Appl. Comput. Harmon. Anal. 27 (2009) 351–366

Contents lists available at ScienceDirect

Applied and Computational Harmonic Analysis

www.elsevier.com/locate/acha

Letter to the Editor

Nonlinear inversion of a band-limited Fourier transform

Gregory Beylkin *, Lucas Monzón

Department of Applied Mathematics, University of Colorado, Boulder, CO 80309-0526, United States

ARTICLE INFO

Article history:
Available online 3 May 2009
Communicated by Thomas Strohmer on 22 July 2008

Keywords:
Band-limited Fourier transform
Discrete Fourier transform
Windows
Filtering
Approximation by exponentials
Approximation by rational functions

ABSTRACT

We consider the problem of reconstructing a compactly supported function with singularities either from values of its Fourier transform available only in a bounded interval or from a limited number of its Fourier coefficients. Our results are based on several observations and algorithms in [G. Beylkin, L. Monzón, On approximation of functions by exponential sums, Appl. Comput. Harmon. Anal. 19 (1) (2005) 17–48]. We avoid both the Gibbs phenomenon and the use of windows or filtering by constructing approximations to the available Fourier data via a short sum of decaying exponentials. Using these exponentials, we extrapolate the Fourier data to the whole real line and, on taking the inverse Fourier transform, obtain an efficient rational representation in the spatial domain. An important feature of this rational representation is that the positions of its poles indicate location of singularities of the function. We consider these representations in the absence of noise and discuss the impact of adding white noise to the Fourier data. We also compare our results with those obtained by other techniques. As an example of application, we consider our approach in the context of the kernel polynomial method for estimating density of states (eigenvalues) of Hermitian operators. We briefly consider the related problem of approximation by rational functions and provide numerical examples using our approach.

1. Introduction

We consider the problem of reconstructing a compactly supported function with singularities from values of its Fourier transform available only in a bounded interval, or from a limited number of its Fourier coefficients. The singularities that we allow include jump discontinuities of a function or its derivatives or, more generally, integrable singularities. In all cases, we assume no prior knowledge of locations of singularities. Specifically, for a compactly supported function $f \in L^2(\mathbb{R})$, we develop an algorithm for finding a rational representation of f which matches a given set of equally spaced samples of the Fourier transform

$$\hat{f}(\xi) = \int_{-\infty}^{\infty} f(x) e^{-2\pi i \xi x}\, dx \tag{1}$$

in some interval $[-a, a]$. Similarly, for a periodic function, given a finite set of its Fourier series coefficients, we obtain a rational trigonometric representation. Our results are based on several observations and algorithms in [5], where we formulate and solve the problem of nonlinear approximation of functions via near-optimal linear combination of exponentials.

* Corresponding author.
E-mail address: beylkin@colorado.edu (G. Beylkin).

1063-5203/$ – see front matter
doi:10.1016/j.acha.2009.04.003

As a representative example, let us consider a compactly supported real-valued function with Fourier transform of the form

$$\hat{f}(\xi) = \sum_{m=1}^{M} a_m |\xi|^{-\alpha_m} e^{i\phi_m \xi}, \tag{2}$$

where $a_m \in \mathbb{C}$, $\xi \in \mathbb{R}$, $\alpha_m > 0$ and $\phi_m \in [0, 2\pi)$. Since f is compactly supported, its Fourier transform $\hat{f}(\xi)$ is infinitely differentiable at $\xi = 0$ in spite of the form of the expression in (2). Thus, the coefficients a_m and exponents α_m and ϕ_m in (2) are interrelated, although the specific form of these relations is not important for our purposes. We note that particular examples of functions f with such Fourier transforms include piecewise polynomials. Since the function f is real, it satisfies $\hat{f}(-\xi) = \overline{\hat{f}(\xi)}$ and, thus, we may assume that we are given $2N+1$ samples $\hat{f}(a\frac{n}{2N})$, where $n = 0, 1, \ldots, 2N$. Using this information, we show how to construct an explicit rational representation of f with a near optimal (minimal) number of poles. We note that a reduction algorithm similar to that in [5] allows us to extend our approach to complex-valued functions (see Remark 1 below).

Reconstructing a function from a finite number of samples of its Fourier transform is an important problem in electrical engineering, usually addressed by windowing the Fourier data to deal with the fact that $\hat{f}$ decays slowly and its abrupt truncation results in Gibbs phenomenon. Indeed, the direct application of the Discrete Fourier Transform (DFT) to the data causes severe over and under shooting near the singularities. We note that more sophisticated filtering techniques that require some prior information on location of singularities have been developed (see [6,8–13,17,18] and references therein). In this paper we describe an approach that avoids both the Gibbs phenomenon and the use of windows/filtering and, effectively, is based on nonlinear extrapolation of Fourier data.

Given only samples of the Fourier transform of a function in a bounded interval, the error of approximating the function may be arbitrarily large due to the unknown Fourier modes outside the interval. In such case we may only talk about representation of the available data. In our case, using a nonlinear approximation of the Fourier data via decaying exponentials, we obtain a rational function which has the property that its Fourier transform matches the available samples within a given accuracy. The advantage of our approach is two-fold: we have a way to estimate the level of noise (similar to other nonlinear approximations, by observing the error decay rate) and we obtain a rational representation of the function that may be useful for further analysis. In this paper, we mostly concentrate on this problem in Sections 2–4, since it is a common situation in signal processing. We note that related approximation problems appear in different practical applications as well. In particular,

(1) Assuming that the function f is in some class and that Fourier values are available in a finite interval of our choice, we may consider the problem of obtaining an accurate rational approximation to that function. The key difference here is that we get either to choose the size of the Fourier interval or may consider a limiting process and resulting convergence rates. We note that different problems and applications of approximating functions with singularities by polynomials may be found in [6,8,9,11–13] (see also references therein). Typical applications include numerical methods and signal processing, for example, problems of edge detection.
(2) Under the same assumptions, we may consider computing the optimal rational approximation of the function in L^∞ norm. The theory of such approximations is the subject of seminal papers [1–3] but the straightforward approach that follows from the theory may fail as a numerical algorithm and the goal is to find a reliable algorithm for this purpose.
(3) Assuming that we are given noisy data, we may consider problems described in (1)–(2) above. Addressing this problem should lead to a useful way of compressing or denoising data using rational approximations as a tool.

To demonstrate that the approach of this paper is useful in these problems, we present some theoretical considerations as well as numerical examples in Section 5, noting that we plan to consider these problems in depth elsewhere. Finally, as an example of application, we consider our approach in the context of the kernel polynomial method (KPM) for estimating the density of states in Section 6.

2. Nonlinear inversion

Our approach relies on the following observations:

- As is shown in [5], for $\alpha_m > 0$, the functions $\xi^{-\alpha_m}$, for $\xi > 0$, have approximations via a short sum of decaying exponentials of the form $\sum_{l=1}^{L_m} w_{m,l} e^{-\tau_{m,l}\xi}$, where all $w_{m,l}$ and $\tau_{m,l}$ are real and positive. Upon substitution into (2), the resulting approximation of the Fourier transform has the form

$$\sum_{m,l} c_{m,l} e^{(-\tau_{m,l} + i\phi_m)\xi}, \tag{3}$$

for some complex coefficients $c_{m,l}$. Hence, it is reasonable to expect an approximation of $\hat{f}$ in (2) as a sum of exponentials with exponents that have negative real part. Thus, instead of looking for $\hat{f}$ in the form (2), we rely on near

optimal approximation by decaying exponentials to approximate $\hat{f}$ in the form (3). We note that trying to recover $\hat{f}$ in the form (2) leads to a complicated system of nonlinear equations [7].

- Given the samples $\hat{f}(a\frac{n}{2N})$, $n = 0, 1, \ldots, 2N$, and assuming that they provide a sufficient oversampling of the function $\hat{f}(\xi)$ in $[-a, a]$, the algorithm in [5] allows us to construct an approximation with a near optimal (minimal) number of terms M,

$$\hat{g}(\xi) = \sum_{m=1}^{M} w_m e^{-\eta_m \xi}, \tag{4}$$

such that

$$\left|\hat{f}(\xi) - \hat{g}(\xi)\right| \leqslant \epsilon, \quad \text{for } 0 \leqslant \xi \leqslant a,$$

where $w_m = u_m + iv_m$, $\eta_m = \tau_m + i\theta_m$ for real u_m, v_m, θ_m and positive τ_m. Here $\epsilon = \epsilon(N, a)$ and our goal is to make the error as small as possible for given N and a.

