) = 0$ , in accordance with the fact that  $\mathcal{V}_6^+ = \{0\}$  (see Fig. 6). Finally, as for the black curve in Fig. 13, the scaled function value around  $20 \cdot 0.445 \approx 8.9$  corresponding to  $\varphi \approx \pi/2 + \varphi_{20}$  is the highest point of the upper shaded rectangle in Fig. 6; the highest point of the lower shaded rectangle is  $20\varrho_{20}^*$  in Table 3.

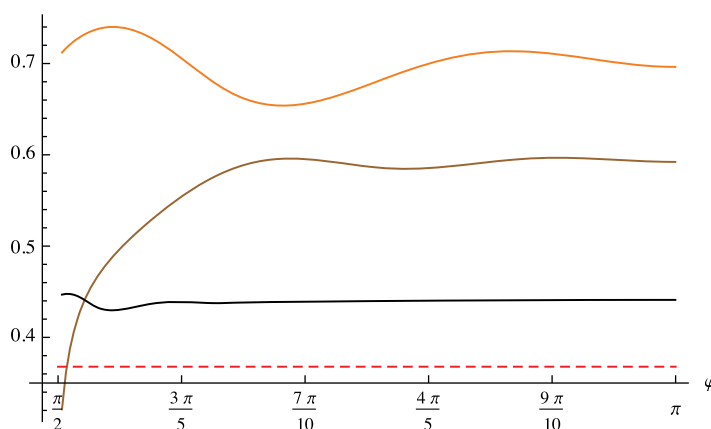


FIG. 13. The extent of  $\mathcal{S}_n$  as a function of the angle  $\varphi$  (see (5.1) and (5.2)). The orange (upper), brown (middle) and black (lower) curves correspond to  $n = 4$ ,  $n = 6$  and  $n = 20$ , respectively. The red line is placed at  $1/e$ .

## 6. Auxiliary lemmas

Below we prove some additional results that were referenced and used in earlier sections.

LEMMA 6.1 For any  $n \in \mathbb{N}^+$  we have  $(n/e)^n \sqrt{2\pi n} < n! \leq e(n/e)^n \sqrt{n}$ .

*Proof.* The proof is a standard monotonicity argument, hence omitted here.  $\square$

In the next lemma and later, we make use of the Lambert  $W$  function (a.k.a. `ProductLog` in *Mathematica*): recall that for  $x \geq -1/e$ , there is a unique  $W(x) \geq -1$  such that

$$x = W(x) e^{W(x)}. \quad (6.1)$$

LEMMA 6.2 The set  $\Sigma_1 \subset \mathbb{C}$  is strictly convex.

*Proof.* By identifying  $\mathbb{C}$  with  $\mathbb{R}^2$ , we see that

$$\Sigma_1 = \{(x, y) \in \mathbb{R}^2 : -W(1/e) \leq x \leq 1, -\sqrt{e^{2x-2} - x^2} \leq y \leq \sqrt{e^{2x-2} - x^2}\}.$$

The proof is finished by checking that  $(\sqrt{e^{2x-2} - x^2})'' < 0$  for  $-W(1/e) < x < 1$ .  $\square$

REMARK 6.3 The value of  $-W(1/e)$  is  $\approx -0.278464543$ . The constant  $W(1/e)$  can be obtained, for example, from the series expansion  $\sum_{n=1}^{+\infty} ((-n)^{n-1}/n!)z^n$  of the  $W$  function about 0. Note that the convergence radius of this alternating series is  $1/e$ . For an account of the  $W$  function, see, for example, Corless et al. (1996).

LEMMA 6.4 The inclusion  $D_{1/4} \subset \Sigma_1$  holds.

