PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 124, Number 12, December 1996, Pages 3819–3833
S 0002-9939(96)03557-5

ASYMPTOTIC ANALYSIS OF DAUBECHIES POLYNOMIALS

JIANHONG SHEN AND GILBERT STRANG

(Communicated by James Glimm)

Dedicated to Gabor Szegö on the 100th anniversary of his birth

Abstract. To study wavelets and filter banks of high order, we begin with the zeros of $\mathbf{B}_p(y)$. This is the binomial series for $(1-y)^{-p}$, truncated after p terms. Its zeros give the $p-1$ zeros of the Daubechies filter inside the unit circle, by $z+z^{-1}=2-4y$. The filter has p additional zeros at $z=-1$, and this construction makes it orthogonal and maximally flat. The dilation equation leads to orthogonal wavelets with p vanishing moments. Symmetric biorthogonal wavelets (generally better in image compression) come similarly from a subset of the zeros of $\mathbf{B}_p(y)$.

We study the asymptotic behavior of these zeros. Matlab shows a remarkable plot for $p=70$. The zeros approach a limiting curve $|4y(1-y)|=1$ in the complex plane, which is the circle $|z-z^{-1}|=2$. All zeros have $|y|\le 1/2$, and the rightmost zeros approach $y=1/2$ (corresponding to $z=\pm i$) with speed $p^{-1/2}$. The curve $|4y(1-y)|=(4\pi p)^{1/2p}\,|1-2y|^{1/p}$ gives a very accurate approximation for finite p.

The wide dynamic range in the coefficients of $\mathbf{B}_p(y)$ makes the zeros difficult to compute for large p. Rescaling y by 4 allows us to reach $p=80$ by standard codes.

1. Introduction

Figure 1 shows the zeros of a particular polynomial of degree 69. The polynomial is the binomial series for $(1-y)^{-70}$, truncated after 70 terms. There is a close connection between those zeros and the 140 coefficients associated with the Daubechies wavelets D_{140}. Our first goal was to find the curve along which the zeros seem to lie.

This is the case $p=70$ of the truncated binomial series for $(1-y)^{-p}$

$$\mathbf{B}_p(y)=1+py+\frac{p(p+1)}{2}y^2+\ldots+\binom{2p-2}{p-1}y^{p-1}. \tag{1}$$

The natural question is the behavior of the zeros as $p\to\infty$. The outstanding contribution to problems of this type was by Szegö [8] in 1924, who studied the truncation of the exponential series. His limiting curve was $|ze^{1-z}|=1$, when the zeros are divided by p. For the truncated binomial $\mathbf{B}_p(y)$, $p>2$, we first prove that every zero satisfies $|Y|<1/2$ and $|4Y(1-Y)|>2^{1/p}$. All the zeros lie outside the limiting curve $|4y(1-y)|=1$. Their convergence to this curve $C=C_\infty$ is slowest near the point $y=1/2$, and we give a more exact expression $Y\approx 1/2+W/2\sqrt{p}$ for

Received by the editors June 25, 1995.
1991 *Mathematics Subject Classification.* Primary 41A58.

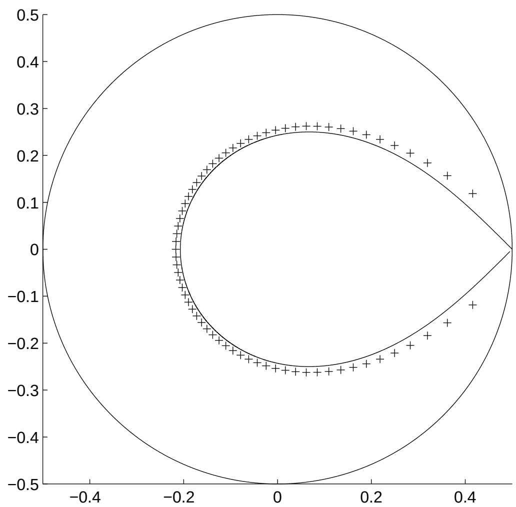

FIGURE 1. The zeros of $\mathbf{B}_{70}(y)$ are close to the curve C_∞.

the location of the rightmost zero. We also find a curve C_p that gives the positions of the other zeros to extra accuracy. The curve C_p lies slightly outside C_∞.

A note about the numerical computation of the zeros. Matlab creates the companion matrix whose characteristic polynomial is $\mathbf{B}_p(y)$. Then it finds the eigenvalues of that matrix. Without scaling, this breaks down at $p = 35$, because of the wide range in the coefficients of $\mathbf{B}_p(y)$. The first coefficient is 1, and by Stirling's formula, the coefficient of y^{p-1} is

$$\binom{2p-2}{p-1} \approx \frac{\sqrt{2\pi(2p-2)}}{2\pi(p-1)} \frac{(2p-2)^{2p-2}}{(p-1)^{2p-2}} = \frac{4^{p-1}}{\sqrt{\pi(p-1)}}. \tag{2}$$

The leading term 4^{p-1} suggests that the variable $4y$ is preferable to y. With this scaling, the Matlab computation remains accurate to $p = 80$. For larger p, a bifurcation (see Figure 2) occurs from roundoff error. The coefficient $\binom{p-1+i}{i}4^{-i}$ of $(4y)^i$ is numbered $b(p-i)$ by Matlab. Then $b(p) = 1$ and the sequence of coefficients is created recursively;

$$\text{for } i = p-1 : -1 : 1 \quad b(i) = b(i+1) * (2p-i-1)/(4*(p-i)). \tag{3}$$

The command "Y = roots$(b)/4$" produces the approximate zeros $Y(1), \ldots, Y(p-1)$.

Experiments with other root-finding algorithms were less successful, even though working with the companion matrix is *a priori* surprising. A polynomial with repeated roots leads to a defective matrix (not diagonalizable). Algorithms based on Newton's method had difficulty with the accurate evaluation of $\mathbf{B}_p$ and $\mathbf{B}_p'$. Lang's algorithm (Lang and Frenzel [4]) is comparable to Matlab 'roots', and probably faster.

