Acta Numerica (2007), pp. 305–379
DOI: 10.1017/S096249290631001X

Filters, mollifiers and the computation of the Gibbs phenomenon

Eitan Tadmor
Department of Mathematics, Institute for Physical Science & Technology
and
Center for Scientific Computation and Mathematical Modeling (CSCAMM)
University of Maryland, College Park, MD 20742, USA
E-mail: tadmor@cscamm.umd.edu

We are concerned here with processing discontinuous functions from their spectral information. We focus on two main aspects of processing such piecewise smooth data: detecting the edges of a piecewise smooth f, namely, the location and amplitudes of its discontinuities; and recovering with high accuracy the underlying function in between those edges. If f is a smooth function, say analytic, then classical Fourier projections recover f with exponential accuracy. However, if f contains one or more discontinuities, its global Fourier projections produce spurious Gibbs oscillations which spread throughout the smooth regions, enforcing local loss of resolution and global loss of accuracy. Our aim in the computation of the Gibbs phenomenon, is to detect edges and to reconstruct piecewise smooth f's, while regaining the high accuracy encoded in the spectral data.

To detect edges, we utilize a general family of edge detectors based on *concentration kernels*. Each kernel forms an approximate derivative of the delta function, which detects edges by *separation of scales*. We show how such kernels can be adapted to detect edges with one- and two-dimensional discrete data, with noisy data, and with incomplete spectral information. The main feature is concentration kernels which enable us to convert global spectral moments into local information in physical space. To reconstruct f with high accuracy we discuss novel families of *mollifiers* and *filters*. The main feature here is making these mollifiers and filters *adapted* to the local region of smoothness while increasing their accuracy together with the dimension of the data. These mollifiers and filters form approximate delta functions which are properly parameterized to recover f with (root-) exponential accuracy.

CONTENTS

1. Introduction

We are interested in processing piecewise smooth functions from their spectral information. The prototype example will be one-dimensional functions that are smooth except for finitely many jump discontinuities. The location and amplitudes of these discontinuities are not correlated. Thus, a piecewise smooth f is in fact a collection of several intervals of smoothness which do *not* communicate among themselves. The jump discontinuities can be viewed as the edges of these intervals of smoothness. Similarly, two-dimensional piecewise smooth functions consist of finitely many edges which lie along simple curves, separating two-dimensional local regions of smoothness. We are concerned here with two main aspects of processing such piecewise smooth data.

(i) Edge detection. Detecting the location and amplitudes of the edges. Often, these are the essential features sought in piecewise smooth data. Moreover, they define the regions of smoothness and are therefore essential for the second aspect.

(ii) Reconstruction. Recovering the underlying function f inside its different regions of smoothness.

There are many classical algorithms to detect edges and reconstruct the data in between those edges, based on *local* information. For example, suppose that the values of a one-dimensional f are given at equidistant grid-points, $f_\nu = f(\nu\Delta x)$. Then, the first-order differences, $\Delta f_\nu := f_{\nu+1} - f_\nu$, can detect edges where $\Delta f_\nu = \mathcal{O}(1)$, by separating them from smooth regions where $\Delta f_\nu = \mathcal{O}(\Delta x)$. Similarly, piecewise linear interpolants can recover the point-values of $f(x)$ up to order $\mathcal{O}\big((\Delta x)^2\big)$. Of course, these are only

asymptotic statements that may greatly vary with the dependence of the $\mathcal{O}$-terms on the *local* smoothness of f in the immediate neighborhood of x. We may do better, therefore, by taking higher-order differences, $\Delta^r f_\nu$, where $\mathcal{O}(1)$-edges are better separated from $\mathcal{O}\big((\Delta x)^r\big)$-regions of smoothness, Similarly, reconstruction of f using r-order approximations, with $r = 2, 3, \ldots$ and so on. In practice, higher accuracy is translated into higher resolution extracted from the information on a given grid. But, as the order of accuracy increases, the stencils involved become wider and one has to be careful not to extract smoothness information *across* edges, since different regions separated by edges are completely independent of each other. An effort to extract information from one region of smoothness into another one, will result in spurious oscillations, spreading from the edges into the surrounding smooth regions, preventing uniform convergence. This is, in general terms, the *Gibbs phenomenon*, which is the starting point of the present discussion.

The prototype for spectral information we are given on f, is the set of its N Fourier coefficients, $\{\widehat{f}(k)\}_{|k|\le N}$. These are *global* moments of f. It is well known that the Fourier projection, $S_N f = \sum_{|k|\le N} \widehat{f}(k) e^{ikx}$ forms a highly accurate approximation of f provided that f is *sufficiently smooth*. In Section 2 we revisit the classical spectral convergence statements and quantify the *actual* exponential accuracy of Fourier projections,

$$|S_N f - f| \lesssim e^{-\eta \sqrt[\alpha]{N}}.$$

Here, the root exponent α is tied to *global* smoothness of f of order $\alpha \ge 1$. But this high accuracy is lost with piecewise smooth f's, due to spurious oscillations which are formed around the edges of f. It is in this context of the Fourier projections, that the formation of spurious oscillations became known as *the* Gibbs phenomenon, originating with Gibbs' letter (Gibbs 1899). This is precisely because of the global nature of $S_N f$, which extracts smoothness information *across* the internal edges of f. The Gibbs' phenomenon is also responsible for a *global* loss of accuracy: first-order oscillations spread *throughout* the regions of smoothness. The highly accurate content in the spectral data, $\{\widehat{f}(k)\}_{|k|\le N}$, is lost in the Fourier projections, $S_N f$. The local and global effects of Gibbs oscillations are illustrated through a simple example in Section 3.

Our aim in the computation of the Gibbs phenomenon, is to detect edges and reconstruct piecewise smooth functions, while regaining the high accuracy encoded in their spectral data. Here, we use two main tools.

(i) Concentration kernels. To detect edges, we employ a fairly general framework based on partial sums of the form

$$K_N^\sigma f(x) := \frac{\pi i}{c_\sigma} \sum_{|k|\le N} sgn(k)\sigma\Big(\frac{|k|}{N}\Big)\widehat{f}(k)e^{ikx}, \qquad c_\sigma := \int_0^1 \frac{\sigma(\xi)}{\xi} d\xi.$$

In Section 4 we show that $K_N^\sigma f(x)$ approximates the *local* jump function, $K_N^\sigma f(x) \approx f(x+) - f(x-)$. Consequently, $K_N^\sigma f$ tends to concentrate near edges, where $f(x+) - f(x-) \neq 0$, which are separated from smooth regions where $K_N^\sigma f \approx 0$. We can express $K_N^\sigma f(x)$ as a convolution with the Fourier projection of f, that is,

$$K_N^\sigma f(x) = \mathcal{K}_N^\sigma * (S_N f)(x), \qquad \mathcal{K}_N^\sigma(x) := -\frac{1}{c_\sigma} \sum_{k=1}^{N} \sigma\Big(\frac{|k|}{N}\Big) \sin kx.$$

Here, $\mathcal{K}_N^\sigma(x)$ are the corresponding *concentration kernels* which enable us to convert the global moments of $S_N f$ into local information about its edges — both their locations and their amplitudes. The choice of concentration factor, σ, is at our disposal. In Section 5 we discuss a few prototype examples of concentration factors and assess the different behavior of the corresponding edge detectors, $K_N^\sigma f$. In Section 6 we turn to a series of *extensions* which show how concentration kernels apply in more general set-ups. In Section 6.1 we discuss the *discrete* framework, applying concentration kernels as edge detectors in the Fourier interpolants, $I_N f = \sum_{|k|\le N} \widehat{f}_k e^{ikx}$. In Section 6.2 we show how concentration kernels can be used to detect edges in *non-periodic* projections, $S_N f = \sum \widehat{f}(k) C_k(x)$, based on general Gegenbauer expansions. In Section 6.3 we show how the concentration factors could be adjusted to deal with *noisy data*, by taking into account the noise variance, $\eta \gg 1/N$, in order to detect the underlying $\mathcal{O}(1)$-edges. Finally, Section 6.4 deals with *incomplete data*: we show how concentration kernels based on partial information, $\{\widehat{f}(k)\}_{k\in K}$, can be complemented by a compressed sensing approach to form effective edge detectors.

Concentration kernels, $\mathcal{K}_N^\sigma(x)$, are approximate *derivatives* of the delta function. Convolution with such kernels, $\mathcal{K}_N^\sigma * (S_N f)$ yield edge detectors by *separation of scales*, separating between smooth and nonsmooth parts of f. In Section 7 we show how to improve the edge detectors by *enhancement* of this separation of scales. In particular, in Section 7.1 we use nonlinear *limiters* which assign low- and high-order concentration kernels in regions with different characteristics of smoothness. The result is *parameter-free*, high-resolution edge detectors for one-dimensional piecewise smooth functions.

In Section 8 we turn to the two-dimensional set-up. Concentration kernels can be used to separate scales in the x_1 and x_2 directions. Enhancements and limiters are shown in Section 8.1 to greatly reduce, though not completely eliminate, the Cartesian staircasing effect. In Section 8.2 we show how concentration kernels are used to detect edges from incomplete two-dimensional data. So far, we have emphasized the role of separation of scales in edge detectors based on concentration kernels, $\mathcal{K}_N^\sigma * (S_N f)(\mathbf{x})$. But how do we actually *locate* those $\mathcal{O}(1)$ edges? In Section 8.3 we discuss the

approach which seeks the zero-level set $\mathbf{x} = (x_1, x_2)$ of $\nabla_{\mathbf{x}} \mathcal{K}_N^\sigma * (S_N f)(\mathbf{x})$. Depending on our choice of the concentration factors, $\sigma(\cdot)$, this leads to a large class of two-dimensional edge detectors which generalize the popular two-dimensional zero-crossing method associated with discrete Laplacian stencils.

Next, we turn our attention to highly-accurate, Gibbs-free *reconstruction* of f inside its regions of smoothness from its (pseudo-)spectral content.

(ii) Mollifiers and filters. We consider two interchangeable processes to recover the values of a piecewise smooth $f(x)$ with high accuracy. These are *mollification*, carried out in the physical space, and *filtering* — carried out in the Fourier space, i.e.,

$$\Phi_{p,\delta} * (S_N f)(x) \longleftrightarrow \sum_{|k|\le N} \varphi_{p,\delta}\Big(\frac{|k|}{N}\Big)\widehat{f}(k)e^{ikx}.$$

Filtering accelerates convergence when premultiplying the Fourier coefficients by a *rapidly decreasing* $\varphi_{p,\delta}(|k|/N)$. The rapid decay of $\varphi_{p,\delta}(|k|/N)\widehat{f}(k)$ as $|k| \uparrow N$ in Fourier space, corresponds to mollification with *highly localized* mollifiers, $\Phi_{p,\delta}(x)$, in physical space.

Section 10 is devoted to mollifiers. There are two free parameters at our disposal. The parameter δ is chosen so that the essential support of $\Phi_{p,\delta} * (S_N f)(x)$ does *not* cross edges of f. To this end we set δ as the distance to the nearest edge, $\delta = d_x := dist\{x, singsuppf\}/\pi$, so that $(x - \pi\delta, x + \pi\delta)$ is the largest interval of smoothness enclosing x. It is here that we use the information on the location of the edges of f. This leads to *adaptive* mollifiers $\Phi_{p,d_x}(x)$. The parameter p is responsible for the *accuracy* of the mollifier. By properly tunning $p = p_N$ to increase with N, one obtains the *root-exponential* accurate mollifiers discussed in Section 10.2:

$$\Phi_{p_N,d_x}(x) := \frac{1}{d_x}\varphi\Big(\frac{\pi x}{d_x}\Big)D_{p_N}\Big(\frac{x}{d_x}\Big), \quad d_x = \frac{1}{\pi}dist\{x, singsuppf\},\ p_N \sim d_x N.$$

Here,

$$D_p(x) := \frac{\sin(p + 1/2)x}{2\pi \sin(x/2)}$$

is the Dirichlet kernel of order p which ensures accuracy by having an increasing number of (almost) vanishing moments, $\int y^n D_{p_N}(y)dy \approx 0$ for $p = 1, 2 \ldots, p_N$, and $\varphi = \varphi_{2q}$ is a proper $C_0^\infty(-1, 1)$ cut-off function[1], e.g.,

$$\varphi_{2q}(y) := e^{\left(\frac{y^{2q}}{y^2 - 1}\right)} 1_{(-1,1)}(y),$$

[1] C_0^∞ denotes the space of compactly supported C^∞-smooth functions.

which ensures that Φ_{p_N,d_x} are properly localized within the d_x-neighborhood of the origin. The result is an adaptive mollifier with *root-exponential accuracy*

$$|\Phi_{p_N,d_x} * (S_N f)(x) - f(x)| \lesssim e^{-\eta\sqrt{d_x N}}.$$

The corresponding root-exponential discrete mollifier is outlined in Section 10.3. The high accuracy of these mollifiers is adapted to the *interior* points, away from the edges where $d_x N \sim 1$. It can be modified to gain polynomial accuracy *up to* the edges. This is described in Section 10.4. In Section 10.5 we discuss mollifiers based on *Gegenbauer expansion* of $S_N f(\pi x)$, with uniform root-exponential accuracy *up to* the edges.

Section 11 is devoted to filters of the form

$$\mathcal{S}^{\varphi}_{p_N} f(x); = \sum_{|k|\le N} \varphi_{p_N}\Big(\frac{|k|}{N}\Big)\widehat{f}(k)e^{ikx}.$$

In Section 11.1 we show that by setting $p_N \sim \sqrt{d_x N}$, the resulting filter is accurate (and hence its associated mollifier satisfies a moment condition) to order p_N. Moreover, the choice of the filter φ_{p_N} yields a highly localized mollifier which is essentially supported in the smoothness interval $(x - \pi d_x, x + \pi d_x)$. This yields the root-exponential convergence rate

$$|\mathcal{S}^{\varphi}_{p_N} f(x) - f(x)| \lesssim e^{-\eta\sqrt{d_x N}}.$$

We conclude, in Section 11.2, revisiting the construction of the adaptive filters and mollifiers with a better localization procedure. Instead of enforcing the compactly supported φ_{2q}'s in either the physical or Fourier space, we appeal to *optimally* space-frequency localized filters

$$\varphi_p(\xi) = \varphi_{p,\delta}(\xi) := e^{-\frac{(\delta\xi)^2}{2}} \sum_{j=0}^{p} \frac{1}{2^j j!}(\delta\xi)^{2j}.$$

We show that an adaptive parameterization, $p = p_N \sim d_x N$ and $\delta_x \sim \sqrt{d_x N}$, yields the *exponentially accurate* mollifier

$$|\mathcal{S}^{\varphi}_{p_N,\delta_x} f(x) - f(x)| \lesssim e^{-\eta d_x N}.$$

There is a rich literature on filters and mollifiers as effective tools for Gibbs-free reconstruction of piecewise smooth functions. Different aspects of this topic are drawn from a variety of sources, ranging from summability methods in harmonic analysis to signal processing – and, in recent years, image processing – and high-resolution spectral computations of propagation of singularities and shock discontinuities. Given the space and time limitations, we are unable to provide a complete road map but we limit ourselves to a few

key references. For general references on harmonic analysis we refer to (Bary 1964, Dym and McKean 1972, Folland 1992, Grafakos 2004, Katznelson 1976, Kröner 1995, Stein 1993, Szego 1958, Torchinsky 1986) and (Zygmund 1959). For applications in signal processing, including wavelets, and recent exciting developments in compressed sensing, we mention (Candes and Romberg 2006, Candes, Romberg, and Tao 2006a, Candes, Romberg and Tao 2006b, Donoho, Elad and Temlyakov 2006, Donoho 2004, Donoho and Tanner 2005, Mallat 1989, Marr and Hildreth 1980, Tao 2005) and the references therein. General reviews on spectral methods, edge detection and the computation of Gibbs phenomenon can be found in (Abarbanel, Gottlieb and Tadmor 1986, Fornberg 1996, Gelb and Tadmor 2000b, Gelb and Gottlieb 2007, Gottlieb and Hesthaven 1998, Gottlieb and Orszag 1977, Gottlieb and Shu 1997, Mhaskar and Prestin 2000, Majda, McDonough and Osher 1978, Tadmor 1989, Maday, Ould-Kaber and Tadmor 1993) and (Tadmor 1994). We also had to leave out numerous other approaches for edge detection and reconstruction of piecewise smooth data. We mention for example, (Eckhoff 1995, Eckhoff 1998, Eckhoff and Wasberg 1995, Bruno, Han and Pohlman 2006, Srinivasa and Rajgopal 1992) and most notably, two-dimensional reconstructions which couple Radon representation with spectral and wavelet-based ridgelets, e.g., (Donoho 1998) and (Candes and Guo 2002).

2. Spectral accuracy

2.1. The spectral Fourier projection

Let $S_N f$ denote the Fourier projection of a 2π-periodic function,

$$S_N f(x) = \sum_{|k|\le N} \widehat{f}(k) e^{ikx}, \qquad \widehat{f}(k) := \frac{1}{2\pi}\int_{-\pi}^{\pi} f(y) e^{-iky} dy. \tag{2.1}$$

It enjoys the well-known *spectral accuracy*, that is, the decay rate of the error, $S_N f - f$, is as rapid as the global smoothness of $f(\cdot)$ permits. The error in this case amounts to the *truncation* error,

$$T_N f(x) := \sum_{|k|>N} \widehat{f}(k) e^{ikx},$$

which is spectrally small in the sense that for *any* $s > 1$ we have

$$|S_N f(x) - f(x)| \le \sum_{|k|>N} |\widehat{f}(k)| \lesssim \|f\|_{C^s} \cdot \frac{1}{N^{s-1}}, \qquad \text{for all } s > 1. \tag{2.2}$$

Here

$$\|f\|_{C^s} := \max_{k\le s} \|f^{(k)}(\cdot)\|_{L^\infty},$$

measures the *global* smoothness of f. The interplay between the global smoothness of f and spectral convergence of its Fourier projection is reflected through Parseval's relation,

$$\|f\|^2_{H^s} = 2\pi \sum_k \big(1+|k|^{2s}\big)|\widehat{f}(k)|^2, \quad \|f\|^2_{H^s} := \int_{-\pi}^{\pi} \big(f(y)\big)^2 + \big(f^{(s)}(y)\big)^2 dy,$$

which in turn, is linked to the *spectral decay* of the Fourier coefficients,

$$|\widehat{f}(k)| \lesssim \|f\|_{C^s} \frac{1}{1+|k|^s}, \qquad s \geq 1. \tag{2.3}$$

Indeed, the latter follows by noting $\|f\|_{C^{s-1}} \lesssim \|f\|_{H^s} \lesssim \|f\|_{C^s}$, or by repeated integration by parts.

The spectral decay rate (2.3) and its related convergence rate (2.2) are asymptotic statements. The *actual* decay rate (as functions of k and N) depends on the growth of $\|f\|_{C^s}$. To quantify the precise spectral accuracy of C^∞ functions, it is therefore convenient to classify such functions according to the growth of their C^s-bounds: f belongs to *Gevrey class* $G_\alpha, \alpha \geq 1$, if there exists $\eta = \eta_f > 0$ such that

$$G_\alpha = \Big\{ f \;\Big|\; \|f\|_{C^s} \lesssim \frac{(s!)^\alpha}{\eta^s}, \qquad s = 1, 2 \ldots \Big\}. \tag{2.4}$$

Two examples of Gevrey classes are in order.

(i) *Analytic functions.* By Cauchy's integral formula, analytic f's belong to G_1, with $2\eta_f$ being the width of their analyticity strip.

(ii) The C^∞_0 *cut-off functions*,

$$\rho_p(x) := e^{\Big(\frac{cx^p}{x^2-\pi^2}\Big)} 1_{(-\pi,\pi)}(x), \qquad c > 0, \quad p \text{ even}, \tag{2.5}$$

belong to G_2. Indeed, a straightforward computation shows that there exists a constant, $\lambda = \lambda_\rho$ (which may depend on p but is otherwise independent of s), such that

$$|\rho_2^{(s)}(x)| \lesssim \frac{s!}{(\lambda_\rho |x^2-\pi^2|)^s} e^{\Big(\frac{cx^p}{x^2-\pi^2}\Big)}, \tag{2.6}$$

and the upper bound on the right, which is maximized at $x = x_{max}$ with $x^2_{max} - \pi^2 \sim -\pi^2 c/s$, implies the G_2-bound (2.4) with $\eta = c\lambda_\rho \pi^2$,

$$\sup_{x\in(-1,1)} |\rho_2^{(s)}(x)| \lesssim s! \Big(\frac{s}{\eta}\Big)^s e^{-s} \lesssim \frac{(s!)^2}{\eta^s}, \quad s = 1, 2, \ldots .$$

We can now combine the spectral decay (2.3) with the G_α-bound (2.4).

It follows that the decay rate of the Fourier coefficients of G_α-functions is exponential to a fractional order,

$$|\widehat{f}(k)| \lesssim \min_s \Big(\frac{s^\alpha}{\eta e^\alpha |k|}\Big)^s \sim e^{-\alpha(\eta|k|)^{1/\alpha}}, \qquad f \in G_{\alpha\geq 1},$$

and consequently the truncation error of their Fourier projection does not exceed,

$$|S_N f(x) - f(x)| \lesssim N e^{-\alpha(\eta N)^{1/\alpha}}, \qquad f \in G_{\alpha\geq 1}.$$

In particular, an analytic $f(\cdot)$, with analyticity strip of width 2η, is *characterized* by an exponential rate corresponding to $\alpha = 1$, (see e.g., (Tadmor 1986)), that is,

$$|\widehat{f}(k)| \lesssim e^{-\eta|k|}, \qquad |S_N f(x) - f(x)| \lesssim N e^{-\eta N}, \qquad f \text{ analytic;} \tag{2.7a}$$

while for G_2-functions, such as the cut-off $\rho_p(x)$, for example, the rate is *root-exponential*, corresponding to $\alpha = 2$:

$$|\widehat{f}(k)| \lesssim e^{-\sqrt{\eta|k|}}, \qquad |S_N f(x) - f(x)| \lesssim N e^{-\sqrt{\eta N}}, \qquad f \in G_2. \tag{2.7b}$$

Remark 2.1. (Notations). Throughout the paper, we use η to denote different Gevrey constants of fractional-exponential orders. The same η in different equations stands for different constants. In section 6.3 η is also used to denote the noise variance.

2.2. Optimal space-frequency decay

The previous examples tell us that if f is a C^∞ compactly supported function, then $|\widehat{f}(k)|$ decay at an exponential rate of a *fractional* order but no faster; indeed, if the $|\widehat{f}(k)|$'s decay exponentially fast then f is analytic and hence it cannot decay sufficiently fast to become compactly supported. The question of optimal joint decay in both physical and Fourier spaces brings us to the classical Heisenberg *uncertainty principle*, which places a lower threshold on the joint space-frequency localization. This lower-threshold manifests itself in a variety of different forms. In the context of the Fourier transform for example, one seeks to minimize the joint variance:

$$\min_{x_0} \|(x-x_0)\Phi(x)\|_{L^2(\mathbb{R}_x)} \cdot \min_{\xi_0} \|(\xi-\xi_0)\varphi(\xi)\|_{L^2(\mathbb{R}_\xi)}, \quad \Phi(x) := \int_{\mathbb{R}} \varphi(\xi) e^{-i\xi x} d\xi.$$

It admits a lower threshold which is achieved by the *quadratic* exponentials $\varphi(\xi) = e^{c(\xi-\xi_0)^2}$. For space-frequency localization in related discrete frameworks we mention recent examples of (Donoho and Huo 2001), (Tao 2005) and (Candes and Romberg 2006). In the present context of Fourier expansions, we now construct a large family of 2π-periodic functions, $\{f_N(x)\}$, with optimal *exponential decay* in both physical and Fourier space; consult (Hoffman and Kouri 2000) and the references therein.

The starting point is the family of functions with quadratic exponential decay

$$\varphi(\xi) := e^{-\frac{\xi^2}{2}} \times \left[\sum_{j=0}^{p} \frac{1}{2^j j!}\xi^{2j}\right]. \tag{2.8a}$$

Their inverse Fourier transform can be expressed in terms of Hermite polynomials, $H_{2j}(x)$, that is,

$$\Phi(x) = e^{-\frac{x^2}{2}} \times \left[\sum_{j=0}^{p} \frac{(-1)^j}{4^j j!} H_{2j}\Big(\frac{x}{\sqrt{2}}\Big)\right]. \tag{2.8b}$$

Observe that with fixed p, each $\Phi(x)$ is an *entire* function and the quadratic exponential decay of its Fourier transform, $\varphi(\xi)$, corresponds to '$\eta_\Phi = \infty$'.

We need to 'tweak' $\Phi(x)$ in two ways.