*Proof.* Since  $|z| \leq \frac{1}{4}$  implies  $\Re(z) \geq -\frac{1}{4}$ , we have  $|ze^{1-z}| \leq \frac{1}{4}e^{5/4} < 1$ .  $\square$

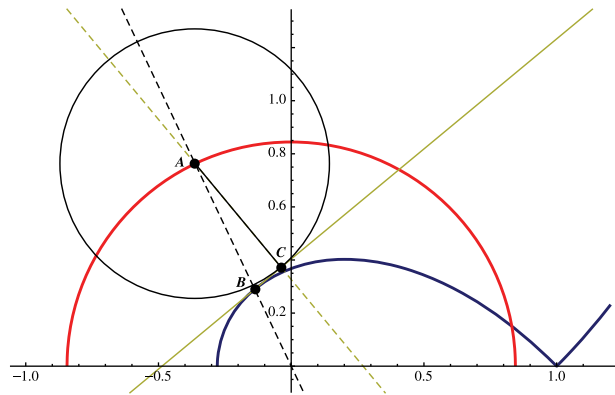


FIG. 14. The construction used in the proof of Lemma 6.5.

LEMMA 6.5 We have  $C_{1/e} \cap \Sigma_1 = \{\pm i/e\}$ . On the other hand, for any  $\varrho > 1/e$ ,  $C_\varrho \cap \Sigma_1 = \emptyset$  and

$$|C_\varrho, \Sigma_1| \geq \frac{\varrho e - 1}{\sqrt{e^2 + 1}}.$$

*Proof.* The intersections  $C_\varrho \cap \Sigma_1$  (for  $\varrho \geq 1/e$ ) are determined by using the explicit representation of  $\Sigma_1$  given in the proof of Lemma 6.2 combined with the inequality

$$\sqrt{x^2 + (\sqrt{e^{2x-2} - x^2})^2} \leq \frac{1}{e}$$

valid for  $-W(1/e) \leq x \leq 0$ , and noting that ‘=’ in ‘ $\leq$ ’ above holds precisely for  $x = 0$ . As for estimating the distance between the sets  $C_\varrho$  and  $\Sigma_1$  (viewed as subsets of  $\mathbb{R}^2$ ), let us fix some  $\varrho > 1/e$  and  $\varphi \in [\pi/2, 3\pi/2]$ , and consider the point  $(\varrho \cos \varphi, \varrho \sin \varphi) \in C_\varrho$ , depicted as point **A** in Fig. 14. Due to symmetry, we can assume  $\varphi \in [\pi/2, \pi]$ . Then the line passing through point **A** and the origin intersects  $\Sigma_1$  at point **B**. We consider the tangent line to  $\Sigma_1$  at **B**. The closest point on this tangent line to **A** is point **C**. Since  $\Sigma_1$  is convex, the distance  $|\mathbf{A}, \mathbf{C}|$  is a lower estimate for  $|\mathbf{A}, \Sigma_1|$ . We keep  $\varrho$  fixed, but vary  $\varphi$ , so  $|\mathbf{A}, \mathbf{C}| \equiv |\mathbf{A}(\varphi), \mathbf{C}(\varphi)|$ . Then clearly,

$$|C_\varrho, \Sigma_1| \geq \inf_{\varphi \in [\pi/2, \pi]} |\mathbf{A}(\varphi), \mathbf{C}(\varphi)|.$$

In Step 1 below, we show that for  $\pi/2 < \varphi < \pi$  we have

$$|\mathbf{A}(\varphi), \mathbf{C}(\varphi)| = \frac{|W((- \cos \varphi)/e) + 1| |W((- \cos \varphi)/e) + \varrho \cos \varphi|}{\sqrt{(W((- \cos \varphi)/e) + \cos^2 \varphi)^2 + \cos^2 \varphi \sin^2 \varphi}}.$$

Obviously, the function  $\varphi \mapsto |\mathbf{A}(\varphi), \mathbf{C}(\varphi)|$  is continuous at  $\varphi = \pi/2$  and  $\varphi = \pi$  as well.