- The approximation $\hat{g}$ in (4) is obtained for $\xi \in [0, a]$ but we can use this expression for any $\xi > 0$ and, furthermore, to be consistent with the goal of g being a real function, extend $\hat{g}$ to the negative semi-axis as the complex conjugate, $\hat{g}(-\xi) = \overline{\hat{g}(\xi)}$. We note that the extended function $\hat{g}(\xi)$ is infinitely differentiable everywhere except at $\xi = 0$. Nevertheless, since the extended function $\hat{g}$ is an accurate approximation to $\hat{f} \in C^\infty[-a, a]$, we use it as an extension of $\hat{f}$ to the whole real line. The inverse Fourier transform of $\hat{g}$ is derived analytically and we obtain an explicit rational representation $g(x)$ of the function $f(x)$,

$$\begin{aligned} g(x) &= \int_{-\infty}^{\infty} \hat{g}(\xi) e^{2\pi i x\xi}\, d\xi = 2\,\mathrm{Re}\left(\int_{0}^{\infty} \hat{g}(\xi) e^{2\pi i x\xi}\, d\xi \right) \\ &= -2\,\mathrm{Re} \sum_{m=1}^{M} \frac{w_m}{2\pi i x - \eta_m} = -\frac{1}{\pi} \sum_{m=1}^{M} \frac{v_m(2\pi x - \theta_m) - u_m \tau_m}{(2\pi x - \theta_m)^2 + \tau_m^2}. \end{aligned} \tag{5}$$

Observe that the function g in (5) is defined by $2M$ poles and corresponding residues. Since g is real on the real line, the poles appear in complex conjugate pairs, z_m and $\bar{z}_m$, $z_m = \eta_m/2\pi i$ with residues $-w_m/2\pi i$ and $\overline{w}_m/2\pi i$. In what follows, when illustrating rational representations of the form (5), we will display poles only in the lower half plane and omit their complex conjugates.

We note that the rational representation (5) not only compresses information describing f into $2M$ complex parameters, since usually $M \ll N$, but also carries additional information via the pole locations in (5). In particular, the accumulation of poles indicate location of singularities (see Figs. 5, 9, 10 and 11).

Remark 1. If the function f is complex-valued, then given $\hat{f}(\xi)$, we use the approach described above on $(\hat{f}(\xi) + \overline{\hat{f}(-\xi)})/2$ and $(\hat{f}(\xi) - \overline{\hat{f}(-\xi)})/(2i)$, which results in obtaining rational representations for the real and imaginary parts of f. Their combination yields a rational representation for f which may not be optimal. However, by applying a reduction algorithm similar to the one in [5, Section 6], we may further reduce the number of terms in the rational representation of f. The details of such algorithm will be described elsewhere.

2.1. Algorithm for approximations by sum of exponentials

We briefly review the algorithm in [5] to approximate functions by sum of exponentials. These approximations, obtained for a finite but arbitrary accuracy, have significantly fewer terms than corresponding Fourier representations (see [4,5]).

Given $\epsilon > 0$ and $2N + 1$ values of the complex-valued function $h(a\xi)$ on a uniform grid in $[0, 1]$, we have developed in [5] an algorithm to obtain complex coefficients (weights) w_m and exponents t_m, and (nearly) minimal $M = M(\epsilon, a)$ such that

$$\left| h\left(a\frac{n}{2N}\right) - \sum_{m=1}^{M} w_m e^{-t_m n} \right| < \epsilon, \tag{6}$$

for any $0 \leqslant n \leqslant 2N$. In this formulation, $a > 0$ scales the problem to the interval $[0, 1]$ and ϵ is the accuracy sought.

Using weights w_m in (6) and exponents $\eta_m = \frac{2N}{a} t_m$, we have an approximation to $h(x)$ by a sum of exponentials valid for all $x \in [0, a]$,

$$\left| h(x) - \sum_{m=1}^{M} w_m e^{-\eta_m x} \right| < \epsilon', \tag{7}$$

where ϵ' is very close to ϵ provided that the function h is appropriately sampled in (6) to justify local interpolation.

The steps to achieve the approximation (6) are as follows:

- Build the $N+1 \times N+1$ Hankel matrix $\mathbf{H}_{kl} = h_{k+l}$ using the samples $h_n = h(a\frac{n}{2N})$, $0 \leqslant n \leqslant 2N$.
- Find a vector $\mathbf{u} = (u_0, \ldots, u_N)$, satisfying $\mathbf{Hu} = \sigma \bar{\mathbf{u}}$, with positive σ close to the target accuracy ϵ. The existence of such vector $\mathbf{u}$ follows from Tagaki's factorization (see [5, p. 22]); the singular value decomposition yields σ as a singular value and $\mathbf{u}$ as a singular vector of $\mathbf{H}$. We label the first $M+1$ singular values of $\mathbf{H}$ in decreasing order $\sigma_0 \geqslant \sigma_1 \geqslant \cdots \geqslant \sigma_M$, where σ_M is chosen so that $\sigma_M/\sigma_0 \approx \epsilon$. Typically, singular values decay rapidly and, thus, $M = \mathcal{O}(\log \epsilon^{-1})$ and $M \ll N$.
- Compute roots γ_m of the polynomial $u(z) = \sum_{n=0}^{N} u_n z^n$ whose coefficients are the entries of the singular vector $\mathbf{u}$ computed in the previous step. The weights w_m are obtained solving the least-squares Vandermonde system

$$\sum_{m=1}^{N} w_m \gamma_m^n = h\left(a\frac{n}{2N}\right), \quad 0 \leqslant n \leqslant 2N.$$

  Typically, only M weights w_m have absolute value larger than the target accuracy ϵ, thus allowing us to remove most roots γ_m from further consideration. It is often possible to predict the approximate location of roots γ_m with significant weights and, thus, find only them. However, if no a priori information on their location is available, then all roots should be computed and selected according to the size of the corresponding weights.
- The exponents t_m in (6) correspond to $t_m = \log \gamma_m$, where log is the principal value of the logarithm.

For more details on this algorithm we refer to [5]. We note that the special case of purely imaginary exponents η_m in (7) is treated in [4].

2.2. Numerical stability and computational cost

The algorithm requires computing singular value decomposition (SVD) of an $(N+1) \times (N+1)$ Hankel matrix, finding $M \ll N$ roots inside the unit disk for a polynomial of degree N and solving a $(2N+1) \times M$ Vandermonde system. We note that it is not necessary to compute a full SVD since (in most circumstances) the singular values have a rapid decay and we are only interested in a single singular value (above a threshold) and the corresponding singular vector. The cost of applying a Hankel matrix is $\mathcal{O}(N \log N)$ and, thus, the cost of this step is less than that of the full SVD. Typically, only M roots are inside the unit disk so the cost of finding them is again less than that of finding all the roots of a polynomial of degree N. Finally, the least squares problem has to recover only M weights, also reducing the computational cost. Notwithstanding these considerations, so far we developed only prototype algorithms and a full development of fast algorithms for this problem is pending.

Addressing numerical stability, we note that care needs to be taken to have extra digits relative to the target accuracy ϵ. Indeed, there are only absolute error estimates for algorithms finding singular values and corresponding singular vectors. For this reason, we carry twice the number of accurate digits relative to ϵ; for example, we use double precision in our computations if the target accuracy corresponds to single precision $\approx 10^{-7}$. These extra digits are typically sufficient to achieve the target accuracy for the final result, as is also the case in computing generalized Gaussian quadratures (see e.g. [4,5]).

3. Reconstructing from band-limited Fourier data

As a representative example, we consider a piecewise polynomial $f(x)$ in the interval $[0, 5]$, with singularities located at $x = 1, 2, 3, 4$. The function is discontinuous at $x = 1$ and has a break in its nth derivative at $x = n+1$, see Fig. 1(a). Specifically, we use f defined as

$$f(x) = \begin{cases} 0, & x \in (-\infty, 1), \\ 1, & x \in [1, 2), \\ (3-x)^2, & x \in [2, 3), \\ 40(3-x)^2(4-x)^3, & x \in [3, 4), \\ 0, & x \in [4, \infty) \end{cases} \tag{8}$$

with Fourier transform

$$\begin{aligned} \hat{f}(\xi) = {} & \frac{75}{\pi^6 \xi^6}\left(e^{-6\pi i\xi} - e^{-8\pi i\xi}\right) - \frac{30i}{\pi^5 \xi^5}\left(3e^{-6\pi i\xi} + 2e^{-8\pi i\xi}\right) \\ & - \frac{15}{\pi^4 \xi^4}\left(3e^{-6\pi i\xi} - e^{-8\pi i\xi}\right) + \frac{i}{4\pi^3 \xi^3}\left(e^{-4\pi i\xi} + 39e^{-6\pi i\xi}\right) + \frac{e^{-4\pi i\xi}}{2\pi^2 \xi^2} - \frac{ie^{-2\pi i\xi}}{2\pi \xi}, \end{aligned}$$

so that, for small ξ,

$$\hat{f}(\xi) = 2 - \frac{127\pi i}{14}\xi + \mathcal{O}(\xi^2).$$

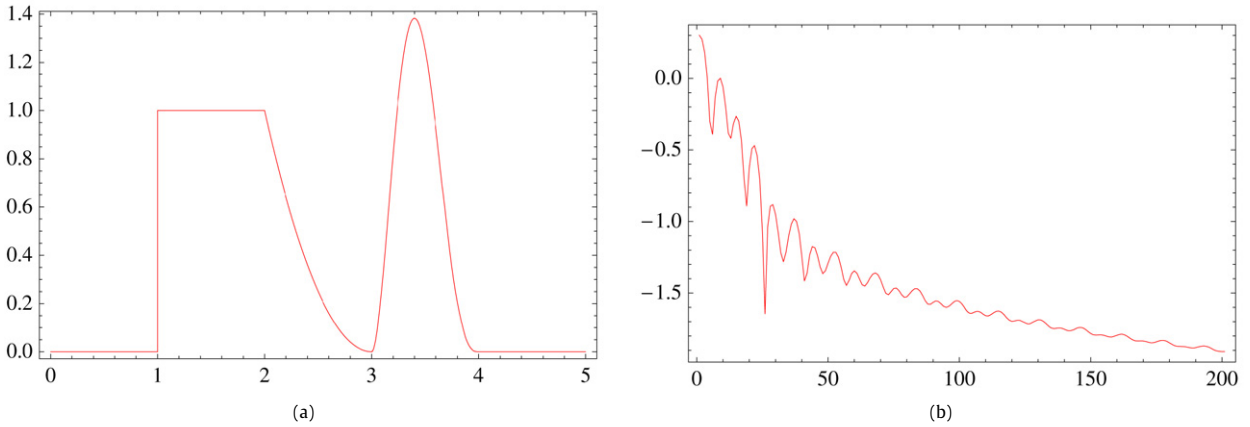

Fig. 1. (a) The compactly supported function f in (8) which has singularities of different types. (b) Absolute values of samples of $\hat{f}$ in (9), where the last sample is of size $1.2 \cdot 10^{-2}$.

