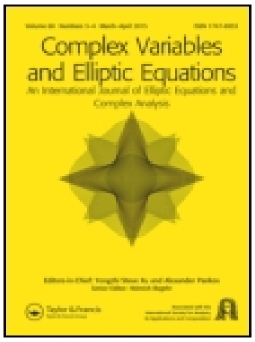

Complex Variables and Elliptic Equations

An International Journal

ISSN: 1747-6933 (Print) 1747-6941 (Online) Journal homepage: http://www.tandfonline.com/loi/gcov20

Approximation of functions by higher order Szegö kernels I. Complex variable cases

Jinxun Wang & Tao Qian

To cite this article: Jinxun Wang & Tao Qian (2015) Approximation of functions by higher order Szegö kernels I. Complex variable cases, Complex Variables and Elliptic Equations, 60:6, 733-747, DOI: 10.1080/17476933.2014.956740

To link to this article: http://dx.doi.org/10.1080/17476933.2014.956740

Published online: 28 Oct 2014.

Complex Variables and Elliptic Equations, 2015
Vol. 60, No. 6, 733–747, http://dx.doi.org/10.1080/17476933.2014.956740

Approximation of functions by higher order Szegö kernels I. Complex variable cases

Jinxun Wang[a,b*] and Tao Qian[b]

[a] *School of Mathematics and Information Science, Guangzhou University, Guangzhou 510006, China;* [b] *Faculty of Science and Technology, Department of Mathematics, University of Macau, Taipa, Macao, China*

Communicated by Y. Xu

(*Received 16 April 2014; accepted 14 August 2014*)

We study the adaptive decomposition of functions in the complex Hardy spaces $\mathcal{H}^2$ by higher order Szegö kernels. The purpose is to treat signals that are essentially of high frequencies. We show that each kernel function (basic function) we use is either a mono-component (as an analytic signal, its instantaneous frequency is positive everywhere), or a sum of two orthogonal mono-components. The proposed decomposition thus belongs to the category of adaptive mono-component decomposition.

Keywords: Hardy space; Szegö kernel; mono-component; adaptive decomposition; dictionary; matching pursuit

AMS Subject Classifications: 30C40; 30H10; 41A20; 46E20

1. Introduction

Let $\mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}$ be the unit disc in the complex plane, and let f be a holomorphic function on $\mathbb{D}$. We say $f \in \mathcal{H}^2(\mathbb{D})$ if

$$\|f\| := \sup_{0\le r<1} \left(\frac{1}{2\pi} \int_0^{2\pi} |f(re^{i\theta})|^2 d\theta \right)^{1/2} < \infty.$$

$\mathcal{H}^2(\mathbb{D})$ is a Hilbert space with the inner product being defined by

$$\langle f, g\rangle := \frac{1}{2\pi} \int_0^{2\pi} f(e^{it})\overline{g}(e^{it})dt,$$

where $f(e^{it}) := \lim_{r\to 1^-} f(re^{it})$ in the L^2-norm sense, as well as in the pointwise convergence sense for almost all $t \in [0, 2\pi)$. A function $f \in \mathcal{H}^2(\mathbb{D})$ if and only if it has the expansion

$$f(z) = \sum_{k=0}^{\infty} c_k z^k$$

*Corresponding author. Email: wjxpyh@gmail.com

in the disc, with $\{c_k\} \in l^2$, namely, $\sum_{k=0}^{\infty} |c_k|^2 < \infty$. In the case, the coefficients c_k's coincide with those of the Taylor series expansion of f, viz.,

$$c_k = \langle f, z^k \rangle = \frac{f^{(k)}(0)}{k!}.$$

A function $f \in L^2(\partial\mathbb{D})$, no matter being real- or complex-valued, is said to be a *mono-component*, if its Hardy space projection, $f + iHf$, where H is the Hilbert transformation on the circle, has a non-negative phase derivative. In other words, this means, writing $f(t) + i(Hf)(t) = \rho(t)e^{i\theta(t)}$, with $\rho(t) \geq 0$, where ρ and θ are called, respectively, "amplitude" and "phase" of $f + iHf$, we have $\theta'(t) \geq 0$ for all $t \in [0, 2\pi)$ ([1]). A function f is a mono-component can also be briefly phrased as its analytic phase derivative is non-negative. Note that a classical derivative of the phase of $f + iHf$ may not exist. For a class of functions, however, this concept can be defined through boundary limits of the same quantity of the corresponding Hardy space projection inside the disc ([2,3]). For functions defined on the real line there is a parallel mono-component function theory. We recall here the definitions of Hilbert transformations in the two contexts. For $s \in L^2(\partial\mathbb{D})$,

$$Hs(t) := \frac{1}{2\pi}\text{p.v.}\int_{-\pi}^{\pi} \cot\left(\frac{t-\mu}{2}\right) s(\mu)d\mu,$$

and for $s \in L^2(\mathbb{R})$,

$$Hs(t) := \frac{1}{\pi}\text{p.v.}\int_{-\infty}^{\infty} \frac{s(\mu)}{t-\mu}d\mu.$$

The definition of mono-component is consistent with the physical requirement that a mono-component should possess a non-negative instantaneous frequency function. A large pool of mono-components has been found, including Möbius transforms (Fourier atoms), Blaschke products of finite and infinite many zeros, and starlike and p-starlike functions ([1,3–8]). Since there exist functions in $\mathcal{H}^2(\mathbb{D})$ that are not mono-components, one naturally seeks for mono-component decompositions. Fourier series is a particular example. One way to identify intrinsic mono-components of a signal is through fast mono-component decomposition.