On the other hand, we show in Step 2 below that  $(\pi/2, \pi) \ni \varphi \mapsto |\mathbf{A}(\varphi), \mathbf{C}(\varphi)|$  is strictly increasing, and

$$\lim_{\varphi \rightarrow \pi/2^+} |\mathbf{A}(\varphi), \mathbf{C}(\varphi)| = \frac{\varrho e - 1}{\sqrt{e^2 + 1}}.$$

Consequently,  $|C_\varrho, \Sigma_1| \geq \inf_{\varphi \in [\pi/2, \pi]} |\mathbf{A}(\varphi), \mathbf{C}(\varphi)| = (\varrho e - 1)/\sqrt{e^2 + 1}$ , so the proof of the lemma is finished by proving Steps 1 and 2.

**Step 1.** Let us fix some  $\varrho > 1/e$  and  $\pi/2 < \varphi < \pi$ . We see from the construction that the coordinates of  $\mathbf{B} = (x, y)$  satisfy  $y = \sqrt{e^{2x-2} - x^2}$  and  $y = x \tan \varphi$ , so (by using  $\tan^2 \varphi + 1 = 1/\cos^2 \varphi$ ) we have  $|x/\cos \varphi| = x/\cos \varphi = e^{x-1}$ , that is,  $x = -W((- \cos \varphi)/e)$ . This yields

$$\mathbf{B} = \left( -W\left(-\frac{\cos \varphi}{e}\right), -W\left(-\frac{\cos \varphi}{e}\right) \tan \varphi \right),$$

so the tangent line to  $\Sigma_1$  at  $\mathbf{B}$  has slope

$$\begin{aligned} & \left( \sqrt{e^{2x-2} - x^2} \right)' \Big|_{x=-W((- \cos \varphi)/e)} \\ &= \frac{e^{-2W((- \cos \varphi)/e)-2} + W((- \cos \varphi)/e)}{\sqrt{e^{-2W((- \cos \varphi)/e)-2} - W((- \cos \varphi)/e)^2}} = -\frac{1}{\sin \varphi \cos \varphi} \left( W\left(-\frac{\cos \varphi}{e}\right) + \cos^2 \varphi \right). \end{aligned}$$

Now we see that the equation for the line passing through points  $\mathbf{B}$  and  $\mathbf{C}$  can be written as  $ax + by + c = 0$  with

$$\begin{aligned} a &= -\frac{1}{\sin \varphi \cos \varphi} \left( W\left(-\frac{\cos \varphi}{e}\right) + \cos^2 \varphi \right), \\ b &= -1 \end{aligned}$$

and

$$\begin{aligned} c &= -W\left(-\frac{\cos \varphi}{e}\right) \tan \varphi - \frac{1}{\sin \varphi \cos \varphi} \left( W\left(-\frac{\cos \varphi}{e}\right) + \cos^2 \varphi \right) W\left(-\frac{\cos \varphi}{e}\right) \\ &= -\frac{1}{\sin \varphi \cos \varphi} W\left(-\frac{\cos \varphi}{e}\right) \left( W\left(-\frac{\cos \varphi}{e}\right) + 1 \right). \end{aligned}$$

Therefore, the distance from point  $\mathbf{A} = (\varrho \cos \varphi, \varrho \sin \varphi)$  to this line is given by

$$\frac{|a\varrho \cos \varphi + b\varrho \sin \varphi + c|}{\sqrt{a^2 + b^2}} = \frac{|W((- \cos \varphi)/e) + 1| |W((- \cos \varphi)/e) + \varrho \cos \varphi|}{\sqrt{(W((- \cos \varphi)/e) + \cos^2 \varphi)^2 + \cos^2 \varphi \sin^2 \varphi}}.$$