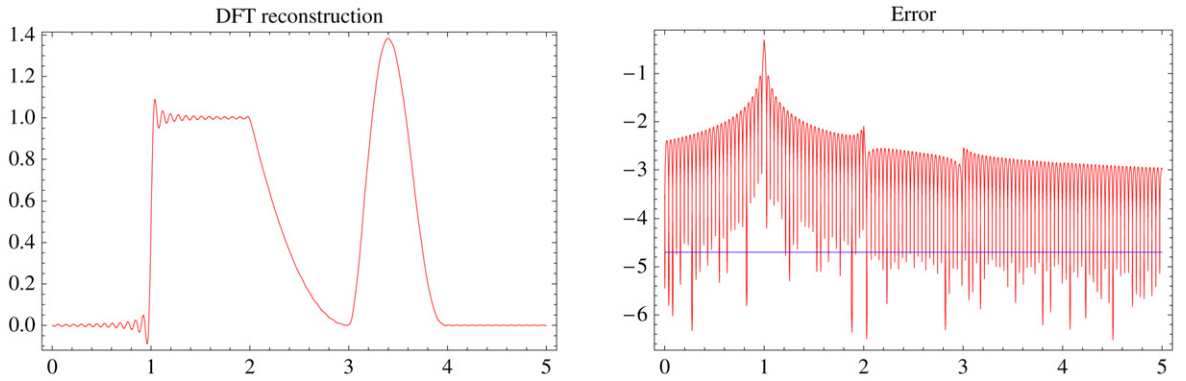

Fig. 2. Direct DFT reconstruction of the function in (8) and its absolute error (using $\log_{10}$ scale on the vertical axis) illustrating Gibbs phenomenon.

For this example we assume that we are given samples

$$h_n = \hat{f}\left(\frac{n}{5\pi}\right), \quad n = 0, \ldots, 200, \tag{9}$$

and a target accuracy ϵ. We assume that the samples have accuracy commensurable with ϵ and that they cover a large enough range as to capture the rate of decay associated with the singularities of f. The absolute values of these 201 samples are depicted in Fig. 1(b), where the magnitude of the Fourier transform decays to $1.2 \cdot 10^{-2}$ at the last available sample. Our goal is to find a rational representation of f in the interval $[0, 5]$ that matches the Fourier data and compare the result with the function f. In the absence of noise, we compare our results with standard filtering to attenuate the Gibbs phenomenon. We then discuss the impact of adding white noise to the Fourier data in Section 3.2. Later, in Section 4, we consider the same function (8) treated as a periodic function. In that section, we also demonstrate our approach on a test problem considered in [10,17,18], where the authors used more sophisticated techniques than standard filtering.

3.1. Windowed vs rational reconstruction in the absence of noise

We start by illustrating Gibbs phenomenon due to the abrupt truncation of the Fourier series. Given Fourier data, we use the DFT to reconstruct f in $[0, 5]$ as

$$\frac{1}{5\pi} \sum_{|n| \leqslant 200} \hat{f}\left(\frac{n}{5\pi}\right) e^{\frac{2\pi i n x}{5\pi}}.$$

The reconstruction and its absolute error are shown in Fig. 2.

The typical way to improve quality of reconstruction is by windowing (filtering) the Fourier data before applying the DFT. Using two different windows, we compare such windowed reconstruction with our approach. We use a Kaiser window with parameters $\alpha = 7.7$ and $N = 401$ (see [14, p. 175]) and we also consider a smooth window constructed to be flat near zero with a smaller transition region (see Fig. 3). The details of the construction of the latter are not important since we

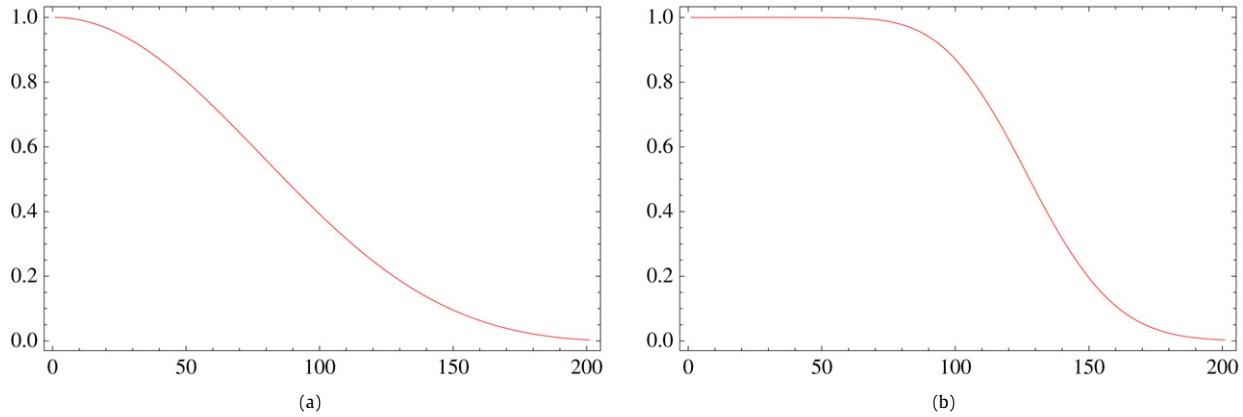

Fig. 3. (a) Kaiser window and (b) a "flat" window for filtering the Fourier data in (9).

emphasize only qualitative differences. The reconstruction error using these windows is illustrated in Fig. 4. Horizontal lines in Figs. 2 and 4 indicate the error level $\approx 2 \cdot 10^{-5}$ (away from the singularities) attained in the second windowed inversion. We observe that windowed reconstruction clearly improve the result vis-à-vis the one shown in Fig. 2. However, the error still spreads quite widely about the singularities of the function.

The key point is that by using a single window our ability to improve this situation and reduce the error is limited since windows always modify the original data, even if slightly (i.e., the reconstructed function does not match the original Fourier values). We provide further comparison with some more sophisticated alternatives in Section 4.2.

In Fig. 4 (bottom) we display our rational reconstruction using the algorithm for a target accuracy of 10^{-8}. In this case we consider the singular value $\sigma_{27} \approx 8.9 \cdot 10^{-9}$, so that the representation in (5) has $M = 27$ terms. Note how not only the overshooting is attenuated but also the error spreads within a much smaller neighborhood about each singularity. Furthermore, we observe that the rational reconstruction achieves the target accuracy away from the singularities. In fact, in the absence of noise, the target accuracy is only limited by the available bandwidth and may be achieved by selecting the appropriate singular value (see Fig. 6).

In Fig. 5, we compare locations of the (isolated) singularities of the original function with the positions of the poles of its rational representation. Note that the poles tend to accumulate as they approach the actual locations of singularities. If we are given additional information about the function, e.g., that it is a piecewise polynomial of low degree, we may then use these poles to estimate the location of the singularities.

3.2. Reconstruction in the presence of noise

Let us now address a more realistic situation in signal processing where the data is affected by noise. We perform the same analysis as before but modify the Fourier data (9) by adding independent Gaussian noise (with the standard deviation $\frac{1}{2} \cdot 10^{-3}$) to both the real and imaginary parts of (9). In this example the l^∞-norm of the noise equals $1.6 \cdot 10^{-3}$, whereas the tail values of the Fourier data are of magnitude $\approx 10^{-2}$.

In Fig. 6 we display the first 28 singular values of the Hankel matrices obtained in the first step of our algorithm. In the presence of noise, we observe a drastic change in the rate of decay around the singular value $\sigma_{10} \approx 0.0123285$. This indicates that choosing a smaller singular value may improve our approximation only marginally and at the expense of significantly more terms. By selecting σ_{10}, we take advantage of a common property of many nonlinear approximations, namely, that a significant reduction of the rate of decay may be interpreted as reaching the level of noise in the data. Since random noise does not have an efficient (short) representation via sums of exponentials, we detect its level by looking for significant changes in the decay of the singular values.

We display in Fig. 7 the error of the rational reconstruction for the selected σ_{10} yielding $M = 10$ terms in (5). At the top of the figure, we display the error obtained using windowed inversion with "flat" window. In both cases the error away from the singularities (plotted as a horizontal line of height $10^{-2.8} \approx 1.58 \cdot 10^{-3}$) is almost identical to the noise level.