(1) Dilation. We need to dilate $\Phi(x)$ in order to control its localization,

$$\Phi_\delta(x) := \frac{1}{\delta}\Phi\Big(\frac{x}{\delta}\Big) \longleftrightarrow \varphi_\delta(\xi) = \varphi(\delta\xi)$$

(2) Periodization. We need a 2π-periodic version of the entire function $\Phi(x)$. To this end, fix N and set[2]

$$\mathcal{S}_N^\varphi(x) := \frac{1}{2\pi}\sum_{k=-\infty}^{\infty} \varphi\Big(\frac{|k|}{N}\Big)e^{ikx}. \tag{2.9a}$$

Another way to express this 'periodization' of Φ is given by the *Poisson summation formula*, (see e.g., (Katznelson 1976), (Torchinsky 1986))

$$\mathcal{S}_N^\varphi(x) = \frac{N}{2\pi}\sum_{j=-\infty}^{\infty} \Phi(N(x+2\pi j)). \tag{2.9b}$$

We can combine both dilation and periodization into one scaling involving N/δ,

$$\mathcal{S}_N^{\varphi_\delta}(x) = \frac{1}{2\pi}\sum_{k=-\infty}^{\infty} \varphi_\delta\Big(\frac{|k|}{N}\Big)e^{ikx} = \frac{1}{2\pi}\sum_{k=-\infty}^{\infty} \varphi\Big(\frac{\delta|k|}{N}\Big)e^{ikx} \equiv \mathcal{S}_{N/\delta}^\varphi(x).$$

We are ready to state our next result.

[2] Observe that the *function* $\mathcal{S}_N^\varphi(x)$ is different from the partial sum *operation*, S_N. The reason for this notation will become clear when we link these two different aspects in our discussion on mollifiers and filters in Section 11.

Lemma 2.2. (Space-frequency exponential decay) Fix p, set $\delta_N := \sqrt{\beta N}$ and consider the 2π-periodic functions,

$$f_N(x) := \mathcal{S}_N^{\varphi_{\delta_N}}(x) = \frac{1}{2\pi}\sum_{k=-\infty}^{\infty} \varphi\Big(\sqrt{\frac{\beta}{N}}|k|\Big)e^{ikx}, \quad \varphi(\xi) = e^{-\frac{\xi^2}{2}}\sum_{j=0}^{p}\frac{\xi^{2j}}{2^j j!}. \tag{2.10}$$

Then, there exists $\eta_1, \eta_2 > 0$ such that, for all $|x| \le \pi$,

$$|f_N(x)| = \Big|\mathcal{S}_N^{\varphi_{\sqrt{\beta N}}}(x)\Big| \lesssim 2^p\sqrt{\frac{N}{\beta}}\left(e^{-\eta_1 N x^2/\beta} + e^{-\eta_2 N/\beta}\right). \tag{2.11a}$$

Moreover, for all $|k| > N$,

$$|\widehat{f}_N(k)| \;\lesssim\; c_{p,N}\cdot e^{-\beta|k|/2}, \qquad c_{p,N} := \sum_{j=0}^{p}\frac{1}{j!}\Big(\frac{\beta N}{2}\Big)^j. \tag{2.11b}$$

Remark 2.3. We conclude, since $c_{p,N}$ has at most pth-order polynomial growth with N, that f_N should have exponential decay in both physical and frequency space. Observe that the detailed structure of the pth-order polynomial factors inside the square brackets on the right of (2.8a) and (2.8b) are not important at this stage, but they will be later on, in Section 11.2 below, when we link the increase of p with N.

Proof. To verify (2.11a) we bound $|H_{2j}(x)| \lesssim j^{j+\frac{1}{2}}(4/e)^j e^{x^2/2}$, in order to estimate the exponential decay of $\Phi_\delta(x) = \frac{1}{\delta}\Phi\big(\frac{x}{\delta}\big)$ in (2.8b),

$$|\Phi_\delta(x)| \lesssim \frac{1}{\delta}e^{-\frac{x^2}{2\delta^2}}\sum_{j=0}^{p}\frac{1}{4^j j!}\cdot\Big|H_{2j}\Big(\frac{x}{\sqrt{2}\delta}\Big)\Big| \lesssim \frac{2^p}{\delta}e^{-\frac{x^2}{4\delta^2}}. \tag{2.12}$$

We appeal to the Poisson representation of $\mathcal{S}_N^{\varphi_\delta}$ in (2.9b)

$$\mathcal{S}_N^{\varphi_\delta}(x) = \frac{N}{2\pi}\Phi_\delta(Nx) + \frac{N}{2\pi}\sum_{j\neq 0}\Phi_\delta(N(x+2\pi j)). \tag{2.13}$$

It follows that all term except the zeroth are exponentially negligible for $|x| \le \pi$,

$$\begin{aligned}\sum_{j\neq 0}|\Phi_{\delta_N}(N(x+2\pi j))| \;&\lesssim\; \frac{2^p}{\delta_N}\sum_{j=1}^{\infty}e^{-\frac{((2j-1)\pi N)^2}{4\delta_N^2}}\\ &\lesssim\; \frac{2^p}{\sqrt{\beta N}}e^{-\eta_2 N/\beta}, \qquad |x|\le\pi.\end{aligned} \tag{2.14a}$$

The zeroth term has double exponential decay (in x)

$$|\Phi_{\delta_N}(Nx)| \lesssim \frac{2^p}{\sqrt{\beta N}} e^{-\eta_1 N x^2/\beta}, \tag{2.14b}$$

and the last two bounds yield the first half of (2.11).

The second half of (2.11) is straightforward:

$$|\widehat{f_N}(k)| = \Big|\varphi\Big(\sqrt{\frac{\beta}{N}}|k|\Big)\Big| \lesssim c_{p,N} \cdot e^{-\Big(\frac{\beta|k|^2}{2N}\Big)},$$

$$c_{p,N} = \max_{\xi \le \delta_N} \sum_{j=0}^{p} \frac{1}{2^j j!} \xi^{2j}, \tag{2.15}$$

and (2.11b) follows for $|k| > N$. ∎

2.3. The pseudo-spectral Fourier projection

If we replace the integrals on the right of (2.1) with the quadrature sampled at the equidistant points,

$$y_\nu := -\pi + \nu h, \qquad h := \frac{2\pi}{2N+1},$$

we obtain the *discrete* Fourier coefficients $\Big\{\widehat{f_k}\Big\}$, which form the N-degree *trigonometric interpolant* of f at these $(2N+1)$ grid-points, that is[3],

$$I_N f(x) := \sum_{|k| \le N} \widehat{f_k} e^{ikx}, \qquad \widehat{f_k} := \frac{h}{2\pi} \sum_{\nu=0}^{2N} f(y_\nu) e^{-iky_\nu}, \tag{2.16}$$

so that $I_N f(x_\nu) = f(x_\nu),\ \nu = 0, \dots 2N$. I_N is known as the *pseudo-spectral* projection. The dual statement of interpolation in physical space is the Poisson summation formula in Fourier space, expressing the $\widehat{f_k}$'s in terms of the $\widehat{f}(k)$'s,

$$\widehat{f_k} = \widehat{f}(k) + \sum_{j \ne 0} \widehat{f}(k + j(2N+1)). \tag{2.17}$$

Thus, the sum of all the Fourier coefficients located at $k[mod(2N+1)]$ have a discrete "alias", $\widehat{f_k}$. This follows at once by substituting $f(y_\nu)$ in (2.16) as the sum $\sum_j \widehat{f}(j) e^{iky_\nu}$. We conclude that the interpolation error

[3] There is a slight difference between the formulae based on an even and an odd number of points; we have chosen to continue with the slightly simpler notation associated with an odd number of points.

$I_N f(x) - f(x)$ consists of two contributions: the truncation error, $T_N f(x) = \sum_{|k|\geq N} \widehat{f}(k) e^{ikx}$, and the *aliasing error*,

$$A_N f(x) = \sum_{|k|\leq N} \Big(\sum_{|j|\geq 1} \widehat{f}(k + j(2N+1)) \Big) e^{ikx}. \tag{2.18}$$

Both $T_N f$ and $A_N f$ involve modes higher than N and if f is sufficiently smooth they have exactly the same spectrally small size, (e.g., (Tadmor 1994)), that is,

$$\begin{aligned} \|A_N f(x)\|_{H^s} &\sim \sum_{|k|\geq N} \big(1 + |k|^{2s}\big) \big| \sum_{j\neq 0} \widehat{f}(k + j(2N+1)) \big|^2 \\ &\leq C_s \sum_{|k|\geq N} \big(1 + |k|^{2s}\big) |\widehat{f}(k)|^2 \lesssim C_s \|T_N f(x)\|_{H^s}^2, \quad s > \frac{1}{2}. \end{aligned} \tag{2.19}$$

We conclude with similar spectral and fractional-exponential convergence rate estimates:

$$\begin{aligned} &|I_N f(x) - f(x)| \\ &\leq \sum_{|k|>N} |\widehat{f}(k)| + \sum_{|k|\leq N} \sum_{|j|\geq 1} |\widehat{f}(k + j(2N+1))| \lesssim \begin{cases} \frac{1}{N^{s-1}}, & f \in C^s \\ \\ N e^{-(\eta N)^{1/\alpha}} & f \in G_\alpha. \end{cases} \end{aligned}$$

We close this section by commenting on the discrete Fourier coefficients of the exponentially localized f_N's in (2.10). The quadratic exponential decay of $\widehat{f_N}(k)$ for $|k| > N$ in (2.15), implies that the aliasing error, $A_N f_N(x)$, is exponentially negligible, and hence $(\widehat{f_N})_k \approx \widehat{f_N}(k)$, that is,

$$|(\widehat{f_N})_k| \lesssim c_{p,N} \cdot \left(e^{-\frac{\beta k^2}{2N}} + \mathcal{O}\Big(e^{-\frac{\beta N}{2}}\Big) \right), \qquad |k| \leq N. \tag{2.20}$$

3. Piecewise smoothness and the Gibbs phenomenon

Both the spectral and the pseudo-spectral Fourier projections, $S_N f$ and $I_N f$, provide highly accurate approximations of f, whose order is limited only by the *global* smoothness of f. What happens when f lacks sufficient smoothness? This will be our main concern in the remaining sections.

We begin with a classical example. Consider an f which is *piecewise smooth* in the sense that it is sufficiently smooth except for finitely many jump discontinuities, say at $x = c_1, c_2, \ldots, c_J$, where

$$[f](c_j) := f(c_j+) - f(c_j-) \neq 0, \quad j = 1, 2, \ldots J.$$

It is natural to measure piecewise smoothness in the space of functions of

bounded variation,

$$\|f\|_{BV} := \|f'\|_{L^1[-\pi,\pi]} < \infty,$$

that is, f' is a smooth function together with finitely many Dirac masses. The finite variation of f implies (via integration by parts) the first-order decay rate of $|\widehat{f}(k)|$,

$$|\widehat{f}(k)| \lesssim \|f\|_{BV} \frac{1}{1+|k|}. \tag{3.1}$$

Indeed, the decay is precisely first order, $|\widehat{f}(k)| \sim 1/|k|$, since a faster decay would imply that f is continuous, (Zygmund 1959). A similar first-order decay occurs in the discrete case: summation by parts of the discrete Fourier coefficients in (2.16) yields

$$\begin{aligned} \widehat{f}_k &= \frac{h}{2\pi} \sum_{\nu=0}^{2N} f(y_\nu) \frac{e^{-iky_\nu} - e^{-iky_{\nu+1}}}{1 - e^{-ikh}} \\ &= \frac{h}{4\pi i \sin\frac{kh}{2}} \sum_{\nu=0}^{2N} \Big(f(y_{\nu+1}) - f(y_\nu) \Big) e^{-iky_\nu}, \end{aligned} \tag{3.2}$$

and hence $\|\widehat{f}_k| \lesssim \|f\|_{TV}/(1+|k|)$, where $\|f\|_{TV}$ denotes the total variation of f.

The first order decay of the (discrete) Fourier coefficients is too weak (as it should be!) to enforce *uniform* convergence of $S_N f(x)$ and $I_N f(x)$. Instead, we turn to examine the *local* convergence of $S_N f(x)$, which is expressed in terms of the *Dirichlet kernel*, $D_N(\cdot)$,

$$S_N f(x) = \int_{-\pi}^{\pi} f(y) D_N(x-y), \quad D_N(y) := \frac{1}{2\pi} \sum_{|k|\le N} e^{iky} \equiv \frac{\sin(N+\frac{1}{2})y}{2\pi \sin(y/2)}.$$

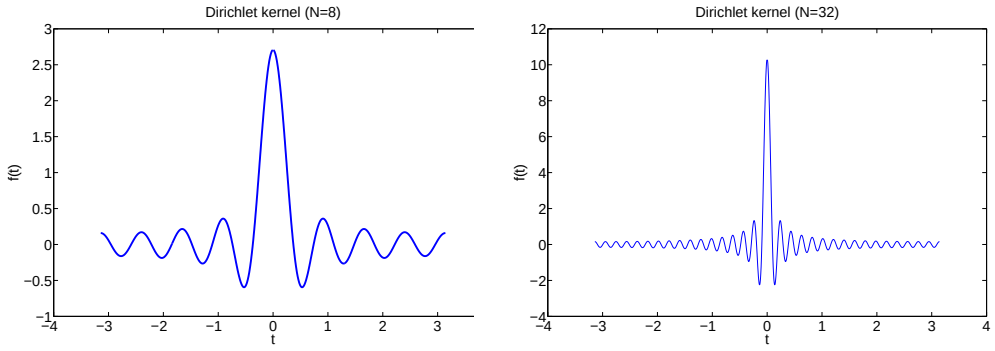

Figure 3.1: Dirichlet kernel with $N = 32$ and $N = 128$ modes.

The Dirichlet kernel $D_N(\cdot)$, is depicted in figure 3.1. It has a sequence of successive local peaks at

$$\frac{(k+1/2)\pi}{N+1/2},\ k=1,2,\dots$$

which accumulate a total mass of diverging order $\|D_N\|_{L^1} \sim \log N$ and which in turn, are responsible for the failure of uniform convergence of $S_N f(x)$. As an example, consider the spectral projection of the Heaviside function $H(x) := sgn(x)$. Since $D_N(\cdot)$ is an even function, we have

$$\begin{aligned} S_N H(x) &= \int_0^\pi \Big[D_N(x-y) - D_N(x+y)\Big]dy = \\ &= \frac{-i}{\pi}\sum_{|k|\le N} e^{ikx}\int_0^\pi \sin(ky)dy = \frac{-2i}{\pi}\sum_{\{|k|\le N: k \text{ odd}\}}\frac{e^{ikx}}{k}. \end{aligned}$$

At $x=0$, we find that $S_N H(x)$ assumes the average value,

$$S_N H(x)\big|_{x=0} = 0.$$

But the convergence is not uniform for $x \approx 0$: for example

$$S_N H(x)\big|_{x=\pm\frac{\pi}{N}} = \frac{\pm 2}{\pi}\sum_{\{1\le k\le N: k \text{ odd}\}}\frac{\sin(k\pi/N)}{k\pi/N}\cdot\frac{2\pi}{N} \approx \frac{\pm 2}{\pi}\int_0^\pi \frac{\sin y}{y}dy.$$

Thus, the spectral projection $S_N H(x)$ magnifies the amplitude of the original jump $[H](0)=2$, forming a 'spurious' oscillation with 18% larger amplitude:

$$S_N H(x)\big|_{x=\frac{\pi}{N}} - S_N H(x)\big|_{x=-\frac{\pi}{N}} \approx \frac{4}{\pi}\int_0^\pi \frac{\sin y}{y}dy = 1.179\times[H](0). \quad (3.3)$$

This behavior is called the *Gibbs phenomenon* after (Gibbs 1899) (consult (Carslaw 1952), (Kröner 1995) or (Hewitt and Hewitt 1979) for a historical perspective). The lack of uniform convergence is depicted in figure 3.2 by spurious oscillations which concentrate near the jumps at $x=0$ and $x=\pm\pi$. For another example, consult figure 5.1(b) below. This is a *local* effect of the Gibbs phenomenon. But the Gibbs phenomenon also has a *global* effect: although the error $S_N H(x) - H(x)$ decays as x moves away from the jumps, the decay rate is limited to first-order, owing to a series of linearly decaying spurious peaks at $k\pi/N$ where

$$\Big(S_N H(x) - H(x)\Big)\Big|_{x=\pm\frac{k\pi}{N}} \sim \frac{1}{N}.$$

Thus, the existence of one or more discontinuities slow down the convergence rate *throughout* the domain. Spectral accuracy is lost.

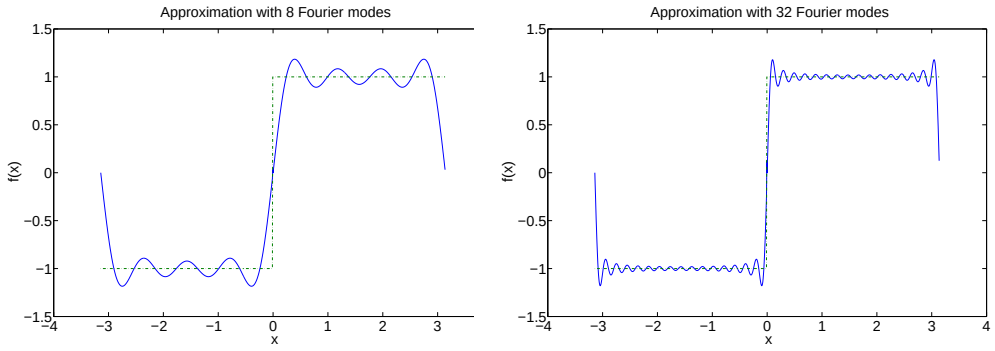

Figure 3.2: Fourier projection of the square-wave function. *Left*: $N = 32$ modes. *Right*: $N = 128$ modes.

4. Detection of edges: concentration kernels

Given the Fourier coefficients, $\{\widehat{f}(k)\}_{k=-N}^{N}$, we are interested in detecting the edges of the underlying piecewise smooth f, namely, to detect their location, $c_1, \ldots, c_J$ and the amplitudes of the jumps, $[f](c_1), \ldots, [f](c_J)$. Extensions to the discrete and non-periodic set-ups will follow in the next sections.

We begin by considering the prototype case of a discontinuous f which is, say, C^2, except for a single jump at $x = c$. The fact that f experiences a jump of size $[f](c)$ dictates the decay of its Fourier coefficients: integration by parts yields,

$$\widehat{f}(k) = [f](c)\frac{e^{-ikc}}{2\pi ik} + \mathcal{O}\left(\frac{1}{|k|^2}\right). \tag{4.1}$$

We want to extract information about the location of the jump from the *phase* of the leading term. To this end we use the localization of the Dirichlet kernel near the origin:

$$\frac{\pi}{N}D_N(x-c) = \frac{1}{2N}\sum_{k=-N}^{N} e^{ik(x-c)} \approx \frac{1}{1+N|x-c|}.$$

It follows that the *derivative* of the Fourier projection $S_N f$, satisfies

$$\begin{aligned}\frac{\pi}{N}S_N(f)'(x) &= \frac{\pi}{N}\sum_{|k|\le N} ik\left\{[f](c)\frac{e^{ik(x-c)}}{2\pi ik} + \mathcal{O}\left(\frac{1}{|k|^2}\right)\right\} \\ &= [f](c)\cdot\frac{1}{2N}\sum_{|k|\le N} e^{ik(x-c)} + \frac{\pi}{N}\sum_{|k|\le N}\mathcal{O}\left(\frac{1}{|k|}\right)\end{aligned}$$

$$= \frac{[f](c)}{1+N|x-c|} + \mathcal{O}\left(\frac{\log N}{N}\right) = \begin{cases} [f](c) + \mathcal{O}\left(\dfrac{\log N}{N}\right), & x \approx c \\ \mathcal{O}\left(\dfrac{\log N}{N}\right), & |x-c| \gg \frac{1}{N}. \end{cases}$$

We see that $\pi S_N(f)'(x)/N$ *concentrates* in the immediate neighborhood of $x = c$ where it approaches the desired amplitude of the jump, $[f](c) \neq 0$, while it decays to order $\mathcal{O}(\log N/N)$ as it moves away from this neighborhood, $|x-c| \gg 1/N$. Thus, we can detect the edge at $x = c$ by separation of scales: a jump of size $|[f](c)| \gg 1/N$ is separated from the region of smoothness where $\pi S_N(f)'(x)/N \approx 1/N$. This result goes back to Fejér (Zygmund 1959, Theorem 9.3).

We turn to consider a general set-up of edge detection based on separation of scales. To this end we introduce a family of *concentration kernels*

$$\mathcal{K}_N^\sigma(y) := -\frac{1}{c_\sigma} \sum_{k=1}^{N} \sigma\left(\frac{k}{N}\right) \sin ky, \qquad \frac{\sigma(\xi)}{\xi} \in C^2[0,1]. \tag{4.2a}$$

Here, c_σ is a normalization constant

$$c_\sigma := \int_0^1 \frac{\sigma(\xi)}{\xi} d\xi, \tag{4.2b}$$

so that, as will be shown in (4.6) below, $\int_0^\pi \mathcal{K}_N^\sigma(y)dy \approx -1$. We set[4]

$$K_N^\sigma f(x) := \mathcal{K}_N^\sigma * f(x) = \frac{\pi i}{c_\sigma} \sum_{|k|\leq N} sgn(k)\sigma\left(\frac{|k|}{N}\right) \widehat{f}(k) e^{ikx}. \tag{4.3}$$

Our purpose is to choose the *concentration factors*, $\sigma(|k|/N)$, such that $K_N^\sigma f$ detects the $\mathcal{O}(1)$-edges, $[f](c_j)$, $j = 1, \ldots, J$, by separating them from a much smaller scale of $K_N^\sigma f(x)$ in regions of smoothness. It turns out that *all* σ's can serve as admissible concentration factors.

Theorem 4.1. (Concentration kernels) (Gelb and Tadmor 1999, Gelb and Tadmor 2000a) Assume that $f(\cdot)$ is piecewise smooth such that

$$\omega_f(y) = \omega_f(y;x) := \frac{f(x+y) - f(x-y) - [f](x)}{y} \in BV[-\pi, \pi] \tag{4.4}$$

Let $\mathcal{K}_N^\sigma(x)$ be an admissible concentration kernel (4.2). Then,

$$K_N^\sigma f(x) = \mathcal{K}_N^\sigma * (S_N f)(x) = \frac{\pi i}{c_\sigma} \sum_{|k|\leq N} sgn(k)\sigma\left(\frac{|k|}{N}\right) \widehat{f}(k) e^{ikx}$$

[4] Observe that $K_N^\sigma f$ is the *operator* associated with, but otherwise different from, the concentration kernel $\mathcal{K}_N^\sigma(x)$.

satisfies the concentration property,

$$K_N^\sigma f(x) \sim \begin{cases} [f](c_j) + \mathcal{O}\Big(\frac{\log N}{N}\Big), & x \sim c_j,\ j = 1, \dots, J \\ \mathcal{O}\Big(\frac{\log N}{N}\Big), & dist\Big\{x, \{c_1, \dots, c_J\}\Big\} \gg \frac{1}{N}. \end{cases} \tag{4.5}$$

Remark 4.2. We will show below that, up to scaling and modulo small "manageable" residual terms, all the different $\mathcal{K}_N^\sigma(y)$'s amount to the same *conjugate Dirichlet kernel*,

$$\mathcal{K}_N^\sigma(y) \sim \frac{\sigma(1)}{c_\sigma}\widetilde{D}_N(y) + \text{lower order terms}, \qquad \widetilde{D}_N(y) := \frac{\cos(N+\frac{1}{2})y}{2\sin(y/2)}.$$

Accordingly, we refer to $K_N^\sigma f(x)$ as a *conjugate sum*. The lemma shows that all these conjugate sums concentrate near the edges. Different σ s yield different concentration kernels $\mathcal{K}_N^\sigma(y)$, and we will explore the role of different σ s in the following sections.

