SURVEY ARTICLE

Positive-instantaneous frequency and approximation

Tao QIAN

Macao Center for Mathematical Science, Macao University of Science and Technology, Macao, China

Abstract   Positive-instantaneous frequency representation for transient signals has always been a great concern due to its theoretical and practical importance, although the involved concept itself is paradoxical. The desire and practice of uniqueness of such frequency representation (decomposition) raise the related topics in approximation. During approximately the last two decades there has formulated a signal decomposition and reconstruction method rooted in harmonic and complex analysis giving rise to the desired signal representations. The method decomposes any signal into a few basic signals that possess positive instantaneous frequencies. The theory has profound relations to classical mathematics and can be generalized to signals defined in higher dimensional manifolds with vector and matrix values, and in particular, promotes kernel approximation for multi-variate functions. This article mainly serves as a survey. It also gives two important technical proofs of which one for a general convergence result (Theorem 3.4), and the other for necessity of multiple kernel (Lemma 3.7).

Expositorily, for a given real-valued signal f one can associate it with a Hardy space function F whose real part coincides with f. Such function F has the form $F = f + iHf$, where H stands for the Hilbert transformation of the context. We develop fast converging expansions of F in orthogonal terms of the form

$$F = \sum_{k=1}^{\infty} c_k B_k,$$

where B_k's are also Hardy space functions but with the additional properties

$$B_k(t) = \rho_k(t)e^{i\theta_k(t)}, \quad \rho_k \geqslant 0, \quad \theta_k'(t) \geqslant 0, \quad \text{a.e.}$$

The original real-valued function f is accordingly expanded

$$f = \sum_{k=1}^{\infty} \rho_k(t) \cos \theta_k(t)$$

which, besides the properties of ρ_k and θ_k given above, also satisfies

$$H(\rho_k \cos \theta_k)(t) = \rho_k(t) \sin \theta_k(t).$$

Real-valued functions $f(t) = \rho(t) \cos \theta(t)$ that satisfy the condition

$$\rho \geqslant 0, \quad \theta'(t) \geqslant 0, \quad H(\rho \cos \theta)(t) = \rho(t) \sin \theta(t)$$

are called mono-components. If f is a mono-component, then the phase derivative $\theta'(t)$ is defined to be instantaneous frequency of f. The above described positive-instantaneous frequency expansion is a generalization of the Fourier series expansion. Mono-components are crucial to understand the concept instantaneous frequency. We will present several most important mono-component function classes. Decompositions of signals into mono-components are called adaptive Fourier decompositions (AFDs). We note that some scopes of the studies on the 1D mono-components and AFDs can be extended to vector-valued or even matrix-valued signals defined on higher dimensional manifolds. We finally provide an account of related studies in pure and applied mathematics.

Keywords: Möbius transform, blaschke product, mono-component, hilbert transform, hardy space, inner and outer functions, adaptive fourier decomposition, rational orthogonal system, nevanlinna factorization, beurling-lax theorem, reproducing kernel hilbert space, several complex variables, clifford algebra, pre-orthogonal AFD

MSC2020 60G35, 30H10, 30J10, 32A50

1 Introduction

It is a common sense among analysts that "On the unit circle it is harmonic analysis; and inside the unit circle is complex analysis" [40], and the same is true for a manifold and its neighborhood regions. In general, the following mechanism may be regarded as complex analysis method of harmonic analysis. When studying analysis on the boundary of a region, say for instance, in an Euclidean space, one can at the cost of one more (or p-more) dimension (dimensions), imbed the region together with its boundary into the larger space, while the latter is equipped with a Cauchy complex structure, including the Cauchy Theorem, Cauchy kernel and the Cauchy formula, etc. That is, one treats the boundary of the region as a co-dimension 1 (or co-dimension p) manifold in the larger space with a Cauchy complex structure. With the complex structure one can define complex Hardy spaces consisting of "good" complex holomorphic functions in the regions divided out by the manifold. By "good" here we

in particular mean the complex-Hardy functions defined on the regions having non-tangential boundary limits as projections into the corresponding function spaces on the manifold. Conversely, each function in an appropriate function class on the manifold can be be made to correspond with one of those Hardy-non-tangential boundary limits, the latter being called analytic signals. Those ideas appeared in the lectures of Cheng and Deng given in Peking University [10], in the book of Gorusin translated by Chen [41], in works of A. McIntosh and, separately, of C. Kenig and other authors, in relation to complex Hardy spaces, singular integrals, boundary value problems and related topics on Lipschitz curves and surfaces. This article serves as a survey on the study that the author and his collaborators have been undertaking by implementing the complex analysis method to harmonic and signal analysis.

The study can be divided into two parts of which one is mono-component function theory, dealing with signals possessing a non-negative instantaneous frequency function; and the other is approximation to arbitrary function with appropriate mono-components. Note that the monomials $z^n, n = 0, 1, \ldots$, are particular mono-components, while the Fourier series expansion is a mono-component approximation. The study that we are to explore is a generalization of the Fourier theory into the scope of the Beurling-Lax Theorem-related direct sum decomposition of the Hardy spaces into the forward shift and backward shift invariant subspaces.

The study at beginning was motivated by the tentative definition of the concept *instantaneous frequency* (IF), or, in brief, the *frequency* function, by Gabor [35]. It is still a controversial concept up to the present time. People tend to believe that for a general signal there is a certain "frequency" at each moment of time. This belief is supported by sinusoidal functions that possess constant frequencies. Justification of existence of a frequency function crucially depends on how to define IF. Unfortunately, except for the sinusoidal functions the IF concept itself appears to be paradoxical: "frenquency" is the oscillation number (or in the averaging sense) per unit time duration, hence a time interval is required in order to determine it; while "instantaneous" involves only a time moment. The concept is also contradictory with the uncertainty principle. There would be no uniformly accepted theory: A variety of engineering definitions of IF have been proposed of which, in the author's opinion, all vague and self-contradictory. None of the theory, nor the applications, are satisfied neither by mathematicians, nor by physicians, and nor by signal analysts as [8, 16]. The author's view is that there may not exist an anticipated IF concept for a general signal. One can, however, propose a mathematical and conceptual definition of instantaneous frequency based on which signals can be effectively analyzed. The proposed definition of IF is based on the Möbius transformation that, by its beauty in mathematics, has been ready to be used for the IF definition. A coherent theory that has delicate and profound relations to classical mathematics and great potential in applications has been initialized. As a new trend of Fourier analysis it emphasizes on non-linear phase phenomenon and consists of two parts: defining the IF concept, and decomposing general

signal into those possessing IF. We call the signals possessing an IF function as *mono-components* (or MCs). Signals that do not possess an IF are called *multi-components*. There are classical function classes belonging to the mono-component function class. There are also newly constructed interesting and significant function classes belonging to the mono-component function class. The mono-component function theory is a combined effort by harmonic and signal analysts (see Section 2 and the related literature in the references).

For mono-component approximation we note that Weiss and Weiss published a paper in 1962 re-proving the Nevanlinna factorization theorem in the complex Hardy spaces of one complex variable. The factorization result is a crucial tool in the complex Hardy spaces theory. Directly related to the approximation, Coifman and Nahon developed the non-linear phase unwinding algorithm (UWA) to expand any analytic signal into a series of Blaschke products [63]. In their 2016 paper [12] Coifman and Steinerberger formally published the UWA theory and algorithm, and further developed some aspects initialized in [63]. A later paper by those authors together with Wu developed the practical use of the unwinding method to functions of finite energy along with their computation method of the IFs [13]. More recently a new paper by Coifman and Peyriére studies invariant subspace decompositions including the Schauder basis property of the unwinding series [11]. Being unaware of Nahon's thesis, Qian independently studied the UWA method and proved its H^2-convergence in Remark 4.4, [71], and independently uses the same English word "unwinding" to call it in [81]. It is noted that UWA is a special case of UWAFD, the latter being incorporated with a sifting process using a generalized backward shifting operator together, as well as a maximal selection principle [71, 84].

As already mentioned unwinding method is only one of the two main ones in the adaptive approximation methodology. The other one is base on the so called maximal selection principle (MSP). The terminology adaptive Fourier decomposition (AFD), as a matter of fact, at the very beginning started from the MSP type [84], and further extended to the UWA type. The 1D maximal selection type AFD heavily depends on the factorization properties of one complex variable. For multi-variables cases, either with the several complex variables or the Clifford algebra settings, the AFD methods are not directly applicable. In our latest studies we extend the AFD idea to reproducing kernel Hilbert spaces with certain boundary vanishing property, called pre-orthogonal AFD, or POAFD in brief. The mono-component function theory and the related AFD approximation theory have found significant applications, including those in system identification, signal and image processing, etc. [27, 59, 61, 95, 125]. We will include some literature with short descriptions on engineering applications.

The writing of the paper is organized as follows (Fig. 1). In Section 2 we present the main results of mono-component function theory, including the definition of mono-component function, the inner function type, the Bedrosian type, and the starlike type mono-components. In Section 3 we give an account on various kinds of AFD algorithms in the classical setting, as well as in reproducing kernel Hilbert spaces. In Section 4 we provide information on related
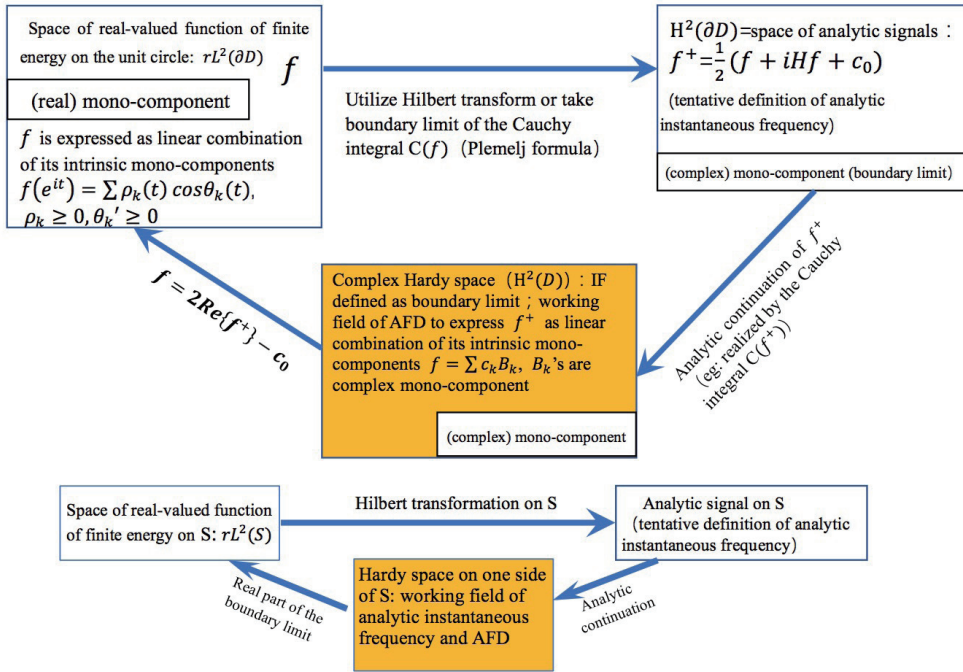
studies and applications.

Fig. 1   Complex harmonic analysis method in analyzing signals

2   Mono-component function theory

2.1   Mono-component and IF

In 1946 Gabor proposed his *analytic signal* approach [38]. Throughout this article we restrict ourselves to only signals with finite energy, or L^2-functions. The theory on the unit circle is parallel with that on the real line. To explain the idea we most time restrict ourselves to the unit circle case. We occasionally jump into the upper-half space context, including for instance when we describe the ideas in relation to the Bedrosian type results in terms of Fourier transform. Let $s(t)$ be a real-valued signal of finite energy on the unit circle $\partial \mathbf{D}$ where $\mathbf{D}$ denotes the unit disc. The associated analytic signal, denoted by $s^+(t)$, is defined

$$s^+(e^{it}) = \frac{1}{2}\left(s(e^{it}) + i\tilde{H}s(e^{it}) + c_0\right), \tag{1}$$

where $\tilde{H}$ is the circular Hilbert transformation, and c_0 is the 0-th Fourier coefficient, or average of s on the circle. That is

$$\tilde{H}s(e^{it}) = \frac{1}{2\pi}\text{p.v.}\int_0^{2\pi} f(e^{iu})\cot\left(\frac{t-u}{2}\right)\mathrm{d}u, \quad c_0 = \frac{1}{2\pi}\int_0^{2\pi} f(e^{iu})\mathrm{d}u.$$

We note s^+ is the non-tangential boundary limit of the Cauchy integral of s (the Plemelj formula):

$$s^+(e^{it}) = \lim_{z \to e^{it}} \frac{1}{2\pi} \int_0^{2\pi} \frac{f(e^{iu})}{z - e^{iu}} e^{iu} du, \qquad \text{a.e.}$$

The fact that s is real-valued makes the Hilbert transform part $\frac{1}{2}\tilde{H}f$ in Eq. (1) the purely imaginary part of s^+, and $s = 2\mathrm{Re}s^+ - c_0$. There also holds the following relation that in the real line context corresponds to the Laplace transform

$$s^+(e^{it}) = \sum_{k=0}^{\infty} c_k e^{ikt}.$$

What is important is that $s^+(e^{it})$ has a holomorphic continuation into the interior of the disc

$$s^+(z) = \sum_{k=0}^{\infty} c_k z^k, \quad |z| < 1,$$

as a Hardy $H^2(\mathbf{D})$ function in the sense that whose non-tangential boundary limit coincides with $s^+(e^{it})$. The Fourier multiplier of the circular Hilbert transformation is $-i\mathrm{sgn}$, that is, if

$$s(e^{it}) = \sum_{k=-\infty}^{\infty} c_k e^{ikt}, \quad c_k = \frac{1}{2\pi} \int_0^{2\pi} s(e^{iu}) e^{-iku} du,$$

where $\mathrm{sgn}(k) = 1$, if $k > 0$, and $\mathrm{sgn}(k) = -1$, if $k < 0$, and $\mathrm{sgn}(0) = 0$, then

$$\tilde{H}s(e^{it}) = \sum_{k=-\infty}^{\infty} (-i)\mathrm{sgn}(k) c_k e^{ikt}.$$

In the sequel we drop the tilde sign above $\tilde{H}$ and write it simply as H.