In [9,10] the authors proposed an adaptive mono-component decomposition method, called *adaptive Fourier decomposition* or AFD. The iterative method produces, as a matter of fact, a Takenaka–Malmquist system (or TM system or rational orthogonal system) expansion. A basic function in a TM system is of the form

$$B_n(z) = B_{\{a_1,\dots,a_n\}}(z) = \frac{\sqrt{1-|a_n|^2}}{z-a_n}\prod_{k=1}^{n}\frac{z-a_k}{1-\overline{a_k}z}, \tag{1}$$

where $a_1 = 0, a_2, \dots, a_n \in \mathbb{D}$ are complex numbers in the unit disc. Note that if all a_k's are zero, then $B_n(z) = z^{n-1}$, and the system reduces to half of the Fourier system. A basic function in a TM system ($a_1 = 0$) is a mono-component. The characteristic properties of AFD are: (i) the parameters are successively selected according to a maximal selection principle that guarantees the maximal gain in energy at every step; (ii) it is different from matching pursuit for the parameters are not selected based on the remainder but on the reduced remainder generated through an application of the so-called *generalized backward shift*, and a parameter can be selected repeatedly; and, (iii) the resulted TM system may not be a basis. These properties distinguish AFD from the traditional studies of TM systems, as well as from matching pursuit. AFD has found applications in control theory ([11,12]), as well as theoretical impacts to approximation theory and operator theory. In AFD, the maximal selection principle is performed as

$$|\langle f, B_{\{a_1,\ldots,a_n\}}\rangle| = \sup_{b\in\mathbb{D}} |\langle f, B_{\{a_1,\ldots,a_{n-1},b\}}\rangle|.$$

The obtained adaptive orthogonal decomposition is of the form

$$f = \sum_{n=1}^{\infty} \langle f, B_{\{a_1,\ldots,a_n\}}\rangle B_{\{a_1,\ldots,a_n\}}. \quad (2)$$

The result was extended to the upper half-plane. They proved that the convergence rate is $1/2$ under certain conditions.

Although AFD can effectively extract characteristic properties of a signal in relation to instantaneous frequencies, it has a drawback point: Applying AFD to a signal of high frequencies means that AFD forces to extract at the first few steps the maximal energy in low frequencies. This is not natural. In practice, there are many signals being of high frequencies by nature, it would be the best if we could faithfully extract at the first steps the components of large energy potions no matter how big or small the corresponding frequencies are.

To treat signals of a great variety of frequency levels, appealing to matching pursuits or greedy algorithms (see [13]), we propose an alternative approach for adaptive mono-component decomposition of functions in the Hardy spaces $\mathcal{H}^2$ for both the unit disc and the upper half-plane, using the higher order complex Szegö kernels (i.e. the higher order partial derivatives of Szegö kernels). Our scheme is simple and easy to implement, since all the kernel functions (basic functions) are rational, and the coefficients in the expansion admit analytic expressions in terms of the higher order partial derivatives of the residues which can be easily handled. Unlike (2), the proposed method does not result in an orthogonal decomposition, as our algorithm is essentially the pure greedy algorithm (for other types of greedy algorithm one can see [14]). It works well for most of the functions (signals), especially for oscillatory signals or those signals of high frequency, because the parameters of the kernel functions can match well with the frequency and amplitude of the original signal. As to the convergence rate, which is assumed to be optimal, waits for further study.

2. The case for the unit disc

It was pointed out in [15] that the TM system (1) is the outcome of the Gram–Schmidt orthogonalization process applying to $\{E_n\}_{n=1}^{\infty}$, where

$$E_n(z) = E_{\{a_1,\ldots,a_n\}}(z) = \begin{cases} \frac{1}{(1-\overline{a_n}z)^{m_n}} & \text{if } a_n \neq 0, \\ z^{m_n-1} & \text{if } a_n = 0, \end{cases}$$

in which m_n is the cardinality of the set $\{i : a_i = a_n, i \leq n\}$. We note that a suitable substitute for E_n is (we still adopt the same notation)

$$E_n(z) = E_{\{a_1,\ldots,a_n\}}(z) = \frac{z^{m_n-1}}{(1-\overline{a_n}z)^{m_n}},$$

since for any $k \in \mathbb{N}$ and $a \neq 0$ there holds the following clear relation

$$\operatorname{span}\left\{\frac{1}{(1-\overline{a}z)^{l+1}} : 0 \leq l \leq k\right\} = \operatorname{span}\left\{\frac{z^l}{(1-\overline{a}z)^{l+1}} : 0 \leq l \leq k\right\}.$$

This observation leads us to consider the following dictionary

$$\mathcal{D} = \left\{e_{k,a} = \frac{\varphi_{k,a}}{\|\varphi_{k,a}\|} : k \in \mathbb{N}, a = a_0 + ia_1 \in \mathbb{D}\right\},$$

where

$$\varphi_{k,a}(z) = \frac{\partial^k}{\partial a_0^k}\left(\frac{1}{1-\overline{a}z}\right) = \frac{k!z^k}{(1-\overline{a}z)^{k+1}}$$

is called the Szegö kernel function of order k for the unit disc, with the parameter $a \in \mathbb{D}$. This terminology comes from the fact that for any $f \in \mathcal{H}^2(\mathbb{D})$ we have by Cauchy's integral formula

$$\langle f, \varphi_{k,a}\rangle = f^{(k)}(a). \tag{3}$$

Obviously, $\{z^n\}_{n=0}^{\infty} = \{e_{k,0}\}_{k=0}^{\infty} \subset \mathcal{D} \subset \mathcal{H}^2(\mathbb{D})$, and any finite subset of $\mathcal{D}$ is linearly independent. Now, let us first show that

PROPOSITION 2.1

$$\|\varphi_{k,a}\|^2 = \frac{(k!)^2}{(1-|a|^2)^{2k+1}} \sum_{l=0}^{k} \binom{k}{l}^2 |a|^{2l}.$$