We now explain how the zeros of $\mathbf{B}_p(y)$ are connected to the coefficients $h(n)$ that generate Daubechies wavelets. It is important to note that the same zeros also lead to biorthogonal filters and symmetric wavelets (cf. Daubechies [2], or Strang and Nguyen [7]). The Daubechies wavelets have orthogonality but not

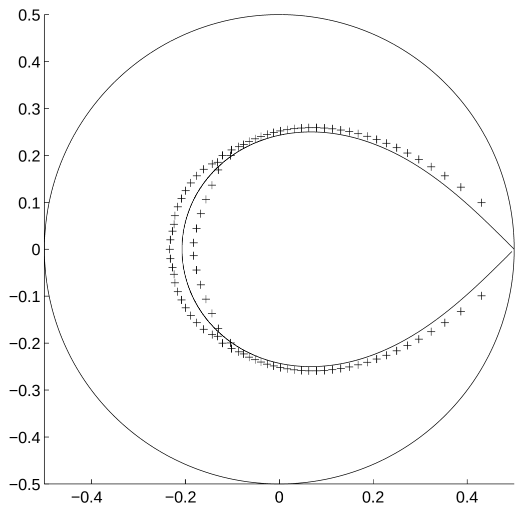

FIGURE 2. A bifurcation occurs from roundoff error, p=100.

symmmetry. The translates and dilates $w(2^j t - k)$ are an orthogonal basis for $L^2(\mathbf{R})$. But the reconstruction of a compressed image is better using symmetric biorthogonal wavelets w and $\tilde{w}$.

2. The Daubechies polynomials $P(z)$

The wavelet coefficients or filter coefficients $h(n)$ are associated with the transfer function $H(z) = \sum_{n=0}^{N} h(n) z^{-n}$. The transpose filter with coefficients $h(-n)$ corresponds to $H(z^{-1})$. The product of the two filters yields a symmetric $P(z)$ that is nonnegative on the unit circle:

$$P(z) = H(z)H(z^{-1}) = \left(\sum_{n=0}^{N} h(n) z^{-n} \right) \left(\sum_{n=0}^{N} h(n) z^{n} \right). \tag{4}$$

The coefficients $h(n)$ of orthogonal filters and wavelets are chosen in two steps:

(1) Select $P(z)$ subject to $P(z) + P(-z) = 1$,
(2) Factor $P(z)$ into $H(z)H(z^{-1})$.

This " spectral factorization " is commonly done by computing the zeros of $P(z)$, which is the problem we study. The zeros come in pairs Z and Z^{-1}. One member of each pair is assigned to $H(z)$—usually the one with $|Z| \leq 1$. The zeros on the unit circle have even multiplicity if and only if $P(z) \geq 0$ on the unit circle. Then this Fejér–Riesz factorization $P(z) = H(z)H(z^{-1})$ will succeed. The coefficients for biorthogonal wavelets come from other factorizations of the same polynomial. For symmetry, the roots Z and Z^{-1} go into the same factor. It is impossible to combine symmetry and orthogonality except in the special case

$$P(z) = \frac{1}{4} z^{-1} + \frac{1}{2} + \frac{1}{4} z = \left(\frac{1 + z^{-1}}{2} \right) \left(\frac{1 + z}{2} \right). \tag{5}$$

This has $P(z) + P(-z) = 1$. The coefficients $\frac{1}{2}, \frac{1}{2}$ in $H(z)$ lead to the Haar wavelet, which has the lowest possible accuracy $p = 1$.

The accuracy p is determined by the number of zeros at $z = -1$. Thus Daubechies considered polynomials of the particular form

$$P(z) = \left(\frac{1+z^{-1}}{2}\right)^p \left(\frac{1+z}{2}\right)^p Q_p(z). \tag{6}$$

She chose the unique $Q_p(z) = cz^{p-1} + \cdots + cz^{-p+1}$ that achieves, with the lowest degree, the condition that gives perfect reconstruction:

$$P(z) + P(-z) = 1. \tag{7}$$

We refer to Daubechies [2] or Strang and Nguyen [7] for the proof that orthogonality of the wavelets requires this condition. The wavelets are constructed from the scaling function that solves the dilation equation

$$\phi(t) = 2\sum_{n=0}^{N} h(n)\ \phi(2t-n). \tag{8}$$

The main point for this paper is the connection of $Q_p(z)$ to $\mathbf{B}_p(y)$.

Theorem 1 (cf. Daubechies [2, page 168]). *The change of variables $z + z^{-1} = 2 - 4y$ yields $Q_p(z) = \mathbf{B}_p(y)$. These are the minimum degree polynomials that produce $P(z) + P(-z) = 1$ or equivalently*

$$(1-y)^p\, \mathbf{B}_p(y) + y^p\, \mathbf{B}_p(1-y) = 1. \tag{9}$$

Proof. First we connect y to z. The factor $[(1+z^{-1})/2]\,[(1+z)/2]$ in $P(z)$ is exactly $1-y$. The factor $[(1-z^{-1})/2]\,[(1-z)/2]$ is y. On the unit circle $z = e^{i\omega}$, the symmetric $Q_p(z)$ reduces to a polynomial in $\cos\omega$, and therefore to some polynomial $\mathbf{B}(y)$ in $y = (1-\cos\omega)/2$. Then $-z$ corresponds to $e^{i(\omega+\pi)}$, thus to $(1-\cos(\omega+\pi))/2 = 1-y$.

With $P(z)$ as in (6), the orthogonality condition (7) is now reduced to

$$(1-y)^p\, \mathbf{B}(y) + y^p\, \mathbf{B}(1-y) = 1. \tag{10}$$

It remains to show that the polynomial $\mathbf{B}(y)$ is the truncated binomial $\mathbf{B}_p(y)$. At $y = 0$ and $y = 1$, equation (10) holds because $\mathbf{B}_p(0) = 1$. The first term has a p-fold zero at $y = 1$ and it is flat at $y = 0$ (with $p-1$ zero derivatives)

$$(1-y)^p\, \mathbf{B}_p(y) = (1-y)^p\, [(1-y)^{-p} + O(y^p)] = 1 + O(y^p). \tag{11}$$

The second term in (10) is the mirror image across $y = 1/2$ of the first, replacing y by $1-y$. The sum has the correct value 1 with $p-1$ zero derivatives at each end. This uniquely determines a polynomial of degree $2p-1$. Therefore (10) is satisfied by $\mathbf{B}_p(y)$. Note that $(1-y)^p\, \mathbf{B}_p(y)$ is the Hermite interpolating polynomial that has maximum flatness at $y = 0$ and $y = 1$ (where it equals 1 and 0). It is the response of a "maxflat lowpass halfband filter".