This Fourier multiplier form of the Hilbert transform gives rise to the Hilbert transform characterization of the Hardy spaces. If restricted to the L^2 cases, it is: A function s of finite energy belongs to the Hardy H^2 space if and only if $Hs = -is$ [67]. This result holds in general contexts including the upper-half space cases in one and higher dimensions ([22, 23, 25]).

In writing $s^+(e^{it}) = \rho(t)e^{i\theta(t)}$, Gabor defined that the derivative of the phase function, $\theta'(t)$, to be the instantaneous frequency of $s(e^{it})$. In commenting on this definition we would say that the definition is "good", because if we take the example that for a positive integer $n, s(e^{it}) = \cos(nt)$, then in such way, $s^+(e^{it}) = e^{int}$, and the phase derivative is n, being complementary with the common sense. Gabor's definition, however, is not valid for general signals $s \in L^2(\partial\mathbf{D})$, but only tentative, due to the following reasons. First s, and thus s^+ as well, is an equivalent class of Lebesgue square-integrable functions that cannot be expected to be smooth and thus have phase derivative; and secondly,

the derivative, if exists, cannot be expected to be non-negative, as required in physics, and thus cannot stand as a qualified instantaneous frequency function. It is the signal analysts who decided that the IFs should be non-negative and thus can be effectively analyzed with applications. The primary importance is that the instantaneous frequency concept is generated from physics practice: it is an extension of the vibrating frequency. In the average sense the phase derivative of an analytic signal is non-negative as read out from the relation

$$\frac{1}{2\pi} \int_0^{2\pi} \theta'(t)|s^+(e^{it})|^2 \mathrm{d}t = \sum_{k=0}^{\infty} k|c_k|^2$$

[16, 29]. Pointwisely, however, the phase derivative of an analytic signal can be negative. For instance, for any non-trivial outer function in the complex Hardy space we have a set of positive Lebesgue measure on which the phase derivatives are strictly less than zero [70].

The strategy is to define a function set that exactly contains the signals having well defined non-negative analytic phase derivatives. It is the set of the functions called mono-components.

Definition (Qian [69]) Let s be a real- or complex-valued signal on the unit circle of finite energy. We call s a mono-component, or real-mono-component, if its analytic signal, or equivalently its projection into the Hardy space H^2, viz., $s^+(t) = \frac{1}{2}\left(s(t) + iHs(t) + c_0\right)$, in its phase-amplitude representation $s^+(t) = \rho(t)e^{i\theta(t)}$ satisfies $\theta'(t) \geqslant 0$, a.e., where the phase derivative $\theta'(t)$ is defined through the non-tangential limit of the same quantity from inside of the region. Precisely, in the unit circle case,

$$\theta'(t) = \lim_{r \to 1-} \theta'_r(t), \quad \text{a.e.,}$$

where $s^+(re^{it}) = \rho_r(t)e^{i\theta_r(t)}$ is the holomorphic continuation inside the unit disc. When s is a mono-component we call s^+ a complex-mono-component, or simply mono-component as well. When and only when s is a mono-component it has an instantaneous frequency function defined as its non-negative analytic phase derivative $\theta'(t)$.

Since s^+ is the non-tangential boundary limit of a Hardy space function inside the unit disc, $\theta'_r(t)$ everywhere exists, and

$$\theta'_r(t) = \mathrm{Re}\{\frac{re^{it}s^{+'}(re^{it})}{s^+(re^{it})}\}.$$

We note that the class of mono-component functions is closed under the multiplication operation but not the addition.

2.2   The inner function type mono-components

The above definition allows a number of interesting mono-component sub-classes. First we will mention the class of inner functions. It is easily observed

that the boundary function of the canonical Möbius transform mapping $a \in \mathbf{D}$ to zero, that is,

$$e_a(e^{it}) = \frac{e^{it} - a}{1 - \overline{a}z} = e^{i\theta_a(t)}$$

is an analytic signal whose phase derivative θ'_a is the Poisson kernel for the disc [38, 68, 78]. This early study along this direction was joined by Chen and Li. This implies that finite Blaschke products (Blaschke products with finitely many zeros) are all mono-components. The question is whether infinite Blaschke products are mono-components. As an application of the Julia-Wolff-Carathéodory Theorem the following result on general inner functions (containing finite and infinite Blaschke products and singular inner functions) is proved [70].

Theorem 2.1 (*Qian* [70]) *Let θ be a real-valued Lebesgue measurable function on the unit circle. Then the phase function $e^{i\theta}$ is a mono-component if and only if $e^{i\theta}$ is the non-tangential boundary limit of an inner function, or, equivalently, if and only if $H(e^{i\theta}) = -ie^{i\theta}$.*

The earlier study in [111] gives good observations and partial results. It is noted that in the earlier digital signal processing (DSP) literature the fact that Blaschke products possess positive phase derivative functions were stated without valid proof, as far as being aware by the author [14]. DSP scholars and engineers have been using the concept physically realizable signals with minimum phase that are outer functions without rigorous proof either. The minimum phase phenomenon is based on the fact that all boundary values of inner functions have non-negative phase derivative. The reference [70] proves the opposite property for outer functions: Under mild conditions that guarantee existence of the phase derivative function $\theta'(t)$ for an outer function there holds

$$\int_0^{2\pi} \theta'(t)\mathrm{d}t = 0.$$

The inner and outer functions are thus characterized in terms of sign properties of their phase derivatives.

2.3   The bedrosian type mono-components

The second class of mono-components is called the Bedrosian type. The Bedrosian theorem asserts the relation

$$H(fg) = fHg$$

under two forms of conditions, both being based on Fourier spectrum property of the functions. The first form of the conditions is that there exists $\sigma > 0$ such that $\mathrm{supp}\hat{f} \subset [-\sigma, \sigma]$, and $\mathrm{supp}\ \hat{g} \subset (-\infty, \sigma] \cap [\sigma, \infty)$. The second form of the conditions is that both functions f and g are in the Hardy H^2 space. In the language of the Fourier spectrum, by recalling the Paley-Wiener Theorem

for the Hardy space functions the second form of the conditions is equivalent with $f, g \in L^2$, supp$\hat{f} \subset [0, \infty)$ and supp $\hat{g} \subset [0, \infty)$. The idea of using the Bedrosian type results is as follows. Suppose that $e^{i\theta}$ is an analytic signal with the property $\theta'(t) \geqslant 0$, a.e. This kind of functions now have all been characterized by Theorem 2.1. One wishes to find a non-negative function $\rho(t)$ that makes the Bedrosian type relation $H(\rho(t)e^{i\theta} = \rho H(e^{i\theta})$ hold. For such a function ρ there holds

$$H(\rho e^{i\theta}) = \rho H(e^{i\theta}) = (-i)\rho e^{i\theta}.$$

By recalling the Hilbert transform characterization of the Hardy space functions the last equality implies that $\rho e^{i\theta}$ is an analytic signal, and, due to the positivity of the phase derivative $\theta'(t)$, it is a mono-component.

However, the original Bedrosian theorem cannot be directly used. The first form of the conditions refers to bandlimiting properties of the functions f and g. That, unfortunately, are not our case: The inner functions g do have the full spectrum range. The second form of the conditions would require that the amplitude function itself is the boundary limit of some Hardy space function. But it is not the case either (see the example given in Eq. (2)).

In order to enrich the mono-component class new conditions for the Bedrosian identity to hold were seeking by mainly a group of Chinese harmonic and signal analysts, including Chen, Tan, Wang, Wang, Xu, Yan, Yang, Yu, Zhang, Yan, etc., as well as the author ([82, 92, 105, 113, 120, 122], etc.), with Fourier analysis methods and complex analysis methods. The most successful result along this line is based on the following observation.

The essential structure of Bedrosian type mono-components is as follows:

$$s(e^{it}) = \left(\frac{1}{1 - \overline{a}_1 e^{it}} + \frac{1}{1 - a_1 e^{-it}} \right) \frac{e^{it} - a_1}{1 - \overline{a}_1 e^{it}} \frac{e^{it} - a_2}{1 - \overline{a}_2 e^{it}}. \tag{2}$$

On the circle it is a real-valued function multiplied with an order-2 Blaschke product. In verifying that $s(e^{it})$ is a Bedrosian type mono-component, the key point is that

$$\frac{1}{1 - \overline{a}_1 z} \frac{z - a_1}{1 - \overline{a}_1 z} \frac{z - a_2}{1 - \overline{a}_2 z}$$

is an analytic function in the disc; and for $|z| = 1$, the product

$$\frac{1}{1 - a_1 \overline{z}} \frac{z - a_1}{1 - \overline{a}_1 z} \frac{z - a_2}{1 - \overline{a}_2 z}$$

has an analytic continuation to the interior part of the disc. As result, $s(z)$ is a bounded analytic function. Since $\frac{1}{1 - \overline{a}_1 e^{it}} + \frac{1}{1 - a_1 e^{-it}}$ is real-valued and has finitely many sign-change pints on $|z| = 1$, it is, therefore, a so called *generalized amplitude* on the circle. We have the following general result ([82], the finite order Blaschke products case is proved in [106]).

Theorem 2.2 *Let $\phi(e^{it})$ be an infinite Blaschke product, where $a_1, \ldots, a_n, \ldots$ are the totality of its zeros, the multiples being all counted. Then (1) $\rho(t)$ is a real-valued function such that $\rho(t)\phi(e^{it}) \in H^p(\partial \mathbf{D}), 1 \leqslant p \leqslant \infty$, if and only if ρ is the real part of some function in the backward-shift-invariant subspace induced by the Blaschke product $\phi(e^{it})$, that is $\rho \in \mathrm{Re}\{H^p(\partial \mathbf{D}) \cap \phi(e^{it})\overline{H^p(\partial \mathbf{D})}\}$; and (2) For $1 < p < \infty$, $\rho \in \mathrm{Re}\{H^p(\partial \mathbf{D}) \cap \phi(e^{it})\overline{H^p(\partial \mathbf{D})}\}$ if and only if, in the L^p norm sense,*

$$\rho(t) = \mathrm{Re}\left\{ \sum_{k=1}^{\infty} c_k B_k(e^{it}) \right\},$$

where $c_k = \langle \rho(t), B_k(e^{it}) \rangle = \frac{1}{2\pi} \int_0^{2\pi} \rho(t)\overline{B_k(e^{it})}\mathrm{d}t, k = 1, 2, \ldots$, and $\{B_k\}_{k=1}^{\infty}$ is the rational orthonormal system (or TM-system) generated by $a_1, \ldots, a_k, \ldots$, that is

$$B_k(z) = \frac{\sqrt{1 - |a_k|^2}}{1 - \overline{a}_k z} \prod_{l=1}^{k-1} \frac{z - a_l}{1 - \overline{a}_l z}. \tag{3}$$

2.4 The non-bedrosian type mono-components: the starlike and boundary starlike type

The third type of mono-components is called the non-Bedrosian type which contains all p-starlike as well as boundary starlike functions in one complex variable. This kind of mono-components explores a different type of connections between mono-components and conformal mappings. Let f denote a univalent conformal mapping that, with $f(0) = 0$, maps the unit disc together with its continuous and thus rectifiable boundary. Obviously $f(e^{it})$ is a complex mono-component as its phase function is increasing along with increasing of the angular variable t. Below we will denote by $\mathscr{S}^*$ the set of such starlike functions. Next we define several other likewise function classes including p-starlike functions as follows.

Definition Let p be any positive integer. Denote by $\mathscr{S}(p)$ the set of p-valent holomorphic functions satisfying the following conditions:

(i) There exists $r: 0 < r < 1$, such that for all $z: r < |z| < 1$, there holds $\mathrm{Re}\{\frac{zf'(z)}{f(z)}\} > 0$; and

(ii) $\int_0^{2\pi} \mathrm{Re}\{\frac{zf'(z)}{f(z)}\}\mathrm{d}t = 2p\pi$ for all $z: r < |z| < 1$. Functions belonging to $\mathscr{S}(p)$ are called p-starlike functions.