Proof. The key to our proof is to observe that $\mathcal{K}_N^\sigma$ is an approximate *derivative* of the delta function. In particular, since $\mathcal{K}_N^\sigma(\cdot)$ is odd

$$\begin{aligned} \mathcal{K}_N^\sigma * f(x) &= -\int_0^\pi K_N^\sigma(y)\Big(f(x+y) - f(x-y)\Big)dy \\ &= -\int_0^\pi \mathcal{K}_N^\sigma(y)\Big(f(x+y) - f(x-y) - [f](x)\Big)dy - [f](x) \times \int_0^\pi \mathcal{K}_N^\sigma(y)dy. \end{aligned}$$

The rectangular quadrature rule and the normalization (4.2b) yield

$$\begin{aligned} \int_0^\pi \mathcal{K}_N^\sigma(y)dy &= \frac{1}{c_\sigma}\sum_{k=1}^N \sigma\left(\frac{k}{N}\right)\frac{(-1)^k - 1}{k} = \\ &= -\frac{1}{c_\sigma}\sum_{k\ \text{odd}\geq 1}^N \frac{\sigma(\xi_k)}{\xi_k}\frac{2}{N} = -1 + \mathcal{O}\left(\frac{1}{N^2}\right), \quad \xi_k := \frac{k}{N}, \end{aligned} \tag{4.6}$$

and we end up with the error estimate

$$\big|K_N^\sigma f(x) - [f](x)\big| \lesssim \left|\int_0^\pi y\mathcal{K}_N^\sigma(y)\omega_f(y;x)dy\right| + \mathcal{O}\left(\frac{1}{N^2}\right). \tag{4.7a}$$

It remains to upperbound the first moment of $\mathcal{K}_N^\sigma\omega_f$. To this end we use the identity $-4\sin^2(y/2)\sin(ky) \equiv \sin\big((k+1)y\big) - 2\sin(ky) + \sin\big((k-1)y\big)$ and twice summation by parts to find

$$4\sin^2\big(\frac{y}{2}\big)c_\sigma\mathcal{K}_N^\sigma(y) \equiv 2\sigma(1)\sin\big(\frac{y}{2}\big)\cos\Big(N+\frac{1}{2}\Big)y$$

$$+\overbrace{\sum_{1\le k\le N-2}\Big(\sigma(\xi_k)-2\sigma(\xi_{k+1})+\sigma(\xi_{k+2})\Big)\sin(k+1)y}^{I_1(y)}$$
$$+\overbrace{\Big(\sigma(\xi_{N-1})-\sigma(1)\Big)\sin Ny}^{I_2(y)}+\overbrace{\Big(\sigma(\xi_2)-2\sigma(\xi_1)\Big)\sin y}^{I_3(y)}.$$

This leads to the corresponding decomposition of $\mathcal{K}_N^\sigma(y)$ as the sum of a conjugate Dirichlet kernel, $\widetilde{D}_N(y)$, plus a residual term, $R_N(y)$ (which is negligible in the precise sense to be outlined below),

$$\mathcal{K}_N^\sigma(y)=\frac{\sigma(1)}{c_\sigma}\overbrace{\frac{\cos(N+\frac12)y}{2\sin(y/2)}}^{\widetilde{D}_N(y)}+\frac{1}{c_\sigma}R_N(y),\quad R_N(y):=\sum_{j=1}^{3}\frac{I_j(y)}{4\sin^2(y/2)}. \tag{4.7b}$$

The conjugate Dirichlet kernel has a small moment due to *cancelation*. Indeed, if we let Ω_f denote

$$\Omega_f(y)\equiv\Omega(y;x):=\frac{y}{4\sin(y/2)}\omega_f(y;x),$$

then the upper-bound

$$\Big|\int_0^\pi y\widetilde{D}_N(y)\omega_f(y)dy\Big|=\Big|\int_0^\pi\cos\Big((N+\frac12)y\Big)\Omega_f(y)dy\Big|\lesssim\frac{\|\omega_f(\cdot)\|_{BV}}{N}, \tag{4.8a}$$

follows from (3.1), since $\|\Omega_f\|_{BV}\lesssim\|\omega_f\|_{BV}$. The logarithmic upper bound of the Dirichlet kernel, $\|D_k\|_{L^1}\sim\log k$, implies

$$\Big|\int_0^\pi y\frac{I_1(y)\omega_f(y)}{4\sin^2(y/2)}dy\Big|\ \lesssim\ \frac{1}{N^2}\|\sigma\|_{C^2}\sum_{k=1}^{N-2}\log k\cdot\|\Omega_f\|_{L^\infty}\lesssim\frac{\log N}{N}, \tag{4.8b}$$

$$\Big|\int_0^\pi y\frac{I_2(y)\omega_f(y)}{4\sin^2(y/2)}dy\Big|\ \lesssim\ \frac{1}{N}\|\sigma\|_{C^1}\cdot\log N\cdot\|\Omega_f\|_{L^\infty}\lesssim\frac{\log N}{N}. \tag{4.8c}$$

Finally, since $|\sigma(\xi)|\lesssim\xi$ we have

$$\Big|\int_0^\pi y\frac{I_3(y)\omega_f(y)}{4\sin^2(y/2)}dy\Big|\ \lesssim\ \Big(\Big|\sigma\Big(\frac1N\Big)\Big|+\Big|\sigma\Big(\frac2N\Big)\Big|\Big)\|\Omega_f\|_{L^\infty}\lesssim\frac1N, \tag{4.8d}$$

and the desired result, (4.5), follows from (4.7a),(4.7b) and (4.8). ■

We conclude this subsection with a couple of remarks.

Remark 4.3. (The behavior of σ and improved concentration) The bounds in (4.8) show that their the overall error does not exceed

$$\frac{\log N}{N}\|\sigma\|_{C^2}+\Big|\sigma\big(\frac1N\big)\Big|+\frac1N|\sigma(1)|. \tag{4.9}$$

Thus, the concentration error (4.5) of order $\mathcal{O}(1/N)$ becomes smaller if $\sigma(\xi)$ decays sufficiently fast at $\xi = 0$ and $\xi = 1$. This issue will be explored in the next section, in the context of the exponential concentration factors, consult (5.4) below.

Remark 4.4. (Concentration kernels: general set-up) The proof of theorem 4.1 reveals that the concentration property holds for arbitrary kernels, $\{\mathcal{K}_N(y)\}$, as long as they satisfy the three key properties:
(i) $\mathcal{K}_N$ are odd, $\mathcal{K}_N(-y) = -\mathcal{K}_N(y)$;
(ii) $\mathcal{K}_N$ are properly normalized so that $\int_{y\geq 0} \mathcal{K}_N(y)dy = -1 + \varepsilon_N$; and
(iii) $\mathcal{K}_N$ has a small first moment of order

$$\Big| \int y\mathcal{K}_N(y)\omega(y)dy \Big| \lesssim \varepsilon_N \|\omega\|_{BV}. \tag{4.10}$$

Here, ε_N is a small scale associated with K_N. If (i)-(iii) hold then we deduce, along the lines of theorem 4.1 (consult (Gelb and Tadmor 2000a, Theorem 2.1)

$$|\mathcal{K}_N * f(x) - [f](x)| \lesssim \varepsilon_N;$$

hence, $\mathcal{K}_N$ detect edges by separating, $|\mathcal{K}_N * f(c_j) \approx [f](c_j)$, from smooth regions where $\mathcal{K}_N * f(x) \approx \varepsilon_N \ll 1$. A few examples are in order.

5. Examples

Compactly supported kernels. We consider a standard mollifier, namely $\phi_{\varepsilon_N}(y) := \frac{1}{\varepsilon_N}\phi(\frac{y}{\varepsilon_N})$, based on an even, compactly supported bump function, $\phi \in C_0^1(-1,1)$ with $\phi(0) = 1$. We then set

$$\mathcal{K}_{\varepsilon_N}(y) = \frac{1}{\varepsilon}\phi'\left(\frac{y}{\varepsilon_N}\right). \tag{5.1}$$

Clearly, $\mathcal{K}_{\varepsilon_N}$ is an odd kernel satisfying the proper normalization

$$\int_{y\geq 0} \mathcal{K}_{\varepsilon_N}(y)dy = -\phi(0) = -1,$$

and its first moment is of order

$$\int_{y\geq 0} |y\mathcal{K}_{\varepsilon_N}(y)|dy = \varepsilon_N \int_0^1 |y| \cdot |\phi'(y)|dy = \mathcal{O}(\varepsilon_N).$$

The concentration property, $\mathcal{K}_{\varepsilon_N} * f(x) = [f](x) + \mathcal{O}(\varepsilon_N)$ follows. As examples we mention edge detectors based on Haar and bi-orthogonal moments; e.g., (Mallat 1989). Localized kernels are limited to finite order of accuracy, no matter how smooth f is.

Polynomial concentration kernels. Set $\sigma(\xi) = \xi$. Then $K_N^\sigma f$ recovers the Fejér conjugate sum

$$K_N^\sigma f(x) = \pi \sum_{|k|\le N} \frac{ik}{N} \widehat{f}(k) e^{ikx} = \frac{\pi}{N} S_N(f)'(x), \qquad \sigma(\xi) = \xi. \tag{5.2}$$

We note in passing that in this case, $\mathcal{K}_N^\sigma(y)$ does not concentrate near the origin as the compactly supported $\mathcal{K}_{\varepsilon_N}$'s do. Instead, (4.10) is fulfilled thanks to the more intricate property of *cancelation of oscillations*. This is the first member in the general family of *polynomial concentration factors*, $\sigma_p(\xi) = \xi^p$, introduced by Golubov (Golubov 1972, Kvernadze 1998, Gelb and Tadmor 1999). Polynomial concentration factors of odd degree, $\sigma_{2p+1}(\xi)$, correspond to differentiation in physical space, that is

$$K_N^{\sigma_{2p+1}} f(x) = (-1)^p \frac{\pi(2p+1)}{N^{2p+1}} \frac{d^{2p+1}}{dx^{2p+1}} S_N f(x), \qquad \sigma_{2p+1}(\xi) = \xi^{2p+1}.$$

Polynomial factors of even degree, $\sigma_{2p}(\xi)$, yield *global* conjugate sums which convolve

$$\widetilde{H}_N(x) := i \sum_{|k|\le N} sgn(k) e^{ikx},$$

that is,

$$K_N^{\sigma_{2p}} f(x) = (-1)^p \frac{2\pi p}{N^{2p}} \widetilde{H}_N * \frac{d^{2p}}{dx^{2p}} S_N f(x), \qquad \sigma_{2p}(\xi) = \xi^{2p}.$$

We shall refer to this family of kernels based on the σ_p-factors as *polynomial concentration kernels*.

Trigonometric concentration kernels. According to (3.3), the difference $S_N f(x+\pi/N) - S_N f(x-\pi/N)$ concentrates near the edges with 18% Gibbs overshoot

$$\frac{S_N f\big(x + \frac{\pi}{N}\big) - f\big(x - \frac{\pi}{N}\big)}{2Si(\pi)/\pi} \approx [f](x), \quad Si(\pi) := \int_0^\pi \frac{\sin x}{x} dx. \tag{5.3}$$

The difference in the numerator amounts to concentration factors $\sigma(\xi) = \sin(\pi\xi)$

$$S_N f(x+\pi/N) - S_N f(x-\pi/N) = 2i \sum_{|k|\le N} \sin\big(\frac{\pi k}{N}\big) \widehat{f}(k) e^{ikx},$$

and the corresponding normalization, $c_\sigma = Si(\pi)$, recovers the denominator in (5.3). This edge detector was advocated by (Banerjee and Geer 1997). It is the first member in the family of *trigonometric concentration factors* $\sigma_\alpha(\xi) = \sin(\alpha\xi)$. We shall have to say more on the relation between concentration factors in Fourier space and their realization as differences in the

physical space when we discuss edge detection in discrete data in Section 6.1.

Exponential concentration factors. Theorem 4.1 provides us with the framework of general concentration kernels which are not necessarily limited to a realization in the physical space. In particular, we seek concentration factors $\sigma(\cdot)$, which vanish at $\xi = 0, 1$ to any prescribed order:

$$\frac{d^j}{d\xi^j}\sigma(\xi)\Big|_{\xi=0} = \frac{d^j}{d\xi^j}\sigma(\xi)\Big|_{\xi=1} = 0, \qquad j = 0, 1, 2, \ldots, p. \tag{5.4}$$

The higher p is, the more localized $\mathcal{K}_N^\sigma(\cdot)$ becomes, since

$$\mathcal{K}_N^\sigma(y_\ell) = -\frac{1}{c_\sigma}\sum_{k=1}^{N}\sigma\left(\frac{k}{N}\right)\sin\frac{2\pi k\ell}{N}, \quad y_\ell := \frac{2\pi\ell}{N}.$$

We observe that $\mathcal{K}_N^\sigma(y_\ell)/N$ coincides with the ℓ-discrete Fourier coefficient of $\sigma(\cdot)$, and since $\sigma(\xi)$ and its first p-derivatives vanish at both ends, $\xi = 0, 1$, the C^p-regularity of σ implies the rapid decay of these discrete Fourier coefficients, $|\widehat{\sigma}_\ell| \lesssim \ell^{-p}$, i.e.,

$$|K_N^\sigma(y_\ell)| \lesssim \|\sigma\|_{C^p[0,1]}\frac{1}{(Ny_\ell)^p}.$$

Thus, for y away from the origin, $K_N^\sigma(y)$ is rapidly decaying for sufficiently large N's. Moreover, we can show that an increasing number of moments of $\mathcal{K}_N^\sigma(\cdot)$ vanish: consult (Gelb and Tadmor 2000a, §2). As an example, consider the *exponential concentration factors*,

$$\sigma_{exp}(\xi) = \xi e^{\frac{1}{\alpha\xi(\xi-1)}}, \tag{5.5}$$

for which (5.4) holds for *all* p's. Indeed, since $\sigma_{\exp}$ is based on a G_2 cut-off function, then $K_N^\sigma f$ becomes root-exponentially small away from the jumps,

$$|K_N^{\sigma_{exp}} f(x)| \lesssim e^{-\sqrt{\eta N}}, \qquad dist\Big\{x, \big\{c_1, \ldots, c_J\big\}\Big\} \gg \frac{1}{N}. \tag{5.6}$$

This leads to an improved separation of edges from regions of smoothness, demonstrated in figure 5.1.

Figure 5.1 compares the Fejér and exponential concentration kernels, $\mathcal{K}_N^{\sigma_1}(x)$ and $\mathcal{K}_N^{\sigma_{exp}}(x)$ for

$$f(x) := \begin{cases} \cos(x - \frac{x}{2}\mathrm{sgn}(|x| - \frac{\pi}{2})), & x < 0, \\ \\ \cos(\frac{5}{2}x + x\mathrm{sgn}(|x| - \frac{\pi}{2})), & x > 0. \end{cases} \tag{5.7}$$

In both cases, $K_N^\sigma f(x)$ concentrates near the two edges at $x = \pm\pi/2$ with

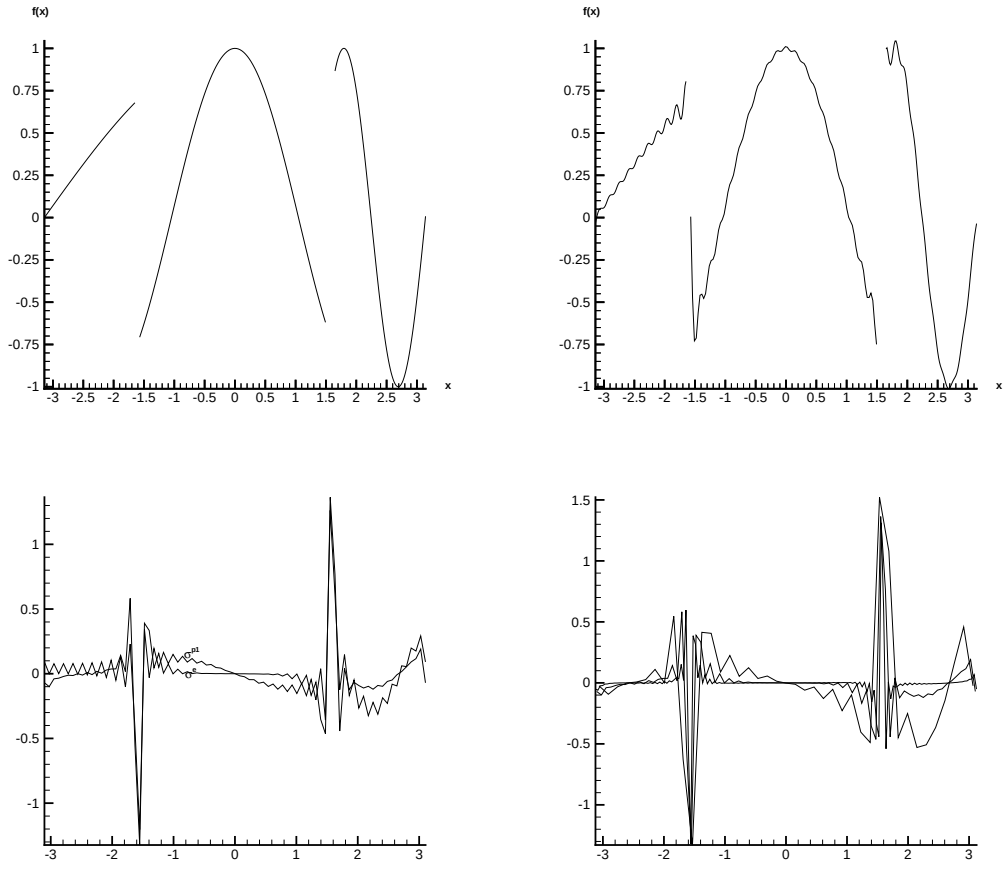

Figure 5.1: *Top left:* Piecewise smooth $f(x)$ in (5.7). *Top right:* Gibbs phenomenon for $S_{40}f(x)$. *Bottom left*: Edge detection in $S_{40}f$ using $K_{40}^{\sigma}f$, comparing the exponential concentration factors $\sigma_{exp}(\xi) = \exp(\frac{1}{6\xi(\xi-1)})$. *vs.* Fejér factors $\sigma(\xi) = \xi$. *Bottom right:* Exponential concentration $K_N^{\sigma_{exp}}f(x)$ with $N = 20, 40, 80$ modes. Observe that the root-exponential decay of $K_{80}^{\sigma_{exp}}f(x)$ becomes almost flat when x is well inside the intervals of smoothness of f, which are well separated from the neighborhoods of the edges.

amplitude $\pm\sqrt{2}$ which are separated from the remaining smooth pieces of $f(x)$. It confirms the improved localization of the exponential concentration factors.

6. Extensions

6.1. Discrete data

We are interested in the recovery of the location and amplitudes of the edges, $[f](c_j),\ j = 1, \dots, J$, from the discrete Fourier coefficients,

$$\widehat{f}_k = \frac{h}{2\pi} \sum_{\nu=0}^{2N} f(y_\nu) e^{-iky_\nu}, \quad h = \frac{2\pi}{2N+1}. \tag{6.1}$$

As before, the regularity of f is revealed by the decay rate of $\widehat{f}_k$: successive summation by parts implies the rapid decay of $\widehat{f}(k)$ for smooth f's, in analogy with (2.3), that is,

$$|\widehat{f}_k| \lesssim \sup_\nu \frac{|\Delta^s f(y_\nu)|}{h^s} \frac{1}{1+|k|^s}, \quad s \geq 1; \tag{6.2}$$

here $h^{-s}\Delta^s$ are the usual divided differences of order s. On the other hand, for the prototype case of an f which experiences a single jump at $x = c$, (3.2) yields

$$\widehat{f}_k = \Big(f(y_{\nu_c+1}) - f(y_{\nu_c})\Big) \frac{e^{-iky_{\nu_c}}}{2\pi ik} + \mathcal{O}\Big(\frac{1}{|k|^2}\Big),$$

where ν_c singles out the cell which encloses the location of the jump discontinuity.

In the discrete case, however, every gridvalue experiences a jump discontinuity: the jumps that are of order $\mathcal{O}(h)$ are acceptable as part of the smooth region, whereas the $\mathcal{O}(1)$ jumps indicate edges of the underlying function $f(x)$. Hence, in the discrete case we can identify a jump discontinuity at $x = c$ by its enclosed gridcell, $[x_{\nu_c}, x_{\nu_c+1}]$, which is characterized by the asymptotic statement

$$f(x_{\nu+1}) - f(x_\nu) = \begin{cases} [f](c) + \mathcal{O}(h), & \text{for } \nu = \nu_c \colon c \in [x_\nu, x_{\nu+1}], \\ \mathcal{O}(h), & \text{for other } \nu's \neq \nu_c. \end{cases} \tag{6.3}$$

Of course, this asymptotic statement (6.3) may itself serve as an edge detector based on the given *gridvalues*, $\{f(x_\nu)\}_{\nu=-N}^{N}$. Higher-order differences, $\Delta^p f(x_\nu)$, yield edge detectors involving increasingly larger, but finite, stencils, with improved separation between cells containing $\mathcal{O}(1)$-scale jumps and smaller, but finite, $\mathcal{O}(h^p)$-cells in regions of smoothness. We now

seek alternative edge detectors based on the discrete Fourier coefficients, $\{\widehat{f}_k\}_{|k|\le N}$. Using proper concentration factors, we shall cover both local and global edge detectors. For example, global edge detectors based on the exponential concentration factors do not lend themselves to local stencils of differences: they enjoy the root-exponential accuracy encountered in (5.6).

Our starting point is the discrete conjugate sum, an analog of (4.3):

$$I_N^\tau f(x) := \frac{\pi i}{c_\tau} \sum_{|k|\le N} sgn(k)\tau\Big(\frac{|k|h}{\pi}\Big)\widehat{f}_k e^{ikx}, \tag{6.4a}$$

where $\tau(\xi)$ are the discrete concentration factors at our disposal and c_τ is the normalization coefficient

$$c_\tau := \frac{\pi}{2}\int_0^1 \frac{\tau(\xi)}{\sin(\pi\xi/2)}d\xi. \tag{6.4b}$$

It is convenient to link the discrete and continuous factors

$$\tau(\xi) = \sigma(\xi)sinc\big(\frac{\pi\xi}{2}\big), \qquad sinc(y) := \frac{\sin y}{y},$$

where the normalization c_τ becomes the usual $c_\tau = c_\sigma = \int_0^1 \sigma(\xi)/\xi d\xi$.

To gain greater insight into the behavior of such detectors we use (6.1) to express the $\widehat{f}_k$'s in terms of the $f(x_\nu)$'s; (6.4a) then reads

$$\begin{aligned} I_N^\tau f(x) &= -\frac{h}{c_\tau}\sum_{\nu=0}^{2N} f(x_\nu)\sum_{k=1}^{N}\sigma\Big(\frac{kh}{\pi}\Big)\frac{\sin kh/2}{kh/2}\sin k(x-x_\nu) \\ &= -\frac{1}{c_\sigma}\sum_{k=1}^{N}\frac{\sigma\big(\frac{kh}{\pi}\big)}{k}\sum_{\nu=0}^{2N} f(x_\nu)2\sin\big(\frac{kh}{2}\big)\sin k(x-x_\nu). \end{aligned}$$

Next, we write the last product on the right as a perfect difference and sum by parts to find

$$I_N^\tau f(x) = \frac{1}{c_\sigma}\sum_{k=1}^{N}\frac{\sigma\big(\frac{kh}{\pi}\big)}{k}\sum_{\nu=0}^{N}\big(f(x_{\nu+1})-f(x_\nu)\big)\cos k(x-x_{\nu+1/2}). \tag{6.5}$$

We claim that the second sum on the right is dominated by the discontinuous cell(s) where $f(x_{\nu_c+1})-f(x_{\nu_c})\sim[f](c)$, while the contributions of the 'smooth' cells are negligible, owing to cancelations of oscillations. To make this statement precise, we first identify the discontinuous cell (and in general, finitely many like it), by its mid-point, $x_{\nu_c+\frac{1}{2}}$. We then find that

$$\sum_{\nu=0}^{2N}\big(f(x_{\nu+1})-f(x_\nu)\big)\sin kx_{\nu+\frac{1}{2}} =$$

$$= \Big([f](c) + \mathcal{O}(h)\Big) \sin kx_{\nu_c+\frac{1}{2}} + \mathcal{O}\Big(\frac{h}{\sin\frac{kh}{2}}\Big). \tag{6.6a}$$

The first term on the right of (6.6a) is the contribution of the single jump at $\nu = \nu_c$. For the remaining terms, $\nu \neq \nu_c$, we use the identity $\sin kx_{\nu+1/2} \equiv -\Big(\cos kx_{\nu+1} - \cos kx_\nu\Big)/2\sin(kh/2)$ to sum by parts once more, accumulating $2N-2 \sim \frac{1}{h}$ terms of order $f(x_{\nu+1}) - 2f(x_\nu) + f(x_{\nu-1}) \sim \mathcal{O}(h^2)$ and two (or finitely many) 'boundary terms' of order $f(x_{\nu+1}) - f(x_\nu) \sim \mathcal{O}(h)$. These amount to the second term on the right of (6.6a). The same argument yields

$$\sum_{\nu=0}^{2N} (f(x_{\nu+1}) - f(x_\nu)) \cos kx_{\nu+\frac{1}{2}} =$$

$$= \Big([f](c) + \mathcal{O}(h)\Big) \cos kx_{\nu+\frac{1}{2}} + \mathcal{O}\Big(\frac{h}{\sin\frac{kh}{2}}\Big). \tag{6.6b}$$

Inserting (6.6) into (6.5) yields

$$I_N^\tau f(x) = [f](c) \times \frac{1}{c_\sigma} \sum_{k=1}^{N} \frac{\sigma\big(\frac{kh}{\pi}\big)}{k} \cos k(x - x_{\nu_c+1/2}) + \mathcal{O}(h|\log h|).$$

Apply theorem 4.1 to the Heaviside function $f(x) = H(x-c)/2$: with $[f](0) = 1$ and $\widehat{f}(k) = -e^{-ikc}/(2\pi ik)$ we find

$$\frac{1}{c_\sigma} \sum_{k=1}^{N} \frac{\sigma(\frac{k}{N})}{k} \cos k(x-c) = \begin{cases} 1 + \mathcal{O}(h|\log h|), & x \approx c, \\ \\ \mathcal{O}(h|\log h|), & dist\Big\{x, c\Big\} \gg h, \end{cases} \tag{6.7}$$

and we obtain the following concentration property.