**Step 2.** It is convenient to set  $w \equiv w(\varphi) := W((- \cos \varphi)/e)$ . Then  $w$  is a strictly increasing bijection, mapping  $[\pi/2, \pi]$  onto  $[0, W(1/e)]$ . For  $\pi/2 < \varphi < \pi$ , that is, for  $0 < w < W(1/e)$ , we have  $\cos \varphi = -w e^{w+1}$  and  $\sin \varphi = \sqrt{1 - e^{2w+2} w^2}$ . Hence, by Step 1 we have

$$|\mathbf{A}(\varphi), \mathbf{C}(\varphi)| = \frac{|w + 1| |1 - \varrho e^{w+1}|}{\sqrt{e^{2w+2}(2w + 1) + 1}}.$$

Now  $w + 1 > w > 0$ , and for  $\varrho > 1/e$  one has  $1 - \varrho e^{w+1} < 0$ , so

$$|\mathbf{A}(\varphi), \mathbf{C}(\varphi)| = \frac{(w + 1)(\varrho e^{w+1} - 1)}{\sqrt{e^{2w+2}(2w + 1) + 1}}.$$

On the other hand, it is trivial that

$$\partial_w \left( \frac{(w+1)(\varrho e^{w+1} - 1)}{\sqrt{e^{2w+2}(2w+1) + 1}} \right) = \frac{e^{2w+2}(2w^2 + 2w + 1) + \varrho e^{3w+3}w + \varrho e^{w+1}(w+2) - 1}{(e^{2w+2}(2w+1) + 1)^{3/2}} > 0,$$

for any  $\varrho > 1/e$  and  $w > 0$ , so  $(\pi/2, \pi) \ni \varphi \mapsto |\mathbf{A}(\varphi), \mathbf{C}(\varphi)|$  is strictly increasing, and

$$\lim_{\varphi \rightarrow \pi/2^+} |\mathbf{A}(\varphi), \mathbf{C}(\varphi)| = \lim_{w \rightarrow 0^+} \frac{(w+1)(\varrho e^{w+1} - 1)}{\sqrt{e^{2w+2}(2w+1) + 1}} = \frac{\varrho e - 1}{\sqrt{e^2 + 1}}.$$

□

REMARK 6.6 For  $\varrho > 1/e$  we have  $i\varrho \in C_\varrho$  and  $i/e \in \Sigma_1$ , so

$$|C_\varrho, \Sigma_1| \leq \varrho - \frac{1}{e}.$$

LEMMA 6.7 For any  $\sigma > 0$  we have

$$\inf \left\{ \left| \frac{w}{w-1} \right| : w \in \mathbb{C}, |w| \geq \sigma \right\} = \frac{\sigma}{1+\sigma}. \quad (6.2)$$

On the other hand, we have

$$\forall \varrho \in \left[ 0.95, \frac{1 + \sqrt{1 + e^2}}{e} \right] : \sup \left\{ \left| \frac{w}{w-1} \right| : w \in \mathbb{C}, |w, \Sigma_1| \geq \delta_\varrho, \Re(w) \leq \delta_\varrho \right\} \leq 1 \quad (6.3)$$

and

$$\forall \varrho \in \left( \frac{1 + \sqrt{1 + e^2}}{e}, 2 \right] : \sup \left\{ \left| \frac{w}{w-1} \right| : w \in \mathbb{C}, |w, \Sigma_1| \geq \delta_\varrho, \Re(w) \leq \delta_\varrho \right\} \leq 1.39, \quad (6.4)$$

where

$$\delta_\varrho := \frac{\varrho e - 1}{2\sqrt{e^2 + 1}}. \quad (6.5)$$

*Proof.* For (6.2), we rewrite the expression as

$$\inf \left\{ \sqrt{x^2 + y^2} / \sqrt{(x-1)^2 + y^2} : x, y \in \mathbb{R}, \sqrt{x^2 + y^2} \geq \sigma \right\},$$

and see that  $\sqrt{(-\sigma)^2 + 0^2} / \sqrt{(-\sigma-1)^2 + 0^2} = \sigma / (1 + \sigma)$ . On the other hand, we directly verify that

$$\frac{\sqrt{x^2 + y^2}}{\sqrt{(x-1)^2 + y^2}} < \frac{\sigma}{1 + \sigma}, \quad \sqrt{x^2 + y^2} \geq \sigma$$

has no real  $x, y$  solutions for any  $\sigma > 0$ .