Definition A function f is said to be a weak p-valent starlike function, and denoted $f \in \mathscr{S}_w(p)$, if and only if it is holomorphic in $\mathbf{D}$ with precisely p zeros in $\mathbf{D}$ (including multiples) and with the expression

$$f(z) = [h(z)]^p \prod_{k=1}^{p} \frac{(z - a_k)(1 - \overline{a}_k z)}{z},$$

where $h \in \mathscr{S}^*$.

With $p = 1$ and $a_1 = 0$ we obtain $\mathscr{S}_w(1) = \mathscr{S}(1)$. The article [43] shows that $\mathscr{S}(p)$ is a proper subset of $\mathscr{S}_w(p)$. The advantage of the latter is that functions in $\mathscr{S}_w(p)$ have an explicit representation formula. In order to reveal the essential structure we assume the convenient property that functions under study have a holomorphic continuation to an open neighborhood of the closed unit disc. Denote by $\mathscr{A}$ the set of such holomorphic functions, one can show $\mathscr{A} \cap \mathscr{S}(p) = \mathscr{A} \cap \mathscr{S}_w(p)$ [82]. To describe the relation between mono-components and the starlike function family we need two more definitions.

Definition [54] A univalent function is said to be a boundary starlike function with respect to the origin if f is holomorphic in $\mathbf{D}$, $\lim_{r \to 1-} f(r) = 0$, $f(\mathbf{D})$ is starlike with respect to the origin, and $\mathrm{Re}\{e^{i\alpha} f(z)\} > 0$ for some real number α and all $z \in \mathbf{D}$. Denote by $\mathscr{G}^*$ the set of all boundary starlike functions with respect to the origin.

The following definition specifies a class of mono-components.

Definition Let $f(e^{it}) = \rho(t)e^{i\theta(t)} \in L^p(\partial\mathbf{D})$, $p \geqslant 1$. Then f is called a Hilbert-n, or H-n atom, if it satisfies the following conditions:

(1) $H(\rho \cos \theta) = \rho \sin \theta$;

(2) $\rho \geqslant 0, \theta' \geqslant 0$ a.e.; and

(3) $\int_0^{2\pi} \theta'(t) dt = n\pi$.

Note that due to (1) f has a holomorphic continuation into the unit disc as a Hardy space function. In (2) the phase derivative θ' takes the sense given in . The condition (3) refers to the multivalent degree of f. The concept H-p atom was first proposed in [69] for $p = 2$ with the result that a function f is a H-2 atom if and only if f is a starlike function about the origin. Some further studies along this line for $p = 2n$ are give in [106]. The following result ultimately presents the relation between the H atoms and the starlike-boundary starlike functions.

Theorem 2.3 *Assume that f is holomorphic in $\overline{\mathbf{D}}$ having p zeros in the open disc $\mathbf{D}$. Then $f(e^{it})$ is a H-n atom, $n \geqslant 1$, if and only if*

$$f^2(z) = \left[\prod_{i=1}^{p} h_i(z)\right]^2 \prod_{j=1}^{n-2p} g_j^2(z)$$

$$= \left[\prod_{k=1}^{p} (z - a_k)\left(\frac{1}{z} - \overline{a}_k\right)\right]^2 \left[\prod_{k=1}^{n-2p} (z - b_k)\left(\frac{1}{z} - \overline{b}_k\right)\right] [h(z)]^n,$$

where $\{a_k\}_{k=1}^{p}$ are the zeros of $f(z)$ inside the unit disc, $\{b_k\}_{k=1}^{n-2p}$ are the zeros of $f(z)$ on the unit circle (both can be with multiples), $h(z) \in \mathscr{S}^*, h_i \in \mathscr{S}_w(1)$, and $g_j(b_j z) \in \mathscr{G}^*$ are all holomorphic in $\overline{\mathbf{D}}$, $i = 1, \ldots, p$, $j = 1, \ldots, n - 2p$.

The results on mono-component functions in all the three categories, viz., the inner function type, the Bedrosian type, and the starlike type, are not only
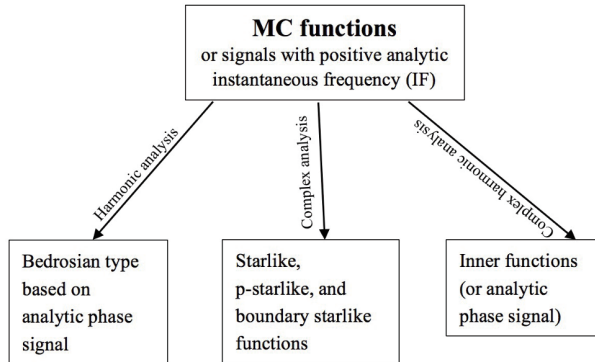
Fig. 2   Basic types of mono-component functions

important results in the mono-component function theory, but also new understanding to related topics in the classical harmonic and complex analysis (Fig. 2).

3   Adaptive fourier approximations

In this part we will give expository descriptions of adaptive Fourier approximations (AFD). In the one complex variable cases AFD gives rise to positive frequency expansions of signals into holomorphic rational functions, while in higher dimensions AFD offers at the moment at least fast converging rational approximations.

3.1   Mono-component decomposition of signals in general

The idea of positive frequency decompositions of signals is not new, it goes back to more than two hundreds years ago in relation to the name Jean Baptiste Joseph Fourier and other names. Fourier series will be exactly a particular case of the general theory that we are now to present.

Let s be a real-valued function defined on the unit circle $\partial \mathbf{D}$ with finite energy. We recall that its Hardy H^2 space projection is $s^+ = \frac{1}{2}(s + iHs + c_0)$. The simple relation $s = 2\mathrm{Re}\{s^+\} - c_0$ implies that a complex-mono-component decomposition $s^+(e^{it}) = \sum_{k=1}^{\infty} \rho_k(t)e^{i\theta_k(t)}$ (The possible coefficients c_k are absorbed by the corresponding amplitudes ρ_k and the phases θ_k) will gives rise to a real-mono-component decomposition, or positive-frequency decomposition of s: $s(e^{it}) = -c_0 + \sum_{k=1}^{\infty} \rho_k(t)\cos\theta_k(t)$. We are hence reduced to decomposing the complex Hardy space function s^+.

The philosophy is to find intrinsic constructing blocks of positive-time varying-instantaneous frequency. Here "intrinsic" has profound meaning. Under the present setting we understand it as "fast converging". In [80] we show that for any Hardy space function s^+ and any $\epsilon > 0$, there exist a constant c and two 1-starlike functions m_1 and m_2 such that

$$\|s^+ - (c + m_1 + m_2)\| \leqslant \epsilon.$$

The two starlike functions m_1 and m_2 are not unique, and, according to their construction, are very irregular. Deeper and more thorough consideration suggests that one would better use function systems of a certain type consisting of well behaved mono-components with explicit expressions such as rational functions, and better with monotonously increasing frequencies. In order to obtain intrinsic decomposition fast convergence would be desired.

3.2 One dimensional core-adaptive fourier decomposition (Core-AFD) and its variations

It did not take much time for the researcher to decide to use the rational orthonormal system or TM system. The difference with the traditional use of the system is that now it has to be adaptive: One must select parameters for the intrinsic construction concern, that means fast convergence. The system was already been introduced in Theorem 2.2. We note that TM systems in general cannot be avoided for they are Gram-Schmidt (G-S) orthogonalization of the partial fractions with poles outside the closed unit disc, the latter being fundamental constructive building blocks of rational functions in the Hardy spaces. TM systems consist of functions of positive frequency due to their construction in (finite) Blaschke products.

In the sequel we change our function notation s^+ in the Hardy $H^2(\mathbf{D})$ to f. In the unit circle context we have $f(z) = \sum_{l=1}^{\infty} c_l z^l, \sum_{l=1}^{\infty} |c_l|^2 < \infty$. Now we seek a decomposition of f into a TM system with adaptively selected parameters. The collection of the functions

$$e_a(z) = \frac{\sqrt{1 - |a|^2}}{1 - \overline{a}z}, \qquad a \in \mathbf{D}$$

consists of normalized Szegö kernels of the disc. Set $f = f_1$. First write

$$f(z) = \langle f_1, e_{a_1} \rangle e_{a_1}(z) + \frac{f_1(z) - \langle f_1, e_{a_1} \rangle e_{a_1}(z)}{\frac{z - a_1}{1 - \overline{a}_1 z}} \frac{z - a_1}{1 - \overline{a}_1 z}.$$

We note that in this stage a_1 can be any complex number in the unit disc and the above is an identity. Denoting

$$f_2(z) = \frac{f_1(z) - \langle f_1, e_{a_1} \rangle e_{a_1}(z)}{\frac{z - a_1}{1 - \overline{a}_1 z}},$$

calling it the *reduced remainder,* the identity is re-written as

$$f(z) = \langle f_1, e_{a_1} \rangle e_{a_1}(z) + f_2(z) \frac{z - a_1}{1 - \overline{a}_1 z}. \tag{4}$$

We call the operator mapping f_1 to f_2 the *generalized a_1-backward shift operator* and f_2 *the generalized a_1-backward shift of* f_1. The terminology is a generalization of the classical backward shift operator

$$S(f)(z) = a_1 + a_2 z + \cdots + c_{k+1} z^k + \cdots = \frac{f(z) - f(0)}{z}.$$

Recognizing that $f(0) = \langle f, e_0 \rangle e_0(z)$, the operator S is the generalized 0-backward shift operator.

Notice that the Szegö kernel is the Cauchy kernel under the arc-length measure and thus has the reproducing kernel property. Due to the orthogonality in the Hilbert space and the modular one property of the Möbius transform, we have the energy relations

$$\|f\|^2 = \|\langle f_1, e_{a_1} \rangle e_{a_1}\|^2 + \|f_2\|^2 = (1 - |a_1|^2)|f_1(a_1)|^2 + \|f_2\|^2.$$

The purpose now is to extract the maximal energy portion from the term $\langle f_1, e_{a_1} \rangle e_{a_1}(z)$. It is reduced to maximize $(1 - |a_1|^2)|f_1(a_1)|^2$ among all $a_1 \in \mathbf{D}$. Although $\mathbf{D}$ is an open set we are fortunate that there exists a_1 in $\mathbf{D}$ such that

$$a_1 = \arg\max\{(1 - |a|^2)|f_1(a)|^2 : a \in \mathbf{D}\}$$

[84]. The existence of such maximal selection is called *Maximal Selection Principle*. Selecting such a_1 and repeating the process for f_2, and so on. We call the process from f_1 to get f_2 through a maximal selection a_1 as *maximal sifting from f_1 to f_2 through a_1*. After n siftings one gets

$$f(z) = \sum_{k=1}^{n} \langle f_k, e_{a_k} \rangle B_k(z) + f_{n+1} \prod_{k=1}^{n} \frac{z - a_k}{1 - \overline{a}_k z},$$

where for $k = 1, \ldots, n$,

$$a_k = \arg\max\{(1 - |a|^2)|f_k(a)|^2 : a \in \mathbf{D}\},$$

$$B_k(z) = B_{\{a_1, \ldots, a_k\}}(z) = \frac{\sqrt{1 - |a_k|^2}}{1 - \overline{a}_k z} \prod_{l=1}^{k-1} \frac{z - a_l}{1 - \overline{a}_l z},$$

and for $k = 2, \ldots, n+1$, f_k is the maximal sifting of f_{k-1} through a_{k-1}, that is,

$$f_k(z) = \frac{f_{k-1}(z) - \langle f_{k-1}, e_{a_{k-1}} \rangle e_{a_{k-1}}(z)}{\frac{z - a_{k-1}}{1 - \overline{a}_{k-1} z}}.$$

We have the following convergent theorem.

Theorem 3.4 For any given function f in the Hardy H^2 space, by applying the maximum sifting process at each step we have

$$f(z) = \sum_{k=1}^{\infty} \langle f_k, e_{a_k} \rangle B_k(z).$$

This result was first proved in [84] based on the complex modula 1 property of the Möbius transform. Below we provide a new proof releasing the modular 1 requirement for the system functions but only based on maximal selections of

the parameters. The essence of the proof is contained in several proofs of [73], also see [96].

Proof We prove the convergence by contradiction. Assume that through a sequence of maximally selected parameters $\mathbf{a} = \{a_1, \ldots, a_n, \ldots\}$ we arrive

$$f = \sum_{k=1}^{\infty} \langle f_k, e_{a_k} \rangle B_k + h, \qquad h \neq 0. \tag{5}$$

The routine argument by using the Riesz-Fisher Theorem shows that both the functions $\sum_{k=1}^{\infty} \langle f_k, e_{a_k} \rangle B_k$ and h are in H^2. We note that from the Hilbert space property h is orthogonal with all B_k, as well as with $\sum_{k=1}^{\infty} \langle f_k, e_{a_k} \rangle B_k$.