Proof

$$\begin{aligned}
\|\varphi_{k,a}\|^2 &= \frac{(k!)^2}{2\pi}\int_0^{2\pi} \frac{d\theta}{\left(1+|a|^2-2|a|\cos\theta\right)^{k+1}} \\
&= \frac{2(k!)^2}{\pi}\int_0^{\infty} \frac{(1+u^2)^k}{\left((1+|a|)^2u^2+(1-|a|)^2\right)^{k+1}}du \\
&= \frac{2(k!)^2}{\pi}\int_0^{\infty} \frac{\left((1+|a|)^2+(1-|a|)^2t^2\right)^k}{(1-|a|^2)^{2k+1}(1+t^2)^{k+1}}dt,
\end{aligned}$$

here we have made the substitutions of variables by setting $u = \tan\frac{\theta}{2}$ and $t = \frac{1-|a|}{1+|a|}u$. And

$$\begin{aligned}
&\int_0^{\infty} \frac{\left((1+|a|)^2+(1-|a|)^2t^2\right)^k}{(1+t^2)^{k+1}}dt \\
&= \int_0^{\infty} \frac{\left((1+t^2)|a|^2+2(1-t^2)|a|+(1+t^2)\right)^k}{(1+t^2)^{k+1}}dt \\
&= \sum_{p=0}^{k}\sum_{q=0}^{k-p}\int_0^{\infty} \frac{\binom{k}{p}\binom{k-p}{q}2^q|a|^{2p+q}(1-t^2)^q}{(1+t^2)^{q+1}}dt \\
&= \sum_{p=0}^{k}\sum_{q=0}^{[\frac{k-p}{2}]} 2^{2q}\binom{k}{p}\binom{k-p}{2q}\int_0^{\infty} \frac{(1-t^2)^{2q}|a|^{2(p+q)}}{(1+t^2)^{2q+1}}dt
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{p=0}^{k} \sum_{q=0}^{[\frac{k-p}{2}]} 2^{2q-1} \binom{k}{p} \binom{k-p}{2q} \frac{\Gamma(q+\frac{1}{2})\Gamma(\frac{1}{2})}{\Gamma(q+1)} |a|^{2(p+q)} \\
&= \sum_{l=0}^{k} \sum_{p=\max(0,2l-k)}^{l} 2^{2l-2p-1} \binom{k}{p} \binom{k-p}{2l-2p} \frac{\Gamma(l-p+\frac{1}{2})\Gamma(\frac{1}{2})}{\Gamma(l-p+1)} |a|^{2l} \\
&= \sum_{l=0}^{k} \sum_{p=2l-k}^{l} 2^{2l-2p-1} \binom{k}{p} \binom{k-p}{2l-2p} \frac{\Gamma(l-p+\frac{1}{2})\Gamma(\frac{1}{2})}{\Gamma(l-p+1)} |a|^{2l} \\
&= \frac{\pi}{2} \sum_{l=0}^{k} \sum_{p=0}^{k-l} \frac{k!}{p!(2l+p-k)!((k-l-p)!)^2} |a|^{2l} \\
&= \frac{\pi}{2} \sum_{l=0}^{k} \binom{k}{l}^2 |a|^{2l},
\end{aligned}
$$

where $[x]$ is the largest integer not greater than x, and we adopt the convention that $\binom{k}{p} = 0$ for $p < 0$. □

The above proposition leads to the following estimate, which is more precise than the usual estimate (see [16]).

COROLLARY 2.1 *If* $f \in \mathcal{H}^2(\mathbb{D})$*, then*

$$
(1-|a|^2)^{k+\frac{1}{2}} |f^{(k)}(a)| \leq \sqrt{(2k)!} \|f\|, \quad \forall a \in \mathbb{D}. \tag{4}
$$

Proof This follows from the Cauchy–Schwarz inequality

$$
|f^{(k)}(a)| = |\langle f, \varphi_{k,a} \rangle| \leq \|f\| \|\varphi_{k,a}\|,
$$

and the fact that

$$
\sum_{l=0}^{k} \binom{k}{l}^2 |a|^{2l} \leq \sum_{l=0}^{k} \binom{k}{l}^2 = \binom{2k}{k}.
$$

□

Next, we shall show that all the elements in $\mathcal{D}$ are mono-components, except for $e_{0,a}$. Before this, we prove a criterion for mono-components which will be convenient for the future study.

THEOREM 2.1 *Suppose* $f \in \mathcal{H}^2(\mathbb{D}) \bigcap C^1(\overline{\mathbb{D}})$ *is the holomorphic extension of some analytic signal, with the expression*

$$
f(z) = f(re^{it}) = \rho(r,t) e^{i\theta(r,t)}.
$$

If for every $t \in [0, 2\pi)$, $\frac{\partial\rho(r,t)}{\partial r} > 0$ *in* $\{r : 1-\delta_t < r < 1\}$ (*which implies that* $|f(z)|$ *increases as r ascends up to* 1), *then* f *is a mono-component, i.e.*

$$\theta'(t) = \frac{d\left(\lim\limits_{r\to 1^-}\theta(r,t)\right)}{dt} \geq 0, \quad \forall t \in [0, 2\pi).$$

Proof The theorem follows from the relation

$$\begin{aligned}
\theta'(t) &= \lim_{r\to 1^-}\frac{\partial\theta(r,t)}{\partial t}\\
&= \lim_{r\to 1^-}\operatorname{Re}\left(\frac{e^{it}f'(z)}{f(z)}\right)\\
&= \lim_{r\to 1^-}\operatorname{Re}\left(\frac{\partial\log f(z)}{\partial r}\right)\\
&= \lim_{r\to 1^-}\frac{\partial\log|f(z)|}{\partial r}\\
&= \lim_{r\to 1^-}\left(\frac{1}{\rho(r,t)}\frac{\partial\rho(r,t)}{\partial r}\right).
\end{aligned}$$