As for (6.3), we have  $0 < \delta_\varrho \leq \frac{1}{2}$  for the given  $\varrho$  values, so

$$\sup \left\{ \left| \frac{w}{w-1} \right| : w \in \mathbb{C}, |w, \Sigma_1| \geq \delta_\varrho, \Re(w) \leq \delta_\varrho \right\} \leq \sup \left\{ \left| \frac{w}{w-1} \right| : w \in \mathbb{C}, \Re(w) \leq \frac{1}{2} \right\} = 1.$$

Finally, for (6.4), we note that, due to  $1 \in \Sigma_1$ ,  $|w, \Sigma_1| \geq \delta_\varrho$  implies  $|w, 1| \geq \delta_\varrho$ , so the supremum is estimated from above by  $\sup\{|w/(w-1)| : w \in \mathbb{C}, |w-1| \geq \delta_\varrho, \Re(w) \leq \delta_\varrho\}$ . Elementary computation shows that for each  $\varrho$  in the given range and for  $w \in \mathbb{C}$  with  $|w-1| \geq \delta_\varrho$  and  $\Re(w) \leq \delta_\varrho$ ,  $|w/(w-1)|$  is maximal when  $|w-1| = \delta_\varrho = \Re(w)$ . For such  $w$ , the maximal value of  $|w/(w-1)|$  is

$$\frac{\sqrt{e^2(\varrho^2 - 4) + 2e(2\sqrt{1+e^2} - 1)\varrho - 4\sqrt{1+e^2} - 3}}{\varrho e - 1} < 1.39.$$

□

## Funding

The three authors were supported by KAUST Award No. FIC/2010/05–2000000231. T.A.K. was also supported by TÁMOP–4.2.2.A–11/1/KONV–2012–0012 *Basic research for the development of hybrid and electric vehicles*. The TÁMOP project is supported by the Hungarian Government and co-financed by the European Social Fund.

## REFERENCES

- BLEHER, P. & MALLISON JR, R. (2006) Zeros of sections of exponential sums. *Int. Math. Res. Not.*, Art. ID 38937, **18**, 49p.
- BUCKHOLTZ, J. D. (1963) Concerning an approximation of Copson. *Proc. Amer. Math. Soc.*, **14**, 564–568.
- BUCKHOLTZ, J. D. (1966) A characterization of the exponential series. *Amer. Math. Monthly*, **73**, 121–123.
- CARPENTER, A. J., VARGA, R. S. & WALDVOGEL, J. (1991) Asymptotics for the zeros of the partial sums of  $e^z$ . *I. Rocky Mountain J. Math.*, **21**, 99–120.
- CORLESS, R. M., GONNET, G. H., HARE, D. E. G., JEFFREY, D. J. & KNUTH, D. E. (1996) On the Lambert  $W$  function. *Adv. Comput. Math.*, **5**, 329–359.
- DIEUDONNÉ, J. (1935) Sur les zéros des polynômes-sections de  $e^x$ . *Bull. Sci. Math.*, **70**, 333–351.
- GARDNER, R. B. & GOVIL, N. K. (1997) Some generalizations of the Eneström–Kakeya theorem. *Acta Math. Hungar.*, **74**, 125–134.
- JELTSCH, R. & NEVANLINNA, O. (1981) Stability of explicit time discretizations for solving initial value problems. *Numer. Math.*, **37**, 61–91.
- JELTSCH, R. & NEVANLINNA, O. (1982) Stability and accuracy of time discretizations for initial value problems. *Numer. Math.*, **40**, 245–296.
- KAPPERT, M. (1996) On the zeros of the partial sums of  $\cos(z)$  and  $\sin(z)$ . *Numer. Math.*, **74**, 397–417.
- KETCHESON, D. I., LÓCZI, L. & PARSANI, M. (2015) *SIAM J. Num. Anal.*, to appear. The corresponding unabridged technical report is available from <http://arxiv.org/abs/1309.1317>. Accessed 25 August 2014.
- MERKLE, M. J. (1995) Inequalities for residuals of power series: a review. *Univ. Beograd. Publ. Elektrotehn. Fak. Ser. Mat.*, **6**, 79–85.
- NEWMAN, D. J. & RIVLIN, T. J. (1972) The zeros of the partial sums of the exponential function. *J. Approx. Theory*, **5**, 405–412.
- NEWMAN, D. J. & RIVLIN, T. J. (1976) Correction to: “The zeros of the partial sums of the exponential function” (J. Approximation Theory **5** (1972), 405–412). *J. Approx. Theory*, **16**, 299–300.