The relation (5) can be re-written

$$f = \left(\sum_{k=1}^{M} + \sum_{k=M+1}^{\infty} \right) \langle f_k, e_{a_k} \rangle B_k + h,$$

where by our notation,

$$g_{M+1} = \sum_{k=M+1}^{\infty} \langle f_k, e_{a_k} \rangle B_k + h = G_{M+1} + h.$$

To proceed we note that

$$\langle f_k, e_{a_k} \rangle = \langle f, B_k \rangle = \langle g_k, B_k \rangle, \tag{6}$$

where $g_k = f - \sum_{l=1}^{k-1} \langle f_l, e_{a_l} \rangle B_l$ is the k-th *standard remainder*.

Therefore, we have

$$g_{M+1} = \sum_{k=M+1}^{\infty} \langle g_k, B_k \rangle B_k + h.$$

Due to the density of the function set $\{e_a\}_{a \in \mathbf{D}}$ in H^2, there exists $a \in \mathbf{D}$ such that $\delta \triangleq |\langle h, e_a \rangle| > 0$. We can in particular choose a to be distinguished from all the selected a_k's. We are now to explore a contradiction in relation to the selections of a_{M+1} for large M. Now, on one hand, by the Bessel inequality applied to the infinite series part in Eq. (5), we have

$$|\langle g_{M+1}, B_{M+1} \rangle| \to 0, \qquad \text{as} \quad M \to 0. \tag{7}$$

On the other hand we will show, for large M,

$$|\langle g_{M+1}, B_{M+1}^a \rangle| > \frac{\delta}{2}. \tag{8}$$

This is then clearly a contradiction.

The rest part of the proof is devoted to showing Eq. (8). Due to the relations

$$|\langle g_{M+1}, B_{M+1}^a \rangle| \geqslant |\langle h, B_{M+1}^a \rangle| - |\langle G_{M+1}, B_{M+1}^a \rangle| \tag{9}$$

and

$$|\langle G_{M+1}, B_{M+1}^a \rangle| \leqslant \|G_{M+1}\| \to 0, \qquad \text{as} \quad M \to \infty, \tag{10}$$

for large M the lower bounds of $|\langle g_{M+1}, B_{M+1}^a \rangle|$ depend on the quantity of $|\langle h, B_{M+1}^a \rangle|$. Now for any positive integer M denote X_{M+1}^a the $(M+1)$-dimensional space spanned by $\{e_a, B_1, \ldots, B_M\}$. We have two ways to compute the energy of the projection of h into X_{M+1}^a, denoted $\|h/X_{M+1}^a\|^2$. One way is based on the orthonormality of $\{B_1, \ldots, B_M, B_{M+1}^a\}$. In such way, due to the orthogonality of h with $B_1, \ldots, B_M$, we have

$$\|h/X_{M+1}^a\|^2 = |\langle h, B_{M+1}^a \rangle|^2.$$

The other way is based on the orthonormalization of the $(M+1)$-tuple in the order $\{e_a, B_1, \ldots, B_M\}$. Then we have

$$\|h/X_{M+1}^a\|^2 \geqslant |\langle h, e_a \rangle|^2 = \delta^2.$$

Hence we have, for any M, $|\langle h, B_{M+1}^a \rangle| \geqslant \delta$. In view of this last estimation and (9), (10), we arrive at the contradiction given by (7), (8). The proof is complete.

Remark It is noted that the selected parameters $a_1, \ldots, a_n, \ldots$ according the maximal principle may not satisfy the hyperbolic non-separable condition

$$\sum_{k=1}^{\infty} (1 - |a_k|) = \infty$$

and thus the generated TM system $\{B_k\}$ may not be a basis. By doing such decomposition one is not interested in whether the resulted TM system is a basis, but only in whether it can effectively expand the given signal f. One is indeed able to do so, and, in fact, achieves fast convergence.

Remark For arbitrary selections of $a_1, \ldots, a_n, \ldots$, we arrive a pre-mono-component decomposition: All entries in the infinite sum after being multiplied by e^{it} become mono-components. If we choose $a_1 = 0$, then all B_k's are mono-components, and AFD offers a mono-component decomposition.

Remark AFD is different from any existing greedy algorithm [24] for the following reasons. (i) On the algorithm side AFD is incorporated with a generalized backward-shift operation, or a sifting process, that changes the standard

remainder to reduced remainder; (ii) The reduced remainder allows multiple selections of parameters for optimal approximation at each of the iteration steps; and (iii) The backward-shift operation automatically generates an orthonormal system, or a TM system without using the Gram-Schmidt process.

Remark Restricted to a practical subclass the convergence rate for AFD is $M/\sqrt{n}$ where n is the order of the AFD partial sum. One has to note that this is a good convergence rate for it is for non-smooth functions (boundary limits of Hardy space functions). The classical convergence theorems may look better but they are for smooth functions.

3.3   Unwinding AFD (UWAFD)

Let $f = hg$, where f, g are Hardy $H^2(\mathbf{D})$ functions, and h is an inner function. Let f and g be expanded into their respective Fourier series, viz.,

$$f(z) = \sum_{k=0}^{\infty} c_k z^k, \qquad g(z) = \sum_{k=0}^{\infty} d_k z^k.$$

The Plancherel Theorem and the modula 1 property of inner functions assert that

$$\sum_{k=0}^{\infty} |c_k|^2 = \|f\|^2 = \|g\|^2 = \sum_{k=0}^{\infty} |d_k|^2.$$

In digital signal processing (DSP) there is following result: for any n,

$$\sum_{k=n}^{\infty} |c_k|^2 \geqslant \sum_{k=n}^{\infty} |d_k|^2$$

(see, for instance [14, 26]).

In DSP this is referred to be energy front loading property of minimum phase signals among physically realizable signals. This amounts to saying that after factorizing out the inner function factor, the convergence rate of the Fourier series of the remaining outer function speeds up. This suggests that the AFD process would be better to incorporate a factorization process for speeding up the convergence. This is reasonable: when a signal by its nature is of high frequency, one should first perform "unwending" before extracting out from it a maximal portion of lower frequency. We proceed it as follows [71, 81]. First we do factorization $f = f_1 = I_1 O_1$, where I_1 and O_1 are, respectively, the inner and outer factors of f. The factorization is based on Nevanlinna's factorization theorem, also see [115]. The outer function has the explicit integral representation

$$O_1(z) = \mathrm{e}^{\frac{1}{2\pi} \int_0^{2\pi} \frac{\mathrm{e}^{it}+z}{\mathrm{e}^{it}-z} \log |f_1(\mathrm{e}^{it})| \mathrm{d}t}.$$

The boundary value of the outer function is computed by using the boundary value of f_1. On the boundary the above integral is taken to be of the principal integral sense. The imaginary part of the integral reduces to the circular Hilbert transform of $\log|f_1(e^{it})|$. Next, we do a maximum sifting to O_1. That gives

$$f(z) = I_1(z)\left[\langle O_1, e_{a_1}\rangle e_{a_1}(z) + f_2(z)\frac{z - a_1}{1 - \overline{a}_1 z}\right],$$

where f_2 is the maximal shifting of O_1 through a_1:

$$f_2(z) = \frac{O_1(z) - \langle O_1, e_{a_1}\rangle e_{a_1}(z)}{\frac{z - a_1}{1 - \overline{a}_1 z}}.$$

By factorizing f_2 into its its inner and outer factors, $f_2 = I_1 O_2$, we have

$$f(z) = I_1(z)\left[\langle O_1, e_{a_1}\rangle e_{a_1}(z) + I_2(z)O_2(z)\frac{z - a_1}{1 - \overline{a}_1 z}\right].$$

We next proceed a maximum sifting to O_2, and so on. In such way we arrive at the unwinding AFD decomposition [71]:

Theorem 3.5 *The above procedure gives rise to the unwinding AFD (UWAFD) decomposition*

$$f(z) = \sum_{k=1}^{n}\prod_{l=1}^{k} I_l(z)\langle O_k, e_{a_k}\rangle B_k(z) + f_{n+1}(z)\prod_{k=1}^{n}\frac{z - a_k}{1 - \overline{a}_k z}\prod_{l=1}^{n} I_l(z),$$

where $f_{k+1} = I_{k+1}O_{k+1}$ is the maximal shifting of O_k through $a_k, k = 1, \ldots, n$, and I_{k+1} and O_{k+1} are respectively the inner and outer functions of f_{k+1}. Furthermore,

$$f(z) = \sum_{k=1}^{\infty}\prod_{l=1}^{k} I_l(z)\langle O_k, e_{a_k}\rangle B_k(z).$$

Remark Like AFD, unwending AFD (UWAFD) is a mono-component or pre-mono-component decomposition. Experiments show that among various AFD type algorithms UWAFD seems to converge most rapidly, that is in particular on singular inner functions [81].

Remark If we do not incorporate the maximal sifting process as in UWAFD, the algorithm falls into UWA, as first developed in [63]. Below we denote by ϕ_k Blaschke products, with finite or infinite zeros, denote by ψ_k products of a singular inner function and an outer function, and denote by f_k H^2 functions, where $f = f_1, f_k(z) = \psi_{k-1}(z) - \psi_{k-1}(0), k = 2, \ldots, c_k = \psi_k(0), k = 1, 2, \ldots$. It proceeds as

$$\begin{aligned}
f(z) = f_1(z) &= \phi_1(z)\psi_1(z)\\
&= \phi_1(z)\left(\psi_1(z) - \psi_1(0) + \psi_1(0)\right)\\
&= c_1\phi_1(z) + \phi_1(z)f_2(z)\\
&= c_1\phi_1(z) + \phi_1(z)\phi_2(z)\left(\psi_2(z) - \psi_2(0) + \psi_1(0)\right)
\end{aligned}$$

$$= c_1\phi_1(z) + c_2\phi_1(z)\phi_2(z) + \phi_1(z)\phi_2(z)f_3(z)$$
$$= \cdots$$
$$= \sum_{k=1}^{\infty} c_k\phi_1(z) \cdots \phi_k(z).$$

The convergence in H^2 was first proved in Remark 4.4, [71], and several generalized convergence results were proved in [12]. In the recent paper [108] we studied the computation aspect of UWA.

3.4 Cyclic AFD for n-best rational approximation

In Core-AFD the parameters $a_1, \ldots, a_k, \ldots$ are selected one by one to construct an optimal sequence of Blaschke forms to approximate the given function

$$\sum_{k=1}^{n} \langle f, B_{\{a_1,\ldots,a_k\}} \rangle B_{\{a_1,\ldots,a_k\}}(z). \tag{11}$$

Now we change the question to the following: Given $f \in H^2(\mathbf{D})$ and a fixed positive integer n, find n parameters $a_1, \ldots, a_n$ such that the associated n-Blaschke form (11) best approximate f, that is

$$\|f - \sum_{k=1}^{n} \langle f, B_{\{a_1,\ldots,a_k\}} \rangle B_{\{a_1,\ldots,a_k\}}(z)\| \tag{12}$$

$$= \min\{\|f - \sum_{k=1}^{n} \langle f, B_{\{b_1,\ldots,b_k\}} \rangle B_{\{b_1,\ldots,b_k\}}(z)\| : \{b_1,\ldots,b_n\} \in \mathbf{D}^n\}. \tag{13}$$

This amounts an optimal but simultaneous selection of n parameters that is obviously better than selections in the one by one manner. Simultaneous selection of the parameters in an approximating n-Blaschke form is equivalent with the so called *optimal approximation by rational functions of order not larger than n*. The latter problem was phrased as *n-best rational approximation*. It has been a long standing open problem, presented as follows.

Let p and q denote polynomials of one complex variable. We say that (p, q) is an *n-pair* if p and q are co-prime, both of degrees less than or equal to n, and q does not have zero in the unit disc. Denote by $\mathscr{R}_n$ the set of all n-pairs. If $(p, q) \in \mathscr{R}_n$, then the rational function p/q is said to be a rational function of degree less or equal n. Let f be a function in the Hardy H^2 space in the unit disc. To find an n-best rational approximation to f is to find an n-pair $p_1, q_1)$ such that

$$\|f - p_1/q_1\| = \min\{\|f - p/q\| : (p, q) \in \mathscr{R}_n\}.$$

Existence of such a minimum solution was proved many decades ago [112], a practical algorithm to get a solution, however, has been an open problem till now. The best n-Blaschke form approximation is essentially equivalent with the n-best rational approximation. There is a separate proof for existence of the

minimum in Eq. (12) [85]. We wish to take the advantages of the Blaschke form to get an practical algorithm for the classical n-best rational approximation problem. By using *cyclic* AFD *algorithm* we can easily get a solution of the above mentioned problem if there is only one critical point for the objective function [76]. In general, cyclic AFD offers a *conditional solution* depending on the initial values to star with. Besides cyclic AFD there previously existed an algorithm, RARL2, by the French institute INRIA, that again can only get a conditional solution [4]. Both the theory and algorithm of cyclic AFD are explicit. It directly finds out the poles of the approximating rational function. The other rational approximation models mostly use the coefficients of p and q as parameters in order to set up and solve the problem. Using coefficients of polynomials, which is double amount of the parameter number of the best-n Blaschke form setting, involves tedious analysis and computation. The ultimate solution of the optimization problem lays on optimal selection of an initial status to start with. Finding an optimal initial status itself is an NP hard problem.