Theorem 6.1. (Discrete concentration kernels) (Gelb and Tadmor 2000b) Assume that $f(\cdot)$ is piecewise C^2-smooth and let $I_N^\tau f(x)$ be an admissible discrete conjugate sum (6.4). Then $I_N^\tau f(x)$ satisfies the concentration property

$$I_N^\tau f(x) \sim \begin{cases} [f](c_j) + \mathcal{O}(h|\log(h)|), & x \sim c_j,\ j = 1, \ldots, J, \\ \\ \mathcal{O}(h|\log(h)|), & dist\Big\{x, \{c_1, \ldots, c_J\}\Big\} \gg h. \end{cases} \tag{6.8}$$

As an example, consider the discrete concentration factors

$$\tau_{2p+1}(\xi) = \xi^{2p+1} sinc\Big(\frac{\pi\xi}{2}\Big), \qquad sinc(y) = \frac{\sin y}{y}.$$

This corresponds to $\sigma_{2p+1}(\xi) = \xi^{2p+1}$ with $c_\tau = c_\sigma = 2p+1$, and (6.5) yields

$$I_N^{\tau_{2p+1}} f(x) \tag{6.9}$$

$$= h\sum_{\nu=0}^{2N}\big(f(x_{\nu+1}) - f(x_\nu)\big)\sum_{k=1}^{N}(2p+1)\Big(\frac{kh}{\pi}\Big)^{2p+1}\cdot\frac{\cos k(x - x_{\nu+\frac{1}{2}})}{kh}.$$

For the first-order method, (6.9) with $p = 0$ reads

$$I_N^\tau f(x) = h\sum_{\nu=0}^{2N}(f(x_{\nu+1}) - f(x_\nu))D_N(x - x_{\nu+\frac{1}{2}}),$$

This tells us that the discrete concentration factors $\tau_1(\xi) = \xi sinc(\pi\xi/2)$ amount to *interpolation* of the first-order local differences, $f(x_{\nu+1}) - f(x_\nu)$, at the intermediate grid points, $x_{\nu+\frac{1}{2}}$. In a similar fashion, concentration kernels associated with the higher order *polynomial factors*, $\tau_{2p+1}(\xi)$, coincide with higher-order derivatives of this interpolant. As p increases, however, the global dependence of interpolation may lead to deterioration of the results, when compared with local edge detectors. An example with *even* order p is illustrated in figure 6.1.

In contrast, if we choose the *trigonometric concentration factor*, $\tau(\xi) = \sin^3(\pi\xi/2)$ it gives the conjugate discrete sum (6.5) corresponding to $\sigma(\xi) = \xi\sin^2(\pi\xi/2)$ with $c_\tau = c_\sigma = 2$,

$$I_N^\tau f(x) = \frac{h}{2\pi}\sum_{\nu=0}^{2N}\big(f(x_{\nu+1}) - f(x_\nu)\big)\sum_{k=1}^{N}\sin^2\big(\frac{kh}{2}\big)\cos k(x - x_{\nu+\frac{1}{2}}).$$

This conjugate discrete sum coincides with the *local* cubic difference (Gelb and Tadmor 2002), $I_N^\tau f(x_{\nu+\frac{1}{2}}) = 8\Delta^3 f(x_{\nu+\frac{1}{2}})$,

$$I_N^\tau f(x_{\nu+\frac{1}{2}}) = 8\Big(-f(x_{\nu+2}) + 3f(x_{\nu+1}) - 3f(x_\nu) + f(x_{\nu-1})\Big).$$

The situation is analogous to the higher-order trigonometric factors $\sigma(\xi) = \xi^{2p+1}$: they have the advantage of being local but their order is finite.

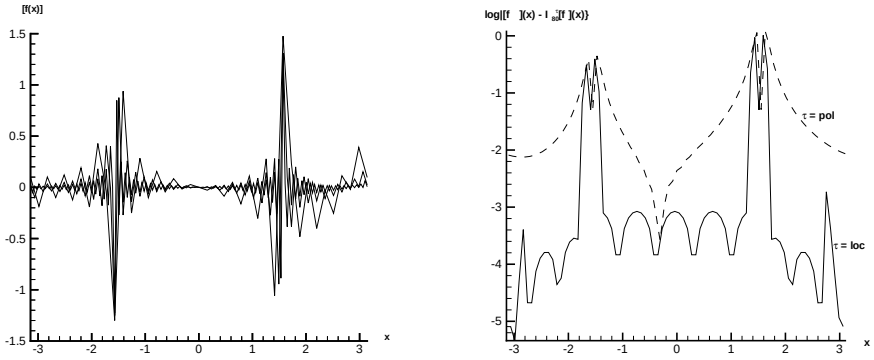

Figure 6.1: *Left:* Detection of edges in the interpolant of f in (5.7) using polynomial concentration factors, $I_N^{\tau_2} f$, with $N = 20, 40$ and 80 modes. *Right:* Logarithmic error of the fifth-order local trigonometric *vs.* the global polynomial concentration factors $I_{80}^{\tau_5} f(x_\nu)$.

6.2. Non-periodic data

We begin with the Gegenbauer expansion of a piecewise smooth $f(\cdot)$:

$$S_N f(x) = \sum_{k=0}^{N} \widehat{f}(k) C_k(x), \qquad \widehat{f}(k) := \int_{-1}^{1} f(x) C_k(x) \omega(x) dx. \tag{6.10}$$

Here $\{C_k(x) = C_k^{(\alpha)}(x)\}_{k\geq 1}$ are orthogonal families of *Gegenbauer polynomials*, associated with different weight functions, $\omega(x) \equiv \omega_\alpha(x) := (1-x^2)^{\alpha-\frac{1}{2}}$,

$$\int_{-1}^{1} C_k^{(\alpha)}(x) C_\ell^{(\alpha)}(x) \omega_\alpha(x) dx = 0, \quad k \neq \ell, \quad \omega_\alpha(x) := (1-x^2)^{\alpha-\frac{1}{2}}. \tag{6.11}$$

They are the eigenfunctions of the singular Sturm–Liouville problem

$$((1-x^2)\omega(x)\big((C_k^{(\alpha)})'(x)\big)' = -a_k \omega(x) C_k^{(\alpha)}(x), \quad -1 \leq x \leq 1, \tag{6.12}$$

with corresponding eigenvalues $a_k = a_k^{(\alpha)} = k(k+2\alpha)$.

As in the periodic case, integration by parts against (6.12) shows that the presence of a single jump discontinuity, $[f](c)$, dictates the linear decay rate of its Gegenbauer coefficients,

$$\widehat{f}(k) = [f](c) \frac{(1-c^2)\omega(c)}{a_k} C_k'(c) + \mathcal{O}\left(\frac{1}{a_k^2}\right). \tag{6.13}$$

To extract information about the location of the jump, we consider the

conjugate sum

$$\frac{\pi\sqrt{1-x^2}}{N}S_N(f)'(x) = \frac{\pi\sqrt{1-x^2}}{N}\sum_{k=1}^{N}\widehat{f}(k)C_k'(x)$$
$$= [f](c)\frac{\pi\sqrt{1-x^2}(1-c^2)\omega(c)}{N}\sum_{k=1}^{N}\left\{\frac{1}{a_k}+\mathcal{O}\left(\frac{1}{a_k^2}\right)\right\}\times C_k'(c)C_k'(x).$$

This is the non-periodic analog of the Fejèr conjugate sum (5.2) in the periodic case.

We want to quantify the localization property of the last summation. To this end, we simplify the computations by making the (non-standard) normalization $\|C_k^{(\alpha)}(x)\|_{\omega_\alpha} = 1$. Integration by parts of (6.12) against $C_k^{(\alpha)}$ then yields $\left(C_k^{(\alpha)}\right)'(x) = \sqrt{a_k}C_{k-1}^{(\beta)}(x)$ with $\beta = \alpha+1$, where the scaling factor $\sqrt{a_k}$ keeps the proper normalization $\|C_k^{(\beta)}(x)\|_{\omega_\beta} = 1$. Substituting in the leading term of the last conjugate sum, we end up with

$$\frac{\pi\sqrt{1-x^2}}{N}S_N(f)'(x) \sim [f](c)\frac{\pi\sqrt{1-x^2}\omega_\beta(c)}{N}\times K_N^{(\beta)}(x,c), \tag{6.14}$$

where

$$K_N^{(\beta)}(x,y) = \sum_{k=1}^{N}C_{k-1}^{(\beta)}(x)C_{k-1}^{(\beta)}(y)$$

is the Christoffel–Darboux kernel (see e.g., (Szego 1958, Thm. 3.2.2)),

$$K_N^{(\beta)}(x,y) = \frac{k_{N-1}}{k_N}\frac{C_N^{(\beta)}(x)C_{N-1}^{(\beta)}(x) - C_N^{(\beta)}(y)C_{N-1}^{(\beta)}(x)}{x-y}, \quad \frac{k_{N-1}}{k_N}\sim\frac{1}{2}. \tag{6.15}$$

The concentration property now depends on the localization of $K_N(c,x)$ (see e.g., (Gelb and Tadmor 2000a, section 3)), i.e.,

$$\frac{\pi\sqrt{1-x^2}\omega_\beta(c)}{N}K_N^{(\beta)}(c,x) \sim \begin{cases} \dfrac{\sqrt{\omega_\alpha(c)}}{\sqrt{\omega_\alpha(x_N)}}\times\dfrac{1}{N|x-c|}\sim\dfrac{1}{N^{1-\alpha}}, & x\neq c, \\ 1, & x=c,\ |c|<1. \end{cases}$$

We summarize.

Corollary 6.2. Let S_Nf denote the truncated Gegenbauer expansion (6.10) of a piecewise smooth f, associated with a weight function $\omega_\alpha = (1-x^2)^{\alpha-\frac{1}{2}}$, $|\alpha|\leq 1/2$. Then $\pi\sqrt{1-x^2}S_N(f)'(x)/N$ admits the concentration

property

$$\left|\frac{\pi\sqrt{1-x^2}}{N}S_N(f)'(x)-[f](x)\right| \lesssim \frac{\log N}{N\sqrt{\omega_\alpha(x)}}, \qquad 1-|x| \lesssim \frac{1}{N^2}.$$

We close this section with the example of a piecewise smooth f with *Chebyshev expansion*,

$$S_N f(x) \sim \sum_k \widehat{f}(k)T_k(x).$$

Using a general family of concentration factors, $\lambda(\xi) \in C^2[0,1]$ (corresponding to $\sigma(\xi)/\xi$ in the periodic case), we end up with

$$\Big|\frac{\pi\sqrt{1-x^2}}{Nc_\lambda}\sum_{k=1}^{N}\lambda\Big(\frac{k}{N}\Big)\widehat{f}(k)T_k'(x)-[f](x)\Big| \lesssim \frac{\log N}{N}, \quad c_\lambda := \int_0^1 \lambda(\xi)d\xi. \tag{6.16}$$

6.3. Noisy data

We consider the problem of detecting edges in a piecewise smooth f from its spectral content, which is assumed to be corrupted by noise. We begin with the simple case of an f which experiences a single jump discontinuity, $[f](c)$. As in (4.1), this implies first-order decay of the Fourier coefficients:

$$\widehat{f}(k) = [f](c)\frac{e^{-ikc}}{2\pi ik} + \widehat{g}(k) + \widehat{n}(k). \tag{6.17}$$

Here, $\widehat{g}(k)$ are associated with the regular part of f after extracting the jump $[f](c)$; their decay is of order $\sim |k|^{-2}$ or faster, depending on the smoothness of the regular part $g(\cdot)$. The new aspect of the problem enters through the $\widehat{n}(k)$'s, which are the Fourier coefficients of the noisy part corrupting the smooth part of the data; we assume $n(\cdot)$ to be white noise with variance $E(|\widehat{n}(k)|^2) = \eta$. With (6.17), the conjugate sum (4.3) becomes

$$\begin{aligned} K_N^\sigma f(x) = [f](c)\frac{2\pi i}{c_\sigma}\sum_{k=1}^{N}\frac{\sigma\big(\frac{k}{N}\big)}{k}\cos k(x-c) \\ -\frac{2\pi}{c_\sigma}\sum_{k=1}^{N}\sigma\Big(\frac{k}{N}\Big)\widehat{g}(k)\sin kx - \frac{2\pi}{c_\sigma}\sum_{k=1}^{N}\sigma\Big(\frac{k}{N}\Big)\widehat{n}(k)\sin kx. \end{aligned}$$

We quantify the "energy" of each of the three sums on the right. E_J and E_R are associated with the discontinuous and regular parts of f,

$$E_J := \sum_{k=1}^{N}\Big(\frac{\sigma\big(\frac{k}{N}\big)}{k}\Big)^2 \approx \frac{1}{N}\int_0^1 \big(\frac{\sigma(\xi)}{\xi}\Big)^2 d\xi, \tag{6.18a}$$

$$E_R := \sum_{k=1}^{N} \sigma^2\Big(\frac{k}{N}\Big)|\widehat{g}(k)|^2 \ll \frac{1}{N^3}\int_0^1 \frac{\sigma^2(\xi)}{\xi^4}d\xi, \tag{6.18b}$$

and E_η is associated with the noisy part of f which was assumed to have variance η:

$$E_\eta := \sum_{k=1}^{N} \sigma^2\Big(\frac{k}{N}\Big)E(|\widehat{n}(k)|^2) \approx \eta N\int_0^1 \sigma^2(\xi)d\xi. \tag{6.18c}$$

Following (Engelberg and Tadmor 2007), the key to detection of edges in such noisy data is to treat the problem as a constrained minimization. We seek a linear combination $a_J E_J + a_R E_R + a_\eta E_\eta$ which minimizes the total energy, thus making the conjugate sum $K_N^\sigma f$ as localized as possible, subject to a prescribed normalization constraint (4.2b)

$$\min\Big\{a_J E_J + a_R E_R + a_\eta E_\eta \;\Big|\; \int_0^1 \frac{\sigma(\xi)}{\xi}d\xi = c_\sigma\Big\}. \tag{6.19}$$

This yields

$$\sigma(\xi) = \frac{C\xi^{-1}}{a_J N^{-1}\xi^{-2} + a_R N^{-3}\xi^{-4} + \eta a_\eta N} = \frac{CN^3\xi^3}{a_J N^2\xi^2 + a_R + \eta a_\eta N^4\xi^4}.$$

We ignore the relatively negligible contribution of the regular part which becomes even smaller as $g(\cdot)$ becomes smoother. Setting $a_R = 0$ we end up with concentration factors of the form

$$\sigma(\xi) = \frac{C}{a_J}\cdot\frac{N\xi}{1+\eta\beta^2N^2\xi^2}, \quad \beta := \sqrt{\frac{a_\eta}{a_J}}. \tag{6.20}$$

It is worthwhile noting that the resulting concentration factor depends on three parameters.
(i) The *relative* size of the amplitudes $\beta = a_\eta/a_J$. Indeed, the normalization factor is given by

$$c_\sigma = \int_0^\pi \frac{\sigma(\xi)}{\xi}d\xi = \frac{C}{a_J\sqrt{\eta}\beta}\tan^{-1}(\sqrt{\eta}\beta N).$$

The corresponding concentration factor

$$\sigma \equiv \sigma_\eta = \frac{\sqrt{\eta}\beta N}{\tan^{-1}(\sqrt{\eta}\beta N)}\cdot\frac{\xi}{1+\eta\beta^2N^2\xi^2}, \tag{6.21a}$$

yields

$$K_N^{\sigma_\eta} f(x) = \frac{\pi i\sqrt{\eta}\beta}{\tan^{-1}(\sqrt{\eta}\beta N)}\sum_{|k|\le N}\frac{|k|}{1+\eta\beta^2k^2}\widehat{f}(k)e^{ikx}, \quad \beta = \sqrt{\frac{a_\eta}{a_J}}. \tag{6.21b}$$

(ii) The number of modes, N. General concentration factors may depend

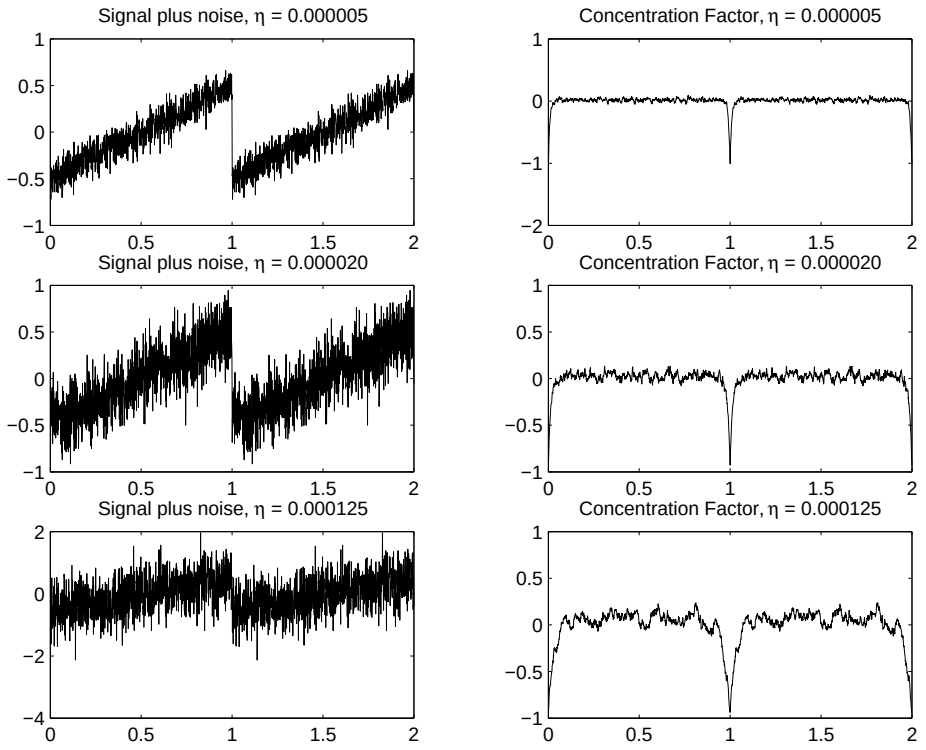

Figure 6.2: Detection of edges in noisy saw-tooth function corrupted with various values of η, using the concentration kernel (6.21a) with $\beta = \pi\eta^{-1/6}$.

on the wave number k and the number of modes N, $\sigma = \sigma_{k,N}$. It is useful to rearrange this dependence, emphasizing the dependence on the relative wave number, $\sigma = \sigma_N\big(\frac{|k|}{N}\big)$. Clearly, theorem 4.1 (and likewise, theorem 6.1) applies to such $\sigma_N(\xi)$'s. Here, one has to verify the precise dependence of the error bound (4.5) on σ_N. In particular, in the present context of noisy data this leads us to consider the following parameterization of the noise.

(iii) The variance of the noise η. We now have *three* scales involved — the small "smoothness" of order $h \sim 1/N$, the noise scale $\sim \eta$ and the $\mathcal{O}(1)$ scale of jump discontinuities. We distinguish between two cases. If η is sufficiently small, $\eta \ll 1/N$ so that $\sqrt{\eta}\beta N \ll 1$, then the noise can be sought as part of the smooth variation of f; indeed, (6.21a) recovers Fejér concentration factor for noise-free data, $\sigma_\eta(\xi) \approx \xi$ (and in particular, $\sigma_\eta(\xi) = \xi$ at the limit of $\eta \downarrow 0$). If on the other hand, $\eta \gtrsim 1/N$, then the $\mathcal{O}(1/N)$-smoothness scale is dominated by the $\mathcal{O}(\eta)$-noise scale, which we assume to be still well-below

the $\mathcal{O}(1)$-scale of the jumps

$$\frac{1}{N} \lesssim \sqrt{\eta}\beta \ll \mathcal{O}(1).$$

In this case, we can ignore the bounded factor $1/\tan^{-1}(\sqrt{\eta}\beta N)$ and using (4.9), we then find an error-bound of order

$$\frac{\log N}{N}\|\sigma_\eta\|_{C^2} + \left|\sigma_\eta\big(\frac{1}{N}\big)\right| + \frac{1}{N}|\sigma_\eta(1)| \lesssim \sqrt{\eta}\beta|\log(\sqrt{\eta}\beta)|.$$

A careful examination of the various error-bounds involved in theorem 4.1 (e.g., (Engelberg and Tadmor 2007)), shows that all other σ_η-dependent contributions to the error do not exceed the small scale of order $\sqrt{\eta}\beta|\log(\sqrt{\eta}\beta)|$,

$$|K_N^{\sigma_\eta} f(x) - [f](x)| \lesssim \sqrt{\eta}\beta|\log(\sqrt{\eta}\beta)|.$$

The resulting concentration kernel, $K_N^{\sigma_\eta} f$ tends to de-emphasize both the low frequencies which are "corrupted" by the jump discontinuity(-ies) and the high frequencies which are corrupted by the noise. Different procedures yield different policies for the choice of β. figure 6.2 demonstrates the edge detected in noisy data using the concentration kernel (6.21) with the advocated $\beta \sim \eta^{-1/6}$.

As an alternative approach, we may replace the L^2-"averaged" effect of the regular part taken in (6.18b) by the BV-like quantity

$$E_R := \sum_{k=1}^{N} \left|\sigma\Big(\frac{k}{N}\Big)\right| \cdot |\widehat{g}(k)|,$$

In this case, the constrained minimization (6.19) with $|\widehat{g}(k)| \sim 1/|k|^2$ yields

$$\sigma_\eta(\xi) = \frac{1}{c_\sigma} \cdot \frac{(N\xi - a_R)_+}{1 + \eta\beta^2 N^2 \xi^2}. \tag{6.22}$$

Figure 6.3 quotes the results of (Engelberg and Tadmor 2007), with the detection of edges in noisy data using these concentration factors which were tuned with $a_R = 6\pi,\ c_\sigma \sim 3$.

6.4. Incomplete data – compressed sensing

We are interested in the detection of edges in a piecewise smooth f from an *incomplete* set of its spectral content, that is, we have access only to $\widehat{f}(k)$'s (or $\widehat{f}_k$'s) for $k \in K$, where K is a strict subset of $\{-N, \dots, N\}$. Our methodology for edge detection in such cases is motivated by the *compressive sensing* approach (Donoho and Tanner 2005, Candes, Romberg and Tao 2006a, Candes, Romberg and Tao 2006b). Equipped with the partial information of $\widehat{f}(k), k \in K$, one can form the incomplete concentration ker-

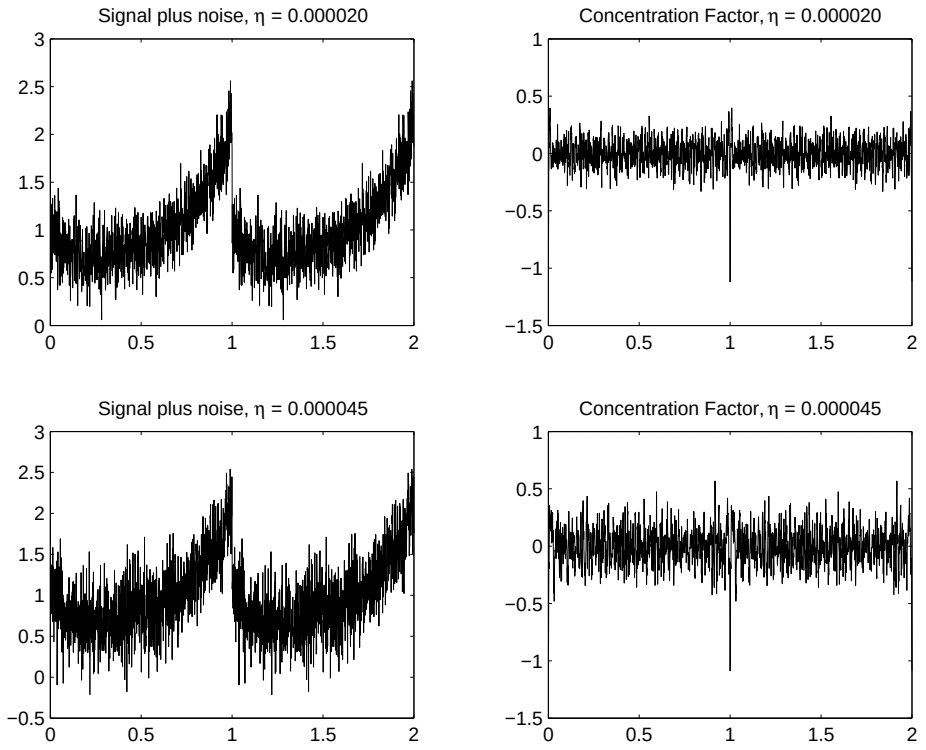

Figure 6.3: Detection of edges in noisy saw-tooth function corrupted with various values of η, using the concentration factors (6.22) with $\beta = \pi\eta^{-1/6}$.

nel

$$K_N^\sigma f(x) = \sum_{k\in K} \widetilde{f}(k)e^{ikx}, \qquad \widetilde{f}(k) := \frac{\pi i}{c_\sigma} sgn(k)\sigma\Big(\frac{|k|}{N}\Big)\widehat{f}(k). \tag{6.23}$$

We follow (Tadmor and Zou 2007), seeking to recover a 'complete' concentration kernel, $g(x) \sim K_N^\sigma f(x)$, of the form

$$g(x) = \overbrace{\sum_{k\in K} \widetilde{f}(k)e^{ikx}}^{\text{prescribed data}} \; + \; \overbrace{\sum_{k\notin K} \widehat{g}(k)e^{ikx}}^{\text{missing data}}.$$

Here, $\widetilde{f}(k)$ are the conjugate coefficients corresponding to the prescribed data for $k \in K$, while the missing conjugate coefficients, $\{\widehat{g}(k)|k \notin K\}$ at our disposal, are sought as minimizers of the total variation $\|g(x)\|_{TV}$,

$$g(x) = argmin\Big\{\|g\|_{TV} \;\Big|\; g(x) = K_N^\sigma f(x) + \sum_{k\notin K} \widehat{g}(k)e^{ikx}\Big\}. \tag{6.24}$$

Similarly, in the discrete case we seek a 'complete' concentration kernel

$$g(x) = \frac{\pi i}{c_\tau} \sum_{k\in K} sgn(k)\tau\Big(\frac{k}{N}\Big)\widehat{f}_k e^{ikx} + \sum_{k\notin K} \widehat{g}_k e^{ikx},$$

which is selected by the TV minimization principle,

$$g(x) = argmin\Big\{\|g\|_{TV} = \sum_{\nu} |g(x_{\nu+1}) - g(x_\nu)| \;\Big|\; \widehat{g}_k = \widetilde{f}_k,\ k \in K\Big\}. \quad (6.25)$$

The complete concentration kernel, $g(x)$, can be viewed as an approximation to the 'ultimate' jump function

$$\Gamma_f(x) := \sum_{j=1}^{J} [f](c_j) 1_{[x_{\nu_{c_j}}, x_{\nu_{c_j}+1}]}(x),$$

where the missing $\{\widehat{g}_k\}_{|k\notin K}$ complement the prescribed $\{\widetilde{f}_k\}_{|k\in K}$ as the approximate Fourier coefficients $(\widehat{\Gamma_f})_k$. The rationale behind the TV minimization in (6.24),(6.25) is to enforce the ℓ_1-minimization of the differences, which imposes sparsity in the sense of maximizing the number of zero differences, (Candes, Romberg and Tao 2006a, Candes, Romberg and Tao 2006b). Hence, it yields $g(x)$ as an approximate jump function with a minimal number of piecewise components. The optimization model (6.25) can be solved by the second order cone programs which takes time $\mathcal{O}(N^3 \log N)$.