□

Corollary 2.2 *If* $k > 0$, *then* $e_{k,a}$ *is a mono-component.*

Proof Firstly, the boundary limit of $e_{k,a}$ is an analytic signal, since $\operatorname{Im}(e_{k,a}(0)) = 0$. Secondly, we note that $\frac{z^k}{(1-\bar{a}z)^{k+1}} = \frac{z}{(1-\bar{a}z)^2}(\frac{z}{1-\bar{a}z})^{k-1}$. Write $z = re^{it}$, then computation yields

$$\frac{\partial}{\partial r}\left(\frac{|z|}{|1-\bar{a}z|^2}\right) = \frac{1-|a|^2|z|^2}{|1-\bar{a}z|^4} > 0,$$

and

$$\frac{\partial}{\partial r}\left(\frac{|z|^2}{|1-\bar{a}z|^2}\right) = \frac{|z|(2-\bar{a}z-\bar{z}a)}{|1-\bar{a}z|^4} > 0.$$

So, $|e_{k,a}(re^{it})| = C_{k,a}\left|\frac{z}{(1-\bar{a}z)^2}\right|\left|\frac{z}{1-\bar{a}z}\right|^{k-1}$ is an increasing function of r for every $t \in [0, 2\pi)$. □

Notice that if $k = 0$, $e_{k,a}$ is not a mono-component. Fortunately, $\frac{1}{1-\bar{a}z} = 1 + \frac{\bar{a}z}{1-\bar{a}z}$, and both 1 and $\frac{z}{1-\bar{a}z}$ are mono-components, with the orthogonal relation $\langle 1, \frac{\bar{a}z}{1-\bar{a}z}\rangle = 0$.

Thus, a given function $f \in \mathcal{H}^2(\mathbb{D})$ can be associated to the mono-component decomposition

$$f = \sum_{l=0}^{n}\langle R^l f, e_{k_l,a_l}\rangle e_{k_l,a_l} + R^{n+1}f = S_n f + R^{n+1}f, \tag{5}$$

where $S_n f := \sum_{l=0}^{n}\langle R^l f, e_{k_l,a_l}\rangle e_{k_l,a_l}$, $R^n f$ is inductively defined by $R^0 f := f$, $R^{n+1}f := R^n f - \langle R^n f, e_{k_n,a_n}\rangle e_{k_n,a_n}$, and we have by (3) that

$$\langle R^l f, e_{k_l,a_l}\rangle = \frac{R^l f^{(k_l)}(a_l)}{\|\varphi_{k_l,a_l}\|} \quad \text{for each } l \in \mathbb{N}. \tag{6}$$

If we are given a real-valued signal $\tilde{f} \in L^2(\partial\mathbb{D})$, then the mono-component decomposition of $\tilde{f}$ can be obtained through the relation ([10])

$$\tilde{f}(e^{it}) = 2\mathrm{Re}\left(f^+(e^{it})\right) - c_0,$$

where

$$f^+(z) = \frac{1}{2\pi}\int_0^{2\pi} \frac{\tilde{f}(e^{it})}{1 - ze^{-it}} dt \in \mathcal{H}^2(\mathbb{D}), \quad c_0 = \frac{1}{2\pi}\int_0^{2\pi} \tilde{f}(e^{it}) dt.$$

Although collectively (5) is not an orthogonal decomposition, there still holds ([13])

$$\|f\|^2 = \sum_{l=0}^{n} |\langle R^l f, e_{k_l,a_l}\rangle|^2 + \|R^{n+1} f\|^2.$$

Usually in pure greedy algorithm, to get a locally optimal projection of each iterated residue onto the atoms in $\mathcal{D}$, at each step e_{k_l,a_l} should be sorted to fulfil the condition

$$|\langle R^l f, e_{k_l,a_l}\rangle| \geq \alpha \sup_{g\in\mathcal{D}} |\langle R^l f, g\rangle|, \tag{7}$$

where the optimality factor α is situated between 0 and 1. The main contribution of this section is to prove that α can attain 1, which is in fact a corollary of the following theorem.

THEOREM 2.2 *Suppose $f \in \mathcal{H}^2(\mathbb{D})$, then*

$$\lim_{k\to\infty} |\langle f, e_{k,a}\rangle| = 0 \tag{8}$$

holds uniformly with respect to $a \in \mathbb{D}$. Write $a = |a|\xi$, then

$$\lim_{|a|\to 1^-} |\langle f, e_{k,a}\rangle| = 0 \tag{9}$$

holds uniformly with respect to $(k, \xi) \in \mathbb{N} \times \partial\mathbb{D}$.

Proof Let

$$T_{f,N}(z) = \sum_{l=0}^{N} \frac{f^{(l)}(0)}{l!} z^l,$$

in which N is large enough such that $\|f - T_{f,N}\| < \epsilon$. Therefore, if $k > N$, we will get

$$\begin{aligned} |\langle f, e_{k,a}\rangle| &\leq |\langle f - T_{f,N}, e_{k,a}\rangle| + |\langle T_{f,N}, e_{k,a}\rangle| \\ &= |\langle f - T_{f,N}, e_{k,a}\rangle| \\ &\leq \|f - T_{f,N}\| \\ &< \epsilon, \end{aligned}$$

that proves (8). To show that (9) is uniform, in view of (8), it suffices to verify the truth for any fixed k. Now, for each $k \in \mathbb{N}$, we have

$$|\langle f, e_{k,a}\rangle| \leq |\langle f - T_{f,N}, e_{k,a}\rangle| + |\langle T_{f,N}, e_{k,a}\rangle| \leq \|f - T_{f,N}\| + \frac{|T_{f,N}^{(k)}(a)|}{\|\varphi_{k,a}\|}.$$

(9) then follows by observing that $T_{f,N}^{(k)}$ is a bounded function, and $\|\varphi_{k,a}\| \to \infty$ as $|a| \to 1^-$.