4. Reconstructing from coefficients of Fourier series

So far we assumed that the Fourier data corresponds to values of the Fourier transform of a compactly supported function. Alternatively, we may consider a case where the function is periodic and we have access to a number of its Fourier series coefficients.

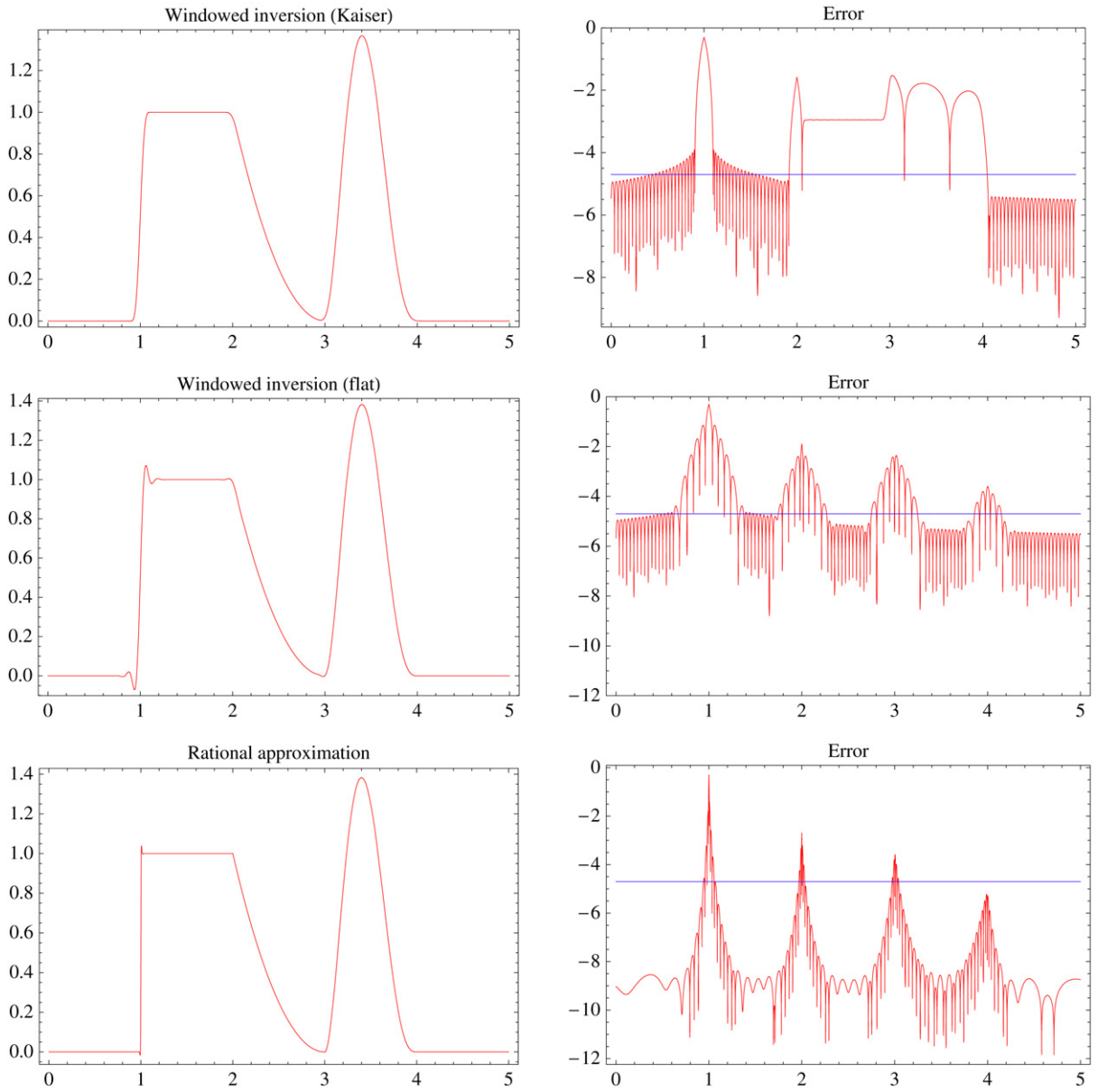

Fig. 4. Comparison of reconstructions with errors displayed using $\log_{10}$ scale on the vertical axis. The horizontal lines indicate the level of $2 \cdot 10^{-5}$ achieved by "flat" window (second from the top). The rational representation with 27 terms is obtained via (5).

Let us assume that the real-valued function to be reconstructed has been rescaled so that its support is in the interval $[0, 1]$. Considering coefficients of the Fourier series,

$$\hat{f}_n = \int_0^1 f(t) e^{-2\pi i n t}\, dt,$$

we have

$$\hat{f}_{-n} = \overline{\hat{f}_n}$$

and, thus,

$$f(x) = \sum_{n\in\mathbb{Z}} \hat{f}_n e^{2\pi i n x} = \hat{f}_0 + 2\,\mathrm{Re} \sum_{n\geqslant 1} \hat{f}_n e^{2\pi i n x}.$$

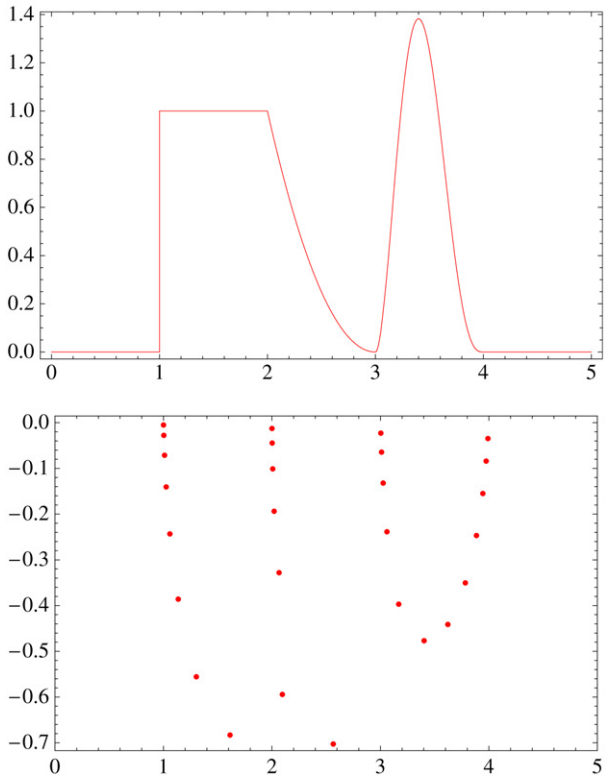

Fig. 5. The top figure displays the piece-wise polynomial function in (8) and, the bottom one, the locations of the 27 poles of its rational representation in (5). We have aligned the horizontal scale of these two figures to illustrate that the positions of the poles in the complex plane arrange themselves in branches as to indicate the location of function singularities (well separated in this case). As poles approach the real axis, their arrangement also corresponds to the type of singularity at that location.

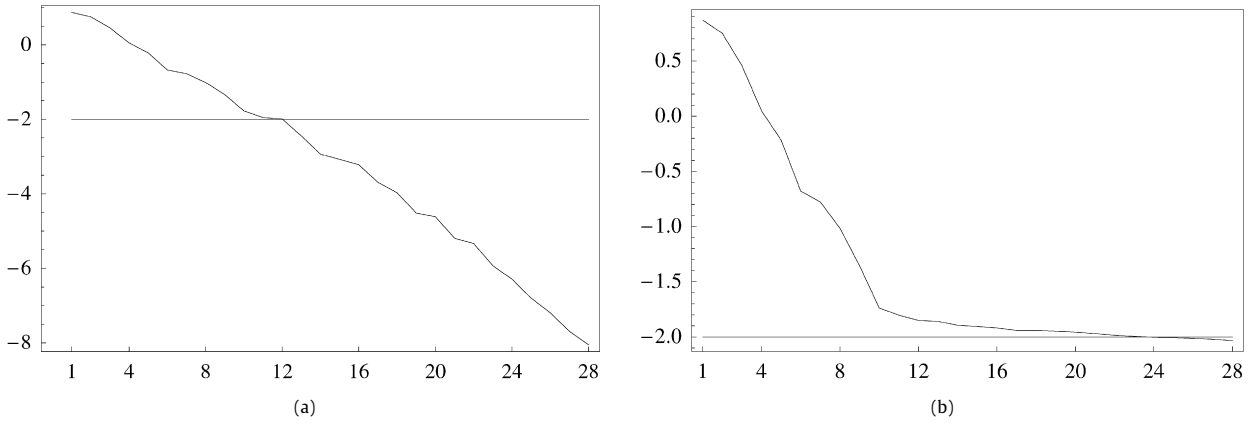

Fig. 6. The first 28 singular values of Hankel matrix with entries in (9) without noise (a) and with Gaussian noise added (b). The singular values are plotted using $\log_{10}$ scale along the vertical axis and their index along the horizontal axis. The change in the rate of decay allows us to establish the level of noise.