The compressed sensing approach can be extended to noisy data (Donoho, Elad and Temlyakov 2006, Candes, Romberg and Tao 2006a). Following (Tadmor and Zou 2007), we assume that the observed (pseudo-)spectral data may be contaminated by white noise with variance $\leq \eta$. To recover edges from such noisy and incomplete data, the following compressed sensing model is sought:

$$g(x) = argmin\Big\{\|g\|_{TV} \;\Big|\; g = \sum_{k=1}^{N} \widehat{g}_k e^{ikx} \ \text{ s.t. } \ \|\widehat{g}_k - \widetilde{f}_k\|_{\ell^2(k\in K)} \leq \eta\Big\}. \quad (6.26)$$

7. Enhancements

The detection of edges in theorems 4.1 and 6.1 is based on the asymptotic behavior of the concentration kernels K_N^σ which separate between the large and small scales as $\varepsilon_N \sim 1/N \downarrow 0$. To improve the edge detection, we want

to enhance the separation of scales in (4.5). To this end we consider

$$N^q\big(K_N^\sigma f(x)\big)^{2q} = \begin{cases} \sim N^q\Big([f](c_j)\Big)^{2q}, & x \approx \{c_1, c_2, \dots, c_J\}, \\ \\ \mathcal{O}(N^{-q}), & dist\Big\{x, \big\{c_1, c_2, \dots, c_J\big\}\Big\} \gg \frac{1}{N}. \end{cases}$$

The exponent $q \geq 1$ is at our disposal: by increasing q, we enhance the separation between the vanishing scale at the points of smoothness (of order $\mathcal{O}(N^{-q})$) and the amplified scale at the jumps (of order $\mathcal{O}(N^q)$).

Next, one must introduce a *critical threshold* to eliminate the unacceptable jumps: only those edges with amplitudes larger than the critical threshold, $[f](x) > J_c^{1/2q}/\sqrt{N}$ will be detected. Here, J_c is a measure which defines the small scale in our computation of edge detection. We note that J_c is data-dependent and is typically related to the variation of the smooth part of f.

Given this critical threshold, we form our enhanced concentration kernel

$$K_{N,J_c}^\sigma f(x) = \begin{cases} K_N^\sigma f(x), & \text{if } N^q\big|K_N^\sigma f(x)\big|^{2q} > J_c, \\ \\ 0, & \text{otherwise.} \end{cases} \tag{7.1}$$

Clearly, with sufficiently large q, one ends up with a sharp edge detector where $K_{N,J_c}^\sigma f(x) = 0$ at all but $\mathcal{O}(1/N)$-neighborhoods of the jumps $x = c_1, c_2 \dots$. In practical applications, $q \leq 3$, will suffice. For example, enhancing the local concentration kernel (5.1) $K_{\varepsilon_N}(y) = \phi'_{\varepsilon_N}(y)$ with $q = 1$ leads to the *quadratic filter* (e.g., (Firoozye and Sverak 1996)) where $(K_{\varepsilon_N} f(x))^2 = (\phi'_{\varepsilon_N} * f(x))^2 \to [f]^2(x)$.

We can apply this nonlinear enhancement in conjunction with discrete concentration kernels $I_N^\tau(y)$. The corresponding *enhanced spectral concentration kernel* amounts to

$$I_{N,J_c}^\tau = \begin{cases} I_N^\tau f(x), & \text{if } N^q\big|I_N^\tau f(x)\big|^{2q} > J_c, \\ \\ 0, & \text{otherwise.} \end{cases} \tag{7.2}$$

Observe that the use of concentration kernels, (4.5) and (6.8), actually detects the $\mathcal{O}(\varepsilon_N)$-*neighborhoods* of jump discontinuities rather than the discontinuities themselves. Figure 7.1 demonstrates how the nonlinear enhancement of concentration kernels helps to pinpoint the location of edges in the discrete and non-periodic set-ups.

7.1. Nonlinear limiter: minmod edge detection

Implementation of the enhanced edge detectors $K_{N,J_c}^\sigma f(x)$ and $I_{N,J_c}^\tau f(x)$ requires an outside threshold parameter, J_c, which should be properly chosen

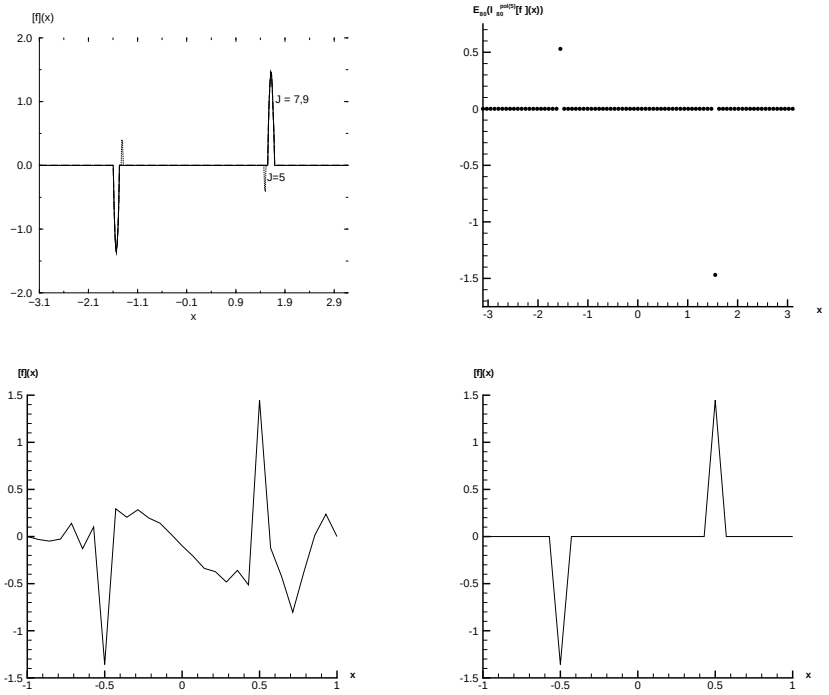

Figure 7.1: *Top left*: Jump value obtained by $K^{\sigma}_{40,J_c} f$ for $f(x)$ in (7.1) with $q = 1$ and $\sigma(\xi) = \sin(\xi)$. *Top right*: $I^{\tau}_{80,J_c} f$ with $\tau(\xi) = \xi^5$ in (7.2). *Bottom*: detection of edges in Chebyshev expansion of $f(x/\pi)$ before and after enhancement with $q = 1$ and $J_c = 5$.

to separate the specific scales associated with f. This becomes an impediment for detecting edges in both small-scale problems and problems with steep gradients and high variation. A second, related difficulty arises when oscillations are formed in the neighborhood of the jump discontinuities. The particular behavior of these oscillations depends on the specific concentration factors used, and it can be difficult to distinguish between a true jump discontinuity and an oscillating artifact, particularly when several jump discontinuities are located in the same neighborhood, i.e. when there is limited resolution for the problem. Wrong parameterization may lead to misidentification of jump discontinuities that are located "too" close together. We discuss an improved enhancement procedure based on the nonlinear limiting of low- and high-order concentration factors. The rationale, outlined in (Gelb and Tadmor 2006), is as follows.

Edge detectors based on a low-order concentration kernel K^{σ}_N with polynomial factors $\sigma_p(\xi), p \sim 1$, or trigonometric factors $\sigma_\alpha(\xi)$, have a relatively slow, $\mathcal{O}(\log(N)/N)$, decay away from the discontinuities, yet they yield only

a few spurious oscillations (if any) in the immediate neighborhoods of the discontinuities. In contrast, highly accurate kernels such as $K_N^{\sigma_{exp}}$ rapidly converge to zero away from the neighborhoods of discontinuities, but suffer from severe oscillations within these immediate neighborhoods. The loss of monotonicity with increasing order is, of course, the canonical situation in many numerical algorithms; the passage from the first-order, monotone Fejér kernel discussed in section 9 below, to the spurious oscillations in the spectrally accurate Dirichlet kernel is a prototypical case.

We therefore take advantage of the different behavior of low- and high-order edge detectors. Away from the jump discontinuities, we let the high-order, possibly exponentially small kernel, dominate, by taking the (signed) minimum,

$$K_N f(x) = s \times \min\big\{|K_N^{\sigma_{high}} f(x)|, |K_N^{\sigma_{low}} f(x)|\big\}, \quad s := sgn\big\{K_N^{\sigma_{high}} f(x)\big\}.$$

As we approach the jump discontinuity, however, high-order methods produce spurious oscillations which should be rejected: this could be achieved through comparison with essentially monotone profiles produced by low-order detectors. Thus, when the two profiles disagree in sign — indicating spurious oscillations — then our detector is set to zero:

$$K_N f(x) = 0, \quad \text{if } sgn\big\{K_N^{\sigma_{high}} f(x)\big\} \neq sgn\big\{K_N^{\sigma_{low}}\big\}.$$

We end up with the so-called *minmod* limiter,

$$K_N^{\sigma_{mm}} f(x) := minmod\big\{K_N^{\sigma_{exp}} f(x), K_N^{\sigma_1} f(x)\big\}, \tag{7.3}$$

which plays a central role in non-oscillatory reconstruction of high-resolution methods for nonlinear conservation laws, (see e.g., (Harten 1983, Tadmor 1998) and the references therein).

This adaptive algorithm can be extended to include several concentration factors, e.g.

$$I_N^{\tau_{mm}} f(x) := minmod\big\{I_N^{\tau_{exp}} f(x), I_N^{\tau_{pol}} f(x), I_N^{\tau_{trig}} f(x)\big\}, \tag{7.4}$$

where the k-tuple *minmod* limiter takes the form

$$minmod\big\{a_1, \ldots, a_k\big\} := \begin{cases} s \times \min\limits_{1\le j\le k} |a_j|, & \text{if } sgn(a_1) = \ldots = sgn(a_k) := s, \\ 0, & \text{otherwise.} \end{cases}$$

It retains the high order in smooth regions while "limiting" the high-order spurious oscillations in the neighborhoods of the jumps by the less oscillatory low-order detectors. By incorporating such a mixture of low-order and high-order methods in different regimes of the computation, the resulting *minmod*-based adaptive detection provides a parameter-free edge detector, which in turn enables more robust nonlinear enhancements.

Figure 7.2 illustrates the improvement in using the *minmod* edge detector (7.4) when applied to a piecewise smooth f exhibited in the top part

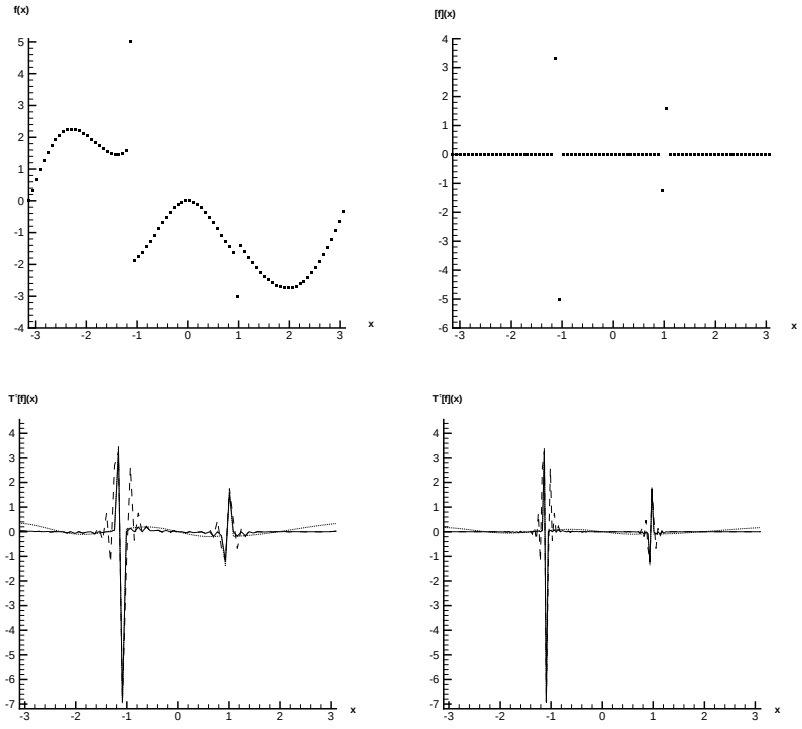

Figure 7.2: *Top left*: Piecewise f with adjacent edge. *Top right*: $K_{80}^{\sigma_{mm}} f(x)$. *Bottom left*: Edge detection using first order $I_N^{\tau_{pol}} f(x)$ (dotted), $I_N^{\tau_{exp}} f(x)$ (dashed) and the $I_N^{\tau_{mm}} f(x)$ algorithm (solid) with $N = 80$ grid-points. *Bottom right*: The same with $N = 160$ points.

of figure 7.2. There is a 'clean' detection of the jump discontinuities which are located close together. It also compares the results of application of the concentration kernels and the *minmod* algorithm. It is evident that the polynomial factor τ_{2p+1} does not converge to zero sufficiently fast away from the discontinuities; hence the steep gradients of the function might be misinterpreted as jump discontinuities. On the other hand, the concentration method using τ_{exp} causes interfering oscillations in the neighborhoods of the discontinuities, making it difficult to determine where the true jumps are. The *minmod* algorithm ensures the convergence to the jump function without interference of the oscillations. An early application of the *minmod* enhancement to non-negative band pass filters can be found in (Bauer 1995, §4).

8. Edge detection in two-dimensional spectral data

8.1. Two-dimensional concentration kernels

Given the two-dimensional spectral data

$$\widehat{f}(\mathbf{k}) := \frac{1}{(2\pi)^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} f(\mathbf{y})e^{-i\mathbf{k}\cdot\mathbf{y}}dy_1dy_2$$

we are interested in detecting the edges of the underlying piecewise smooth $f(\cdot)$. We assume the generic case where these edges lie along simple curves, and we proceed with a straightforward application of one-dimensional concentration kernels which apply dimension-by-dimension. Accordingly, we have a 2-vector concentration kernel $\mathbf{K}_N^\sigma$, of the form,

$$\mathbf{K}_N^\sigma f(\mathbf{x}) = \left[\begin{array}{c} K_{N,x_1}^\sigma \\ K_{N,x_2}^\sigma \end{array}\right] f(\mathbf{x}) := \frac{\pi i}{c_\sigma}\sum_{|k_1|\le N}\sum_{|k_2|\le N}\left[\begin{array}{c} sgn(k_1) \\ sgn(k_2) \end{array}\right]\sigma\Big(\frac{|\mathbf{k}|}{N}\Big)\widehat{f}(\mathbf{k})e^{i\mathbf{k}\cdot\mathbf{x}}. \tag{8.1}$$

Edges along the x_1-axis with each fixed $x_2 \in [-\pi,\pi]$ are sought as extremal values of the first component $K_{N,x_1}^\sigma f(x_1,\cdot)$, while $K_{N,x_2}^\sigma f(\cdot,x_2)$ process edges along the x_2-axis. Similarly, the discrete set-up is based on the two-dimensional conjugate sums

$$\mathbf{I}_N^\tau f(\mathbf{x}) = \frac{\pi i}{c_\tau}\sum_{|k_1|\le N}\sum_{|k_2|\le N}\left[\begin{array}{c} sgn(k_1) \\ sgn(k_2) \end{array}\right]\tau\Big(\frac{|\mathbf{k}|h}{\pi}\Big)\widehat{f}(\mathbf{k})e^{i\mathbf{k}\cdot\mathbf{x}}. \tag{8.2}$$

The approach is simple to implement, although it may suffer from the Cartesian preference when the edges lie along curves which do not align with the axis, as illustrated in figure 8.1 for $f(\mathbf{x})$ whose edges lie along the circle $|\mathbf{x}| = 0.7\pi$. We observe the familiar Cartesian-based phenomenon of staircasing; much of it is removed by nonlinear enhancement.

A parameter-free enhancement based on the minmod limiter (7.4) yields improved results for edge detection in the Shepp-Logan brain image, shown in figure 8.2.

8.2. Incomplete data

The extension of edge detection for *incomplete* data in two dimensions, $\big\{\widehat{f}(\mathbf{k})\big\}_{\mathbf{k}\in K}$ where $K \subsetneq [-N,N]^2$, is straightforward. We shall focus on the discrete case, where we set a rectangular grid $\mathbf{x}_{\nu,\mu} := (\nu\Delta x_1, \mu\Delta x_2)$. We seek $\widehat{\mathbf{g}}_\mathbf{k}|\mathbf{k}\notin K$ which produces an approximate concentration kernel

$$\mathbf{g}(\mathbf{x}) = \frac{\pi i}{c_\tau}\sum_{\mathbf{k}\in K}\left[\begin{array}{c} sgn(k_1) \\ sgn(k_2) \end{array}\right]\tau\Big(\frac{|\mathbf{k}|}{N}\Big)\widehat{f}_\mathbf{k}e^{i\mathbf{k}\cdot\mathbf{x}} + \sum_{\mathbf{k}\notin K}\widehat{\mathbf{g}}_\mathbf{k}e^{i\mathbf{k}\cdot\mathbf{x}}, \tag{8.3a}$$

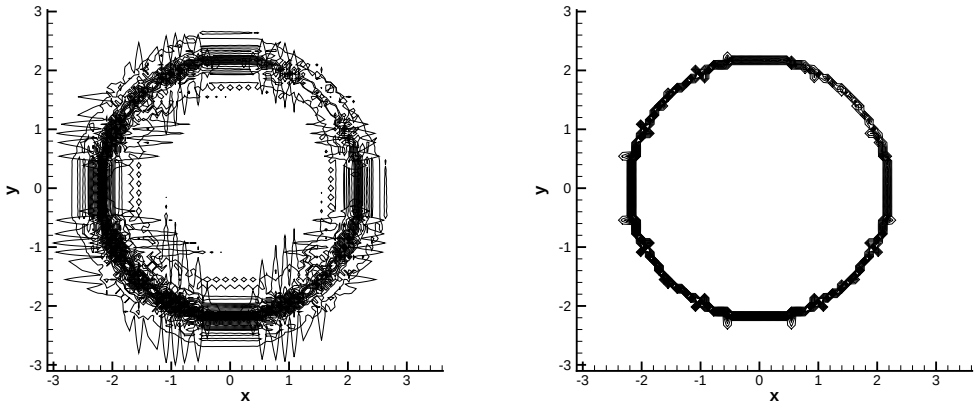

Figure 8.1: The two-dimensional detection $\mathbf{I}_{80}^{\tau_{exp}} f(\mathbf{x})$ for a circular edge, before and after enhancement

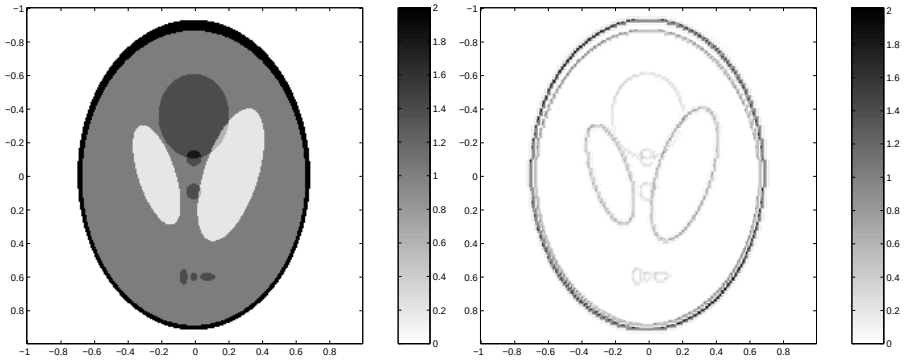

Figure 8.2: *Left*: Contour plot of the Shepp-Logan brain image. *Right*: Nonlinear enhancement procedure (7.4) applied to the Shepp-Logan brain phantom image.

with minimal total variation

$$\|\mathbf{g}\|_{TV} = \sum_{\nu,\mu} |\mathbf{g}(\mathbf{x}_{\nu+1,\mu}) - \mathbf{g}(\mathbf{x}_{\nu,\mu})|\Delta x_1 + |\mathbf{g}(\mathbf{x}_{\nu,\mu+1}) - \mathbf{g}(\mathbf{x}_{\nu,\mu})|\Delta x_2. \quad (8.3b)$$

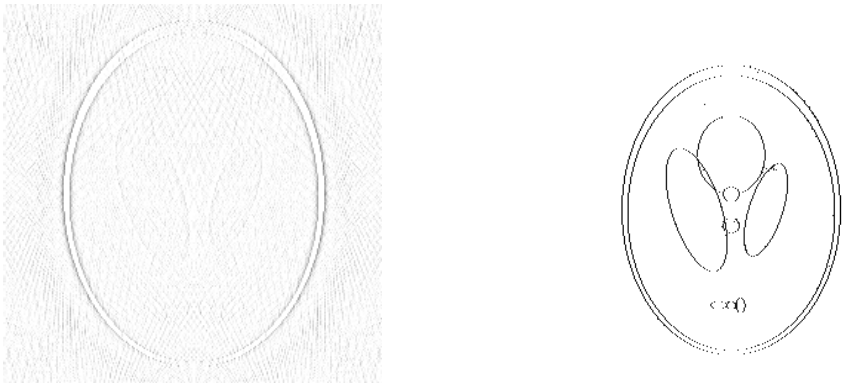

Figure 8.3: Recovered phantom image from incomplete spectral data. *Left:* The result by the back projection. *Right:* the recovered edges of Shepp-phantom graph by compressive sensing edge detection (8.3).

Figure 8.3 illustrates edge detection for use of the compressive sensing model of $\mathbf{K}_N^\sigma f(\mathbf{x})$ for incomplete data of the two-dimensional Shepp-Logan phantom image. Here, $N = 256$, we use Fejér concentration factors, $\sigma(\xi) = \xi$ and partial data are gathered along each of 100 *radial* lines in the spectral domain.

8.3. Concentration kernels and zero-crossing

The zero-crossing method is one of the popular methods in edge detection in two-dimensional data: consult (Marr and Hildreth 1980) and the references therein. It searches for zero crossings in the discrete Laplacian of the function f, in order to find the underlying edges. This is intimately connected with the conjugate kernels. To clarify this point, we begin with the one-dimensional example of the Fejér conjugate sum (5.2), $K_N^{\sigma_1} f = \pi S_N(f)'(x)/N$. Recall that edges are sought as extremal values of $K_N^{\sigma_1} f(x)$, and we therefore seek the zeros of $K_N^{\sigma_1}(f)''(x)$, i.e.,

$$\Big\{ c_j \;\Big|\; \frac{d}{dx} K_N^{\sigma_1} f(x)|_{x=c_j} \propto \frac{d^2}{dx^2} S_N f(x)|_{x=c_j} = 0 \Big\}. \quad (8.4)$$

This is the one-dimensional zero-crossing. We note that unlike the edges sought as the extrema of $K_N^{\sigma_1} f(x)$, the zero-crossing in (8.4) may introduce additional spurious inflection points. A similar situation occurs with general concentration factors. Setting $\sigma(\xi) := \xi\lambda(\xi)$, then (4.3) amounts to

$$K_N^\sigma f(x) = \frac{\pi}{Nc_\lambda} \sum_{|k|\le N} ik\lambda\Big(\frac{|k|}{N}\Big)\widehat{f}(k)e^{ikx}, \qquad c_\lambda = \int_0^1 \lambda(\xi)d\xi.$$

Thus $K_N^\sigma f(x)$ with $\sigma(\xi) = \xi\lambda(\xi)$ is merely the derivative of a mollified version of $S_N(f)(x)$,

$$K_N^\sigma f(x) = \frac{d}{dx}\Lambda_N * S_N(f)(x), \qquad \Lambda_N(x) := \frac{1}{2Nc_\lambda} \sum_{|k|\le N} \lambda\Big(\frac{|k|}{N}\Big)e^{ikx},$$

and the zero-crossing procedure amounts to identifying edges as zeros of this mollified spectral projection,

$$\Big\{c_j \;\Big|\; \frac{d^2}{dx^2}\Lambda_N * S_N f(x)\Big|_{x=c_j} = 0, \quad \Lambda_N(x) := \frac{1}{2Nc_\lambda} \sum_{|k|\le N} \lambda\Big(\frac{|k|}{N}\Big)e^{ikx}\Big\}. \tag{8.5}$$

By suitable choice of λ, we obtain a large class of "regularized" zero-crossings. But once again, we need to augment (8.5) with a procedure to rule out inflection points which otherwise could be detected as spurious edges.