For any given natural number n the objective function for the n-Blaschke optimization problem is

$$A(f; a_1, \ldots, a_n) = \|f\|^2 - \sum_{k=1}^{n} |\langle f, B_k \rangle|^2. \tag{14}$$

Definition *An n-tuple $(a_1, \ldots, a_n)$ is said to be a coordinate-minimum point (CMP) of the objective function $A(f; z_1, \ldots, z_n)$ if for any chosen k among $1, \ldots, n$, whenever we fix the rest $n - 1$ variables, being $z_1 = a_1, \ldots, z_{k-1} = a_{k-1}, z_{k+1} = a_{k+1}, \ldots, z_n = a_n$, and select the kth variable z_k to minimize the objective function, we have*

$$a_k = \arg\min\{A(f; a_1, \ldots, a_{k-1}, z_k, a_{k+1}, \ldots, a_n) \colon z_k \in \mathbf{D}\}.$$

In the Core-AFD algorithm we progress the following procedure: For a $(k - 1)$-tuple $\{a_1, \ldots, a_{k-1}\}$ in $\mathbf{D}$ we produce the reduced remainders $f_2, \ldots, f_k$, and for f_k we apply the Maximal Selection Principle to find an a_k giving rise to $\max\{|\langle f_n, e_a \rangle| \colon a \in \mathbf{D}\}$. The proposed cyclic AFD algorithm repeats such procedure always for $k = n$: For any permutation P of $1, \ldots, n$, for the first $(n-1)$ parameters in the order $a_{P(1)}, \ldots, a_{P(n-1)}$ we produce the corresponding reduced remainders $f_2, \ldots, f_n$, and then use the Maximal Selection Principle to select a new and optimal $a_{P(n)}$.

The proposed cyclic AFD algorithm is validated in the following theorem.

Theorem 3.6 *Suppose that f is not an m-Blaschke form for any $m < n$. Let $s_0 = \{b_1^{(0)}, \ldots, b_n^{(0)}\}$ be any n-tuple of parameters inside $\mathbf{D}$. Fix some $n - 1$ parameters of s_0 and make an optimal selection of the single remaining parameter according to the Maximal Selection Principle based on the objective function (14). Denote the obtained new n-tuple of parameters by s_1. We repeat*

this process and make cyclic optimal selections over the n parameters. We thus obtain a sequence of n-tuples $s_0, s_1, \ldots, s_l, \ldots$, with decreasing objective function values d_l that tend to a limit $d \geqslant 0$, where, in the notation and formulation of (14),

$$d_l = A(f; b_1^{(l)}, \ldots, b_n^{(l)}) = \|f\|^2 - \sum_{k=1}^{n} (1 - |b_k^{(l)}|^2)|f_k^{(l)}(b_k^{(l)})|^2. \qquad (15)$$

Then, (i) If $\overline{s}$, as an n-tuple, is a limit of a subsequence of $\{s_l\}_{l=0}^{\infty}$, then $\overline{s}$ is in **D**; *(ii) $\overline{s}$ is a CMP of $A(f; \ldots)$; (iii) If the correspondence between a CMP and the corresponding value of $A(f; \ldots)$ is one to one, then the sequence $\{s_l\}_{l=0}^{\infty}$ itself converges to the CMP, being dependent of the initial n-tuple s_0; (iv) If $A(f; \ldots)$ has only one CMP, then $\{s_l\}_{l=0}^{\infty}$ converges to a limit $\overline{s}$ in* **D** *at which $A(f; \ldots)$ attains its global minimum value.*

We refer the reader to [76] for further details and examples of cyclic AFD. In a recent paper the algorithm is further developed [88].

3.5 Pre-Orthogonal adaptive Fourier decomposition (POAFD) for reproducing kernel Hilbert spaces

The theory and algorithm that will be developed in this section, as a matter of fact, can be extended to more general contexts. To explain just the idea we restrict ourselves to the simplest cases, including the weighted Bergman spaces and weighted Hardy spaces, etc. In the said simple setting the Hilbert space $\mathscr{H}$ consists of functions defined in an open connected region $\mathscr{E}$ (can be unbounded) in the complex plane, and the reproducing kernel k_a is a real-analytic function of the variable $\overline{a}$ in $\mathscr{E}$ satisfying the relation

$$f^{(l)}(a) = \langle f, \left(\frac{\partial}{\partial \overline{a}}\right)^l k_a \rangle, \quad l = 1, 2, \ldots. \qquad (16)$$

Let $\{a_1, \ldots, a_n, \ldots\}$ be a finite or infinite sequence. For a fixed n we define the multiple of a_n, denoted by $l(a_n)$, to be the repeating times of a_n in the n-tuple $\{a_1, \ldots, a_n\}$. With this definition, for instance, the multiple of a_1 is just 1, and the multiple of a_2 will depend on whether $a_2 = a_1$. If yes, then $l(a_2) = 2$, and, if not, $l(a_2) = 1$, and so on. Note that it is a little abuse of notation for it is not dependent on the value of a_n but on the repeating times of a_n in the corresponding n-tuple. We accordingly define

$$\tilde{k}_{a_n} \triangleq \left[\left(\frac{\partial}{\partial \overline{a}}\right)^{l(a_n)-1} k_a\right]_{a=a_n} \triangleq \left(\frac{\partial}{\partial \overline{a}}\right)^{l(a_n)-1} k_{a_n}. \qquad (17)$$

We further assume the following boundary vanishing condition, implying the Maximal Selection Principle in every individual context, as follows: Let $a_1, \ldots,$

a_{n-1} be previously given, and $\{B_1, \ldots, B_{n-1}\}$ be the Gram-Schmidt orthonormalization of $\{\tilde{k}_{a_1}, \ldots, \tilde{k}_{a_{n-1}}\}$, then for every $f \in \mathscr{H}$, the pre-orthogonal system has the property

$$\lim_{a \to \partial \mathscr{H}} \langle f, B_n^a \rangle = 0, \tag{18}$$

where $\{B_1, \ldots, B_{n-1}, B_n^a\}$ is the Gram-Schmidt orthonormalization of $\{\tilde{k}_{a_1}, \ldots, \tilde{k}_{a_{n-1}}, k_a\}$, with $a \neq a_k, k = 1, \ldots, n-1$. We note (4) if $a \to \partial \mathscr{H}$ then a is different from any $a_k, k = 1, \ldots, n-1$, when a gets close to the boundary; and (5) in any case the limit $a \to \partial \mathscr{H}$ is in the sense of the topology of the one-point-compactification of the complex plane while the "one point" takes to be ∞. With this boundary vanishing assumption we conclude the Maximal Selection Principle of POAFD: Under the assumption (18), through a compact argument using the Bolzano-Weierstrass theorem, there exists a sequence $\{b_j\}_{j=1}^\infty$ such that none of the b_j's take any values $a_1, \ldots, a_{n-1}$, and $\lim_{j \to \infty} b_j = a_n \in \mathscr{E}$, and

$$\lim_{j \to \infty} |\langle f, B_n^{b_j} \rangle| = \max\{|\langle f, B_n^a \rangle| \colon a \in \mathscr{H}\}. \tag{19}$$

Under those conditions we can prove the following lemma.

Lemma 3.7

$$\lim_{l \to \infty} B_n^{b_j} = B_n^{a_n},$$

where $\{B_1, \ldots, B_{n-1}, B_n^{a_n}\}$ is the Gram-Schmidt orthonormalization of $\{\tilde{k}_{a_1}, \ldots, \tilde{k}_{a_{n-1}}, \tilde{k}_{a_n}\}$.

Proof If a_n does not coincide with any $a_k, k = 1, \ldots, a_{n-1}$, then $\lim_{j \to \infty} B_n^{b_j} = B_n^{a_n}$, where $\{B_1, \ldots, B_{n-1}, B_n^{a_n}\}$ is the Gram-Schmidt orthonormalization of $\{\tilde{k}_{a_1}, \ldots, \tilde{k}_{a_{n-1}}, k_{a_n}\} = \{\tilde{k}_{a_1}, \ldots, \tilde{k}_{a_{n-1}}, \tilde{k}_{a_n}\}$. Now consider the case that a_n coincides with some of the earlier $a_1, \ldots, a_{n-1}$, or in other words, $l(a_n) > 1$. That means that, in the notation (17), the $(l-1)$ functions $k_{a_n}, \frac{\partial}{\partial \bar{a}} k_{a_n}, \ldots, \left(\frac{\partial}{\partial \bar{a}}\right)^{(l-2)} k_{a_n}$ have already appeared in the sequence $\{\tilde{k}_{a_1}, \ldots, \tilde{k}_{a_{n-1}}\}$. As a consequence, the function

$$T_{l-2}(b_j, a_n) = k_{a_n} + \frac{\frac{\partial}{\partial \bar{a}} k_{a_n}}{1!}(\overline{b_j} - \overline{a}_n) + \cdots + \frac{\left(\frac{\partial}{\partial \bar{a}}\right)^{(l-2)} k_{a_n}}{(l-2)!}(\overline{b_j} - \overline{a}_n)^{l-2},$$

as the order-$(l-2)$ Taylor expansion of the function $k_a(z)$ in $\overline{b}_j$ about $\overline{a}_n$, is already in the linear span of $B_1, \ldots, B_{n-1}$. This last atatement amounts to the relation

$$T_{l-2}(b_j, a_n) - \sum_{k=1}^{n} \langle T_{l-2}(b_j, a_n), B_k \rangle B_k = 0. \tag{20}$$

Since b_j's are different from all $a_k, k = 1, \ldots, n$, we have,

$$B_n^{b_j}(z) = \frac{k_{b_j}(z) - \sum_{k=1}^{n-1}\langle k_{b_j}, B_k\rangle B_k(z)}{\|k_{b_j} - \sum_{k=1}^{n-1}\langle k_{b_j}, B_k\rangle B_k\|}. \tag{21}$$

Inserting (20) into (21), and dividing by $(\overline{b_j} - \overline{a}_n)^{l-1}$ and $|\overline{b_j} - \overline{a}_n|^{l-1}$ to the numerator and the denominator parts, respectively, we have

$$B_n^{b_j}(z) = e^{-(l-1)\theta} \frac{\frac{k_{b_j}(z) - T_{l-2}(k_{b_j}, a_n)(z)}{(\overline{b_j} - \overline{a}_n)^{l-1}} - \sum_{k=1}^{n-1}\left\langle \frac{k_{b_j} - T_{l-2}(b_j, a_n)(z)}{\overline{w} - \overline{a}_n)^{l-1}}, B_k\right\rangle B_k(z)}{\left\| \frac{k_{b_j}(z) - T_{l-2}(b_j, a_n)(z)}{(\overline{b_j} - \overline{a}_n)^{l-1}} - \sum_{k=1}^{n-1}\left\langle \frac{k_{b_j}(z) - T_{l-2}(b_j, a_n)(z)}{(\overline{b_j} - \overline{a}_n)^{l-1}}, B_k\right\rangle B_k\right\|}, \tag{22}$$

where $e^{i\theta}$ is the tangential direction of the limiting $b_j \to a_n$. We can, in fact, take any direction, including $\theta = 0$. Letting $b_j \to a_n$ with $\theta = 0$, and using the Lagrange type remainder of the Taylor expansion, we obtain

$$\lim_{j\to\infty} B_n^{b_j}(z) = \frac{\tilde{k}_{a_n}(z) - \sum_{k=1}^{n-1}\langle \tilde{k}_{a_n}, B_k\rangle B_k(z)}{\|\tilde{k}_{a_n} - \sum_{k=1}^{n-1}\langle \tilde{k}_{a_n}, B_k\rangle B_k\|}.$$

Therefore, $\{B_1, \ldots, B_{n-1}, B_n^{a_n}\}$ is the Gram-Schmidt orthonormalization of $\{\tilde{k}_{a_1}, \ldots, \tilde{k}_{a_{n-1}}, \tilde{k}_{a_n}\}$, as desired.

Remark The essence of the proof is contained in [72] and [73]. The proof for the one complex variable case as presented here first appears in [96].

We have the pre-orthogonal adaptive Fourier decomposition (POAFD) convergence theorem as follows.

Theorem 3.8 *Selecting $\{a_1, \ldots, a_n, \ldots\}$ according to the Maximal Selection Principle set by (19), we have*

$$f = \sum_{k=1}^{\infty}\langle f, B_n\rangle B_n,$$

where for any positive integer n, $\{B_1, \ldots, B_{n-1}, B_n\}$ is the Gram-Schmidt orthonormalization of $\{\tilde{k}_{a_1}, \ldots, \tilde{k}_{a_{n-1}}, \tilde{k}_{a_n}\}$.

One can adopt the same proof for the AFD convergence (Theorem 3.4) in which only the standard remainders g_k's are concerned. As a matter of fact, the sifting process and the role of the induced remainders are taken place by the pre-orthogonal process.