Given coefficients $\hat{f}_n$ for $n = 0, 1, \ldots, 2N$, using the algorithm of Section 2.1, we approximate them as

$$\left| \hat{f}_n - \sum_{m=1}^{M} w_m e^{-\eta_m n} \right| < \epsilon,$$

where $\mathrm{Re}(\eta_m) > 0$. As a result, we represent $f(x)$ by the trigonometric rational function $g(x)$ defined as

$$g(x) = \mathrm{Re} \sum_{m=1}^{M} w_m + 2\,\mathrm{Re} \sum_{n \geqslant 1} \sum_{m=1}^{M} w_m e^{-\eta_m n} e^{2\pi i n x} = \mathrm{Re} \sum_{m=1}^{M} w_m + 2\,\mathrm{Re} \sum_{m=1}^{M} \frac{w_m}{e^{-2\pi i x + \eta_m} - 1}, \tag{10}$$

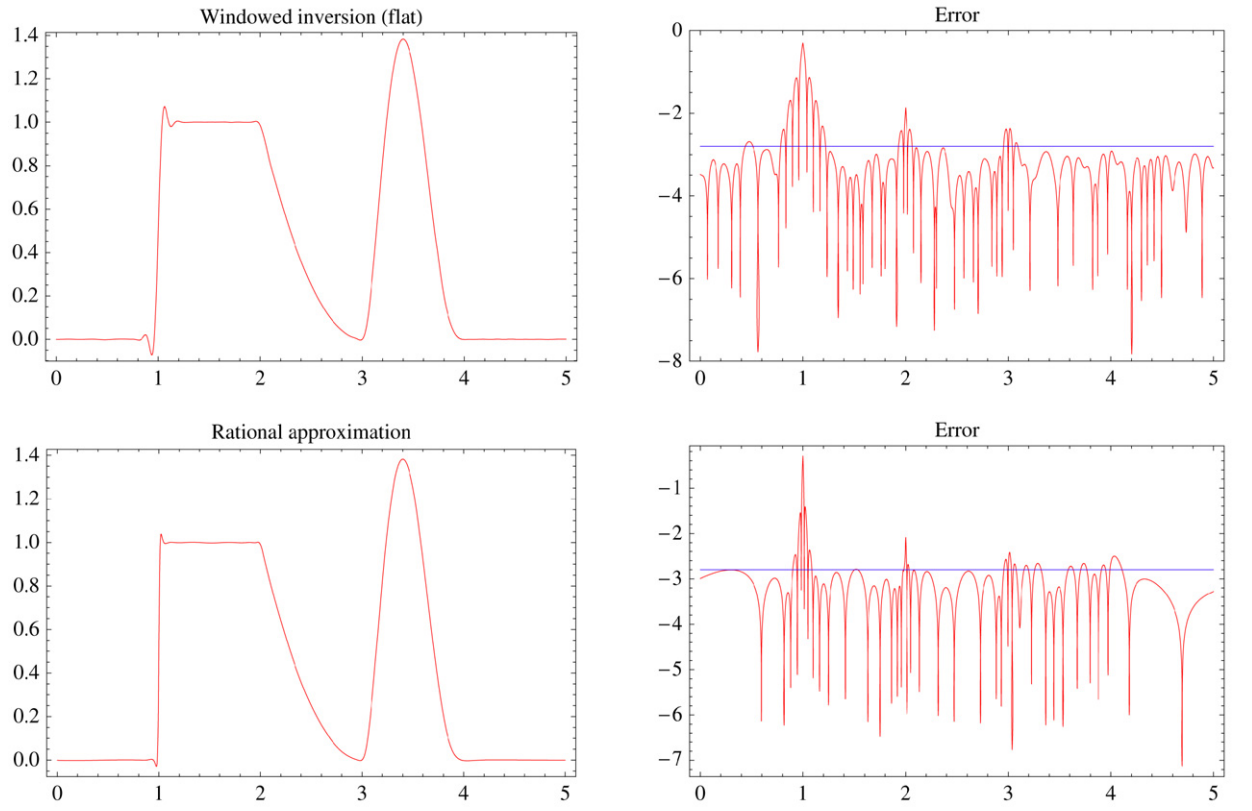

Fig. 7. Comparison of reconstructions in the presence of Gaussian noise with standard deviation $\frac{1}{2}\cdot 10^{-3}$. Absolute errors are displayed on a $\log_{10}$ scale along the vertical axis and the horizontal lines indicate the error level $1.6\cdot 10^{-3}$. We note that the error using the Kaiser window is already at this level for the noiseless case (see Fig. 4) and is not displayed here for that reason.

where we use $\mathrm{Re}\sum_{m=1}^{M} w_m$ instead of $\sum_{m=1}^{M} w_m$ since it approximates the real value $\hat{f}_0 = \int_0^1 f(t)\,dt$.

Again the function g in (10) is defined by $2M$ poles and corresponding residues. Let us consider the rational function $\tilde{g}(z)$ such that $\tilde{g}(e^{-2\pi i x}) = g(x)$. The poles of $\tilde{g}$ appear in pairs, z_m and $1/\bar{z}_m$, $z_m = e^{-\eta_m}$ with residues $w_m e^{-\eta_m}$ and $-\overline{w}_m e^{\bar{\eta}_m}$ and zero constant term. In illustrating rational representations of the form in (10), we display poles of $g(x)$ to interpret the representation on $[0,1]$, i.e., we display $\eta_m/2\pi i$ rather than the poles z_m of $\tilde{g}$.

4.1. Reconstruction of compactly supported functions considered as periodic functions

For f in (8), let us consider $f(5x)$ as a periodic function in $[0,1]$ and use as input its Fourier coefficients $\frac{1}{5}\hat{f}(n/5)$, $n = 0, 1, \ldots, 2N$, instead of values of the Fourier transform (9) in Section 3. We reconstruct f using 63 and 127 of its Fourier coefficients (i.e. $N = 31$ and $N = 63$ above). For $N = 31$ we select the singular value $\sigma_{23} \approx 1.33357\cdot 10^{-8}$, leading to a representation with 23 terms. The resulting accuracy away from singularities of the function is approximately 7.6 digits. For $N = 63$ we select $\sigma_{46} \approx 5.10495\cdot 10^{-16}$, leading to 46 terms and accuracy away from singularities of about 15 digits. The errors of reconstruction are displayed in Fig. 8.

We note that by using 63 Fourier coefficients, we roughly cover the same frequency interval of the Fourier transform as in (9) and, as expected, the error of reconstruction is similar in both cases (see bottom figure in Fig. 4).

4.2. An example from [10,17,18].

Although our approach is substantially different than in e.g. [10,17,18], we consider an example used in these papers to provide a way to compare our method with other available techniques. Since our nonlinear approximation is specific for each function to which it is applied, our algorithm is adaptive and does not require any prior assumption on e.g., distances between singularities of the function. On the other hand, the approach in e.g. [10,17,18] is based on parametrized optimization for a class of functions and, thus, is adaptive only in a limited sense.

As in [10,17,18], we consider the periodic function

$$f(x) = \begin{cases} (2e^{4\pi x} - 1 - e^{\pi})/(e^{\pi} - 1), & \text{for } 0 \leqslant x < \frac{1}{4}, \\ -\sin(\frac{4}{3}\pi x - \frac{1}{3}\pi), & \text{for } \frac{1}{4} \leqslant x < 1, \end{cases} \tag{11}$$

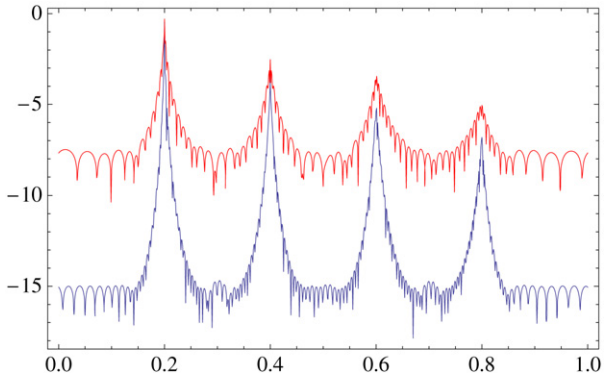

Fig. 8. Reconstruction errors (using $\log_{10}$ scale on vertical axis) for the periodic function $f(5x)$ on $[0,1]$, where f is given in (8). The upper (lower) curve shows the error using as input 63 (127) of its Fourier coefficients.

where we rescaled the interval to $[0,1]$ to be consistent with our notation. We apply our algorithm to the first 63 and 127 Fourier coefficients of f and display the results in Fig. 9. In the case of 63 coefficients, we select $\sigma_{15} \approx 6.2928 \cdot 10^{-9}$ leading to 15 terms in the reconstruction and accuracy away from singularities of the function of about 8 digits. For 127 coefficients, we chose $\sigma_{27} \approx 3.24298 \cdot 10^{-16}$ leading to 27 terms and accuracy away from singularities of about 15 digits. As an improvement over parametric techniques of [10,17,18], we note that, away from singularities of the function, our approach achieves the desirable features of uniform L^∞-type approximation.

5. On rational approximation of functions

So far we constructed our rational representations assuming that we have knowledge of either the values of the Fourier transform of a function on some interval or, alternatively, a limited number of its Fourier coefficients. In this section we consider a related approximation problem, where we control the size of the Fourier interval or the number of the Fourier coefficients. Specifically, given a compactly supported function with integrable singularities, we are interested in constructing a rational approximation and quantifying its properties given its Fourier transform in a bandlimited interval. A theoretical foundation of our approach lies in interpreting results in [5] as an extension of the theory developed by Krein et al. in [1–3]. We plan to address this problem at length elsewhere but, in this section, we present some observations for the purpose of illustration.