A similar set-up holds in the two-dimensional case. Here, we consider the generic case of a piecewise smooth $f(\mathbf{x})$ whose edges lie along simple curves to be detected by the 2-vector of concentration kernels, $\mathbf{K}_N^\sigma f(\mathbf{x})$, in (8.1). To simplify matters, we set $\sigma(\xi) = \xi\lambda(\xi)$. The two-dimensional detection based concentration approach now seeks the edges as extremal values of

$$\nabla_{\mathbf{x}}\Lambda_N * S_N f(\mathbf{x}), \qquad \Lambda_N(\mathbf{x}) = \frac{1}{2Nc_\lambda} \sum_{|\mathbf{k}|\le N} \lambda\Big(\frac{|\mathbf{k}|}{N}\Big)e^{i\mathbf{k}\cdot\mathbf{x}}. \tag{8.6}$$

The zero-crossing method realizes these extremal values as the zeros of

$$\Big\{\mathbf{c} \;\Big|\; \Delta_{\mathbf{x}}\Lambda_N * S_N f(\mathbf{x})\Big|_{\mathbf{x}=\mathbf{c}} = 0\Big\}, \tag{8.7a}$$

We observe that the two-dimensional zero-crossing could add a considerable number of spurious edges: e.g., (Ulupinar and Medioni 1988, Clark 1989). We can improve this deficiency, by augmenting (8.7a) with a more careful zero-crossing criterion, e.g.,

$$\Big\{\mathbf{c} \;\Big|\; \frac{\partial^2}{\partial x_1^2}\Lambda_N * S_N f(\mathbf{x}) \text{ or } \frac{\partial^2}{\partial x_2^2}\Lambda_N * S_N f(\mathbf{x}) \text{ changes sign at } \mathbf{x} = \mathbf{c}\Big\}. \tag{8.7b}$$

Following (Tadmor and Zou 2007), we can now combine the improved

zero crossing (8.7b) with compressed sensing, in order to deal with incomplete (and possibly noisy...) data. Given the partial spectral information, $\big\{\widehat{f}(\mathbf{k})\big\}_{\mathbf{k}\in K}$, we seek to complement the missing data by the usual TV-minimization

$$g(\mathbf{x}) = argmin\big\{\|g\|_{TV} \;\Big|\; g(\mathbf{x}) = \sum_{\mathbf{k}\in K} |\mathbf{k}|^2 \widehat{f}(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}} + \sum_{\mathbf{k}\notin K} |\mathbf{k}|^2 \widehat{g}(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}}\big\}. \tag{8.8}$$

Observe the *sparsity* of the TV-based compressive sensing of the zero-crossings in the minimizer (8.8): it is tied to minimizing the number of zero components of $g \approx \Delta_{\mathbf{x}} K_N^\sigma f(\mathbf{x})$, that is, minimizing the number of piecewise linear components of $K_N^\sigma f(\mathbf{x})$.

Figure 8.4: Edge detection in incomplete spectral data by zero crossing. *Left:* Original image. *Center:* Image recovered from incomplete data zero crossing. *Right:* Zero crossing combined with compressive sensing (8.8), with Gaussian concentration factor Λ^1 and threshold $\zeta = 3$.

Figure 8.4, from (Tadmor and Zou 2007), illustrates edge detection from incomplete data using (improved) zero-crossing with compressed sensing. Here we use the normalized Gaussian,

$$\Lambda^\beta(\mathbf{x}) := \frac{1}{2\pi\beta^2} e^{-\frac{|\mathbf{x}|^2}{2\beta^2}},$$

which is a typical choice of zero-crossing mollifier, Λ_N.

9. Reconstruction of piecewise smooth data

We want to reconstruct a piecewise smooth f from its spectral coefficients $\{\widehat{f}(k)\}_{|k|\le N}$. To avoid the spurious Gibbs oscillations formed by the spectral projection $S_N f$, one may consider the classical Fejér partial sums

$$S_N^F f(x) := \sum_{|k|\le N} \Big(1 - \frac{|k|}{N}\Big)\widehat{f}(k)e^{ikx},$$

which amount to a convolution against the *Fejér kernel*,

$$S_N^F f(x) = (F_N * f)(x), \quad F_N(y) := \frac{1}{2\pi}\sum_{|k|\le N}\Big(1-\frac{|k|}{N}\Big)e^{iky} \equiv \frac{1}{2\pi N}\Big(\frac{\sin\big(\frac{Ny}{2}\big)}{\sin\big(\frac{y}{2}\big)}\Big)^2.$$

Since $F_N \ge 0$, it follows that $-1 \le S_N^F H(x) \le 1$; moreover, since $H'(x) \ge 0$ implies $S_N^F(H)'(x) \ge 0$, it follows that $S_N^F H(x)$ increases *monotonically* between -1 and 1. Thus, Fejér partial sums avoid spurious oscillations; in fact they are monotone, and converge uniformly whenever f is continuous. But the monotonicity of Fejér partial sums comes at a price: according to a classical theorem of Korovkin (e.g., (DeVore and Lorentz 1993)), every family of linear *positive operators* such as the S_N^F is at most second-order accurate, that is

$$|S_N^F f(x) - f(x)| \lesssim \frac{1}{N^2}, \quad f \in C^1$$

and this second-order convergence rate estimate does *not* improve for more regular f's (since it is essentially dictated by $f(x) = 1, x$ and x^2).

It is possible to utilize the Fejér sums to regain spectral accuracy, while still avoiding Gibbs oscillations. To this end, we consider the partial sum

$$S_N^\varphi f(x) := \sum_{|k|\le N} \varphi\Big(\frac{|k|}{N}\Big)\widehat{f}(k)e^{ikx}, \quad \varphi\Big(\frac{|k|}{N}\Big) = \begin{cases} 1, & |k| \le \frac{N}{2}, \\ 2 - \frac{2|k|}{N}, & \frac{N}{2} \le |k| \le N. \end{cases}$$

Although S_N^φ is no longer positive, it is the difference of two positive Fejér sums,

$$S_N^\varphi f(x) \equiv 2S_N^F f(x) - S_{N/2}^F f(x),$$

and as such, it converges uniformly whenever f is merely continuous. At the same time, the convergence rate of S_N^φ increases together with the global smoothness of f, and we have the spectral error estimate

$$\begin{aligned} &|S_N^\varphi f(x) - f(x)| \\ &\quad \le \sum_{\frac{N}{2}\le|k|\le N} \Big|1 - \frac{2|k|}{N}\Big| \cdot |\widehat{f}(k)| + \sum_{|k|>N} |\widehat{f}(k)| \lesssim \|f\|_{C^s}\frac{1}{N^{s-1}}, \quad \text{for all } s > 1. \end{aligned}$$

Observe that the spectral accuracy of S_N^φ lies in the fact that the first $N/2$ coefficients in S_N^φ are left unchanged,

$$\varphi(\xi) = \begin{cases} 1, & 0 \le \xi \le \frac{1}{2}, \\ \\ 2 - 2\xi, & \frac{1}{2} \le \xi \le 1. \end{cases} \tag{9.1}$$

But what happens when we apply $S_N^\varphi f$ to piecewise smooth f's? As we shall explore in the next few sections, the answer lies with the smoothness of $\varphi(\cdot)$ or, equivalently, the decay behavior of the mollifier associated with (9.1), $\mathcal{S}_N^\varphi(x) = 2F_N(x) - F_{\frac{N}{2}}(x)$.

The preceding examples demonstrate two interchangeable processes which are available for recovering the rapid convergence in the piecewise smooth case. These are *mollification*, carried out in the physical space, and *filtering*, carried out in the Fourier space, i.e.,

$$\Phi * (S_N f)(x) \longleftrightarrow \sum_{|k|\le N} \varphi\Big(\frac{|k|}{N}\Big)\widehat{f}(k)e^{ikx}.$$

Filtering accelerates convergence when pre-multiplying the Fourier coefficients by a *rapidly decreasing* $\varphi(|k|/N)$, as $|k| \uparrow N$. This rapid decay in Fourier space corresponds to mollification with *highly localized* mollifiers, $\Phi(x) = \mathcal{S}_N^\varphi(x)$, in physical space[5]:

$$\mathcal{S}_N^\varphi(x) = \frac{1}{2\pi} \sum_{|k|\le N} \varphi\Big(\frac{|k|}{N}\Big)e^{ikx}.$$

There is a rich literature on filters and mollifiers as effective tools for Gibbs-free reconstruction of piecewise smooth functions. Different aspects of this topic are drawn from a variety of sources, ranging from summability methods in harmonic analysis to signal processing – and in recent years, image processing and high-resolution spectral computations of propagation of singularities and shock discontinuities.

Classical mollifiers of finite polynomial order, $\mathcal{O}(N^{-p})$, are dictated by a moment condition of order p, (10.1), discussed in Section 10 below. By properly tunning $p = p_N$ to increase with N, one obtains *spectrally accurate* mollifiers, (Gottlieb and Tadmor 1985) and spectrally accurate filters (Majda, McDonough and Osher 1978, Vandeven 1991). Improved results are obtained by a further adaptation of p_N to the distance from the edges (Boyd 1995, Boyd 1996). By carefully tuning p_N together with proper G_2 cut-off functions, we obtain improved *root-exponential* accurate mollifiers,

[5] Observe that $\mathcal{S}_N^\varphi(x)$ is the mollifier function associated with, but otherwise different from, the filtered sum, $S_N^\varphi f$.

(Tadmor and Tanner 2002), which are discussed in Section 10.2 below, together with the corresponding discrete mollifiers in Section 10.3. The roo-exponential accuracy of these mollifiers is adapted to the *interior* points, away from the vicinity of the edges. It can be modified to gain polynomial accuracy *up to* the edges; the details are outlined in Section 10.4. Finally, in Section 10.5 we discuss mollifiers based on Gegenbauer expansion (Gottlieb, Shu, Solomonoff and Vandeven 1992, Gottlieb and Shu 1998, Gelb and Tanner 2006), with uniform root-exponential accuracy *up to* the edges. In Section 11 we revisit the construction of accurate mollifiers based on the corresponding filters. We conclude in Section 11.2 with *exponentially* accurate mollifiers (Tanner 2006), based on optimally space-frequency localized filters.

We now turn to discuss these mollifiers and filters which enable the highly-accurate, Gibbs-free reconstruction of f from its (pseudo-)spectral content.

10. Spectral mollifiers

10.1. Compactly supported mollifiers

We begin with classical compactly supported mollifiers. Fix $p < q$ and let $\Phi = \Phi_p \in C_0^q(-\pi, \pi)$, be a unit mass kernel which possesses $p-1$ *vanishing moments*,

$$\int_{-\pi}^{\pi} x^n \Phi(x) dx = \begin{cases} 1, & n = 0, \\ 0, & n = 1, \dots, p. \end{cases} \tag{10.1}$$

Example 10.1. (Mollifiers satisfying the moment condition). It is easy to construct such Φ's satisfying the moment constraints for small p's. As an example for arbitrary p, we can set Φ_p to be the ω_α-weighted Gegenbauer polynomial of degree p, (see (6.12)),

$$\Phi_p(x) = c_{\alpha,p}\Big(1 - \Big(\frac{x}{\pi}\Big)^2\Big)^{\alpha-\frac{1}{2}} C_p^{(\alpha)}\Big(\frac{x}{\pi}\Big) 1_{(-\pi,\pi)}(x), \quad \alpha > q.$$

(Clearly, such a $\Phi_p(x)$ is a C_0^p-function, which can be normalized to have a unit mass by a proper choice of $c_{\alpha,p}$.) The ω_α-orthogonality of the $C_k^{(\alpha)}$'s, (see (6.11)), implies that $\Phi_p^{(\alpha)}$ satisfies the moment condition (10.1).

Next, given a mollifier $\Phi(x) = \Phi_p(x)$ satisfying (10.1), we form the family of *dilated mollifiers*

$$\Phi_{p,\delta}(x) := \frac{1}{\delta}\Phi_p\Big(\frac{x}{\delta}\Big),$$

with δ being a free dilation parameter at our disposal; by tuning δ we can adjust the support of $\Phi_{p,\delta}$ over the symmetric interval $(-\pi\delta, \pi\delta)$. Observe that $\Phi_{p,\delta}$ retains the same p vanishing moments (10.1) over its restricted support $(-\pi\delta, \pi\delta)$. To reconstruct f from its spectral projection, we consider the mollified Fourier projection

$$\Phi_{p,\delta} * (S_N f)(x) \approx f(x).$$

We now turn to examine the error $\Phi_{p,\delta} * (S_N f) - f$. By orthogonality, $\Phi_{p,\delta} * (S_N f) = S_N(\Phi_{p,\delta}) * f$, hence we can express the error as the sum of two terms:

$$\Phi_{p,\delta} * (S_N f)(x) - f(x) \equiv \overbrace{\Big(S_N(\Phi_{p,\delta}) - \Phi_{p,\delta}\Big) * f(x)}^{\text{truncation}=T_N(\Phi_{p,\delta})*f} + \overbrace{\Big(\Phi_{p,\delta} * f(x) - f(x)\Big)}^{\text{regularization}} \tag{10.2}$$

The first term on the right is the usual *truncation error*, which, by (2.2), does not exceed

$$|T_N(\Phi_{p,\delta}) * f(x)| \lesssim \|f\|_{L^1} \cdot \|\Phi_{p,\delta}\|_{C^q} \frac{1}{N^{q-1}} \quad \lesssim \quad \alpha_{p,q} \frac{1}{\delta^{q+1} N^{q-1}}, \qquad \text{where } \alpha_{p,q} = \|\Phi_p\|_{C^q}. \tag{10.3a}$$

The second term on the right of (10.2) represents the *regularization error*,

$$\Phi_{p,\delta} * f(x) - f(x) = \int_{-\pi}^{\pi} \big[f(x - \delta y) - f(x)\big] \Phi_p(y) dy;$$

It does not involve any spectral content of f but depends solely on the regularity of f in the interval $(x - \pi\delta, x + \pi\delta)$. The moment condition implies that Φ_p is orthogonal to the first p terms in the Taylor expansion of

$$f(x - \delta y) - f(x) = \sum_{n=1}^{p} \frac{(-1)^n}{n!} \delta^n f^{(n)}(x) y^n + \frac{1}{(p+1)!} \delta^{(p+1)} f^{(p+1)}(\dots) y^{p+1},$$

and we are left with the following bound on the *regularization error*:

$$|\Phi_{p,\delta} * f(x) - f(x)| \quad \lesssim \quad \beta_p \|f\|_{C^{p+1}(x-\delta, x+\delta)} \delta^{p+1}, \qquad \text{where } \beta_p = \frac{1}{p!} \int_{-\pi}^{\pi} |y|^p |\Phi_p(y)| dy. \tag{10.3b}$$

There are two ways to make both error bounds, (10.3a) and (10.3b), small.

(i) Fix $p \sim q$ and set $\delta = \delta_N \sim 1/\sqrt{N}$. In this case, we recover $f(x)$ from the δ_N-neighborhood of it spectral projection $S_N f(x)$,

$$|\Phi_{p,\delta_N} * (S_N f)(x) - f(x)| \lesssim \gamma_p \big(1 + \|f\|_{C^p(x-\delta_N, x+\delta_N)}\big) \frac{1}{N^{p/2}}, \quad \delta_N = \frac{1}{\sqrt{N}}.$$

Here, $\gamma_p = \alpha_{p,p} + \beta_p$. This yields the locally supported mollifiers, Φ_{p,δ_N}. Their finite accuracy is determined by the finitely many vanishing moments in (10.1). To gain spectral accuracy requires an increasingly smooth mollifier Φ_p with an increasing number of (almost) vanishing moments. The result is a family of *local* mollifiers, Φ_{p_N,δ_N}, whose degree $p = p_N$ is adjusted as an increasing function of N. Local mollifiers do not make use of all the information available in the interval of smoothness enclosing x, and are therefore replaced by global mollifiers.

(ii) Fix δ by setting

$$\delta = d_x, \quad d_x := \frac{1}{\pi} dist\Big\{x, \{c_1, \ldots, c_J\}\Big\} [mod\ \pi], \tag{10.4}$$

so that $(x - \pi\delta, x + \pi\delta)$ is the largest interval of smoothness enclosing x. We pause here to make the following remark.

Remark 10.2. Note that d_x can be calculated from the given (pseudo-)spectral data. It is here that we use the information about the edges, $\{c_1, \ldots, c_J\}$, detected from $S_N f$ and $I_N f$.

Once we set $\delta = d_x$, we let p vary as an increasing function of N: by choosing $p = p_N$ (which necessarily has an increasing order of smoothness, $q_N > p_N$), we can try to enforce a spectrally small β_{p_N} in (10.3b) while balancing a spectrally small ratio $\alpha_{p_N,q_N} N^{1-q_N}$ in (10.3a). This balancing act depends of course on a careful study of the asymptotic behavior of $\|f\|_{C^p}$ and Φ_p, as p increases. The result is a family of *adaptive* mollifiers, Φ_{p_N,d_x}, whose degree $p = p_N$ is adapted as an increasing function of N while their support, d_x, is adapted to the largest interval of smoothness enclosing x. The Φ_{p_N,d_x} are *global* mollifiers; they achieve (root-)exponential convergence rate by *cancelation*.

10.2. Adaptive mollifiers: root-exponential accuracy

Following (Gottlieb and Tadmor 1985), we consider the compactly supported mollifiers

$$\Phi_p(x) := \rho_2(x) D_p(x), \qquad \rho_2(x) = e^{\left(\frac{cx^2}{x^2 - \pi^2}\right)} 1_{(-\pi,\pi)}(x), \tag{10.5}$$

where ρ_2 is the G_2 cut-off function imported from (2.5) to *localize* the Dirichlet kernel, D_p. Recall that after dilation, this family of mollifiers takes the form

$$\Phi_{p,d_x}(x) = \frac{1}{d_x} \rho_2\Big(\frac{x}{d_x}\Big) D_p\Big(\frac{x}{d_x}\Big)$$

where d_x is set by (10.4) and the degree p is at our disposal.

Let us estimate the error, $\Phi_{p,d_x} * (S_N f)(x) - f(x)$. Following (10.2),

we proceed in two steps, starting with the *truncation error*, $T_N(\Phi_{p,d_x})(x)$. According to (10.3a), its decay is controlled by the C^q-regularity of Φ_{p,d_x}. As the product of a G_2-function and the analytic Dirichlet kernel, we deduce $\Phi_{p,d_x} \in G_2$. But we still need to quantify the dependence of its G_2-bound on *both* p and d_x, as we are going to let p increase with N. To this end, we use the Leibniz rule and (2.6)

$$\begin{aligned}
|\Phi_p^{(s)}(x)| &\leq \sum_{k=0}^{s} \left(\begin{array}{c} s \\ k \end{array} \right) |\rho_2^{(k)}(x)| \cdot |D_p^{(s-k)}(x)| \\
&\lesssim s! \left(\sum_{k=0}^{s} \frac{p^{s-k}}{(s-k)!} \frac{1}{(\eta|x^2-\pi^2|)^k} \right) \cdot e^{\left(\frac{cx^2}{x^2-\pi^2} \right)} \\
&\lesssim \frac{s!}{(\lambda_\rho|x^2-\pi^2|)^s} e^{\left(p\lambda_\rho|x^2-\pi^2| + \frac{cx^2}{x^2-\pi^2} \right)},
\end{aligned}$$

which after dilation reads

$$|\Phi_{p,d_x}^{(s)}(x)| \lesssim s! \left(\frac{d_x}{|c(x)|} \right)^s \cdot e^{\left(\frac{p|c(x)|}{{d_x}^2} + \frac{c\lambda_\rho x^2}{c(x)} \right)}, \quad c(x) := \lambda_\rho\big(x^2 - \pi^2 {d_x}^2\big).$$

The upper bound on the right hand side is maximized at $x = x_{max}$ with $x_{max}^2 - \pi^2 {d_x}^2 \sim -c\pi^2 {d_x}^2/s$, which leads to the G_2-regularity bound for Φ_{p,d_x} (here, $\eta := c\lambda_\rho \pi^2$)

$$\sup_{x\in(-1,1)} |\Phi_{p,d_x}^{(s)}(x)| \lesssim s! \left(\frac{s}{\eta d_x e} \right)^s e^{\left(\frac{p\eta}{s}\right)} \lesssim \frac{(s!)^2}{(\eta d_x)^s} e^{\left(\frac{p\eta}{s}\right)} \quad s = 1, 2, \ldots \quad (10.6)$$

Equipped with (10.6), we find that the truncation error (10.3a) does not exceed

$$|T_N(\Phi)_{p,d_x}(x)| \lesssim \frac{(s!)^2}{(\eta d_x N)^s} e^{\left(\frac{p\eta}{s}\right)} \sim \left(\frac{s^2}{\eta d_x e^2 N} \right)^s e^{\left(\frac{p\eta}{s}\right)} =: M(s,p), \quad (10.7)$$

for all $s > 1$. We seek the minimizer, $s = s_{min}$, such that

$$\partial_s (\log M(s,p))_{|s=s_{min}} = \log(\frac{s_{min}^2}{\eta d_x N}) - \frac{p\eta}{s_{min}^2} = 0.$$

This yields a rather precise bound on s_{min} which turns out to be essentially independent of p. Indeed, for the first expression on the right to be positive we set $s_{min} = \sqrt{\beta \eta d_x N}$ with a free $\beta > 1$ at our disposal. The corresponding optimized $p = p_N(x) = \frac{s^2}{\eta} \log \frac{s^2}{\eta \delta N}_{|s=s_{min}}$ amounts to

$$p_N(x) = \beta \log \beta \cdot d_x N. \quad (10.8)$$

We conclude with an optimized choice of $p = p_N(x)$ of order $\mathcal{O}(d_x N)$, which is *adapted* to the distance between x and the singular support of f. The resulting exponentially small truncation error bound, (10.7), now reads

$$|T_N(\Phi_{p_N,d_x})(x)| \lesssim \frac{(s!)^2}{(\eta d_x N)^s} e^{\left(\frac{p_N \eta}{s}\right)}{}_{|s=\sqrt{\beta\eta d_x N}} \sim \sqrt{d_x N}\left(\frac{\beta}{e}\right)^{2\sqrt{\beta\eta d_x N}}. \tag{10.9a}$$

In the second step, we turn our attention to the *regularization error*

$$\begin{aligned}&\Phi_{p,d_x} * f(x) - f(x)\\ &\quad = \int_{-\pi}^{\pi} \overbrace{f(x - d_x y)\rho_2(y)}^{F(y;x)} D_p(y)dy - f(x) \equiv S_p F(y;x)\big|_{y=0} - F(0;x)\end{aligned}$$

in (10.2). Assume that $f(\cdot)$ is piecewise smooth: to simplify matters, we match its piecewise smoothness with that of ρ_2, assuming that f is a piecewise-G_2 function. Our choice of $\delta = d_x$ in (10.4) guarantees that $f(x - \delta y)$ is G_2 in the range $|y| \le \pi$ and, hence, so is its product with $\rho_2(y)$, implying the G_2-regularity of $F(y;\cdot) = f(\cdot - d_x y)\rho_2(y)$. When dealing with local mollifiers, we use the moment condition (10.1) to bound their regularization error in (10.3b). Instead, dealing with global mollifiers, we now use the global root-exponential decay (2.7b), which yields $\eta = \eta_{\rho,f}$ such that

$$|\Phi_{p_N,d_x} * f(x) - f(x)| = |\big(S_p F(y;x) - F(y;x)\big)_{|y=0}| \lesssim p\ e^{-\alpha\sqrt{\eta p}}.$$

The same choice of an adaptive $p = p_N$ made in (10.8) yields essentially the same exponentially small bound on the regularization error,

$$|\Phi_{p_N,d_x} * f(x) - f(x)| \lesssim d_x N \cdot e^{-2\sqrt{\beta \log\beta \cdot \eta d_x N}}. \tag{10.9b}$$

The free β can now be further optimized by equilibrating the truncation and regularization error bounds (10.9a) and (10.9b) where $\beta_{opt}\log\beta_{opt} \sim 1/\sqrt{e}$. We summarize with the following theorem.