Remark For Repeating selection of parameters and POAFD we refer the reader to the references [72, 73, 85, 90]. was previously called POGA or PreOGA, etc.
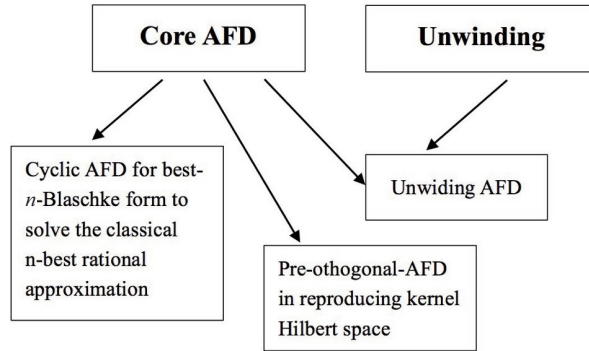
Fig. 3   Various Adaptive Fourier Decomposition Methods

The previous names do not reflect the crucial role of the complex holomorphic function methods (Fig. 3).

4   Related studies and applications

4.1   Aspects in relation to beurling-lax shift-invariant subspaces

The AFD type expansions is, in a great extent, related to the Beurling-Lax shift-invariant subspaces of the Hardy H^2 spaces. In the unit disc case,

$$H^2(\mathbf{D}) = \overline{\mathrm{span}}\{B_k\}_{k=1}^{\infty} \oplus \phi H^2(\mathbf{D}), \tag{23}$$

where $\{B_k\}_{k=1}^{\infty}$ is the TM system generated by a sequence $\{a_1, \ldots, a_n, \ldots\}$, where multiples are counted, and ϕ is the Blaschke product with the zeros $\{a_1, \ldots, a_n, \ldots\}$ including the multiples. Note that when ϕ can be defined with the a_k's as its zeros, there holds the condition

$$\sum_{k=1}^{\infty}(1 - |a_k|) < \infty,$$

and the associated TM system is not a basis. Although this has been well know its relations with adaptive expansions, as far as being aware by the author, are for the first time explored. Lately TM systems being Schauder system were proved [77]. The space decomposition relation (23) was extended to H^p spaces, where $p \neq 2$. Relations between backward shift invariant subspaces and bandlimited functions and Bedrosian identity [83] were studied. There are open problems such as whether for $p \neq 2$ there exist adaptive and fast converging expansions by using TM systems, and for $p = 2$ how far one can extend Eq. (23) to higher dimensions. The study has a great room to be further developed.

4.2 Extra-strong uncertainty principle

The phase and frequency studies in the mono-component function theory lay necessary foundations for digital signal processing. In the related studies what is called extra-strong uncertainty principle

$$\sigma_t^2 \sigma_\omega^2 \geqslant \frac{1}{4} + \left(\int_{-\infty}^{\infty} |t - \langle t \rangle| |\phi(t) - \langle \omega \rangle| |f(t)|^2 \mathrm{d}t \right)^2 \tag{24}$$

was established, where f is a real-valued signal, σ_t^2 and σ_ω^2 are the standard deviations with respect to, respectively, the time and the Fourier frequency, and $\langle t \rangle$ and $\langle \omega \rangle$ are the corresponding means [20]. An uncertainty principle of the same type was given by L. Cohen

$$\sigma_t^2 \sigma_\omega^2 \geqslant \frac{1}{4} + | \int_{-\infty}^{\infty} (t - \langle t \rangle)(\phi(t) - \langle \omega \rangle) |f(t)|^2 \mathrm{d}t|^2,$$

which is obliviously weaker. We further extend the above result to multi-dimensional contexts [20, 21, 28, 29, 30].

4.3 The dirac-type time-frequency distributions based on mono-component decompositions

The Dirac type time-frequency distribution (D-TFD) of the form

$$P(t, \omega) = \rho^2(t)\delta(\omega - \theta'(t)) \tag{25}$$

has been ultimate destiny of signal analysts. There have existed several time-frequency distributions, including windowed Fourier transform and Wagner-Ville transform, etc., of which none are satisfied. The existing time-frequency distributions do not give explicit and clear frequency components, and often depend on parameter selections. Positive-frequency decompositions of signals offered by the AFD decompositions naturally give rise to Dirac-type time-frequency distributions. For a single mono-component $m_1(t) = \rho_1(t) \cos \theta_1(t)$ the corresponding D-TFD according to Eq. (25) is the graph $(t, \theta_1'(t))$ of the function $\omega = \theta_1'(t)$ in the ω-t plane, while the weight $\rho_1^2(t)$ may be represented by colors continuously changing along with changing of the values $\rho_1^2(t)$. If a signal f is expanded into a series consisting of its "intrinsic composing" mono-components, then its D-TFD is the bunch of color-weighted graphs of which each is made from a composing mono-component [27, 125]. This definition has been interested and being paid attention by prominent signal analysts including Leon Cohen and Lorenzo Galleani, etc., and has been used in practice (see below the application section).

4.4 Higher dimensional AFDs

To develop an AFD like approximation theory in higher dimensions a Cauchy type structure is necessary, that is mainly for use of the reproducing kernel property in deducing Maximal Selection Principle in the underlying space. By using the Cauchy structure in Clifford algebra or in several complex variables

we achieved the AFD type theories and algorithms for functions of several real variables on the plane (Clifford Hardy spaces and Hardy spaces on tubes), and on the real spheres, and of several complex variables on the n-torus, and on the n-complex spheres [90, 91, 102, 114]. With Alpay, Colombo, Sabadini we achieved analogous theory involving matrix valued Blaschke products [1, 2]. This study has impact in general to rational approximation in a number of spaces [6].

4.5 Fourier spectrum characterization of Hardy spaces: analytic signals revised

The Paley-Wiener Theorem for the classical Hardy H^2 space over the upper-half complex space addresses the fact that if $f \in L^2(\mathbf{R})$, then further $f \in H^2(\mathbf{C}^+)$ if and only if $\mathrm{supp}\hat{f} \subset [0, \infty)$. This result is systematically extended to $H^p(\mathbf{C}^+)$ for all $p \in [1, \infty]$, where the Fourier transform can be defined at least in the distribution sense [67, 93]. The results are summarized as: Letting $f \in L^p(\mathbf{R})$, then f is further the non-tangential boundary limit of some function in the complex Hardy space $H^p(\mathbf{C}^+)$ if and only if $\hat{f} = \chi_+ \hat{f}$, where is the indicator (characteristic) function of the right-half-real line, and the Fourier transform may take the distribution sense. The generalization to the Hardy spaces on tubes (extending the $p = 2$ case in [104] to $1 \leqslant p \leqslant \infty$) was published in [47]. The results of the same type but with the Clifford algebra setting is now on line [25] proving that a Clifford-valued function $f \in L^p(\mathbf{R}^n)$ is the non-tangential boundary limit of some Clifford-valued Hardy space function in the upper-half space if and only if $\hat{f} = \chi_+ \hat{f}$, where $\chi_+(\underline{\xi}) = \frac{1}{2}\left(1 + i\frac{\underline{\xi}}{|\underline{\xi}|}\right), \underline{\xi} = \xi_1 e_1 + \cdots + \xi_n e_n$ (the Hardy space projection function). This extends some partial cases proved in [104] for conjugate harmonic systems. In various contexts Fourier spectrum characterizations give rise to Hardy spaces decompositions for L^p, $1 < p \leqslant \infty$ that further induce Hardy space decompositions of L^p, $0 < p < 1$ [31, 48]. Hardy space decomposition is the strategy that we have been using to study Lebesgue spaces of various integrability. The strategy is extensively implemented along with the mono-component and AFD approximation theories. In particular, for any signal f, by multiplying $\hat{f}$ with χ_+ and then taking the inverse Fourier transform, we obtain the associated analytic signal. This is philosophically valid in any context. We finally note that the Hardy space decomposition issue has been extended to the L^p-vector fields and one obtains the Hardy-Hodge decomposition [5].

4.6 Hilbert transforms as singular integral operators: analytic signals revised

As in the 1-D case [7], on higher dimensional manifolds one defines the non-scalar part of the non-tangential boundary limit of a hyper-complex holomorphic function to be the Hilbert transform of the scalar part of it [94]. Hilbert transform therefore is a particular singular integral. It is, in particular, not the singular Cauchy transform. One must study singular integrals to understand Hilbert transform. On one dimensional manifolds, including Lipechitz perturbations of the real line and the circle, certain singular integrals of holomorphic

kernels form an operator algebra as studied in a series of work of McIntosh, Li, Semmes, Qian, Long and Wang [37, 55, 56, 66]. The theory on the plane was earlier established in the work or under the influence of A. McIntosh [36, 51, 52]. Through first generalizing the results of Fueter and Sce to arbitrary Euclidean spaces (as technical necessity) the author established the theory of the operator algebra of singular integrals of Clifford monogenic kernels on Lipschitz perturbations of the unit sphere for any dimension [75]. Based on the established singular integral theory Hilbert transformations of the plane and of the sphere became well understood. Analytic signals on the sphere, for instance, are constructed as follows (Fig. 4). Let f be a real-valued signal of finite energy on a manifold $\mathscr{S}$. Denote by $H_{\mathscr{S}}$ the Hilbert transform of f on the manifold. Then the analytic signal on $\mathscr{S}$ is defined to be $f^+ = f + H_{\mathscr{S}}f$, where $H_{\mathscr{S}}f$ is the non-scalar part (on sphere it is a 2-form valued function). f^+ has Clifford monogenic extension to one of the two regions divided by $\mathscr{S}$. One can derive, if ζ is on the plane or on the sphere,

$$f^+(\zeta) = \rho_f(\zeta) \left(\frac{f(\zeta)}{\rho_f(\zeta)} + \frac{H_{\mathscr{S}}(f)(\zeta)}{|H_{\mathscr{S}}(f)(\zeta)|} \frac{|H_{\mathscr{S}}(f)(\zeta)|}{\rho_f(\zeta)} \right) \tag{26}$$

$$= \rho_f(\zeta) \left(\cos\theta(\zeta) + \frac{H_{\mathscr{S}}(f)(\zeta)}{|H_{\mathscr{S}}(f)(\zeta)|} \sin\theta(\zeta) \right) \tag{27}$$

$$= \rho_f(\zeta) e^{\frac{H_{\mathscr{S}}(f)(\zeta)}{|H_{\mathscr{S}}(f)(\zeta)|}\theta(\zeta)}, \tag{28}$$

where $\rho_f(\zeta) = \sqrt{|f(\zeta)|^2 + |H_{\mathscr{S}}f(\zeta)|^2}$, and $\left(\frac{H_{\mathscr{S}}(f)(\zeta)}{|H_{\mathscr{S}}(f)(\zeta)|} \right)^2 = -1$, the latter being a varying imaginary element just like the complex imaginary element with the property $i^2 = -1$. The instantaneous frequency is defined, as in the classical case through the monogenic continuation, but formally read

$$\theta'(\zeta) = \mathrm{Re}\{ [(\Gamma_\zeta - I)f^+(\zeta)] [(f^+(\zeta))^{-1}] \},$$

the latter can be expressed in terms of the angle $\theta(\zeta)$ [94, 121], where Γ_ζ is the surface Dirac operator on the manifold. In such format the basic idea of IF and the related approximation in higher dimensions make sense. The related studies published, in respectively 2015 and 2017 with Science Press two monographs books [73] and [74]. We finally note that Hilbert transformation may be characterize by commutativity with the affine groups in the underlying symmetric manifold which shows that the three objects the Hilbert transformation, the Dirac differential operator and the group representation theory have intimate relations [22, 23].

4.7 Applications

AFD has demonstrative applications in system identification and signal analysis. Applications in system identification include [17, 59, 60, 61]. A number of signal analysts promoted the AFD methods. Below we summarize part of the applications found in the literature.
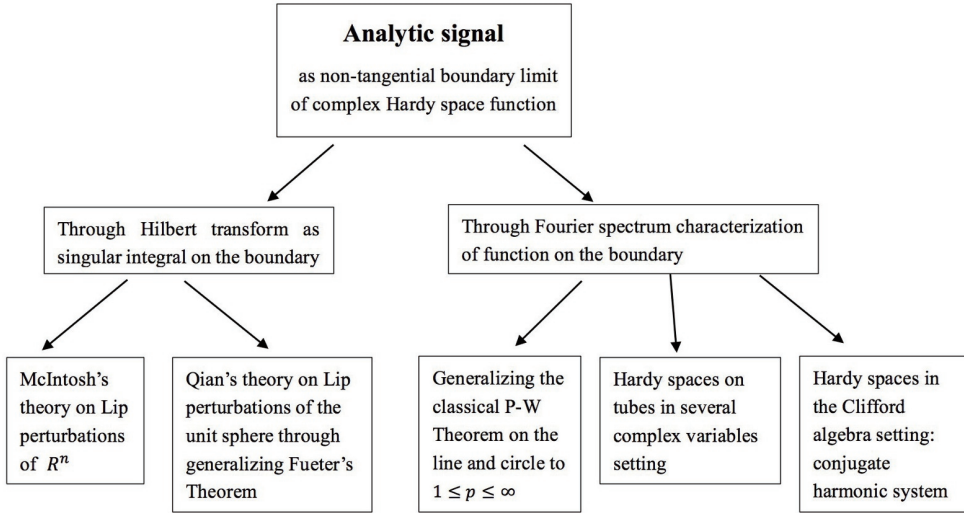
Fig. 4   Analytic signals in various contexts

It is commented in [116] that as a new method AFD was proposed in the recent years that could be used to decompose and reconstruct signals. It contains the classical Fourier method as a particular case. Experiments show that the 1D AFDs achieve excellent signal decomposition and reconstruction results. The article [W1] compares 2D AFD with the traditional frequency digital watermark methods, including discrete cosine transform DCT, discrete wavelet transform DWT, discrete Fourier transform DFT, etc., and concludes that 2D AFD has better transparency and robustness under attacking. The article [W2] revises the 2D AFD algorithm and, as a result, increases its speed, and uses it in denoising.