We prefer to work with $\mathbf{B}_p(y)$ instead of $Q_p(z)$ for two reasons. $\mathbf{B}_p(y)$ is an ordinary polynomial of degree $p-1$, with convenient coefficients, while $Q_p(z)$ is a Laurent polynomial of the same degree. Each zero of $\mathbf{B}_p(y)$ gives two zeros of $Q_p(z)$ from the rule $Z + Z^{-1} = 2 - 4Y$. From that pair, we choose the zero Z_n inside the unit circle to go into the Daubechies polynomial

$$H_p(z) = \left(\frac{1+z^{-1}}{2}\right)^p \prod_{n=1}^{p-1} \frac{1-z^{-1}Z_n}{1-Z_n}. \tag{12}$$

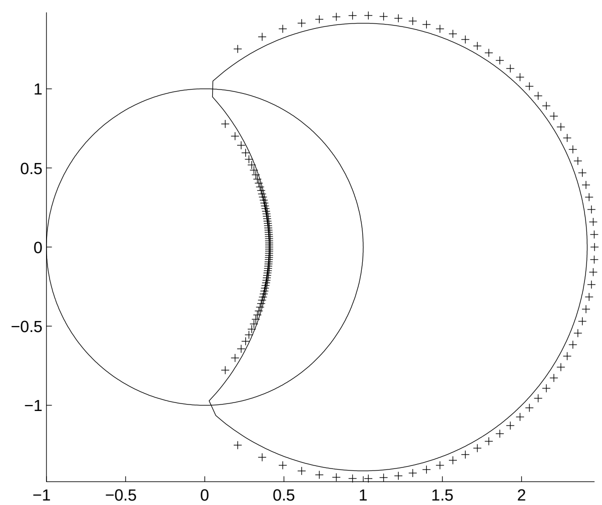

FIGURE 3. The 138 zeros of $Q_{70}(z)$ are close to the limiting curve.

The p zeros at $z = -1$ give high accuracy. For the wavelets, they give p vanishing moments (cf. Daubechies [2], or Strang and Nguyen [7]). If the factor with the Z's is omitted, the dilation equation produces spline functions —with accuracy p but not orthogonal to their translates. It is these extra zeros $Z_1, \dots, Z_{p-1}$ of $Q_p(z)$, coming from the zeros $Y_1, \dots, Y_{p-1}$ of $\mathbf{B}_p(y)$, that achieve condition (7) and yield orthogonal wavelets.

The next section will show that the zeros approach the curve $|4y(1-y)| = 1$ in the complex y-plane. In the z-plane, this curve becomes $|z - z^{-1}| = 2$, and this figure looks like a moon (It consists of two circular arcs, $|z-1| = \sqrt{2}$ and $|z+1| = \sqrt{2}$, meeting at $z = \pm i$). By Theorem 2 below, the zeros Z_n lie in the right halfplane $\mathrm{Re}(z) > 0$. Figure 3 shows the 138 zeros of $Q_{70}(z)$; each pair Z and Z^{-1} corresponds to one point Y in Figure 1. The zeros are outside the limiting curve, by Theorem 3. They approach the curve most slowly near $z = \pm i$ (which corresponds to $y = 1/2$). The limiting curve retains the special property of each $Q_p(z)$, that the zeros come in pairs Z and Z^{-1}.

□

3. THE POSITION OF THE ZEROS OF $\mathbf{B}_p(y)$

The first step is to prove that $|Y| < 1/2$ (Figure 4) and that $|4Y(1-Y)| > 2^{1/p}$ (Figure 5). The former is easy, and the latter begins with Szegö's key idea — to represent the remainder between $(1-y)^{-p}$ and $\mathbf{B}_p(y)$ by Taylor's integral formula.

Theorem 2. *For $p = 2$, the only zero is $Y = -1/2$. For $p > 2$ all the zeros satisfy $|Y| < 1/2$. Therefore each Z has $Re(Z) > 0$.*

Before proving it, we need a theorem due to Eneström and Kakeya (cf. Marden [5]):

EK Theorem. *Let $p(y)$ be a polynomial of degree n with all coefficients a_i real and positive. Define $r_i = a_i/a_{i+1}$, $0 \le i \le n-1$. Then all zeros of $p(y)$ must lie in the closed annulus: $\min_i r_i \le |y| \le \max_i r_i$.*

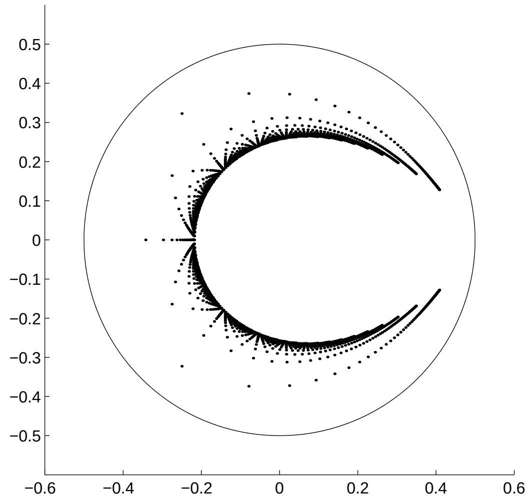

FIGURE 4. All zeros lie inside the circle $|y| = 1/2$, $p = 1 : 1 : 60$.

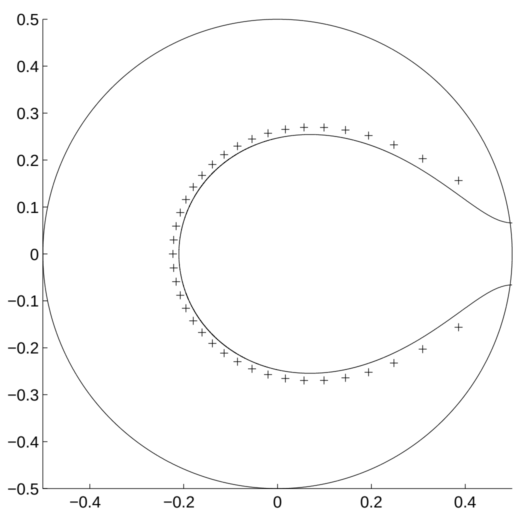

FIGURE 5. All zeros lie outside the curve $|4y(1-y)| = 2^{1/p}$, p=40.

The details about when and how the zeros can really lie on the border of the prescribed annulus is discussed by Anderson, Saff, and Varga [1]. Their sharpened form gives the strict inequality $|Y| < 1/2$ for $p > 2$.