The proof of the theorem is complete. □

Let the adaptive mono-component decomposition (5) subject to the selection criterion (7) with $\alpha = 1$, we have

THEOREM 2.3

$$\|S_n f - f\| = \|R^{n+1} f\| \to 0 \ (n \to \infty), \tag{10}$$

and for any $k \in \mathbb{N}$ and any compact subset Ω of $\mathbb{D}$, the pointwise convergence

$$\left|(S_n f)^{(k)}(z) - f^{(k)}(z)\right| = \left|(R^{n+1} f)^{(k)}(z)\right| \to 0 \ (n \to \infty) \tag{11}$$

uniformly holds with respect to $z \in \Omega$.

Proof (10) is a consequence of [13, Theorem 1] and the fact that $\overline{\text{span}\mathcal{D}} = \overline{\text{span}\{z^n\}_{n=0}^{\infty}} = \mathcal{H}^2(\mathbb{D})$. In addition, from (4) we know that

$$\left|(R^{n+1} f)^{(k)}(z)\right| \le \frac{\sqrt{(2k)!}\|R^{n+1} f\|}{(1 - |z|^2)^{k+\frac{1}{2}}},$$

which, together with (10), gives (11). □

Example Let $f(z) = \frac{z^5}{(2-z)^4}$, then $f \in \mathcal{H}^2(\mathbb{D})$ and $\|f\|^2 \approx 0.112026$. The adaptive mono-component decomposition of f by the higher order Szegö kernels is given by

$$f(z) = c_0 e_{k_0,a_0}(z) + c_1 e_{k_1,a_1}(z) + c_2 e_{k_2,a_2}(z) + R^3 f(z), \quad z \in \mathbb{D}, \tag{12}$$

where the parameters and coefficients are listed in Table 1. So, $\|R^3 f\|^2 = \|f\|^2 - |c_0|^2 - |c_1|^2 - |c_2|^2 \approx 2.13 \times 10^{-4}$.

If we use the adaptive Fourier decomposition (2), then

$$f(z) = \sum_{l=1}^{6} \langle f, B_{\{a_1,\ldots,a_l\}}\rangle B_{\{a_1,\ldots,a_l\}}(z) + R^6 f(z) = \sum_{l=1}^{6} c_l B_{\{a_1,\ldots,a_l\}}(z) + R^6 f(z).$$

The associated parameters and coefficients are shown in Table 2. By computation, $\|R^6 f\|^2 = \|f\|^2 - \sum_{l=1}^{6} |c_l|^2 \approx 0.004640$.

Table 1. Parameters and coefficients for higher order Szegö kernels.

i	Parameter k_i	Parameter a_i	Coefficient c_i	Energy $\|c_i\|^2$
0	5	$0.3882 - 3.028 \times 10^{-6}i$	$0.333881 - 8.390 \times 10^{-6}i$	0.111476
1	3	$0.3687 - 0.4120i$	$-2.664 \times 10^{-3} - 0.013160i$	1.803×10^{-4}
2	7	$0.5111 + 0.04472i$	$0.009354 + 0.008321i$	1.567×10^{-4}

Table 2. Parameters and coefficients for TM system.

i	Parameter a_i	Coefficient c_i	Energy $\|c_i\|^2$
1	0	0	0
2	$0.9397 + 1.927 \times 10^{-6}i$	$0.210993 + 3.26454 \times 10^{-6}i$	0.044518
3	$0.8158 + 0.3376i$	$0.166207 + 0.064296i$	0.031759
4	$0.8582 - 0.3233i$	$-0.091105 + 0.098243i$	0.017952
5	$0.9027 + 0.03013i$	$-0.098430 - 0.004388i$	0.009708
6	$0.7291 - 0.5073i$	$0.047801 + 0.034118i$	0.003449

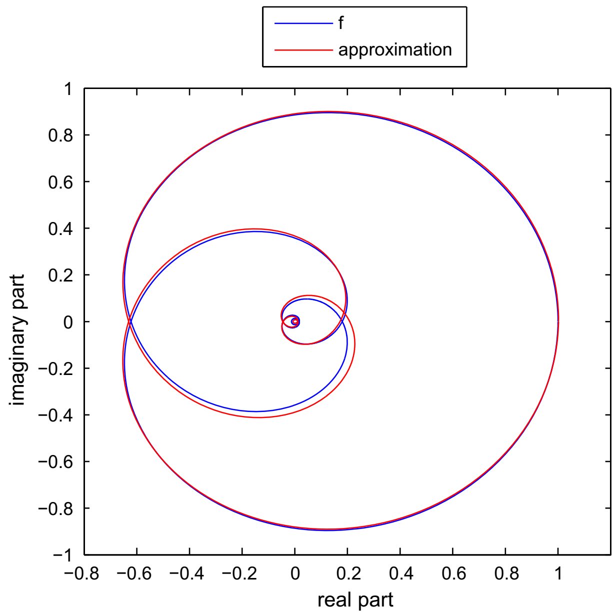

Figure 1. By higher order Szegö kernels.

If we choose the Fourier system for decomposition, then

$$f(z) = \frac{z^5}{16} + \frac{z^6}{8} + \frac{5z^7}{32} + \frac{5z^8}{32} + \frac{35z^9}{256} + \frac{7z^{10}}{64} + R^{11} f(z),$$

with the energy of the remainder term being $\|R^{11} f\|^2 \approx 0.112026 - 0.099014 = 0.013012$.