In [50] Liang et al. at Beijing Jiaotong University propose a new fault diagnosis method of rolling bearing based on AFD. They show that AFD can avoid using band-pass filters, the latter often suffering from the difficulty of algorithm parameter selection, they show that AFD adaptively, efficiently, and accurately diagnose all kinds of rolling bearing problems.

In [117] the authors study interference and separation between the lung sound (LS) and the heart sound (HS) signals. Due to the overlap in their frequency spectra, it is difficult to separate them. The article proposes a novel separation method based on AFD. This AFD-based separation method is validated on real HS signals from the University of Michigan Heart Sound and Murmur Library, as well as real LS signals from the 3M repository. Simulation results indicate that the proposed method is more effective than the extraction methods based on the recursive least square (RLS), than the standard empirical mode decomposition (EMD) and its various extensions, including the ensemble EMD (EEMD), the multivariate EMD (M-EMD) and the noise assisted M-EMD (NAM-EMD).

Over the years people have made unremitting studies in predicting the stock

price movements. In [124] a novel automatic stock movement forecasting system is proposed, which is based on the newly developed signal decomposition approach — adaptive Fourier decomposition (AFD). AFD can effectively extract the signal primary trend, which is specifically suitable in the Dow Theory based automatic technique analysis. Effectiveness of the proposed approach is assessed through the comparison with the direct BP approach and manual observation. The result is proved to be promising.

In [123] an AFD based time-frequency speech analysis approach is proposed. Given the fact that the fundamental frequency of speech signals often undergo fluctuation, the classical short-time Fourier transform (STFT) based spectrogram analysis suffers from the difficulty of window size selection. AFD is a newly developed signal decomposition theory. The outstanding characteristic of AFD is to provide instantaneous frequency for each decomposed component, so the time-frequency analysis becomes accessible. Experiments are conducted based on the sample sentence in TIMIT Acoustic-Phonetic Continuous Speech Corpus. The results show that the AFD based time-frequency distribution outperforms the STFT.

AFD has already been employed to the productions of CASA Environmental Technology Co., Ltd, including the second generation of BEWs (Biological Early Warning System) and ETBEs (Ecological Toxicity Biological Exposed System). The two systems takes advantages of AFD and Unwinding to analyze the biological behavioral signal. Compared with the traditional Fourier, wavelet and EMD algorithms, the AFD approach can efficiently solve early warning judgment for low concentration pollutants and disturbance of fish biological clock and other problems.

In control theory the the authors of the article [49] introduce an AFD algorithm to eliminate the channel noise superimposed on the output signal in the wireless transmission process. In the frequency domain, based on AFD, an ILC method for discrete linear system with wireless transmission is proposed. Simulation results show that the AFD algorithm is able to achieve signal denoising well in the case of small decomposition threshold compared with Fourier decomposition. Thus the goal that the output signal of ILC system can track the desired signal is better achieved.

Indian researchers in their article [42] assert that to analyze biomedical signals in relation to e-health devices the frequency domain method outperforms the time domain method, and among numerate frequency domain methods (Hermit, Fourier, Karhunen-Loeve, Wavelet) AFD appears to have features of a greater variety, and more stable for the data compression. Based on compression using AFD they started to manufacture economic, accurate and stable domestic e-health devices.

Apart from China and Asia, AFD has also achieved international influence. Interests, studies and applications of AFD are found in relevant literature, by Ph.D. thesis of Fulle at Michigan University on oxygenic photosynthesis; by Kirkbas et al. on optimal basis pursuit based on jaya optimization for adaptive Fourier decomposition [45]; by Vatchev, on a class of intrinsic trigonometric

mode polynomials [110]; by Mashreghi et al. on Blaschke Products and Applications [57]; by Krausshar et al. on Clifford and harmonic analysis on cylinders and tori [46]; by Colombo et al. on the Fueter mapping theorem in integral form and the $\mathscr{F}$-functional calculus [19]; by Falcão et al. on remarks on the generation of monogenic functions; by Colombo et al. on the Fueter primitive of bi-axially monogenic functions [18]; by Salomon on analysis of the anisotropy in image textures [97]; by Sakaguchi on the related integral-type method in higher order differential equations [98, 99, 100, 101]; by León on instantaneous frequency estimation and representation of the audio signal through complex wavelet additive synthesis [9]; by Mozes on computing the instantaneous frequency for an ECG signal [62]; by Gomes, as Doctoral dissertation, on compressive sensing in Clifford analysis; by Eisner et al. on discrete orthogonality of the Malmquist Takenaka system on the upper half plane and rational approximation [32]; and by Perotti on his article in directional quaternionic Hilbert operators [64].

References

1. Alpay D, Colombo F, Qian T, Sabadini I. Adaptive orthonormal systems for matrix-valued functions. Proceedings of the American Mathematical Society, 2017, 145(5): 2089–2106

2. Alpay D, Colombo F, Qian T, Sabadini I. Adaptive Decomposition: The Case of the Drury-Arveson Space. Journal of Fourier Analysis and Applications, 2017, 23(6), 1426–1444

3. Axelsson A, Kou K I, Qian T. Hilbert transforms and the Cauchy integral in Euclidean space. Studia Mathematica, 2009, 193(2): 161–187

4. Baratchart L, Cardelli M, Olivi M. Identification and rational L^2 approximation, a gradient algorithm. Automatica, 1991, 27: 413–418

5. Baratchart L, Dang P, Qian T. Hardy-Hodge Decomposition of Vector Fields in R^n. Transactions of the American Mathematical Society, 2018, 370: 2005–2022

6. Baratchart L, Mai W X, Qian T. Greedy Algorithms and Rational Approximation in One and Several Variables. In: Bernstein S., Kaehler U., Sabadini I., Sommen F. (eds) Modern Trends in Hypercomplex Analysis. Trends in Mathematics, 2016: 19–33

7. Bell S. The Cauchy Transform, Potential theory and Conformal Mappings. CRC Press, Boca, Raton, 1992

8. Boashash B. Estimating and interpreting the instantaneous frequency of a signal-Part 1: Fundamentals. Proceedings of The IEEE, 1992, 80(4): 520–538

9. Beltrán J R, de León P. Instantaneous frequency estimation and representation of the audio signal through Complex Wavelet Additive Synthesis. International Journal of Wavelets, Multiresolution and Information Processing, 2014, 12(03): 1450030

10. Cheng M T, Deng G T. Lecture notes on harmonic analysis. Peking University, 1979

11. Coifman R, Peyriére J. Phase unwinding, or invariant subspace decompositions of Hardy spaces. Journal of Fourier Analysis and Applications, 2019, 25: 684–695

12. Coifman R, Steinerberger S. Nonlinear phase unwinding of functions. J Fourier Anal Appl, 2017, 23: 778–809

13. Coifman R, Steinerberger S, Wu H T. Carrier frequencies, holomorphy and unwinding. SIAM J. Math. Anal., 2017, 49(6): 4838–4864

14. Cheng Q S. Digital Signal Processing. Peking University Press, 2003, in Chinese

15. Chen Q H, Qian T, Tan L H. Constructive Proof of Beurling-Lax Theorem, Chin. Ann. of Math., 2015, 36: 141–146

16. Cohen L. Time-Frequency Analysis: Theory and Applications. Prentice Hall, 1995

17. Chen Q H, Mai W X, Zhang L M, Mi W. System identification by discrete rational atoms. Automatica, 2015, 56: 53–59

18. Colombo F, Sabadini I, Sommen F. The Fueter primitive of bi-axially monogenic functions. Communications on Pure and Applied Analysis, 2014, 13(2): 657–672

19. Colombo F, Sabadini I, Sommen F. The Fueter mapping theorem in integral form and the $\mathscr{F}$-functional calculus. Mathematical Methods in the Applied Sciences, 2010, 33(17): 2050–2066

20. Dang P, Deng G T, Qian T. A Sharper Uncertainty principle. Journal of Functional Analysis, 2013, 265(10): 2239–2266

21. Dang P, Deng G T, Qian T. A Tighter Uncertainty Principle For Linear Canonical Transform in Terms of Phase Derivative. IEEE Transactions on Signal Processing, 2013, 61(21): 5153–5164

22. Dang P, Liu H, Qian T. Hilbert Transformation and Representation of $ax + b$ Group. Canadian Mathematical Bulletin, 2018, 61(1): 70–84

23. Dang P, Liu H, Qian T. Hilbert Transformation and $r\mathrm{Spin}(n)+\mathbf{R}^n$ Group. arXiv:1711.04519v1[math.CV], 2017

24. Davis G, Mallet S, Avellaneda M. Adaptive Greedy Approximations, Constr. Approxi., 1997, 13: 57–98

25. Dang P, Mai W X, Qian T. Fourier Spectrum Characterizations of Clifford H^p Spaces on R_+^{n+1} for $1 \leqslant p \leqslant \infty$. Journal of Mathematical Analysis and Applications, 2020, 483: 123598

26. Dang P, Qian T. Analytic Phase Derivatives, All-Pass Filters and Signals of Minimum Phase. IEEE Transactions on Signal Processing, 2011, 59(10): 4708–4718

27. Dang P, Qian T. Transient Time-Frequency Distribution based on Mono-component Decompositions. International Journal of Wavelets, Multiresolution and Information Processing, 2013, 11(3): 1350022

28. Dang P, Qian T, Chen Q H. Uncertainty Principle and Phase Amplitude Analysis of Signals on the Unit Sphere. Advances in Applied Clifford Algebras, 2017, 27(4): 2985–3013

29. Dang P, Qian T, You Z. Hardy-Sobolev spaces decomposition and applications in signal analysis. J. Fourier Anal. Appl., 2011. 17(1): 36–64

30. Dang P, Qian T, Yang Y. Extra-strong uncertainty principles in relation to phase derivative for signals in Euclidean spaces. Journal of Mathematical Analysis and Applications, 2016, 437(2): 912–940

31. Deng G T, Qian T. Rational approximation of Functions in Hardy Spaces. Complex Analysis and Operator Theory, 2016, 10(5): 903–920

32. Eisner T, Pap M. Discrete orthogonality of the Malmquist Takenaka system of the upper half plane and rational interpolation. Journal of Fourier Analysis and Applications, 2014, 20(1): 1–16

33. Falcão M I, Cruz J F, Malonek H R. Remarks on the generation of monogenic functions. International Conference on the Applications of Computer Science and Mathematics in Architecture and Civil Engineering, 17, Weimar, 2006

34. Fulcheri P, Olivi P. Matrix rational H^2 approximation: a gradient algorithm based on schur analysis. SIAM I. Control Optim., 1998, 36(6): 2103–2127

35. Gabor D. Theory of communication. J. IEE., 1946, 93: 429–457

36. Gaudry G I, Long R, Qian T. A Martingale proof of L^2-boundednessof Clifford-Valued Singular Integrals. Annali di Mathematica Pura Ed Applicata, 1993, 165: 369–394

37. Gaudry G, Qian T, Wang S L, Boundedness of singular integrals with holomorphic kernels on star-shaped closed Lipschitz curves. Colloquium Mathematicum, 1996: 133–150

38. Garnett J B. Bounded Analyic Functions. Academic Press, 1981

39. Gomes N R. Compressive sensing in Clifford analysis. Doctoral Dissertation, Universidade de Aveiro (Portugal), 2015

40. Gong S. Private commimication, 2002

41. Gorusin G M. Geometrical Theory of Functions of One Complex Variable. translated by Jian-Gong Chen, 1956

42. Ganta P, Manu G, Anil Sooram S. New Perspective for Health Monitoring System. International Journal of Ethics in Engineering and Management Education, 2016

43. Hummel J A. Multivalent starlike function. J. d' analyse Math., 1967, 18: 133–160

44. Huang N E. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proc. R. Soc. Lndon, 1998: 903–995

45. Kirkbas A, Kizilkaya A, Bogar E. Optimal basis pursuit based on jaya optimization for adaptive fourier decomposition. Telecommunications and Signal Processing, 2017 40th International Conference on IEEE: 538–543

46. Krausshar R S, Ryan J. Clifford and harmonic analysis on cylinders and tori. Revista Matematica Iberoamericana, 2005, 21(1): 87–110