Proof of Theorem 2. By (1), $\mathbf{B}_p(y)$ satisfies the condition of the EK Theorem. And in this case, $r_i = (i+1)/(p+i)$ for $0 \le i \le p-2$. Thus $\min_i r_i = r_0 = 1/p$, and $\max_i r_i = r_{p-2} = 1/2$. Then the truth of the statement on Y follows immediately from the EK Theorem. Therefore $Z + Z^{-1} = 2 - 4Y$ lies in the right halfplane, which implies that $\mathrm{Re}(Z) > 0$. $\square$

Theorem 3. *The zeros of* $\mathbf{B}_p(y)$ *satisfy* $|4Y(1-Y)| > 2^{1/p}$.

Proof. $\mathbf{B}_p(y)$ is the truncated Taylor series for the function $(1-y)^{-p}$. The pth derivative of this function is $p(p+1)\dots(2p-1)(1-y)^{-2p}$. Then Taylor's integral formula for the remainder $\mathbf{R}_p(y) = (1-y)^{-p} - \mathbf{B}_p(y)$ is

$$\mathbf{R}_p(y) = (2p-1)\binom{2(p-1)}{p-1}\int_0^y (y-s)^{p-1}\,(1-s)^{-2p}\,ds \tag{13}$$

$$= (2p-1)\binom{2(p-1)}{p-1}\cdot y^p \cdot \int_0^1 (1-t)^{p-1}\,(1-yt)^{-2p}\,dt. \tag{14}$$

Call this last integral $\mathbf{I}_p(y)$. Since each zero has $|Y| < 1/2$, we have $|1-Yt|^{-1} < (1-t/2)^{-1}$, for any $t \in (0\,,\,1]$. Then

$$|\mathbf{I}_p(Y)| < \int_0^1 (1-t)^{p-1}\,(1-t/2)^{-2p}\,dt = \mathbf{I}_p(\frac{1}{2}). \tag{15}$$

At $y = 1/2$, equation (9) gives $\mathbf{B}_p(1/2) = 2^{p-1}$. Thus the remainder is

$$\mathbf{R}_p(\frac{1}{2}) = (1-\frac{1}{2})^{-p} - 2^{p-1} = 2^{p-1}. \tag{16}$$

At each zero of $\mathbf{B}_p$, we know that $\mathbf{R}_p(Y) = (1-Y)^{-p}$. Now (14)–(16) combine into

$$|4Y(1-Y)|^{-p} = |4^{-p}\,Y^{-p}\,\mathbf{R}_p(Y)| < |4^{-p}\,(\frac{1}{2})^{-p}\,\mathbf{R}_p(\frac{1}{2})| = \frac{1}{2}.$$

This is the bound $|4Y(1-Y)| > 2^{1/p}$ that puts Y outside the limiting curve, and completes Theorem 3. □

Now we describe more precisely the location of the zeros of $\mathbf{B}_p(y)$. As in Szegö's problem for the exponential series (see the new methods and additional results in Varga [9]), there are two regions to consider: near $y = 1/2$ and away from $y = 1/2$. Suppose D is a circle around $y = 1/2$, with fixed small radius δ. Theorem 6 studies the zeros inside D, and Theorems 4 and 5 study the zeros outside. Together they prove that the zeros approach the limiting curve $|4y(1-y)| = 1$.

Lemma. *At any point with $|y| < 1/2$ and $|y - 1/2| > \delta$,*

$$\mathbf{I}_p(y) = \frac{1}{p(1-2y)} + O(p^{-2}). \tag{17}$$

Proof. In the integral $\mathbf{I}_p$, change variables from t to $w = (1-t)/(1-yt)^2$. Then w goes from 1 to 0 and the derivative is $dw/dt = (2y - yt - 1)/(1-yt)^3$. We leave part of the integral in terms of t

$$\mathbf{I}_p(y) = -\int_0^1 w^{p-1}\cdot\frac{1-yt}{2y-1-yt}\cdot dw. \tag{18}$$

As $p \to \infty$ the power w^{p-1} is concentrated near $w = 1$. Around that endpoint the leading term of the fraction in the integral is $(2y-1)^{-1}$. The integration of w^{p-1} gives (17) and proves the lemma. □

Suppose that $\mathbf{B}_p(Y) = 0$ and thus $\mathbf{R}_p(Y) = (1-Y)^{-p}$. By (14) and the lemma,

$$
\begin{aligned}
[4Y(1-Y)]^{-p} &= 4^{-p}(2p-1)\binom{2(p-1)}{p-1}\mathbf{I}_p(Y) \\
&= 4^{-p}\binom{2p-2}{p-1}\frac{2}{1-2Y}(1+O(p^{-1})) \qquad \text{from (17)} \\
&= \frac{1}{(1-2Y)\sqrt{4\pi p}}(1+O(p^{-1})) \qquad \text{from (2).} \qquad (19)
\end{aligned}
$$

The pth root displays the equation of the approximate curve C_p and the error term

$$|4Y(1-Y)| = |1-2Y|^{\frac{1}{p}}\,(4\pi p)^{\frac{1}{2p}}\,(1+O(p^{-2})). \tag{20}$$

Theorem 4. *All zeros outside the circle $|y-1/2| = \delta$ are not farther than $c(\delta)p^{-2}$ from the curve C_p:*

$$|4y(1-y)| = |1-2y|^{\frac{1}{p}} \cdot (4\pi p)^{\frac{1}{2p}}. \tag{21}$$

Proof. Let y be the point on C_p nearest to Y and $\epsilon = Y - y$. We must show that ϵ is $O(p^{-2})$. Since $|1+\epsilon|^{1/p} = 1 + O(|\epsilon|/p)$ for complex ϵ, one has

$$
\begin{aligned}
|1-2Y|^{\frac{1}{p}} &= |1-2y|^{\frac{1}{p}} \cdot \left|1 + \frac{\epsilon}{1-2y}\right|^{\frac{1}{p}} \\
&= |1-2y|^{\frac{1}{p}} \cdot (1 + O(\frac{|\epsilon|}{p})), \\
|4Y(1-Y)| &= |4y(1-y)| \cdot \left|1 + \frac{1-2y}{y(1-y)} \cdot \epsilon + O(\epsilon^2)\right| \\
&= |4y(1-y)| \cdot |1 + E\epsilon + O(\epsilon^2)|
\end{aligned}
$$

where $E = (1-2y)/(y(1-y))$. Since y is on the curve C_p, division yields

$$
\begin{aligned}
\frac{|4Y(1-Y)|}{|1-2Y|^{\frac{1}{p}}(4\pi p)^{\frac{1}{2p}}} &= \frac{|1+E\epsilon+O(\epsilon^2)|}{1+O(\frac{|\epsilon|}{p})} \\
&= |1 + E\epsilon + o(|\epsilon|)| \\
&= 1 + O(p^{-2}) \qquad \text{using (20).}
\end{aligned}
$$