The approximation of f on the boundary $\partial\mathbb{D}$ by these methods is shown in Figures 1–3, where the horizontal axis and vertical axis correspond, respectively, to the real part and imaginary part of the boundary limit of the objective function. The *composing-transient-time-frequency distribution* (CTTFD) (cf. [17]) based on the density of instantaneous frequency (cf. [18,19]) for decomposition (12) is displayed in Figure 4. Comparing these three methods, we could see clearly that the first method outperforms others for f.

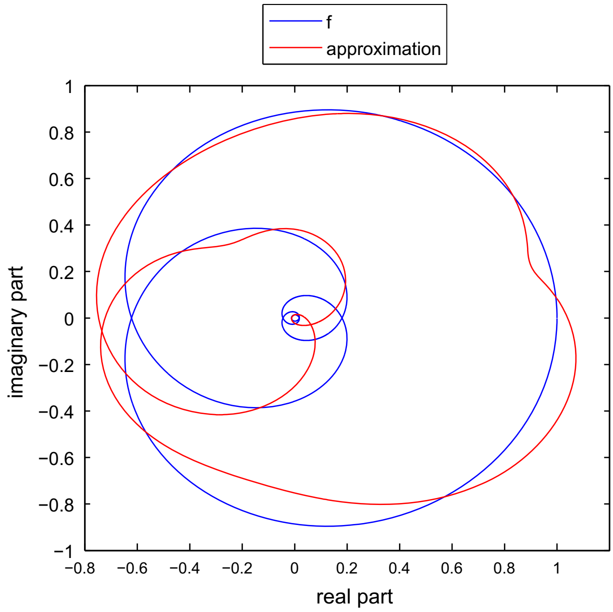

Figure 2. By adaptive Fourier decomposition.

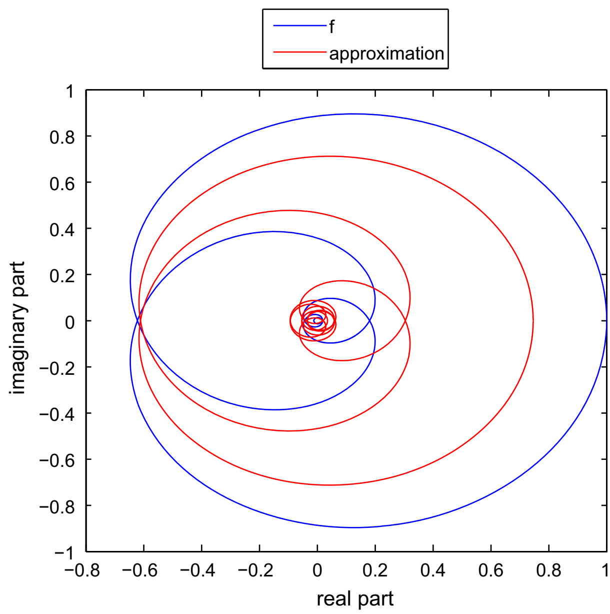

Figure 3. By Fourier system.

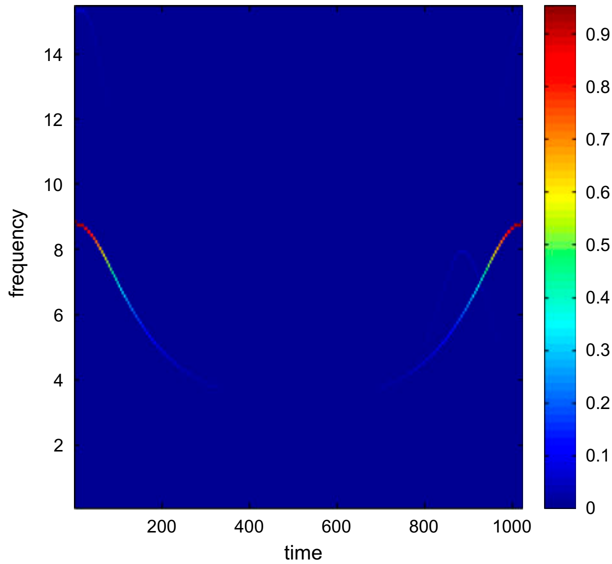

Figure 4. CTTFD for decomposition (12).

3. The case for the upper half-plane

Denote the upper half-plane $\{z \in \mathbb{C} : \operatorname{Im} z > 0\}$ by $\mathbb{C}^+$. Let f be holomorphic on $\mathbb{C}^+$, we say $f \in \mathcal{H}^2(\mathbb{C}^+)$ if

$$\|f\| := \sup_{y>0} \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} |f(x+iy)|^2 dx \right)^{1/2} < \infty.$$

$\mathcal{H}^2(\mathbb{C}^+)$ is endowed with the inner product

$$\langle f, g \rangle := \frac{1}{2\pi} \int_{-\infty}^{\infty} f(x)\overline{g}(x) dx,$$

where $f(x) := \lim_{y \to 0^+} f(x+iy)$ for almost every $x \in \mathbb{R}$.

$\mathcal{H}^2(\mathbb{D})$ is isometrically isomorphic to $\mathcal{H}^2(\mathbb{C}^+)$ under the map

$$(Tf)(z) = \frac{\sqrt{2}}{z+i} f\left(\frac{i-z}{i+z} \right), \quad f \in \mathcal{H}^2(\mathbb{D}).$$

However, the image of a higher order Szegö kernel function in $\mathcal{H}^2(\mathbb{D})$ under this map is no longer a higher order Szegö kernel function in $\mathcal{H}^2(\mathbb{C}^+)$, except for the Szegö kernel function of order 0. The main purpose of this section is to investigate for $\mathcal{H}^2(\mathbb{C}^+)$ the analogue of what we have done in the unit disc.