47. Li H C, Deng G T, Qian T. Fourier Spectrum Characterizations of H^p Spaces on Tubes Over Cones for $1 \leqslant p \leqslant \infty$. Complex Analysis and Operator Theory, 2018, 12: 1193–1218

48. Li H C, Deng G T, Qian T. Hardy space decomposition of on the unit circle: 0¡p¡1. Complex Variables and Elliptic Equations: An International Journal, 2016, 61(4): 510–523

49. Lei Y, Fang Y, Zhang L M. Iterative learning control for discrete linear system with wireless transmission based on adaptive fourier decomposition. Control Conference (CCC), 2017 36th Chinese IEEE

50. Liang Y, Jia L M, Cai G. A new approach to diagnose rolling bearing faults based on AFD. Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation-Volume II, Springer

51. Li C, McIntosh A, Qian T. Clifford algebras, Fourier transforms, and singular Convolution operators on Lipschitz surfaces. Revista Matematica Iberoamericana, 1994, 10(3): 665–695

52. Li C, McIntosh A, Semmes S. Convolution Singular Integrals on Lipschitz Surfaces. Journal of the American Mathematical Society, 1992: 455–481

53. Li S, Qian T, Mai W X. Sparse Reconstruction of Hardy Signal And Applications to Time-Frequency Distribution. International Journal of Wavelets, Multiresolution and Information Processing, 2013

54. Lyzzaik A. On a conjecture of M.S. Robertson. Proc. Am. Math. Soc., 1984, 91: 108–210

55. McIntosh A, Qian T. Convolution singular integrals on Lipschitz curves. Springer-Verlag, Lecture Notes in Maths, 1991, 1494: 142–162

56. McIntosh A, Qian T. L^p Fourier multipliers along Lipschitz curves. Transactions of The American Mathematical Society, 1992, 333(1): 157–176

57. Mashreghi J, Fricain E. Blaschke products and their applications. Springer, 2013

58. Mai W X, Qian T, Saitoh S. Adaptive Decomposition of Functions with Reproducing Kernels. in preparation

59. Mi W, Qian T. Frequency Domain Identification: An Algorithm Based On Adaptive Rational Orthogonal System. Automatica, 2012, 48(6): 1154–1162

60. Mo Y, Qian T, Mi W. Sparse Representation in Szego Kernels through Reproducing Kernel Hilbert Space Theory with Applications. International Journal of Wavelet. Multiresolution and Information Processing, 2015, 13(4): 1550030

61. Mi W, Qian T, Wan F. A Fast Adaptive Model Reduction Method Based on Takenaka-Malmquist Systems. Systems and Control Letters, 2012, 61(1): 223–230

62. Mozes F E, Szalai J. Computing the instantaneous frequency for an ECG signal. Scientific Bulletin of the "Petru Maior" University of Targu Mures, 2012, 9(2): 28

63. Nahon M. Phase Evaluation and Segmentation. Ph.D. Thesis, Yale University, 2000

64. Perotti A. Directional quaternionic Hilbert operators. Hypercomplex analysis, Birkhüser Basel, 2008: 235–258

65. Picinbono B. On instantaneous amplitude and phase of signals. IEEE Transactions on Signal Processing, 1997, 45(3): 552–560

66. Qian T. Singular integrals with holomorphic kernels and Fourier multipliers on star-shape Lipschitz curves. Studia Mathematica, 1997, 123(3): 195–216

67. Qian T. Characterization of boundary values of functions in Hardy spaces with applications in signal analysis. Journal of Integral Equations and Applications, 2005, 17(2): 159–198

68. Qian T. Analytic Signals and Harmonic Measures. Journal of Mathematical Analysis and Applications, 2006, 314(2): 526–536

69. Qian T. Mono-components for decomposition of signals. Mathematical Methods in the Applied Sciences, 2006, 29(10): 1187–1198

70. Qian T. Boundary Derivatives of the Phases of Inner and Outer Functions and Applications. Mathematical Methods in the Applied Sciences, 2009, 32: 253–263

71. Qian T. Intrinsic mono-component decomposition of functions: An advance of Fourier theory. Mathematical Methods in Applied Sciences, 2010, 33: 880–891

72. Qian T. Two-Dimensional Adaptive Fourier Decomposition. Mathematical Methods in the Applied Sciences, 2016, 39(10): 2431–2448

73. Qian T. Adaptive Fourier Decomposition: A Mathematical Method Through Complex Analysis. Harmonic Analysis and Signal Analysis, Science Press (in Chinese), 2015

74. Qian T, Li P T. Singular Integrals and Fourier Theory. Science Press (in Chinese), 2017

75. Qian T. Fourier analysis on starlike Lipschitz surfaces. Journal of Functional Analysis, 2001, 183: 370–412

76. Qian T. Cyclic AFD Algorithm for Best Approximation by Rational Functions of Given Order. Mathematical Methods in the Applied Sciences, 2014, 37(6): 846–859

77. Qian T, Chen Q H, Tan L H. Rational Orthogonal Systems are Schauder Bases. Complex Variables and Elliptic Equations, 2014, 59(6): 841–846

78. Qian T, Chen Q H, Li L Q. Analytic unit quadrature signals with non-linear phase. Physica D: Nonlinear Phenomena, 2005, 303: 80–87

79. Qian T, Huang J S. AFD on the n-Torus. in preparation

80. Qian T, Ho I T, Leong I T, Wang Y B. Adaptive decomposition of functions into pieces of non-negative instantaneous frequencies. International Journal of Wavelets, Multiresolution and Information Processing, 2010, 8(5): 813–833

81. Qian T, Li H, Stessin M. Comparison of Adaptive Mono-component Decompositions. Nonlinear Analysis: Real World Applications, 2013, 14(2): 1055–1074

82. Qian T, Tan L H. Characterizations of Mono-components: the Blaschke and Starlike types. Complex Analysis and Operator Theory, 2015: 1–17

83. Qian T, Tan L H. Backward shift invariant subspaces with applications to band preserving and phase retrieval problems. Mathematical Methods in the Applied Sciences, 2016, 39(6): 1591–1598

84. Qian T, Wang Y B. Adaptive Fourier Series-A Variation of Greedy Algorithm. Advances in Computational Mathematics, 2011, 34(3): 279–293

85. Qian T, Wegert E. Optimal Approximation by Blaschke Forms. Complex Variables and Elliptic Equations, 2013, 58(1): 123–133

86. Qian T, Sproessig W, Wang J X. Adaptive Fourier decomposition of functions in quaternionic Hardy spaces. Mathematical Methods in the Applied Sciences, 2012, 35(1): 43–64

87. Qian T, Tan L H, Wang Y B. Adaptive Decomposition by Weighted Inner Functions: A Generalization of Fourier Serie. J. Fourier Anal. Appl., 2011, 17(2): 175–190

88. Qian T, Wang J X. Adaptive Decomposition of Functions by Higher Order Szegö Kernels I: A Method for Mono-component Decomposition. submitted to Acta Applicanda Mathematicae

89. Qian T, Wang J Z. Gradient Descent Method for Best Blaschke-Form Approximation of Function in Hardy Space. http://arxiv.org/abs/1803.08422

90. Qian T, Sproessig W, Wang J X. Adaptive Fourier decomposition of functions in quaternionic Hardy spaces. Mathematical Methods in the Applied Sciences, 2012,35: 43–64

91. Qian T, Wang J X, Yang Y. Matching Pursuits among Shifted Cauchy Kernels in Higher-Dimensional Spaces. Acta Mathematica Scientia, 2014, 34(3): 660–672

92. Qian T, Wang R, Xu Y S, Zhang H Z. Orthonormal Bases with Nonlinear Phase. Advances in Computational Mathematics, 2010, 33: 75–95

93. Qian T, Xu Y S, Yan D Y, Yan L X, Yu B. Fourier Spectrum Characterization of Hardy Spaces and Applications. Proceedings of the American Mathematical Society, 2009, 137(3): 971–980

94. Qian T, Yang Y. Hilbert Transforms on the Sphere With the Clifford Algebra Setting. Journal of Fourier Analysis and Applications, 2019, 15: 753–774

95. Qian T. Zhang L M, Li Z X. Algorithm of Adaptive Fourier Decomposition. IEEE Transaction on Signal Processing, Dec., 2011, 59(12): 5899–5902

96. Qu W, Dang P. Rational Approximation in a Class of Weighted Hardy Spaces. Complex Analysis and Operator Theory volume, 2019, 13: 1827–1852

97. Salomon L. Analyse de l'anisotropie dans des images texturées, 2016

98. Sakaguchi F, Hayashi M. General theory for integer-type algorithm for higher order differential equations. Numerical Functional Analysis and Optimization, 2011, 32(5): 541–582

99. Sakaguchi F, Hayashi M. Differentiability of eigenfunctions of the closures of differential operators with rational coefficient functions, arXiv:0903.4852, 2009

100. Sakaguchi F, Hayashi M. Practical implementation and error bound of integer-type algorithm for higher-order differential equations. Numerical Functional Analysis and Optimization, 2011, 32(12): 1316–1364

101. Sakaguchi F, Hayashi M. Integer-type algorithm for higher order differential equations by smooth wavepackets. arXiv:0903.4848, 2009

102. Schepper D, Qian T, Sommen F, Wang J X. Holomorphic Approximation of L_2-functions on the Unit Sphere in R3. Journal of Mathematical Analysis and Applications, 2014, 416(2): 659–671

103. Sharpley R C, Vatchev V. Analysis of intrinsic mode functions. Constructive Approximation, 2006, 24: 17–47

104. Stein E M, Weiss G. Introduction to Fourirer Analysis on Euclidean Spaces. Princeton University Press, Princeton, New Jersey, 1971

105. Tan L H, Shen L X, Yang L H. Rational orthogonal bases satisfying the Bedrosian Identity. Advances in Computational Mathematics, 2010, 33: 285–303

106. Tan L H, Yang L H, Huang D R. The structure of instantaneous frequencies of periodic analytic signals. Sci. China Math., 2010, 53(2): 347–355

107. Tan L H, Qian T. Backward Shift Invariant Subspaces With Applications to Band Preserving and Phase Retrieval Problems

108. Tan L H, Qian T. Extracting Outer Function Part from Hardy Space Function. Science China Mathematics, 2017, 60(11): 2321–2336

109. Tan L H, Qian T, Chen Q H. New aspects of Beurling Lax shift invariant subspaces. Applied Mathematics and Computation, 2015: 257–266

110. Vatchev V. A class of intrinsic trigonometric mode polynomials. International Conference Approximation Theory. Springer, Cham, 2016: 361–373

111. Vliet D V. Analytic signals with non-negative instantaneous frequency. Journal of Integral Equations and Applications, 2009, 21: 95–111

112. Walsh J L. Interpolation and Approximation by Rational Functions in the Complex Plane. American Mathematical Society: Providence, RI, 1969

113. Wang S L. Simple Proofs of the Bedrosian Equality for the Hilbert Transform. Science in China, Series A: Mathematics, 2009, 52(3): 507–510

114. Wang J X, Qian T. Approximation of monogenic functions by higher order Szegö kernels on the unit ball and the upper half space. Sciences in China: Mathematics, 2014, 57(9): 1785–1797

115. Weiss G, Weiss M. A derivation of the main results of the theory of Hp-spaces. Rev. Un. Mat. Argentina, 1962, 20: 63–71
116. Wu W. Applications in Digital Image Processing of Octonions Analysis and the Qian Method. South China Normal University, 2014
117. Wang Z, da Cruz J N, Wan F. Adaptive Fourier decomposition approach for lung-heart sound separation. Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA). 2015 IEEE International Conference on. IEEE, 2015
118. Wu M Z, Wang Y, Li X M. Fast Algorithm of The Qian Method in Digital Watermarking. Computer Engineering and Desining, 2016
119. Wu M Z, Wang Y, Li X M. Improvement of 2D Qian Method and its Application in Image Denoising. South China Normal University, 2016
120. Xu Y S. Private comminication, 2005
121. Yang Y, Qian T, Sommen F. Phase Derivative of Monogenic Signals in Higher Dimensional Spaces. Complex Analysis and Operator Theory, 2012, 6(5): 987–1010
122. Yu B, Zhang H Z. The Bedrosian Identity and Homogeneous Semi-convolution Equations. Journal of Integral Equations and Applications, 2008, 20: 527–568
123. Zhang L. A New Time-Frequency Speech Analysis Approach Based On Adaptive Fourier Decomposition. World Academy of Science, Engineering and Technology, International Journal of Electrical, Computer, Energetic, Electronic and Communication Engineering, 2013
124. Zhang L M, Liu N, Yu P. A novel instantaneous frequency algorithm and its application in stock index movement prediction. IEEE Journal of Selected Topics in Signal Processing, 2012, 6(4): 311–318
125. Zhang L M, Qian T, Mai W X, Dang P. Adaptive Fourier decomposition-based Dirac type time-frequency distribution. Mathematical Methods in the Applied Sciences, 2017, 40(8): 2815–2833