Since δ is fixed, $E = O(1)$. Therefore $\epsilon = O(p^{-2})$. $\square$

Corollary. *All zeros outside the circle $|y-1/2| = \delta$ are not farther than $c'(\delta)p^{-1}$ from the curve D_p drawn in Figure 6:*

$$|4y(1-y)| = 1 + \epsilon_p, \quad \textit{where } \epsilon_p = \frac{\log(4\pi p)}{2p}.$$

A further argument directly based on (19) provides a more detailed information about these regular zeros, which is given in our next theorem:

Theorem 5. *Let $u = 4y(1-y)$, and $r_p = 1 + \epsilon_p$ as defined in the corollary above. Then for any fixed small positive number α,*

$$U_k = r_p \exp(2\pi i \frac{k}{p}), \qquad p\alpha \le k \le p(1-\alpha), \quad k \in \mathbb{N},$$

$$Y_k = \frac{1+\sqrt{1-U_k}}{2}, \qquad \textit{(take the negative real part branch of } \sqrt{}\,\textit{)}$$

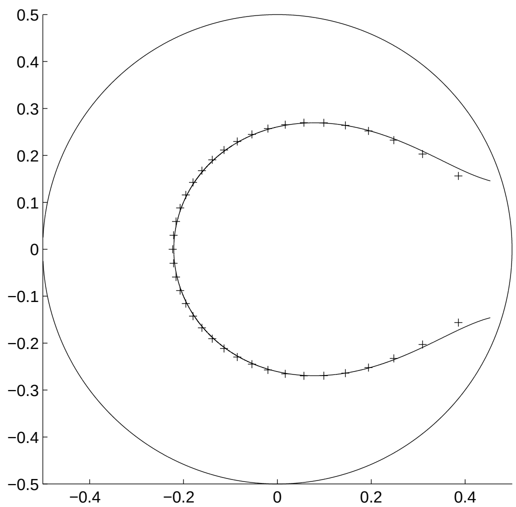

FIGURE 6. D_p is a first order approximation curve for 'regular' zeros, p=40.

gives a first order approximation (i.e. with error of order $O(p^{-1})$) to the regular zeros lying outside the circle $|y-1/2|=\delta(\alpha)$, where, $\delta(\alpha)\to 0$ as $\alpha\to 0$.

Proof. We leave the details of the proof to readers. The idea is to write the exact zeros on u–plane in a form of $r\exp(i\theta)$, and then use (19) and asymptotic analysis to find r and θ. Please note that the theorem says that on u–plane, the regular zeros are asymptotically equidistributed. □

Remark. Using the result of the coming theorem on singular zeros and (19), one can show that: $U_k=\exp(2\pi i\frac{k}{p})$, $k=0,1,...,p-1$, together with the the same Y_k defined in Theorem 5, gives a global approximation to the zeros of $\mathbf{B}_p(y)$ with error of order $O(p^{-1/2})$.

The value of $y=1/2$ is in every respect a singular point for this problem. It corresponds to points $z=i$ and $z=-i$ on the unit circle. We now prove that the zeros Y approach $1/2$ at speed $p^{-1/2}$, as Moler discovered by Matlab experiment. Surprisingly, the coefficient of $p^{-1/2}$ comes from a zero W of the complementary error function

$$\mathrm{erfc}(w)=1-\mathrm{erf}(w)=\frac{2}{\sqrt{\pi}}\int_w^\infty e^{-s^2}\,ds.$$

The corollary will improve slightly a known result for the location of these zeros.

Theorem 6. *If W is a zero of erfc(w), there is a zero Y of $\mathbf{B}_p(y)$ and a zero Z of $Q_p(z)$ such that*

$$Y=\frac{1}{2}+\frac{W}{2\sqrt{p}}+O(p^{-\frac{3}{2}}), \tag{22}$$

$$Z=i-\frac{W}{\sqrt{p}}-\frac{iW^2}{2p}+O(p^{-\frac{3}{2}}). \tag{23}$$

Proof. We introduce a new expression for $P(y) = (1-y)^p \mathbf{B}_p(y)$, which is exactly $P(z)$ defined in (6) with $z + z^{-1} = 2 - 4y$. As a function of y, this is a polynomial of degree $2p-1$ whose derivative has $p-1$ zeros both at $y=0$ and $y=1$ (see (11)). Therefore the derivative is a multiple of $y^{p-1}(1-y)^{p-1}$, and we have an *incomplete beta function*

$$P(y) = (1-y)^p \, \mathbf{B}_p(y) = 1 - c_p^{-1} \, 2^{2p-1} \int_0^y t^{p-1} \, (1-t)^{p-1} dt. \tag{24}$$

The number c_p is determined by setting $y = 1$:

$$c_p = 2^{2p-1} \int_0^1 t^{p-1} \, (1-t)^{p-1} \, dt = 2^{2p-1} \frac{\Gamma(p)^2}{\Gamma(2p)} = 2^{2p-1} \left((2p-1) \binom{2p-2}{p-1} \right)^{-1}.$$

By Stirling's formula or using the result of (2), we have

$$c_p = \sqrt{\frac{\pi}{p}} \, (1 + O(p^{-1})). \tag{25}$$

By symmetry, the value of the integral in (24) at $y = 1/2$ should be $2^{1-2p} c_p/2$. Therefore $P(1/2) = 1/2$. In order to see the detail of the zeros of $\mathbf{B}_p(y)$ near $y = 1/2$, we introduce a new variable by $y - 1/2 = w/(2\sqrt{p})$. Then

$$\begin{aligned}
P(y) = P(\frac{1}{2} + \frac{w}{2\sqrt{p}}) &= P(\frac{1}{2}) - c_p^{-1} \, 2^{2p-1} \int_0^{\frac{w}{2\sqrt{p}}} (\frac{1}{2} + t)^{p-1} (\frac{1}{2} - t)^{p-1} \, dt \\
&= \frac{1}{2} - 2 \, c_p^{-1} \int_0^{\frac{w}{2\sqrt{p}}} (1 - 4t^2)^{p-1} \, dt \\
&= \frac{1}{2} - \frac{2\sqrt{p}}{\sqrt{\pi}} \int_0^{\frac{w}{2\sqrt{p}}} e^{-4pt^2} \, dt \, (1 + O(p^{-1})) \qquad (26) \\
&= \frac{1}{2} - \frac{1}{\sqrt{\pi}} \int_0^{w} e^{-s^2} ds \, (1 + O(p^{-1})) \qquad (27) \\
&= \frac{1}{2} \text{erfc}(w) + O(p^{-1}).
\end{aligned}$$

The third step (26) used (25) and $e^{-4t^2} = 1 - 4t^2 + O(t^4)$, and in (27) $s = 2\sqrt{p}\, t$.