The dictionary consists of the higher order Szegö kernels for this case is

$$\mathcal{D} = \left\{ e_{k,a} = \frac{\varphi_{k,a}}{\|\varphi_{k,a}\|} : k \in \mathbb{N}, a = a_0 + ia_1 \in \mathbb{C}^+ \right\},$$

where

$$\varphi_{k,a}(z) = \frac{\partial^k}{\partial a_0^k}\left(\frac{1}{z-\overline{a}}\right) = \frac{k!}{(z-\overline{a})^{k+1}},$$

and

$$\|\varphi_{k,a}\|^2 = \frac{(k!)^2}{2\pi}\int_{-\infty}^{\infty}\frac{dx}{\left((x-a_0)^2+a_1^2\right)^{k+1}} = \frac{(2k)!}{(2a_1)^{2k+1}}.$$

For any $f \in \mathcal{H}^2(\mathbb{C}^+)$, there holds

$$\langle f, e_{k,a}\rangle = \frac{i(2a_1)^{k+\frac{1}{2}} f^{(k)}(a)}{\sqrt{(2k)!}}.$$

Consequently,

$$(2a_1)^{k+\frac{1}{2}}|f^{(k)}(a)| \le \sqrt{(2k)!}\|f\|, \quad \forall a \in \mathbb{C}^+.$$

Similarly, we have

PROPOSITION 3.1 *For any $(k,a) \in \mathbb{N}\times\mathbb{C}^+$, $e_{k,a}$ is a mono-component.*

This can be easily verified in view of the following theorem.

THEOREM 3.1 *Suppose $f \in \mathcal{H}^2(\mathbb{C}^+)\bigcap C^1(\overline{\mathbb{C}^+})$, with the representation*

$$f(z) = f(x+iy) = \rho(x,y)e^{i\theta(x,y)}.$$

If for every $x \in \mathbb{R}$, $\frac{\partial\rho(x,y)}{\partial y} < 0$ in $\{y : 0 < y < \delta_x\}$, then f is a mono-component, i.e.

$$\theta'(x) = \frac{d\left(\lim\limits_{y\to 0^+}\theta(x,y)\right)}{dx} \ge 0, \quad \forall x \in \mathbb{R}.$$

Proof

$$\begin{aligned}\theta'(x) &= \lim_{y\to 0^+}\frac{\partial\theta(x,y)}{\partial x}\\ &= \lim_{y\to 0^+}\mathrm{Re}\left(\frac{-if_x'(x+iy)}{f(x+iy)}\right)\\ &= -\lim_{y\to 0^+}\mathrm{Re}\left(\frac{f_y'(x+iy)}{f(x+iy)}\right)\\ &= -\lim_{y\to 0^+}\mathrm{Re}\left(\frac{\partial\log f(z)}{\partial y}\right)\\ &= -\lim_{y\to 0^+}\frac{\partial\log|f(z)|}{\partial y}\\ &= -\lim_{y\to 0^+}\left(\frac{1}{\rho(x,y)}\frac{\partial\rho(x,y)}{\partial y}\right).\end{aligned}$$

□

The adaptive mono-component decomposition for $f \in \mathcal{H}^2(\mathbb{C}^+)$ is given by

$$f = \sum_{l=0}^{n} \langle R^l f, e_{k_l,a_l}\rangle e_{k_l,a_l} + R^{n+1} f, \tag{13}$$

subject to the condition

$$|\langle R^l f, e_{k_l,a_l}\rangle| = \sup_{(k,a)\in\mathbb{N}\times\mathbb{C}^+} |\langle R^l f, e_{k,a}\rangle|, \tag{14}$$

which is guaranteed by

THEOREM 3.2 *Suppose* $f \in \mathcal{H}^2(\mathbb{C}^+)$*, then*

$$\lim_{k\to\infty} |\langle f, e_{k,a}\rangle| = 0 \tag{15}$$

holds uniformly with respect to $a \in \mathbb{C}^+$*;*

$$\lim_{a_1\to 0^+} |\langle f, e_{k,a}\rangle| = \lim_{a_1\to\infty} |\langle f, e_{k,a}\rangle| = 0 \tag{16}$$

holds uniformly with respect to $(k, a_0) \in \mathbb{N}\times\mathbb{R}$*, and*

$$\lim_{|a_0|\to\infty} |\langle f, e_{k,a}\rangle| = 0 \tag{17}$$

holds uniformly with respect to $(k, a_1) \in \mathbb{N}\times(0,\infty)$.

Proof Since $\overline{\mathrm{span}\{e_{0,b} : b \in \mathbb{C}^+\}} = \mathcal{H}^2(\mathbb{C}^+)$, for any $\epsilon > 0$ there exists

$$T_{f,N}(z) = \sum_{l=1}^{N} \frac{c_l}{z - \overline{b_l}}$$

such that $\| f - T_{f,N}\| < \epsilon$. Thus

$$|\langle f, e_{k,a}\rangle| \le |\langle f - T_{f,N}, e_{k,a}\rangle| + |\langle T_{f,N}, e_{k,a}\rangle| < \epsilon + \sum_{l=1}^{N} |c_l| \left|\left\langle \frac{1}{z-\overline{b_l}}, e_{k,a}\right\rangle\right|.$$

For each $b_l \in \mathbb{C}^+$ $(1 \le l \le N)$, we have

$$\begin{aligned}\left|\left\langle \frac{1}{z-\overline{b_l}}, e_{k,a}\right\rangle\right| &= \left| \frac{i(-1)^k k!(2a_1)^{k+\frac{1}{2}}}{\sqrt{(2k)!}(a-\overline{b_l})^{k+1}} \right| \\ &= \sqrt{\frac{(2k)!!}{(2k-1)!!}} \frac{\sqrt{2}a_1^{k+\frac{1}{2}}}{|a-\overline{b_l}|^{k+1}} \\ &\le \sqrt{\frac{(2k)!!}{(2k-1)!!}} \frac{\sqrt{2}a_1^{k+\frac{1}{2}}}{(a_1+\mathrm{Im}\, b_l)^{k+1}}.\end{aligned}$$