Theorem 10.3. (Root-exponential accurate mollifiers) (Tadmor and Tanner 2002) Given the Fourier projection, $S_N f(\cdot)$ of a piecewise smooth function, $f(\cdot) \in \text{piecewise} - G_2$, we consider the 2-parameter family of spectral mollifiers

$$\Phi_{p,\delta}(x) := \frac{1}{\delta}\rho_2\Big(\frac{x}{\delta}\Big)D_p\Big(\frac{x}{\delta}\Big), \quad \rho_2 := e^{\left(\frac{cx^2}{x^2-\pi^2}\right)} 1_{(-\pi,\pi)}(x),\ c > 0. \tag{10.10a}$$

Fix x inside one of the smoothness intervals of f and set the adaptive pa-

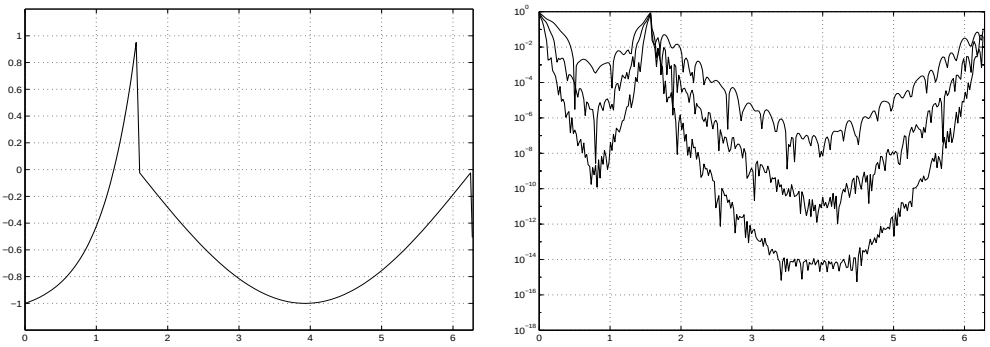

Figure 10.1: *Left:* Function $f(x)$ in (10.12). *Right*: Log error of its reconstruction from $S_N f$, $N = 32, 64, 128$ using the adaptive mollifier $\Phi_{p_N,\delta}$ in (10.10) with $p_N = d_x N/\sqrt{e}$.

rameterization

$$\delta = d_x := \frac{1}{\pi} dist\Big\{x, \{c_1, \ldots, c_J\}\Big\} [mod\ \pi], \tag{10.10b}$$

$$p = p_N(x) \sim d_x N/\sqrt{e}. \tag{10.10c}$$

Then, there exists a constant $\eta = \eta_{\rho,f}$ such that $\Phi_{p_N,d_x} * S_N f$ recovers $f(x)$ with the following root-exponential accuracy:

$$|\Phi_{p_N,d_x} * (S_N f)(x) - f(x)| \lesssim d_x N e^{-0.84\sqrt{\eta d_x N}}. \tag{10.11}$$

Figure 10.1 illustrates the reconstruction of

$$f(x) = \begin{cases} (2e^{2x} - 1 - e^{\pi})/(e^{\pi} - 1), & x \in [0, \pi/2) \\ -\sin(2x/3 - \pi/3), & x \in [\pi/2, 2\pi) \end{cases} \tag{10.12}$$

using the mollifier (10.10). Although the error estimates which lead to theorem 10.3 serve only as upperbounds for the errors, it is still remarkable that the (close to) optimal parameterization of the adaptive mollifier is found to be essentially independent of the properties of $f(\cdot)$.

We conclude this section with several remarks on the root-exponential accuracy behind the mollifiers Φ_{p_N,d_x} in (10.10).

Remark 10.4. (Spectral *vs.* root-exponential decay) The two-parameter family of mollifiers, $\Phi_{p,\delta}$, was introduced by (Gottlieb and Tadmor 1985). They used *spectral decay* bounds of the regularization and truncation errors,

$$|\Phi_{p,\delta} * f(x) - f(x)| \lesssim \|\rho_2\|_{C^s} \|f\|_{C^s(x-\delta,x+\delta)} \Big(\frac{2}{p}\Big)^s,$$

$$|T_N \Phi_{p,\delta}(x)| \lesssim \|\rho_2\|_{C^s} \Big(\frac{1+p}{\delta N}\Big)^s,$$

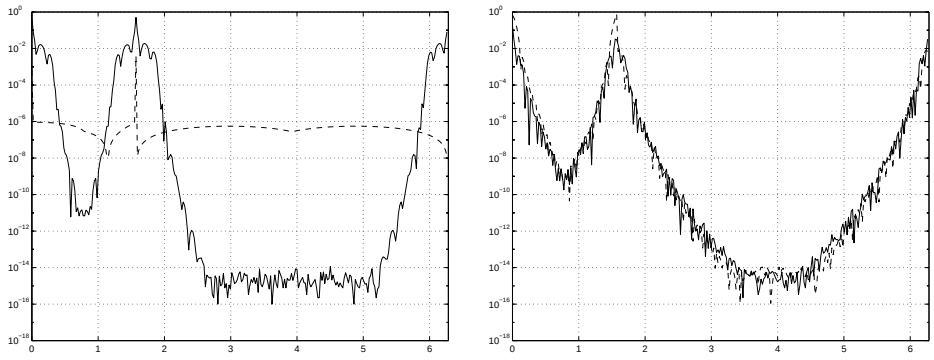

Figure 10.2: *Left:* Regularization errors (dashed) and truncation errors (solid) using the spectral mollifier $\Phi_{p,\delta}$ of degree $p \sim \sqrt{N}$ with δ. *Right:* The same using the adaptive mollifier Φ_{p_N,d_x} with $p_N = d_x N/\sqrt{e}$.

which led to the spectral decay

$$|\Phi_{p,\delta} * S_N f(x) - f(x)|_{p\sim\sqrt{N}} \leq Const_{s,d_x} \frac{1}{N^{s/2}}. \tag{10.13}$$

Although this estimate yields the desired spectral convergence rate sought for by (Gottlieb and Tadmor 1985), it suffers from coupling the regularization and truncation through the *same* dependence of p on s and on δ, which in turn leads to their balance at the pessimistic estimate of $p_N \sim \sqrt{N}$. Indeed, the results in figure 10.2 clearly indicate that the contributions of the truncation and regularization terms are equilibrated only when $p \sim N$. Moreover, figure 10.2 illustrates that a non-adaptive choice of $p = p_N$ which is independent of d_x, e.g., $p \sim \sqrt{N}$, leads to a loss of convergence in a larger $\mathcal{O}(N^{-1/2})$-neighborhood of the discontinuity, compared with the adaptive parameterization $p_N \sim d_x N$, which achieves, in figure 10.1(b), root-exponential accuracy up to the immediate, $\mathcal{O}(1/N)$ vicinity of these discontinuities.

Remark 10.5. (Piecewise smooth f's) The root-exponential error estimate (10.11) originates with the (piecewise-)smoothness of f and ρ_2 measures in G_2. Similar results apply in general cases when $f \in$ piecewise$-G_\alpha$ or $f \in$ piecewise$-C^s$. In this case, the convergence rate is, respectively, root-α exponential and s-order polynomial.

Remark 10.6. (Gevrey regularity) Optimal parameterization of Φ_{p_N,d_x} depends in an essential way on the *Gevrey regularity* of the cut-off function $\rho_2(\cdot) \in G_2$. This G_2-regularity is reflected for example, in the log-error in figure 10.1.

Remark 10.7. (Approximate moment condition) To achieve root-exponential

accuracy, we give up the *exact* moment condition. Instead, it is satisfied up to an exponentially small error (see (10.17)). Relaxing the side constraints imposed on an "approximate identity" will turn out to be a key feature which enables us to maintain (root-)exponential accuracy.

10.3. Root-exponential accurate reconstruction of pseudo-spectral data

We are interested in reconstruction of a piecewise smooth $f(x)$ from its Fourier interpolant, $I_N f(x)$. To this end we consider the discrete convolution

$$\Phi_{p_N,d_x} * I_N f(x) = h \sum_{\nu=0}^{2N} \Phi_{p_N,d_x}(x - y_\nu) f(y_\nu), \tag{10.14}$$

corresponding to (10.11). The overall error, $\Phi_{p_N,d_x} * I_N f(x) - f(x)$, consists of aliasing and regularization errors. According to (2.18), the former is upper bounded by the truncation of f', which retains the same piecewise analyticity properties as f does. We conclude with the following result.

Theorem 10.8. (Reconstruction of piecewise smooth discrete data) (Tadmor and Tanner 2002). Given equidistant grid-values, $\{f(x_\nu)\}_{\nu=0}^{2N}$ of $f(\cdot) \in$ piecewise $- G_2$, we consider the spectral mollifiers (10.10),

$$\Phi_{p,\delta}(x) := \frac{1}{\delta}\rho_2\Big(\frac{x}{\delta}\Big) D_p\Big(\frac{x}{\delta}\Big), \quad \rho_2 := e^{\left(\frac{cx^2}{x^2-\pi^2}\right)} 1_{(-\pi,\pi)}(x),\ c > 0,$$

with adaptive parameterization, $\delta = d_x := \frac{1}{\pi} dist\Big\{x, \big\{c_1, \ldots, c_J\big\}\Big\} [mod\ \pi]$ and $p_N(x) \sim d_x N/\sqrt{e}$. Then, there exists a constant $\eta = \eta_{\rho,f}$, such that the discrete convolution (10.14) recovers $f(x)$ with the following root-exponential accuracy:

$$\Big| h \sum_{\nu=0}^{2N} \Phi_{p_N,d_x}(x - y_\nu) f(y_\nu) - f(x) \Big| \lesssim (d_x N)^2 e^{-0.84\sqrt{\eta d_x N}}. \tag{10.15}$$

Observe that, by forming the discrete convolution (10.14), we completely bypass the need to compute the discrete Fourier coefficients, $\widehat{f_k}$. Instead, we recover $f(x)$ with root-exponential accuracy directly from the given gridvalues in the d_x-smooth neighborhood enclosing x,

$$\big\{f(y_\nu) \mid |x - y_\nu| \le d_x\big\}.$$

Thus, by relaxing the property of exact interpolation, $I_N f(x)_{|x=x_\nu} = f(x_\nu)$, the Fourier interpolant, $I_N f(x)$, is replaced here by what we might call the Fourier "expolant", $\Phi_{p_N,d_x} * I_N f$, a root-exponentially accurate approximant which recovers smooth as well as piecewise smooth f's.

10.4. Polynomial accuracy up to the edges

What happens in the neighborhood of jump discontinuities where $d_x N \sim 1$? We observe that the error bound (10.11) is of order $\mathcal{O}(1)$ as reflected in figure 10.1. To understand the source of this loss of accuracy, we note that the two ingredients involved in $\Phi_{p,\delta}$, namely, $\rho_2(x)$ and $D_p(x)$, have essentially different roles, associated with the two independent parameters δ and p: the role of $\rho_2\left(\frac{x}{\delta}\right)$ is to *localize* the support of $\Phi_{p,\delta}(x)$ to the δ-neighborhood of x; the Dirichlet kernel $D_p(x)$ is charged, by varying p, with controlling the increasing number of *near vanishing moments* of $\Phi_{p,\delta}$, and hence the overall superior accuracy of our mollifier. Indeed, since

$$\rho(0) = 1, \tag{10.16}$$

we find that the moments of Φ_{p_N,d_x} are of order

$$\begin{aligned}\int_{-\pi d_x}^{\pi d_x} y^n \Phi_{p_N,d_x}(y)dy &= \int_{-\pi}^{\pi} (yd_x)^n \rho_2(y) D_{p_N}(y)dy = \left(d_x\right)^n D_{p_N} * (y^n\rho_2)(y)_{|_{y=0}} \\ &\approx \delta_{n0} + \left(d_x\right)^n \inf_n \left\{ \|y^n\rho_2(y)\|_{C^n} \frac{1}{p_N^{n-1}} \right\} \approx \delta_{n0} + \left(d_x\right)^n e^{-\sqrt{\eta d_x N}}. \end{aligned} \tag{10.17}$$

Consequently, Φ_{p_N,d_x} possesses exponentially small moments at all x's except for the immediate vicinity of the jumps where $d_x N \sim 1$, the same $\mathcal{O}(1/N)$ neighborhoods where the error bound (10.11) is of order $\mathcal{O}(1)$. To enforce a faster convergence in these neighborhoods of the jumps, we ask that finitely many moments of $S_N\Phi_{p_N,d_x}$ vanish *exactly*,

$$\int_{-\pi}^{\pi} y^n (S_N\Phi_{p_N,d_x})(y)dy = \begin{cases} 1, & n = 0, \\ \\ 0, & n = 1, 2, \ldots, r. \end{cases}$$

This amounts to the vanishing moment constraint

$$\overbrace{\int_{-\pi}^{\pi} \Phi_{p_N}(y)dy = 1}^{\text{unit mass}}, \quad \overbrace{\int_{-\pi}^{\pi} S_N(y^n)\Phi_{p_N}(y)dy = 0}^{\text{vanishing moments}}, \ n = 1, 2, \ldots, r. \tag{10.18}$$

It follows that adaptive mollifiers satisfying (10.18) recover $f(x)$ with the desired polynomial order $\mathcal{O}(d_x)^r$, i.e.,

$$\Phi_{p_N,d_x} * S_N f(x) = f(x) + \log(p_N)\mathcal{O}(d_x)^{r+1}.$$

The point to note here is that this error estimate holds *up to* the edges. Indeed, noting that, for each x, the function $f(x - d_x y)$ remains smooth in the neighborhood $|y| \leq \pi$, the vanishing moments (10.18) imply

$$\Phi_{p_N,d_x} * S_N f(x) - f(x) = \int_{-d_x}^{d_x} \Phi_{p_N,d_x} S_N f(x-y)dy - f(x)$$
$$= \int_{-\pi}^{\pi} \Phi_{p_N}(y) S_N f(x-d_x y)dy - f(x) = \int_{-\pi}^{\pi} \Big[f(x-d_x y) - f(x)\Big](S_N \Phi_{p_N})(y)dy$$
$$\sim (d_x)^{r+1} \int_{-\pi}^{\pi} S_N(y^{r+1})\Phi_{p_N}(y)dy \lesssim \log(p_N)(d_x)^{r+1}.$$

To enforce (10.18) we modify the cut-off ρ_2, setting

$$\widetilde{\rho}_2(x) = \mathcal{M}_0 \rho_2(x), \qquad \mathcal{M}_0 := \frac{1}{\int_{\pi}^{\pi} \Phi_{p_N}(y)dy}. \tag{10.19a}$$

Observe that this normalizes $\widetilde{\rho}_2$ so that

$$\widetilde{\Phi}_{p_N,d_x} := \frac{1}{d_x}(\widetilde{\rho}_2\Big(\frac{x}{d_x}\Big) D_{p_N}\Big(\frac{x}{d_x}\Big), \qquad \widetilde{\rho}_2(x) = \mathcal{M}_0\rho_2(x), \tag{10.19b}$$

has a unit mass and hence (10.18) holds, at the expense of an exponentially negligible rescaling of ρ_2 in (10.16):

$$\widetilde{\rho}_2(0) = \mathcal{M}_0 = \frac{1}{(D_{p_N} * \rho_2)(0)} = 1 + d_x N e^{-2\sqrt{\eta p_N}}, \quad p_N \sim d_x N/\sqrt{e}.$$

Moreover, since Φ_{p_N,d_x} is even, its odd moments vanish, i.e., (10.18) holds for $r = 1$. We end up with the following corollary.

Corollary 10.9. (Uniformly quadratic, root-exponential mollifiers) The adaptive mollifier $\widetilde{\Phi}_{p_N,d_x}$ in (10.19) recovers piecewise smooth $f(x)$ with root-exponential accuracy at interior points of smoothness, and with quadratic accuracy in the vicinity of jump discontinuities, i.e.,

$$\big|\widetilde{\Phi}_{p_N,d_x}(x) * (S_N f)(x) - f(x)\big| \lesssim \log(d_x N)(d_x)^2 \cdot e^{-0.84\sqrt{\eta d_x N}}. \tag{10.20}$$

In a similar manner, we can enforce higher vanishing moments by proper *normalization* of the cut-off $\rho_2(\cdot)$. To enforce (10.18) with $r = 2$, for example, we use a rescaled cut-off, $\widetilde{\rho}_2(x)$, given by

$$\widetilde{\rho}_2(x) = \mathcal{M}_2\rho_2(x), \quad \mathcal{M}_2(x) := \frac{m_2(x)}{\int_{-\pi}^{\pi}, m_2(y)\Phi_{p_N}(y)dy} \tag{10.21a}$$

where

$$m_2(x) = 1 + a_2 x^2, \quad a_2 := \frac{-\int_{-\pi}^{\pi} S_N\big(y^2\big)\Phi_{p_N}(y)dy}{\int_{-\pi}^{\pi} S_N\big(y^2\big)y^2\Phi_{p_N}(y)dy}. \tag{10.21b}$$

As before, the resulting mollifier

$$\widetilde{\Phi}_{p_N,d_x}(x) := \frac{1}{d_x}\widetilde{\rho}_2\Big(\frac{x}{d_x}\Big)D_{p_N}\Big(\frac{x}{d_x}\Big), \qquad \widetilde{\rho}_2(x) = \mathcal{M}_2(x)\rho_2(x), \tag{10.21c}$$

is admissible in the sense of satisfying the normalization (10.16) modulo an exponentially small error term, since the pre-factor $\mathcal{M}_2(0) - 1$ is equally negligible. Now, $\widetilde{\Phi}_{p_N,d_x}$ has a unit mass; moreover, the S_N-projection of $\widetilde{\Phi}_{p_N,d_x}$ satisfies the *exact* second vanishing moment (10.18) with $r = 2$, for

$$\begin{aligned}\int_{-\pi}^{\pi} S_N\big(y^2\big)\Big(1 + a_2y^2\Big)\Phi_{p_N}(y)dy \\ = \int_{-\pi}^{\pi} S_N\big(y^2\big)\Phi_{p_N}(y)dy + a_2\int_{-\pi}^{\pi} S_N\big(y^2\big)y^2\Phi_{p_N}(y)dy = 0.\end{aligned}$$

Finally, since $S_N\widetilde{\Phi}_{p_N,d_x}$ is even, its third moment vanishes as well, which implies t the normalized mollifier (10.21).

Corollary 10.10. (Uniformly quartic, root-exponential mollifiers) The adaptive mollifier $\widetilde{\Phi}_{p_N,d_x}$ in (10.21) recovers piecewise smooth $f(x)$ with root-exponential accuracy at interior points of smoothness while maintaining a fourth-order convergence rate in the immediate vicinity of the jump discontinuities,

$$\big|\widetilde{\Phi}_{p_N,d_x}(x) * (S_Nf)(x) - f(x)\big| \lesssim \log(d_xN)(d_x)^4 \cdot e^{-0.84\sqrt{\eta d_xN}}. \tag{10.22}$$

We can implement a similar upgrade up to the edges in the discrete case. To this end, we consider the normalized mollifier

$$\widetilde{\Phi}_{p_N,d_x}(x) = \frac{1}{d_x}\widetilde{\rho}_2\Big(\frac{x}{d_x}\Big)D_{p_N}\Big(\frac{x}{d_x}\Big), \quad \widetilde{\rho}_2(x) := \mathcal{M}_r(x)\rho_2(x). \tag{10.23a}$$

Here, $\mathcal{M}_r(x)$ is a pre-factor of the form

$$\mathcal{M}_r(x) = \frac{m_r(x)}{\displaystyle\sum_{\nu} m_r\Big(\frac{x - y_\nu}{d_x}\Big)\Phi_{p_N,d_x}(x - y_\nu)h}, \quad m_r(x) = 1 + a_1x + \ldots + a_rx^r, \tag{10.23b}$$

whose r free parameters are sought so that the first r *discrete* moments of $\widetilde{\Phi}_{p_N,d_x}(y)$ vanish,

$$\sum_{\{y_\nu\,:\,|x-y_\nu|\le d_x\}} (x - y_\nu)^n\Phi_{p_N,d_x}(x - y_\nu)h = \begin{cases} 0, & n = 0, \\ \\ 0, & n = 1, \ldots, r. \end{cases} \tag{10.23c}$$

Observe that, unlike the moment constraint (10.18) associated with the continuous case, the discrete constraint (10.23c) is not translation invariant and

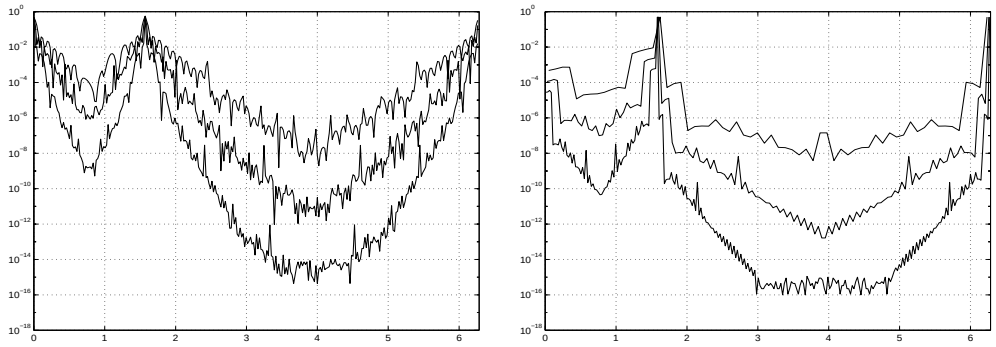

Figure 10.3: *Left*: Log error of reconstructing f from $S_N f,\ N = 32, 64, 128$ using the 4th-order normalized adaptive mollifier (10.21c) with $p_N = d_x N/\sqrt{e}$. *Right*: Reconstruction of f from its discrete data, $I_N f,\ N = 32, 64, 128$, using the 4th-order normalized mollifier (10.23).

hence requires x-dependent normalizations. The additional computational effort is minimal, however, due to the discrete summations which are localized in the immediate vicinity of x.

Then, (10.15) is replaced by the improved error estimate

$$\left| h \sum_{\nu=0}^{2N} \widetilde{\Phi}_{p_N, d_x}(x - y_\nu) f(y_\nu) - f(x) \right| \lesssim (d_x)^{r+1} e^{-0.84\sqrt{\eta d_x N}}, \quad r \sim N d_x. \tag{10.24}$$

Figure 10.3 illustrates the improvement using the normalized adaptive mollifier (10.21) and its discrete version (10.23), in reconstructing the same $f(x)$ used in (10.12)

$$f(x) = \begin{cases} (2e^{2x} - 1 - e^{\pi})/(e^{\pi} - 1), & x \in [0, \pi/2), \\ -\sin(2x/3 - \pi/3), & x \in [\pi/2, 2\pi). \end{cases}$$

Compared with the adaptive mollifier (10.10a) used in figure 10.1, the improvement of the error *up to* the edges is evident.

10.5. Gegenbauer-based mollifiers: exponential accuracy up-to the edges

We want to recover the values of a piecewise analytic $f(x)$ inside each interval of smoothness, with exponential accuracy, *uniformly* in $x \in (c_{j-1}, c_j),\ j = 1, \ldots, J$. After a proper translation and dilation of each interval, we may assume that f experiences a single jump discontinuity at $|x| = \pi$ and we seek exponential recovery of $f(x),\ |x| \le \pi$ *up to* the boundary. We have now come full circle, returning to out starting point, the Gegenbauer polynomials

$C_k^{(\alpha)}(x)$, which formed the moment-satisfying mollifiers in example 10.1. Let

$$G_N^{(\alpha)} f(x) := \sum_{k=0}^{N} \langle f, C_k^{(\alpha)} \rangle_{\omega_\alpha} C_k^{(\alpha)}(x)$$

denote the truncated Gegenbauer expansion of $f(x)$, $x \in (-1,1)$, where $\langle f, C_k^{(\alpha)} \rangle_{\omega_\alpha}$ are normalized moments of f with respect to the weight function $\omega_\alpha(x) = (1-x^2)^{\alpha-\frac{1}{2}}$. The *Gegenbauer reconstruction* of f, (see (Gottlieb and Shu 1998) and the references therein) is the reprojection of $S_N f(x)$,

$$G_p^{(\alpha)}\big(S_N f\big)_\pi(x), \qquad g_\pi := g(\pi x),$$

with a proper parameterization of $p = p_N$ and $\alpha = \alpha_N$.

Remark 10.11. Observe that the Gegenbauer reconstruction can be evaluated in terms of a (non-translatory) convolution with the corresponding Christoffel-Darboux mollifier (6.15),

$$G_p^{(\alpha)}\big(S_N f\big)_\pi(x) \sim \int_{-1}^{1} K_p^{(\alpha)}(x,y)(S_N f)_\pi(y)dy.$$

To determine these parameters, we upper-bound the error in the standard fashion, (10.2), by the sum of regularization and truncation errors,

$$G_p^{(\alpha)}\big(S_N f\big)_\pi(x) - f_\pi(x) = \overbrace{G_p^{(\alpha)} f_\pi(x) - f_\pi(x)}^{\text{regularization}} + \overbrace{G_p^{(\alpha)}\big(S_N f\big)_\pi(x) - G_p^{(\alpha)} f_\pi(x)}^{\text{truncation}}.$$

The regularization error does not involve any spectral information of f, but depends solely on the regularity of $f(x)$ over the interval $(-\pi,\pi)$. Since f_π is assumed to be analytic inside $(-1,1)$, its Gegenbauer projection is exponentially accurate *up to* the boundary,

$$|G_p^{(\alpha)} f_\pi(x) - f_\pi(x)| \le c_\alpha e^{-\eta p}, \qquad \text{for all } x \in (-1,1). \tag{10.25}$$

We now come to the truncation error which was shown to be upper-bounded by (Gottlieb, Shu, Solomonoff and Vandeven 1992):

$$\big\| G_p^{(\alpha)}\big(S_N f\big)_\pi(x) - G_p^{(\alpha)} f_\pi(x) \big\|_{L^\infty(-1,1)} \lesssim \left(\frac{\eta p}{N}\right)^{\alpha}.$$

Thus, to upgrade this polynomial decay in N, one has to increase $\alpha = \alpha_N$ while carefully balancing the growth of c_{α_N} in (10.25) by adjusting $p = p_N$. To this end, one sets $\alpha = \theta p \sim N$, to obtain exponentially small regularization and truncation errors (Gottlieb and Shu 1997).