Let us show that by an appropriate selection of bandlimit, our approach yields an accurate rational approximation away from singularities. Let f be a real, compactly supported, bounded function. We assume that f or its derivatives have a finite number of jump discontinuities. Let us describe a collection of points (within the support of f) surrounded by intervals of length $2d$ where the function has at least J bounded derivatives. We denote by $\mathcal{S}_R(J,d)$ all points such that

$$\mathcal{S}_R(J,d) = \left\{ x^* \in \operatorname{supp}(f) \colon \sup_{\substack{x \in (x^*-d, x^*+d) \\ j=0,\dots,J}} \left| \frac{f^{(j)}(x)}{j!} \right| \leqslant R \right\}.$$

Assuming that we know $\hat{f}$ in the interval $[0,a]$, our algorithm constructs an approximation of $\hat{f}$ by a sum of decaying exponentials $\hat{g}$ as in (4) so that

$$\left| \hat{f}(\xi) - \hat{g}(\xi) \right| \leqslant \epsilon,$$

for $\xi \in [-a,a]$. As is described previously, we use g, the inverse Fourier transform of $\hat{g}$, as our rational approximation to f. In order to elucidate the resulting approximation away from the singularities of f, we recast our construction as that of using an auxiliary window (filter) specific to the function f. We claim that there exists a real, even function $w \in C^\infty(\mathbb{R})$, such that its Fourier transform $\hat{w} \in C^\infty(\mathbb{R})$ satisfies

$$\left| \hat{f}(\xi)\hat{w}(\xi) - \hat{g}(\xi) \right| \leqslant 2\epsilon \quad \text{for } \xi \in [-a,a] \tag{12}$$

and

$$\left| \hat{f}(\xi)\hat{w}(\xi) - \hat{g}(\xi) \right| \leqslant Ce^{-\tau\xi} \quad \text{for } |\xi| > a, \tag{13}$$

where τ is the smallest real part of η_m in (4). Observe that we may choose $\hat{w}(\xi) = 1$ on an interval around zero to assure vanishing moments of w,

$$\int_{-\infty}^{\infty} w(y) y^j \, dy = \delta_{j0}, \tag{14}$$

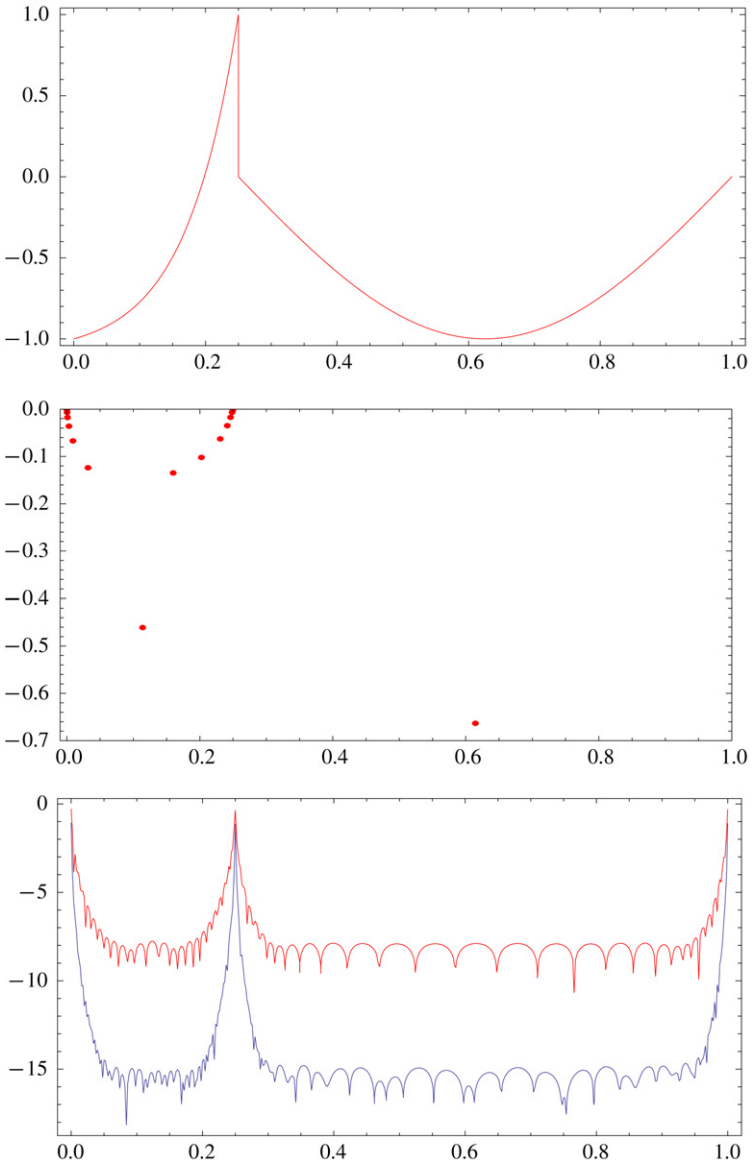

Fig. 9. The function in (11) (top). Locations of poles in our reconstruction using 63 Fourier coefficients of the function (middle). The error of reconstruction (on $\log_{10}$ scale) using 63 (upper curve) and 127 (lower curve) Fourier coefficients (bottom).

for $j = 0, 1, \ldots, J$. Furthermore, since w is in the Schwartz class, let us define its essential support as an interval $[-\mu, \mu]$ such that

$$\left| \int\limits_{|y| \geqslant \mu} y^j w(y)\, dy \right| \leqslant \epsilon \tag{15}$$

for $j = 1, \ldots, J$ and

$$\int\limits_{|y| \geqslant \mu} \left| w(y) \right| dy \leqslant \epsilon. \tag{16}$$

Let us consider a point $x^* \in \mathcal{S}_R(J, d)$ and estimate $|f(x^*) - g(x^*)|$ using an auxiliary window $w(y)$ and $\mu \leqslant d$ in (15) and (16).

Table 1
Poles and residues for the rational approximation of cubic B-spline in (19). Additional poles and residues are obtained using $\mathcal{R}e(x_{15+j}) = -\mathcal{R}e(x_j)$, $\mathcal{I}m(x_{15+j}) = \mathcal{I}m(x_j)$, $\mathcal{R}e(r_{15+j}) = -\mathcal{R}e(r_j)$, and $\mathcal{I}m(r_{15+j}) = \mathcal{I}m(r_j)$, $j = 1, \ldots, 11$.

m	Poles x_m	Residues r_m
1	$-1.997053825738 - 0.041848574989i$	$-9.514094765917 \cdot 10^{-8} + 2.949432764529 \cdot 10^{-7}i$
2	$-1.985093835924 - 0.116247747709i$	$-7.398150516982 \cdot 10^{-6} + 1.085431852158 \cdot 10^{-5}i$
3	$-1.942318359793 - 0.249192697117i$	$-2.153244131832 \cdot 10^{-4} + 1.135567803227 \cdot 10^{-4}i$
4	$-1.818635287314 - 0.451483609710i$	$-2.650691827111 \cdot 10^{-3} - 7.203148810670 \cdot 10^{-4}i$
5	$-1.523713033381 - 0.684208929095i$	$-4.513365554706 \cdot 10^{-3} - 2.124200560033 \cdot 10^{-2}i$
6	$-0.9967286925548 - 0.0333709389124i$	$2.054545128838 \cdot 10^{-7} - 4.459263040031 \cdot 10^{-7}i$
7	$-0.9850608105725 - 0.0867942045402i$	$1.064196040376 \cdot 10^{-5} - 1.035546311131 \cdot 10^{-5}i$
8	$-0.9664953591574 - 0.1635030210839i$	$9.294751012114 \cdot 10^{-5} - 1.106384074126 \cdot 10^{-4}i$
9	$-0.9662797372494 - 0.3188343019160i$	$7.644802613666 \cdot 10^{-4} - 2.371705375296 \cdot 10^{-3}i$
10	$-0.8701051439415 - 0.8240451711313i$	$0.2090163034563 - 0.1095421067525i$
11	$-0.2301421634957 - 1.0390253662364i$	$0.4669140825381 + 0.3676149993496i$
12	$-0.3989519124449i$	$9.498041476592 \cdot 10^{-3}i$
13	$-0.1899831616952i$	$4.587246718446 \cdot 10^{-4}i$
14	$-0.08695259008980i$	$2.317706263913 \cdot 10^{-5}i$
15	$-0.03208572889710i$	$6.199165261451 \cdot 10^{-7}i$

Denoting by $\tilde{f}$ the convolution of w and f and using that w is an even function, we have

$$\tilde{f}(x) = \int_{-\infty}^{\infty} f(x-y)w(y)\,dy = \int_{-\infty}^{\infty} f(x+y)w(y)\,dy,$$

and

$$\tilde{f}(x^*) - f(x^*) = \int_{|y|<\mu} \big(f(x^*+y) - f(x^*)\big)w(y)\,dy + \int_{|y|\geqslant\mu} \big(f(x^*+y) - f(x^*)\big)w(y)\,dy. \tag{17}$$

Using the Taylor expansion of f and (14), the first term in (17) equals

$$\sum_{j=1}^{J-1} \frac{f^{(j)}(x^*)}{j!}\left(-\int_{|y|\geqslant\mu} y^j w(y)\,dy\right) + \int_{|y|<\mu} \frac{f^{(J)}(\alpha)}{J!} y^J w(y)\,dy$$

for some α in the interval $(x^*-\mu, x^*+\mu) \subseteq (x^*-d, x^*+d)$. Let us also assume that

$$\left|\int_{|y|<\mu} \frac{f^{(J)}(\alpha)}{J!} y^J w(y)\,dy\right| \leqslant R\epsilon, \tag{18}$$

which is the case if f is a piecewise polynomial of degree less than J.