By Stirling's formula,

$$\frac{(2k)!!}{(2k-1)!!} \sim \sqrt{k\pi},$$

and we note that $\frac{\sqrt{2}a_1^{k+\frac{1}{2}}}{(a_1+\operatorname{Im} b_l)^{k+1}}$ takes the maximum at $a_1 = (2k+1)\operatorname{Im} b_l$. Hence,

$$\left|\left\langle \frac{1}{z-\overline{b_l}}, e_{k,a}\right\rangle\right| \leq C\frac{k^{\frac{1}{4}}(2k+1)^{k+\frac{1}{2}}}{\sqrt{\operatorname{Im} b_l}(2k+2)^{k+1}} \leq \frac{C}{\sqrt{\operatorname{Im} b_l}}k^{-\frac{1}{4}},$$

that proves (15). (16) and (17) then follow from the estimate for any fixed $k \in \mathbb{N}$:

$$\left|\left\langle \frac{1}{z-\overline{b_l}}, e_{k,a}\right\rangle\right| = \sqrt{\frac{2(2k)!!}{(2k-1)!!}}\frac{a_1^{k+\frac{1}{2}}}{|a-\overline{b_l}|^{k+1}} \leq \begin{cases} C_{k,b_l}a_1^{k+\frac{1}{2}} & \text{if } a_1 \text{ is small,} \\ \frac{C_k}{\sqrt{a_1}} & \text{if } a_1 \text{ is large,} \\ \frac{C_{k,b_l}a_1^{k+\frac{1}{2}}}{|a_0|^{k+1}} & \text{if } |a_0| \text{ is large.} \end{cases}$$

□

Analogous to Theorem 2.3, we have for the adaptive decomposition (13) that

THEOREM 3.3

$$\|R^{n+1}f\| \to 0 \ (n \to \infty), \tag{18}$$

and for any $k \in \mathbb{N}$ *and any* $t > 0$,

$$(R^{n+1}f)^{(k)}(z) \to 0 \ (n \to \infty) \tag{19}$$

uniformly holds on $\{z \in \mathbb{C} : \operatorname{Im} z \geq t\}$.

Funding

This work was supported by Macao FDCT 098/2012/A3, research grant of the University of Macau No. UL017/08-Y4/MAT/QT01/FST, and Multi-Year Research Grant of the University of Macau No. MYRG116(Y1-L3)-FST13-QT.

References

[1] Qian T. Mono-components for decomposition of signals. Math. Meth. Appl. Sci. 2006;29:1187–1198.

[2] Dang P, Qian T, You Z. Hardy-Sobolev spaces decomposition in signal analysis. J. Fourier Anal. Appl. 2011;17:36–64.

[3] Qian T. Boundary derivatives of the phases of inner and outer functions and applications. Math. Meth. Appl. Sci. 2009;32:253–263.

[4] Qian T, Chen QH, Li LQ. Analytic unit quadrature signals with nonlinear phase. Phys. D: Nonlinear Phenomena. 2005;203:80–87.

[5] Tan LH, Shen LX, Yang LH. Rational orthogonal bases satisfying the Bedrosian identity. Adv. Comput. Math. 2010;33:285–303.

[6] Xu Y, Yan D. The Bedrosian identity for the Hilbert transform of product functions. Proc. Amer. Math. Soc. 2006;134:2719–2728.

[7] Yang LH, Zhang HZ. The Bedrosian identity for $\mathbf{H}^p$ functions. J. Math. Anal. Appl. 2008;345:975-984.

[8] Yu B, Zhang HZ. The Bedrosian identity and homogeneous semi-convolution equations. J. Integral Equ. Appl. 2008;20:527–568.
[9] Qian T. Intrinsic mono-component decomposition of functions: an advance of Fourier theory. Math. Meth. Appl. Sci. 2010;33:880–891.
[10] Qian T, Wang YB. Adaptive Fourier series-a variation of greedy algorithm. Adv. Comput. Math. 2011;34:279–293.
[11] Mi W, Qian T. Frequency-domain identification: an algorithm based on an adaptive rational orthogonal system. Automatica. 2012;48:1154–1162.
[12] Mi W, Qian T, Wan F. A fast adaptive model reduction method based on Takenaka-Malmquist systems. Syst. Control Lett. 2012;61:223–230.
[13] Mallat SG, Zhang Z. Matching pursuits with time-frequency dictionaries. IEEE Trans. Signal Process. 1993;41:3397–3415.
[14] Temlyakov VN. Greedy approximation. Vol. 20, Cambridge monographs on applied and computational mathematics. Cambridge: Cambridge University Press; 2011.
[15] Qian T, Wegert E. Optimal approximation by Blaschke forms. Complex Var. Ellipt. Equ. 2013;58:123–133.
[16] Garnett JB. Bounded analytic functions. Revised 1st ed. New York (NY): Springer-Verlag; 2007.
[17] Dang P, Qian T, Guo YY. Transient time-frequency distribution based on mono-component decompositions. Int. J. Wavelets Multiresolut. Inf. Process. 2013;11:1350022. 24 p.
[18] Cohen L. Time-frequency analysis. NewYork (NY): Prentice-Hall; 1995.
[19] Cohen L, Lee C. Instantaneous frequency, its standard deviation and multicomponent signals. Proc. SPIE. 1988;975:186–208.