Let W be a zero of erfc(w). All zeros are simple, because the derivative e^{-w^2} is never zero. The fundamental theorem of complex analysis says that as $p \to \infty$, $P(1/2 + w/(2\sqrt{p}))$ is zero at some point $w = W + O(p^{-1})$. In terms of y, $Y = 1/2 + W/(2\sqrt{p}) + O(p^{-3/2})$, which is (22), because $\mathbf{B}_p(y)$ shares every zero with $P(y)$ except $y = 1$. □

Corollary. *Every zero of erfc(w) has* $|\arg \; W| < 3\pi/4$.

Proof. The corresponding Y lies outside the limiting curve $|4y(1-y)| = 1$, which intersects itself at $y = 1/2$ with slopes ± 1. In the limit, $W = (Y - 1/2)/\sqrt{p} + O(p^{-1})$ must have $|\arg \; W| \le 3\pi/4$. If equality held, W^2 would be purely imaginary. Then Theorem 6 would give

$$|4Y(1-Y)| = |1 - W^2 p^{-1} + O(p^{-2})| = 1 + O(p^{-2}).$$

This contradicts the inequality $|4Y(1-Y)| > 2^{1/p}$ in Theorem 2, proving the corollary. □

Fettis, Caslin, and Cramer [3] computed the zeros of $\mathrm{erfc}(w)$ to very high accuracy. They also proved an asymptotic form of the statement $|\arg\ W| \leq 3\pi/4$. It is interesting to see the complete statement (which their numerical table confirms) proved by such an indirect argument involving the zeros of $\mathbf{B}_p(y)$.

These zeros approach $1/2$ at order $p^{-1/2}$, close to the line $Y-1/2 = W/2\sqrt{p}$. By the corollary, the slope of this line is not ± 1. Therefore the distance from Y_p to the limiting curve C is of strict order $p^{-1/2}$ near $y = 1/2$. In this region, the error order in equation (20) rises to p^{-1}. This applies in particular to the rightmost zero, which comes from the first W tabulated in [3], $Y \approx 1/2 + (-1.3548\ldots + i1.9914\ldots)/2\sqrt{p}$.

4. Steepness at $\omega = \frac{\pi}{2}$

A change of variables $t = (1-\cos\theta)/2$ in (24) produces the integral of $\sin^{2p-1}\theta$. The limits of integration are related by $y = (1-\cos\theta)/2$, which is exactly the change associated with $z = e^{i\omega}$ in the proof of Theorem 1. Thus (24) is Meyer's form (cf. Meyer [6, page 43]) of the halfband filter $P(z)$ in equation (6)

$$P(e^{i\omega}) = 1 - c_p^{-1} \int_0^{\omega} \sin^{2p-1}\theta \, d\theta. \tag{28}$$

The zero at $y = 1$ becomes the celebrated "zero at π" for the frequency response $P(e^{i\omega})$. This zero at $\omega = \pi$ is of order $2p$, from the power of $\sin\theta$ in (28) and the form of $P(z)$ in (6). Factorization gives pth order zeros for the Daubechies polynomials in $P(z) = H(z)H(z^{-1})$. That zero at $\omega = \pi$ and $z = -1$ is responsible for the p vanishing moments in the wavelets.

The trigonometric polynomial $P(e^{i\omega})$ drops monotonically from one to zero on $0 \leq \omega \leq \pi$ (see (28)). The first $2p-1$ derivatives are zero at $\omega = 0$, and $\omega = \pi$, from the vanishing of $\sin^{2p-1}\theta$. Furthermore this integral of $(1-\cos\theta)^{p-1}\sin\theta$ involves only odd powers of $\cos\theta$, and the only even power is the constant term. $P(e^{i\omega})$ is odd around its value $1/2$ at $\omega = \frac{\pi}{2}$, and it is called "halfband".

An important question for such a filter is the slope at $\omega = \frac{\pi}{2}$. This slope determines the width of the frequency band, in which P drops from 1 to 0. An ideal filter has a jump; its graph is a brick wall (however, this ideal is not a polynomial). An optimally designed polynomial of order N has slope nearly $O(N^{-1})$. There will be ripples in the graph of $P(e^{i\omega})$—a monotonic polynomial cannot provide such a sharp cutoff. The Daubechies filters are necessarily less sharp: $O(N)$ becomes $O(\sqrt{N})$.

Theorem 7. *The slope of $P(e^{i\omega})$ in (28) is approximately $\sqrt{p/\pi}$ at $\omega = \pi/2$. The transition from nearly 1 to nearly 0 is over an interval (i.e. transition band) of width $2\sqrt{2/p}$.*

Proof. The integral in (28) has derivative $\sin^{2p-1}(\pi/2) = 1$ at $\omega = \pi/2$. The slope of $P(e^{i\omega})$ is exactly the constant $-c_p^{-1}$. By (25) this is $-\sqrt{p/\pi} + O(p^{-\frac{3}{2}})$. To measure the drop in $P(e^{i\omega})$ around $\omega = \pi/2$, we integrate from $\pi/2 - \sigma/\sqrt{p}$ to $\pi/2 + \sigma/\sqrt{p}$. Shifting by $\pi/2$ to center the integral, and scaling by $\theta = \tau/\sqrt{p}$, the drop is

$$c_p^{-1}\int_{-\sigma/\sqrt{p}}^{\sigma/\sqrt{p}} \sin^{2p-1}\theta d\theta \approx \frac{1}{c_p\sqrt{p}}\int_{-\sigma}^{\sigma}(1-\frac{\tau^2}{2p})^{2p-1}d\tau$$

$$\approx \frac{1}{\sqrt{\pi}}\int_{-\sigma}^{\sigma} e^{-\tau^2}\,d\tau. \tag{29}$$

Thus 95% of the drop comes for $\sigma = \sqrt{2}$ (within two standard deviations from the mean, for the normal distribution). This transition interval has width $\Delta\omega = 2\sqrt{2/p}$, as the theorem predicts. That rule was found experimentally by Kaiser and Reed at the beginning of the triumph of digital filters. □

5. The existence of the limiting multiresolution—independence on the spectral factorization procedure

For large p, there are many different ways to make spectral factorization (SF) from $P(z)$ (see section 2). Consequently, there are many orthogonal scaling function–wavelet pairs generated from one $P(z)$. Two choices often made are the 'min–phase' factorization mentioned in section 2 and the 'least asymmetric' factorization discussed in Daubechies' book [2]. In this section, we will prove that different SFs do give different multiresolutions, and on the other hand, asymptotically (with respect to p), the multiresolutions generated from different SFs tend to be the same. We introduce a 'distance' for two arbitrary subspaces in a separable Hilbert space.