Thus, by definition of $\mathcal{S}_R(J,d)$, and assumptions (15), (16) and (18), we bound

$$\big|\tilde{f}(x^*) - f(x^*)\big| \leqslant JR\epsilon + 2\|f\|_\infty \epsilon.$$

On the other hand, choosing the right balance between parameters ϵ, a and τ, we find that $\tilde{f}$ is close to g. In fact, we have

$$\big|\tilde{f}(x^*) - g(x^*)\big| \leqslant \int_{|\xi|\leqslant a} \big|\hat{f}(\xi)\hat{w}(\xi) - \hat{g}(\xi)\big|\,d\xi + \int_{|\xi|>a} \big|\hat{f}(\xi)\hat{w}(\xi) - \hat{g}(\xi)\big|\,d\xi,$$

so that by using (12) and (13), we arrive at

$$\big|\tilde{f}(x^*) - g(x^*)\big| \leqslant 4a\epsilon + \frac{2}{\tau} e^{-a\tau}.$$

5.1. Rational approximation of the cubic B-spline

As an illustration, let us construct a rational approximation of the cubic B-spline,

$$\beta(x) = \begin{cases} \frac{3}{4}|x|^3 - \frac{3}{2}x^2 + 1, & |x| \leqslant 1, \\ \frac{1}{4}(2-|x|)^3, & 1 < |x| \leqslant 2, \\ 0, & 2 < |x|. \end{cases}$$

Using samples of its Fourier transform,

$$h_n = \hat{\beta}\left(\frac{n}{16}\right), \quad n = 0, \ldots, 400,$$

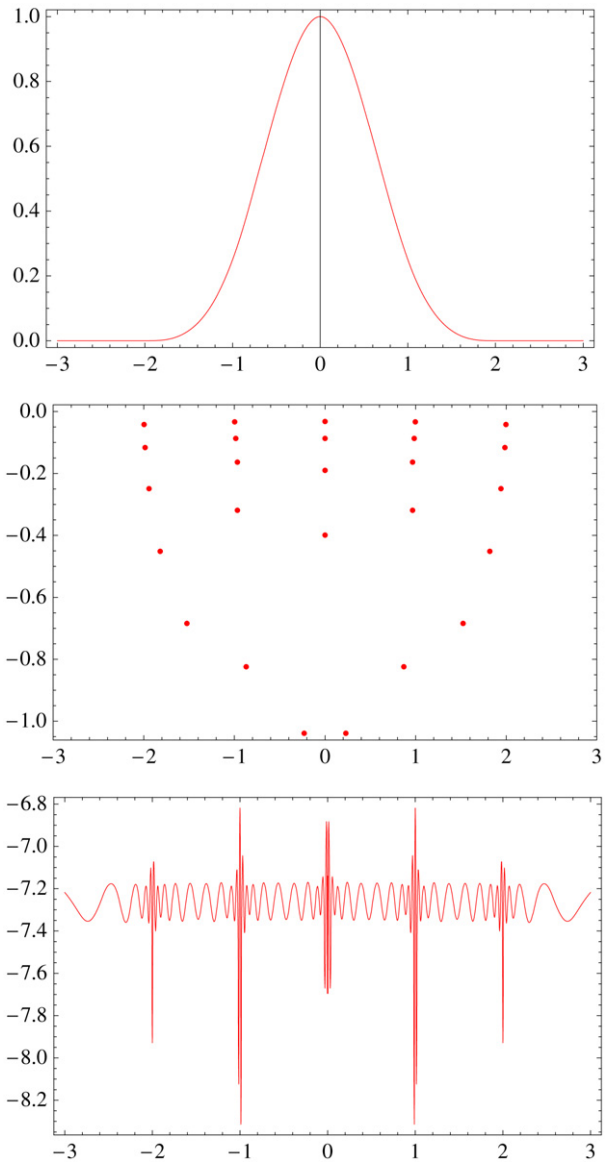

Fig. 10. Cubic B-spline considered in the interval $[-3, 3]$ (top). Locations of the 26 poles (in the lower half-plane) of its rational approximation (middle). Log-error of its reconstruction over the interval $[-3, 3]$ (bottom). We aligned the figures to illustrate the arrangement of the poles in the complex plane vs. singularities of the function being approximated.

where

$$\hat{\beta}(\xi) = \frac{3}{2}\left(\frac{\sin \pi\xi}{\pi\xi}\right)^4,$$

we approximate $\beta(x)$ by a rational function

$$g(x) = \mathrm{Re}\left(\sum_{m=1}^{M} \frac{r_m}{x - x_m}\right), \tag{19}$$

where, in this example, $M = 26$ and poles x_m and residues r_m are given in Table 1. The positions of the poles and the error of approximation are shown in Fig. 10. In this example, samples h_n were approximated with maximum absolute error $\approx 2.5 \cdot 10^{-8}$ leading to an approximation of $\hat{\beta}$ anywhere in the interval $[0, 25]$ with maximum absolute error $\approx 6 \cdot 10^{-7}$and that of β on $[-3, 3]$ with maximum absolute error $\approx 1.5 \cdot 10^{-7}$.

5.2. Rational approximation of functions with integrable singularities

Although we have considered piecewise polynomials, our approach is not limited to constructing rational approximations of such functions. In fact, by interpreting one of our examples in [5], we obtain a rational approximation of

$$f(x) = \begin{cases} \frac{2}{\sqrt{1-x^2}}, & |x| < 1, \\ 0, & \text{elsewhere.} \end{cases}$$

Indeed, we have

$$\hat{f}(\xi) = \left(\frac{2}{\sqrt{1-x^2}} 1_{[-1,1]}(x) \right)\hat{}\,(\xi) = 2\pi \mathrm{J}_0(2\pi\xi),$$

and in [5] we show how to efficiently approximate the Bessel function J_0 by a near optimal sum of exponentials. The number of terms needed for such approximation increases only logarithmically with the accuracy sought and with the size of the interval where the approximation is valid.

Using the samples,

$$h_n = \hat{f}\left(\frac{n}{15}\right) = 2\pi \mathrm{J}_0\left(2\pi \frac{n}{15}\right), \quad n = 0, \ldots, 180,$$

we select the singular value $\sigma_{18} \approx 7.28541 \cdot 10^{-9}$ leading to an approximation with maximum absolute error $\epsilon \approx 3.76 \cdot 10^{-9}$. We display our results in Fig. 11. At the top we display the function $f(x)$ on the interval $[-2, 2]$. In the middle, we display the locations of the 18 poles of its rational approximation pointing to the singularities and, at the bottom, the approximation error. In this case, the maximum absolute error away from the singularities is exactly ϵ.

6. An application to estimation of density of states

As an application, we propose to use our approach within the so-called kernel polynomial method (KPM) for estimating the density of states (eigenvalues) of a Hermitian operator, see e.g. [15,16]. The purpose of the method is to estimate the density of states without explicitly computing the eigenvalues. The KPM requires only matrix-vector multiplication and, thus, it is a method of choice if matrices representing the operator are large but, perhaps, sparse. However, since the density of states may have a discontinuous first derivative, the straightforward application of KPM leads to Gibbs phenomenon and we suggest our approach as an alternative to the use of filtering [15,16].

Using KPM, we first scale the given operator $\mathcal{H}$ as $\mathcal{X} = a\mathcal{H} + b\mathcal{I}$, so that the spectrum of $\mathcal{X}$ is within the interval $[-1, 1]$. The density of states is defined as

$$D(x) = \frac{1}{N} \sum_{n=1}^{N} \delta(x - x_n), \tag{20}$$

where x_n are the eigenvalues of $\mathcal{X}$ and N is the total number of eigenvalues counting their multiplicity. The density of states may be written as

$$D(x) = \frac{1}{\pi\sqrt{1-x^2}} \left(\mu_0 + 2 \sum_{m=1}^{\infty} \mu_m T_m(x) \right), \tag{21}$$

where T_m are the Chebyshev polynomials and μ_m are

$$\mu_m = \int_{-1}^{1} T_m(x) D(x)\, dx = \frac{1}{N} \int_{-1}^{1} \sum_{n=1}^{N} \delta(x - x_n) T_m(x)\, dx = \frac{1}{N} \operatorname{Tr}\big(T_m(\mathcal{X})\big),$$

where Tr denotes the trace of the matrix. An approximation to $D(x)$ is then obtained by truncating the series (21) and estimating the coefficients μ_m by, for example, computing

$$\tilde{\mu}_m = \frac{1}{L} \sum_{r} r^t T_m(\mathcal{X}) r, \tag{22}$$

using L Gaussian random vectors r (see [15]). The algorithm takes advantage of the three term recurrence available for the Chebyshev polynomials,

$$T_{m+1}(\mathcal{X}) = 2\mathcal{X} T_m(\mathcal{X}) - T_{m-1}(\mathcal{X}),$$

and, thus, requires only matrix vector multiplication to compute the approximate coefficients in (22).