- PRITSKER, I. E. & VARGA, R. S. (1997) The Szegő curve, zero distribution and weighted approximation. *Trans. Amer. Math. Soc.*, **349**, 4085–4105.
- SAFF, E. B. & VARGA, R. S. (1975) On the zeros and poles of Padé approximants to  $e^z$ . *Numer. Math.*, **25**, 1–14.
- SAFF, E. B. & VARGA, R. S. (1976) Zero-free parabolic regions for sequences of polynomials. *SIAM J. Math. Anal.*, **7**, 344–357.
- SANZ-SERNA, J. M. (1980) Some aspects of the boundary locus method. *BIT*, **20**, 97–101.
- SIEMIENIUCH, J. L. (1976) Properties of certain rational approximations to  $e^{-z}$ . *BIT*, **16**, 172–191.
- SZEGŐ, G. (1924) Über eine Eigenschaft der Exponentialreihe. *Sitzungsber. Berl. Math. Ges.*, **23**, 50–64.
- TITCHMARSH, E. C. (1939) *The Theory of Functions*, 2nd edn. Oxford: Oxford University Press.
- VARGA, R. S. & CARPENTER, A. J. (1990) Asymptotics for the zeros of the partial sums of  $e^z$ . II. *Computational Methods and Function Theory (Valparaíso, 1989)*. Lecture Notes in Mathematics, vol. 1435. Berlin: Springer, pp. 201–207.
- VARGA, R. S. & CARPENTER, A. J. (2000) Zeros of the partial sums of  $\cos(z)$  and  $\sin(z)$ . I. *Numer. Algorithms*, **25**, 363–375.
- VARGA, R. S. & CARPENTER, A. J. (2001) Zeros of the partial sums of  $\cos(z)$  and  $\sin(z)$ . II. *Numer. Math.*, **90**, 371–400.
- VARGA, R. S. & CARPENTER, A. J. (2010) Zeros of the partial sums of  $\cos(z)$  and  $\sin(z)$ . III. *Appl. Numer. Math.*, **60**, 298–313.
- VARGA, R. S., CARPENTER, A. J. & LEWIS, B. W. (2008) The dynamical motion of the zeros of the partial sums of  $e^z$ , and its relationship to discrepancy theory. *Electron. Trans. Numer. Anal.*, **30**, 128–143.
- VARGAS, A. R. (2012) Zeros of sections of some power series. *MSc Thesis*, Dalhousie University, Halifax.
- WALKER, P. (2003) The zeros of the partial sums of the exponential series. *Amer. Math. Monthly*, **110**, 337–339.
- WANNER, G., HAIRER, E. & NØRSETT, S. P. (1978) Order stars and stability theorems. *BIT Numer. Math.*, **18**, 475–489.
- YILDIRIM, C. Y. (1994) On the tails of the exponential series. *Canad. Math. Bull.*, **37**, 278–286.
- ZEMYAN, S. M. (2005) On the zeroes of the Nth partial sum of the exponential series. *Amer. Math. Monthly*, **112**, 891–909.