Definition. Let H be a separable Hilbert space, and $H^{(1)}$, $H^{(2)}$ its two closed subspaces. Define:

$$d(H^{(1)}, H^{(2)}) = \inf_{(E^{(1)},E^{(2)})} d(E^{(1)}, E^{(2)}), \tag{30}$$

where $E^{(i)}$ stands for any ordered orthonormal basis of $H^{(i)}$, or $E^{(i)} = (e_1^{(i)}, e_2^{(i)}, ...)$, and $d(E^{(1)}, E^{(2)})$ is defined by:

$$d(E^{(1)}, E^{(2)}) = \begin{cases} \sqrt{2}, & \text{if } E^{(1)\#} \neq E^{(2)\#}, \\ \max_n |e_n^{(1)} - e_n^{(2)}|, & \text{otherwise.} \end{cases}$$

Now let's state two useful lemmas slightly different from those at the beginning of Chapter 8 of Daubechies' book [2].

Lemma 1. *The functions $\phi_1(t-k)$ and $\phi_2(t-k)$ are orthonormal bases for the same subspace of $L^2(\mathbb{R})$ if and only if there exists a 2π–periodic function $\alpha(\omega)$ in $L^2[0, 2\pi]$ such that:*

(1) $\hat{\phi}_2 = \alpha\hat{\phi}_1$;

(2) $|\alpha| = 1$, *a.e.*

Lemma 2. *Suppose the conditions of Lemma 1 are satisfied, and both ϕ_1 and ϕ_2 are compactly supported. With the convention $\int \phi_i = 1$, $i = 1, 2$, ϕ_2 must be a shifted (by an integer) copy of ϕ_1.*

Now we can state our next two theorems:

Theorem 8. *For a given order p, different SFs from $P(z)$ give different multiresolutions.*

Theorem 9. *For any fixed order p, let $P(z) = H_i(z)H_i(z^{-1})$, $i = 1, 2$, be any two different SFs, and ϕ_i the corresponding scaling functions, and $V_0(\phi_i)$ the spaces spanned by their translates $\phi_i(t-k)$. Then:*

$$\lim_{p\to\infty} \max_{(H_1,H_2)} d(V_0(\phi_1), V_0(\phi_2)) = 0. \tag{31}$$

The proof of Theorem 8 can be done quickly. If two SFs of $P(z)$ produce a same multiresolution, then by Lemma 2, the associated scaling functions must be the shifted copies of each other. Therefore, there exists an integer k, s.t. $H_2(z) = z^{-k}H_1(z)$. k must be zero, since we always assume that $H_i|_{z^{-1}=0} \neq 0$ and H_i are FIR filters. This means that H_1 and H_2 are identical. Now we turn to the proof of Theorem 9. (For simplicity, we use $f(\omega)$ to denote the function $f(e^{-i\omega})$.)

Proof of Theorem 9. (1) Since $\overset{\wedge}{\phi}_i = \frac{1}{\sqrt{2\pi}} \prod_{n\geq 1} H_i(\frac{\omega}{2^n})$, we have:

$$|\overset{\wedge}{\phi}_i|^2 = \frac{1}{2\pi}\prod_{n\geq 1}|H_i(\frac{\omega}{2^n})|^2 = \frac{1}{2\pi}\prod_{n\geq 1}P(\frac{\omega}{2^n}). \tag{32}$$

which implies that the local spectral energy of the scaling functions depends only on the order p.

(2) Define

$$\alpha(\omega) = \begin{cases} \dfrac{\overset{\wedge}{\phi}_2(\omega)}{\overset{\wedge}{\phi}_1(\omega)}, & \omega \in [-\pi, \pi], \\ \text{periodic extension}, & \text{otherwise}. \end{cases}$$

Then α is well-defined and in $L^2[-\pi, \pi]$ since for large p, the local spectral energy (i.e. $|\overset{\wedge}{\phi}_i|^2$) is always positive on $[-\pi, \pi]$. By the argument in (1), $|\alpha| = 1$, a.e.

(3) Now define

$$\phi_{(1)} = (\alpha(\omega)\overset{\wedge}{\phi}_1(\omega))^{\vee}.$$

Then: 1. $\overset{\wedge}{\phi}_{(1)} = \overset{\wedge}{\phi}_2$, for $\omega \in [-\pi, \pi]$; 2. By Lemma 1, $V_0(\phi_{(1)}) = V_0(\phi_1)$.

(4) $\phi_{(1)}$ is 'close' to ϕ_2:

$$\begin{aligned} ||\phi_{(1)} - \phi_2|| &= ||\overset{\wedge}{\phi}_{(1)} - \overset{\wedge}{\phi}_2|| \\ &= ||\overset{\wedge}{\phi}_{(1)} - \overset{\wedge}{\phi}_2||_{L^2(R\backslash[-\pi,\pi])} \\ &\leq ||\overset{\wedge}{\phi}_{(1)}||_{L^2(R\backslash[-\pi,\pi])} + ||\overset{\wedge}{\phi}_2||_{L^2(R\backslash[-\pi,\pi])} \\ &= ||\overset{\wedge}{\phi}_1||_{L^2(R\backslash[-\pi,\pi])} + ||\overset{\wedge}{\phi}_2||_{L^2(R\backslash[-\pi,\pi])} \\ &= 2(1 - ||\overset{\wedge}{\phi}_i||_{L^2[-\pi,\pi]}) = \epsilon_p. \end{aligned} \tag{33}$$

As $p \to \infty$, $P(\omega)$ converges to the indicator of $[-\frac{\pi}{2}, \frac{\pi}{2}]$. Therefore, by (32), $|\overset{\wedge}{\phi}_i|$ converges to the box function on $[-\pi, \pi]$ with height $\frac{1}{\sqrt{2\pi}}$. Thus $||\overset{\wedge}{\phi}_i||_{L^2[-\pi,\pi]} \to 1$, which implies that in (33), $\epsilon_p \to \infty$ as $p \to \infty$.