Since the density of states $D(x)$ may have a discontinuous first derivative, the coefficients $\tilde{\mu}_m$ may have slow decay. Consequently, KPM also incorporates the windowing of $\tilde{\mu}_m$ to avoid Gibbs phenomenon as in [15,16].

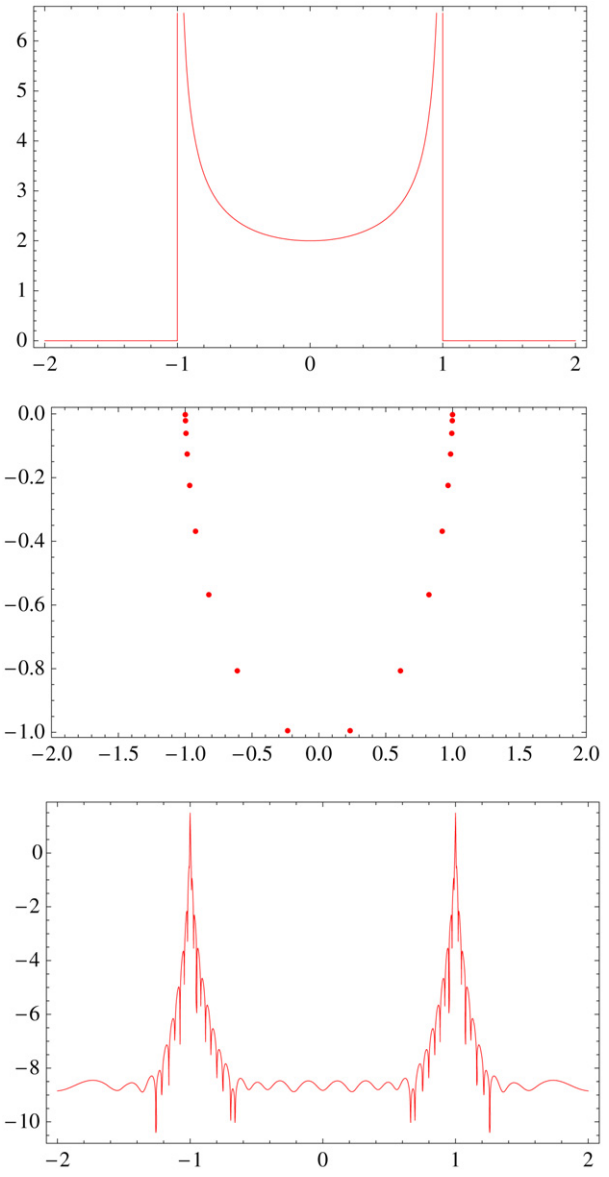

Fig. 11. The function $\frac{2}{\sqrt{1-x^2}} 1_{[-1,1]}(x)$ considered in the interval $[-2, 2]$ (top). Locations of the 18 poles in our reconstruction if we use 181 Fourier coefficients of the function (middle). Log-error of its reconstruction over the interval $[-2, 2]$ (bottom).

We note that μ_m are the coefficients of the periodic function

$$d(\theta) = \sin\theta D(\cos\theta) \tag{23}$$

in the expansion

$$d(\theta) = \frac{1}{\pi}\left(\mu_0 + 2\sum_{m=1}^{\infty} \mu_m \cos(m\theta)\right), \quad 0 < \theta < \pi,$$

which follows by changing variables $x = \cos\theta$ in

$$\int_0^{\pi} d(\theta)\cos m\theta\, d\theta = \int_{-1}^{1} T_m(x) D(x)\, dx = \mu_m.$$

Thus, by considering the computed $\tilde{\mu}_m$ in (22) as Fourier coefficients of the function $d(\theta)$ in (23), we may apply the algorithm of Section 4 to obtain an approximation to $d(\theta)$ for $0 < \theta < \pi$ and, therefore, to $D(x)$ for $-1 < x < 1$. As a result,

we not only obtain an accurate approximation of $D(x)$ but also an a posteriori estimate of the noise level that we read off the change in the rate of decay of the singular values of the corresponding Hankel matrix.

7. Conclusions

The nonlinear inversion of the bandlimited Fourier transform of this paper avoids Gibbs phenomenon without resorting to traditional windowing or more sophisticated filtering techniques. We show how to apply our method when the measured or computed data are either values of the Fourier transform or coefficients of the Fourier series. The nonlinear approximation of the Fourier data by a sum of exponentials is near optimal, i.e. requires a near minimal number of terms, and, hence, yields an efficient rational representation. The poles of the rational representation arrange themselves in patterns indicating location of singularities of the function. In the examples where the underlying function is known, we observe that the error of approximation away from singularities behaves as that of uniform L^∞ approximation, which we believe follows from interpreting results in [5] as an extension of the theory in [1–3]. One of the conclusions is that, in the absence of noise, the accuracy of approximation is only limited by the available bandwidth and may be achieved by selecting the appropriate singular value of corresponding Hankel matrix.

We foresee many practical applications of our method as an alternative to usual data windowing/filtering techniques. As an example, we consider KPM method for estimating the density of states. We also anticipate applications in solving a variety of inverse problems where the input data may be reduced to samples of the Fourier transform. An important example of such situation arises in Magnetic Resonance Imaging. However, this application requires a much more detailed analysis than the length of this article permits and we plan to address it elsewhere. To this end, we are working on a higher dimensional extension of our method.

Acknowledgments

This research was partially supported by NSF grant DMS-0612358, DOE/ORNL grant 4000038129, DOE grant DE-FG02-03ER25583 and AFOSR grant FA9550-07-1-0135.

References

[1] V.M. Adamjan, D.Z. Arov, M.G. Kreĭn, Infinite Hankel matrices and generalized Carathéodory–Fejér and I. Schur problems, Funktsional. Anal. i Prilozhen. 2 (4) (1968) 1–17.
[2] V.M. Adamjan, D.Z. Arov, M.G. Kreĭn, Infinite Hankel matrices and generalized problems of Carathéodory–Fejér and F. Riesz, Funktsional. Anal. i Prilozhen. 2 (1) (1968) 1–19.
[3] V.M. Adamjan, D.Z. Arov, M.G. Kreĭn, Analytic properties of the Schmidt pairs of a Hankel operator and the generalized Schur–Takagi problem, Mat. Sb. (N.S.) 86 (128) (1971) 34–75.
[4] G. Beylkin, L. Monzón, On generalized Gaussian quadratures for exponentials and their applications, Appl. Comput. Harmon. Anal. 12 (3) (2002) 332–373.
[5] G. Beylkin, L. Monzón, On approximation of functions by exponential sums, Appl. Comput. Harmon. Anal. 19 (1) (2005) 17–48.
[6] W. Cai, D. Gottlieb, C.-W. Shu, On one-sided filters for spectral Fourier approximations of discontinuous functions, SIAM J. Numer. Anal. 29 (4) (1992) 905–916.
[7] K.S. Eckhoff, Accurate reconstructions of functions of finite regularity from truncated Fourier series expansions, Math. Comp. 64 (210) (1995) 671–690.
[8] A. Gelb, A hybrid approach to spectral reconstruction of piecewise smooth functions, J. Sci. Comput. 15 (3) (2000) 293–322.
[9] A. Gelb, E. Tadmor, Detection of edges in spectral data, Appl. Comput. Harmon. Anal. 7 (1) (1999) 101–135.
[10] A. Gelb, J. Tanner, Robust reprojection methods for the resolution of the Gibbs phenomenon, Appl. Comput. Harmon. Anal. 20 (1) (2006) 3–25.
[11] D. Gottlieb, C.-W. Shu, On the Gibbs phenomenon and its resolution, SIAM Rev. 39 (4) (1997) 644–668.
[12] H.N. Mhaskar, J. Prestin, On the detection of singularities of a periodic function, Adv. Comput. Math. 12 (2000) 95–131.
[13] H.N. Mhaskar, J. Prestin, Polynomial operators for spectral approximation of piecewise analytic functions, Appl. Comput. Harmon. Anal. 26 (2009) 121–142.
[14] B. Porat, A Course in Digital Signal Processing, John Wiley, 1997.
[15] R.N. Silver, H. Roder, Calculation of densities of states and spectral functions by Chebyshev recursion and maximum entropy, Phys. Rev. E 56 (4) (1997) 4822–4829.
[16] S. Sota, M. Itoh, Fast and accurate scheme for Green functions and eigenvectors: Regulated polynomial expansion without Gibbs oscillation, J. Phys. Soc. Jpn. 76 (5) (2007).
[17] E. Tadmor, Filters, mollifiers and the computation of the Gibbs phenomenon, Acta Numer. 16 (2007) 305–378.
[18] E. Tadmor, J. Tanner, Adaptive mollifiers for high resolution recovery of piecewise smooth data from its spectral information, Found. Comput. Math. 2 (2) (2002) 155–189.