(5) Now $\{\phi_{(1)}(x-n) | \quad n \in \mathbb{Z}\}$ and $\{\phi_2(x-n) | \quad n \in \mathbb{Z}\}$ are the orthonormal bases of $V_0(\phi_1)$ and $V_0(\phi_2)$ respectively, and

$$||\phi_{(1)}(x-n) - \phi_2(x-n)|| = ||\phi_{(1)}(x) - \phi_2(x)|| = \epsilon_p. \tag{34}$$

By definition,

$$d(V_0(\phi_1), V_0(\phi_2)) \le \epsilon_p$$

which finishes the proof. □

Suppose that in (32), $P(\omega)$ is exactly the indicator of $[-\frac{\pi}{2}, \frac{\pi}{2}]$. Then the corresponding $|\overset{\wedge}{\phi_i}|$ will be exactly a box function on $[-\pi, \pi]$ with height $1/\sqrt{2\pi}$. The proof of Theorem 9 shows that the multiresolution is somehow independent of the phase. Thus one can conjecture that the limiting multiresolution $V = V(\phi)$ exists and $\phi = (\frac{1}{\sqrt{2\pi}}\mathrm{Ind}_{[-\pi,\pi]})^{\vee}$. That is, $\phi = \mathrm{sinc}(x) = \frac{\sin \pi x}{\pi x}$. Actually a similar argument to that of Theorem 9 does lead us to the following discovery:

Theorem 10. *For any fixed order p, let $P(z) = H^{(p)}(z)H^{(p)}(z^{-1})$ be an arbitrary SF, and $\phi^{(p)}$ the consequent mother scaling function. Then:*

$$\lim_{p\to\infty} \max_{H^{(p)}} d(V_0(\phi^{(p)}), V_0(sinc)) = 0. \tag{35}$$

Now we can stand at the limiting end $\phi = \mathrm{sinc}(x)$ to watch the long multiresolution series generated from finite p. We list some of its properties:

(1) $V_0(\mathrm{sinc})$ is exactly the function space of all $L^2(\mathrm{R})$ functions with spectral band limited from $-\pi$ to π. Thus any function in this space can be analytically extended to be an entire function on the complex plane.
(2) For any $f \in V_0(\mathrm{sinc})$, its component along $\mathrm{sinc}(x-n)$ is simply given by its value at $x = n$ (Shannon Sampling Theorem). The importance of this fact is that it reduces the task of computing an inner product to a simple evaluation.
(3) The wavelet analysis coming from the limiting multiresolution $V(\mathrm{sinc})$ turns out to be very simple. The projection operator associated with the wavelets space W_j at level j, is nothing but the spectral truncation operator associated wih the union of the spectral intervals $\pm[2^{-j}, 2^{-j+1}]\pi$. That is, spectral truncation is the simplest wavelet analysis.

Added in proof

We have learned that before our work in 1995, Kateb and Lemarié found the asymptotic behavior of the zeros. Their results are summarized in Comptes Rendus (vol. 320, 1995, pp. 5–8) and in Applied and Computational Harmonic Analysis **4** (vol. 2, 1995, pp. 398–399). The complete results in the 1994 Orsay report "The phase of the Daubechies filters" will be published in Revista Matematica. This important work goes further toward the goal of understanding the asymptotics of the Daubechies wavelets.

It might be useful to identify the four steps to be analyzed:

(1) The $2p-1$ zeros of $H_p(z) = \sum_k h_p[k]z^{-k}$.
(2) The phase of $H_p(z)$ on the unit circle $z = e^{i\omega}$.
(3) The scaling function $\phi_p(t)$ with Fourier transform $\prod_{k=1}^{\infty} H_p(\omega/2^k)$.
(4) The wavelet $w_p(t) = \sum_k (-1)^k h_p[2p-1-k]\phi_p(2t-k)$.

Our present paper takes step 1, while Kateb and Lemarié took both steps 1 and 2. They found the leading term $pg(\omega)$ in the phase of $H_p(\omega)$. All difficulty is with the phase; the amplitude $|H_p(\omega)|$ approaches an ideal filter. Building on the first two steps, we recently found the asymptotic forms of the scaling function and the wavelet. The former involves a dilated and shifted Airy function $\phi_p(t) \approx a_p \mathrm{Ai}(a_p(t-t_p))$ up to near t_p, where a_p and t_p depend only on p. This is matched to regions of damped oscillation. The main tool in our preprint "Asymptotic structures of Daubechies scaling functions and wavelets" is the method of stationary phase.

References

1. N. Anderson, E.B. Saff and R.S. Varga, *On the Eneström–Kakeya Theorem and its sharpness*, Linear Algebra Appl. **28** (1979), 5–16. MR **81i:**26011
2. I. Daubechies, *Ten Lectures on Wavelets*, SIAM, Philadelphia, 1992. MR **92e:**42025
3. H.E. Fettis, J.C. Caslin and K.R. Cramer, *Complex zeros of the error function and of the complementary error function*, Math. Comp. **27** (1973), 401–407. MR **48:**5333
4. M. Lang and B.C. Frenzel, *Polynomial Root Finding*, preprint, Rice University, 1994.
5. M. Marden, *Geometry of Polynomials*, Mathematical Surveys No. 3, AMS, Providence, 1966, p. 137. MR **37:**1562
6. Y. Meyer, *Wavelets: Algorithms and Applications*, SIAM, Philadelphia, 1993. MR **95f:**94005
7. G. Strang and T. Nguyen, *Wavelets and Filter Banks*, Wellesley–Cambridge Press, Wellesley, 1996.
8. G. Szegö, *Über eine eigenschaft der exponentialreihe*, Sitzungsber. Berlin Math. Ges. **23** (1924), 50–64.
9. R.S. Varga, *Scientific Computation on Mathematical Problems and Conjectures*, SIAM, Philadelphia, 1990. MR **92b:**65012

Department of Mathematics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

E-mail address: jhshen@math.mit.edu

E-mail address: gs@math.mit.edu