The superior accuracy of the resulting Gegenbauer reconstruction is illustrated in figure 10.4, from (Gelb and Gottlieb 2007). It comes with a price, however: a sufficiently small θ needs to be *carefully tuned* (e.g., (Gottlieb and Shu 1997)) so that $\theta^{-\theta}\eta \lesssim 1$, where $\eta = \eta_f$ measures the width of the *ellipse* of analyticity of f in the complex plane (corresponding to the width of

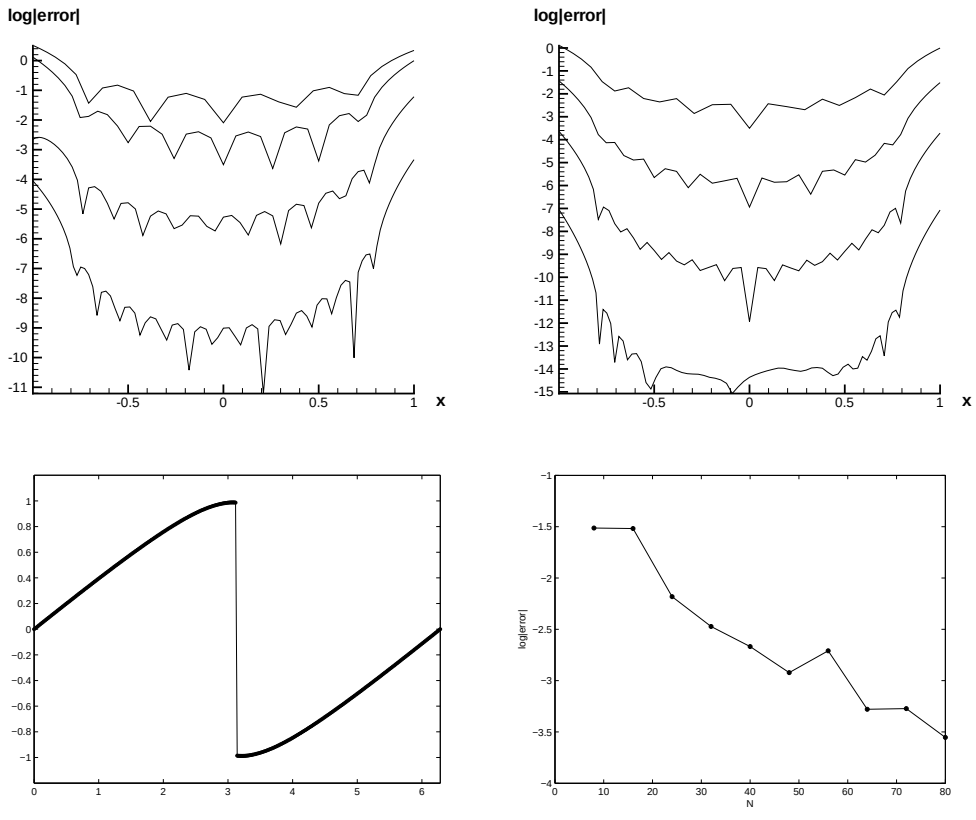

Figure 10.4: *Top*: Error in log-log scale of the Gegenbauer reconstruction $G_p^{(\alpha)}(S_N f)_\pi(x)$ with $N = 16, 24, 36$ and 52 modes for $f(x) = \cos(1.4\pi(x - 1))$, $x \in (-1, 1)$. *Top left*: $\alpha_N = p_N = N/4$. *Top right*: $\alpha_N = p_N = N/5$. *Bottom left*: Gegenbauer reconstruction of $I_{40}f(x)$, where $I_N f = u_N(x, t = 1.5)$ is a steady discontinuous solution of inviscid Burgers' equation, computed using a smoothed pseudo-spectral Fourier projection, $\partial_t u_N(x_\nu, t) + \partial_x\Big(\frac{I_N(u_N)^2}{2}\Big)(x_\nu, t) = 0$ and subject to $u_N(x_\nu, t = 0) = \sin(x_\nu)$. *Bottom right*: Log-log plot of the error.

the analyticity *strip* in the periodic case: e.g., (Tadmor 1986)). This translates into tuning of p_N, α_N, depending on the different analyticity regions for each smoothness interval of f. The overall Gegenbauer reconstruction method becomes rather sensitive to its parameterization (Boyd 2005). The superior accuracy is achieved at the expense of losing the robustness we had with the reconstruction methods based on adaptive mollifiers. A more robust reconstruction was offered recently in (Gelb and Tanner 2006), where the strongly peaked Gegenbauer weight, $\omega_\alpha(x) = (1-x^2)^{\alpha-\frac{1}{2}}$, is replaced by the Froud weight $\omega_m(x) = e^{-cx^{2m}}$.

11. Spectral filters

11.1. Adaptive filters: root-exponential accuracy

In Section 10 we showed how to parameterize an optimal mollifier, $\Phi_{p_N,d_x}(x)$, in order to gain the root-exponential convergence for piecewise analytic f's. The key ingredient in our approach was *adaptivity*, where the $\delta = d_x$ and $p_N \sim d_x N$ were adapted to the maximal region of local smoothness. Here we continue the same line of thought by introducing *adaptive filters*, which allow the same root-exponential recovery of piecewise analytic functions.

We consider a family of general filters $\varphi(\cdot) \in C^q(\mathbb{R})$, operating in Fourier space:

$$S_N^\varphi(x) := \sum_{|k|\le N} \varphi\left(\frac{|k|}{N}\right) \widehat{f}(k) e^{ikx}. \tag{11.1}$$

They are characterized by two main properties.

(i) First, we seek the *rapid smooth decay* of $\varphi(\xi)$ as ξ moves away from the origin. Translated from Fourier to physical space, the operation of $S_N^\varphi f$ corresponds to mollification against the smoothing kernel $\mathcal{S}_N^\varphi(x)$[6]:

$$S_N^\varphi f(x) = \mathcal{S}_N^\varphi * (S_N f)(x), \qquad \mathcal{S}_N^\varphi(x) := \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \varphi\Big(\frac{|k|}{N}\Big) e^{ikx}.$$

Then the rapid smooth decay of $\varphi(\cdot)$ is responsible for $\mathcal{S}_N^\varphi(x)$, which is *strongly localized* around $x = 0$.

(ii) Second, the mollifier $\mathcal{S}_N^\varphi(x)$ associated with the filter $\varphi(\xi)$ is required to satisfy a moment condition, (10.1). This property drives the *accuracy* by annihilating an increasing number of the moments of $\mathcal{S}_N^\varphi$. As observed in remark 10.7, however, a key ingredient in the construction of exponentially accurate mollifiers in Section 10.2 was giving up the exactness of (10.1). In a

[6] As before, the *function* $\mathcal{S}_N^\varphi(x)$ represents a *smoothing kernel* associated with but otherwise different from the corresponding filtering *operator*, $S_N^\varphi f$.

similar manner, in our quest for exponentially accurate filters, the moment condition (10.1) is replaced with the *accuracy condition*

$$\varphi^{(n)}(0) = \delta_{n0}, \qquad n = 0, 1, \ldots, p. \tag{11.2}$$

It follows that if the filter φ is p-order accurate, then its associated mollifier, $\mathcal{S}_N^\varphi$, satisfies the moment condition to order p.

We can quantify the above statements in a more precise manner, with a couple of examples which are summarized in the following two claims.

Claim 11.1. (About the rapid decay of $\mathcal{S}_N^\varphi(x)$) Let φ be a $C_0^q[-1,1]$-filter. Then, its associated mollifier,

$$\mathcal{S}_N^\varphi(x) := \frac{1}{2\pi} \sum_{|k|\le N} \varphi\Big(\frac{|k|}{N}\Big) e^{ikx},$$

is strongly localized near the origin in the sense that

$$|\mathcal{S}_N^\varphi(x)| \lesssim N\|\varphi\|_{C^q} \frac{1}{(|x|N)^q}, \quad 0 < |x| \le \pi. \tag{11.3a}$$

Thus, the smoother φ, the better $\mathcal{S}_N^\varphi$ is localized. As an example, we state as immediate consequence of (11.3a).

Example 11.2. If $\varphi \in G_\alpha$ then $\mathcal{S}_N^\varphi(x)$ experiences the root-exponential decay, namely, there exists $\eta = \eta_1$ (depending on φ) such that, for all $|x| \le \pi$,

$$|\mathcal{S}_N^\varphi(x)| \lesssim N \min_q \frac{(q!)^\alpha}{(\eta_\varphi |x| N)^q} \lesssim (1 + |x|N) e^{-\eta_1 \sqrt[\alpha]{|x|N}}, \quad \varphi \in G_\alpha. \tag{11.3b}$$

At the end of the 'smoothness scale', we find the entire function φ with quadratic exponential decay (2.8a); the mollifier $\mathcal{S}_N^{\varphi_{\delta_N}}(x)$, with $\delta_N = \sqrt{\beta N}$, admits exponential decay (2.11a).

We turn to verify claim 11.1 in two different ways. First, we rewrite $\mathcal{S}_N^\varphi(x)$ in the form

$$\mathcal{S}_N^\varphi(x) = \sum_{|k|\le N} \varphi(\xi_k) \frac{e^{i(k+1)x} - e^{ikx}}{e^{ix} - 1}, \quad \xi_k := kh,\ h = \frac{1}{N}.$$

Summation by parts yields

$$\mathcal{S}_N^\varphi(x) = \frac{1}{e^{ix} - 1} \sum_{|k|\le N} \Delta_h \varphi(|\xi_k|) e^{ikx} + \text{a couple of boundary terms},$$

and by repeating this argument,

$$\mathcal{S}_N^\varphi(x) = \frac{h^q}{(e^{ix} - 1)^q} \sum_{|k|\le N} h^{-q} \Delta_h^q \varphi(|\xi_k|) e^{ikx} + 2q \text{ boundary terms}.$$

We can safely neglect the small boundary terms (precisely because of (11.2)) and the C_0^q-regularity of $\varphi(\cdot)$ implies (11.3a).

A second approach is to use the Poisson summation formula (2.9b), expressing $\mathcal{S}_N^\varphi(x)$ in terms of Φ, the inverse Fourier transform of φ,

$$\mathcal{S}_N^\varphi(x) \equiv \frac{N}{2\pi} \sum_{j=-\infty}^{\infty} \Phi(N(x+2\pi j)), \qquad \Phi(x) = \int_{\mathbb{R}} \varphi(\xi) e^{i\xi x} d\xi.$$

This, together with the spectral decay estimate (2.3), yields (11.3a)

$$|\mathcal{S}_N^\varphi(x)| \lesssim N^{1-q} \|\varphi\|_{C^q} \sum_{j=-\infty}^{\infty} \frac{1}{|x+2\pi j|^q} \lesssim N\|\varphi\|_{C^q} \frac{1}{(N|x|)^q}, \quad 0 < |x| \le \pi. \qquad \blacksquare$$

Claim 11.3. (Accuracy and moment condition) If $\varphi \in C_0^q[-1,1]$ satisfies the *accuracy condition* of order $p < q$,

$$\varphi^{(n)}(0) = \delta_{n0}, \qquad n = 0, 1, \ldots, p,$$

then $\mathcal{S}_N^\varphi(x)$ satisfies the moment condition to order p,

$$\int_{-\pi}^{\pi} y^n \mathcal{S}_N^\varphi(y) dy = \delta_{n0}, \quad n = 0, 1, \ldots, p. \tag{11.4a}$$

Moreover, $\mathcal{S}_N^\varphi(x)$ concentrates in a neighborhood of the origin in the sense that the contribution to its moments outside such a neighborhood is negligible,

$$\Big| \int_{|y|\ge r}^{\pi} y^n \mathcal{S}_N^\varphi(y) dy \Big| \lesssim \|\varphi\|_{C^p} \frac{1}{(rN)^{p-1}}, \quad n = 0, 1, \ldots, p. \tag{11.4b}$$

Thus, the more accurate φ is, the better $\mathcal{S}_N^\varphi$ satisfies the moment condition. As an example we state the following immediate consequence of (11.4).

Example 11.4. If $\varphi = \varphi_p$ is a G_α-filter which is accurate of order $p = p_N \sim (rN)^{1/\alpha}$, then the unit mass $\mathcal{S}_N^\varphi$ has vanishing moments to order p_N. Moreover, there exists $\eta_2 > 0$ (depending on φ) such that

$$\int_{|y|\ge r}^{\pi} y^n \mathcal{S}_N^\varphi(y) dy \;=\; \delta_{n0} + \mathcal{O}\left(\min_{p \le \sqrt[\alpha]{rN}} \frac{(p!)^\alpha}{(\eta_2 rN)^p} \right) = \delta_{n0} + \mathcal{O}\Big(e^{-\eta_2 p_N}\Big), \tag{11.5}$$
$$\text{for } n \le p_N \sim (rN)^{1/\alpha}, \quad \varphi \in G_\alpha.$$

To verify the first part of claim 11.3, we appeal again to the Poisson formula (2.9b), expressing $\mathcal{S}_N^\varphi(x)$ in terms of translates of $\Phi(x)$:

$$\int_{-\pi}^{\pi} y^n \mathcal{S}_N^\varphi(y) dy = \frac{N}{2\pi} \int_{-\pi}^{\pi} y^n \Phi(Ny) dy \;+\; \frac{N}{2\pi} \sum_{j\ne 0} \int_{-\pi}^{\pi} y^n \Phi(N(y+2\pi j)) dy$$

$$= \overbrace{\frac{N}{2\pi}\int_{-\infty}^{\infty} y^n \Phi(Ny)dy}^{\mathcal{I}_1} - \overbrace{\frac{N}{2\pi}\int_{|y|\ge\pi} y^n \Phi(Ny)dy}^{\mathcal{I}_2}$$
$$+ \overbrace{\frac{N}{2\pi}\sum_{j\neq 0}\int_{-\pi}^{\pi} y^n \Phi(N(y+2\pi j))dy}^{\mathcal{I}_3}.$$

The first term on the right equals $\mathcal{I}_1 = (iN)^{-n}\varphi^{(n)}(0)$, since $\Phi(x)$ is the inverse Fourier transform of $\varphi(\xi)$, and therefore, since φ is p-order accurate, $\mathcal{I}_1 = \delta_{n0},\ n \le p$. The second and third terms cancel, and (11.4a) follows. To verify the second part of the claim, we use the Poisson summation formula again to write

$$\int_{r\le|y|\le\pi} y^n \mathcal{S}_N^{\varphi}(y)dy \;=\; \overbrace{\frac{N}{2\pi}\int_{r\le|y|\le\pi} y^n \Phi(Ny)dy}^{\mathcal{I}_1}$$
$$+ \overbrace{\frac{N}{2\pi}\int_{r\le|y|\le\pi} y^n\big(\sum_{j\neq 0}\Phi(N(y+2\pi j))\big)dy}^{\mathcal{I}_2}.$$

The usual spectral decay rate $|\Phi(y)| \lesssim \|\varphi\|_{C^p}\cdot|y|^{-p}$ implies

$$|\mathcal{I}_1| \lesssim N^{1-p}\|\varphi\|_{C^p}\int_{|y|\ge r} y^{n-p}dy \lesssim \|\varphi\|_{C^p}\frac{1}{(rN)^{p-1}} \qquad n = 0, 1, \ldots, p.$$

Similarly,

$$|\mathcal{I}_2| \lesssim N\|\varphi\|_{C^p}\sum_{j\neq 0}\frac{\pi^n}{\big((2j-1)N\pi\big)^p} \lesssim \|\varphi\|_{C^p}\frac{1}{N^{p-1}}, \qquad n = 0, 1, \ldots, p,$$

and (11.4b) follows. ■

We note that it is rather simple to construct admissible filters satisfying the last requirement for an *arbitrary* p; a prototype example is given by the G_2-filters

$$\varphi_p(\xi) = e^{\left(\frac{\xi^p}{\xi^2-1}\right)}1_{(-1,1)}(\xi). \tag{11.6}$$

This should be contrasted with the more intricate construction of mollifiers satisfying the exact moment condition in example 10.1. We are now ready to state a key result.

Theorem 11.5. (Root-exponential filters) (Tadmor and Tanner 2005)

Assume that $f(\cdot)$ is piecewise analytic and let $S_N^{\varphi_p}$ denote the filtered sum

$$S_N^{\varphi_p} f(x) := \sum_{|k|\le N} \varphi_p\left(\frac{|k|}{N}\right) \widehat{f}(k) e^{ikx}, \quad \varphi_p(\xi) = e^{\left(\frac{\xi^p}{\xi^2-1}\right)} 1_{(-1,1)}(\xi). \tag{11.7}$$

We set the order $p = p_N(x) \sim \sqrt{d_x N}$ (p_N even) where, as usual,

$$d_x := \frac{1}{\pi} dist\Big\{x, \{c_1, \dots, c_J\}\Big\} [mod\ \pi],$$

so that $(x - \pi d_x, x + \pi d_x)$ is the largest interval of analyticity enclosing x. Then, the adaptive filter $S_N^{\varphi_{p_N}} f$ recovers the point values $f(x)$ within the following root-exponential accuracy:

$$|S_N^{\varphi_{p_N}} f(x) - f(x)| \lesssim d_x N \cdot e^{-\eta\sqrt{d_x N}}. \tag{11.8}$$

Here, the constant $\eta = \eta_{\varphi,f}$ is dictated by the specific Gevrey and piecewise analyticity properties of φ and f.

Proof. We begin by decomposing the filtering error into the usual truncation and regularization term (compare (10.2)),

$$S_N^{\varphi_{p_N}} f(\cdot) - f(\cdot) = \overbrace{\mathcal{S}_{p_N}^{\varphi} * S_N f - \mathcal{S}_{p_N}^{\varphi} * f}^{\text{truncation}} + \overbrace{\mathcal{S}_{p_N}^{\varphi} * f - f}^{\text{regularization}}.$$

Here, $\mathcal{S}_{p_N}^{\varphi}(x) \equiv \mathcal{S}_N^{\varphi_{p_N}}(x)$, is an abbreviated notation for the mollifier associated with φ_{p_N},

$$\mathcal{S}_{p_N}^{\varphi}(x) := \frac{1}{2\pi} \sum_{|k|\le N} \varphi_{p_N}\Big(\frac{|k|}{N}\Big) e^{ikx}.$$

Since $\mathcal{S}_{p_N}^{\varphi}$ is a trigonometric polynomial of degree $\le N$, the truncation error vanishes:

$$\mathcal{S}_{p_N}^{\varphi} * S_N f - \mathcal{S}_{p_N}^{\varphi} * f = (S_N \mathcal{S}_{p_N}^{\varphi} - \mathcal{S}_{p_N}^{\varphi}) * f \equiv 0.$$

We remain with the regularization error, which we split into two terms:

$$\begin{aligned} \overbrace{\mathcal{S}_{p_N}^{\varphi} * f(x) - f(x)}^{\text{regularization}} \;=\; & \overbrace{\int_{\theta d_x \le |y| \le \pi} \mathcal{S}_{p_N}^{\varphi}(y)\Big[f(x) - f(x-y)\Big]dy}^{\mathcal{I}_1} + \\ & + \overbrace{\int_{|y|\le \theta d_x} \mathcal{S}_{p_N}^{\varphi}(y)\Big[f(x) - f(x-y)\Big]dy}^{\mathcal{I}_2}. \end{aligned} \tag{11.9}$$

Here $\theta < 1$ is a free parameter at our disposal. The first term on the right of (11.9) is straightforward: the G_2-regularity of φ_p implies the root-

exponential decay of $\mathcal{S}^{\varphi}_{p_N}$, namely, (11.3b) with $p_N = \sqrt{d_x N}$ implies

$$|\mathcal{I}_1| \lesssim (1 + \theta d_x N) \cdot e^{-\eta_1 \sqrt{\theta d_x N}}, \quad \eta_1 = \eta_\varphi > 0. \tag{11.10}$$

We turn to the second error term, $\mathcal{I}_2$. As before, it will be shown to be small due to *cancelation* of oscillations with the increasing order p of $\mathcal{S}^{\varphi}_{p_N}$. To this end we use Taylor's expansion of $f(\cdot) - f(\cdot - y)$ to express $\mathcal{I}_2$ as

$$\mathcal{I}_2 = \overbrace{\sum_{1 \le n < \theta p_N} \frac{(-1)^n}{n!} f^{(n)}(x) \int_{|y| \le \theta d_x} y^n \mathcal{S}^{\varphi}_{p_N}(y) dy}^{\mathcal{I}_{21}} + \overbrace{\frac{(-1)^{\theta p_N}}{(\theta p_N)!} f^{(\theta p_N)}(\cdot) \int_{|y| \le \theta d_x} y^{\theta p_N} \mathcal{S}^{\varphi}_{p_N}(y) dy}^{\mathcal{I}_{22}}. \tag{11.11}$$

But since φ_{p_N} is accurate to order p_N, (11.4a) and (11.5) with $r = \theta d_x$ and $p_N = \sqrt{rN}$, tell us that

$$\int_{|y| \le \theta d_x} y^n \mathcal{S}^{\varphi}_{p_N}(y) dy = -\int_{|y| \ge r}^{\pi} y^n \mathcal{S}^{\varphi}_{p_N}(y) dy = \delta_{n0} + \mathcal{O}(e^{-\eta_2 \sqrt{\theta d_x N}}),$$

and hence

$$|\mathcal{I}_{21}| \lesssim \sum_{1}^{\theta p_N} \frac{\pi^n}{\eta_f^n} e^{-\eta_2 \sqrt{d_x N}} \lesssim e^{\sqrt{d_x N}(\kappa_1 \theta - \eta_2 \sqrt{\theta})}, \qquad \kappa_1 := \log(\pi/\eta_f). \tag{11.12a}$$

Finally, the term $\mathcal{I}_{22}$ is exponentially small since near the origin, $|\mathcal{S}^{\varphi}_N(y)| \lesssim 2^{p_N}$, e.g., by (2.14), and by choosing sufficiently small θ,

$$|\mathcal{I}_{22}| \lesssim \frac{1}{(\eta_f)^{\theta p_N}} (\theta d_x)^{\theta p_N} 2^{p_N} \lesssim \left(2\Big(\frac{\theta\pi}{\eta_f}\Big)^{\theta}\right)^{p_N} \lesssim e^{-\eta \sqrt{d_x N}}. \tag{11.12b}$$

Result, (11.8), follows from (11.10) and (11.12) by choosing appropriately small $\theta = \theta(\eta_f, \eta_1, \eta_2)$. ∎

Figure 11.1 illustrates the use of the adaptive filter (11.7) to reconstruct the same function dealt earlier with the adaptive mollifier (10.10) illustrated in figure 10.1.

11.2. Optimal filters: exponentially accurate reconstruction

To reconstruct piecewise analytic f's with *exponential* accuracy, we turn to the filters based on the exponential optimality of space-frequency localiza-

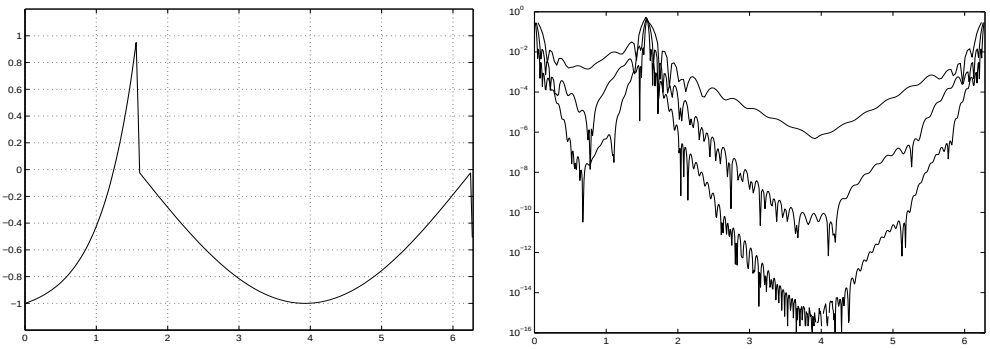

Figure 11.1: The function $f(x)$ in (10.12) and the log-error in its reconstruction from $S_N f,\ N = 32, 64, 128$ using the adaptive filter $S_N^{\varphi_{p_N}}$ in (11.7) with $p_N = \max(2, \sqrt{d_x N})$. .

tion discussed in Section 2.2,

$$\varphi_{p,\delta}(\xi) := e^{-\frac{(\delta\xi)^2}{2}} \sum_{j=0}^{p} \frac{1}{2^j j!} (\delta\xi)^{2j}.$$

The $\varphi_{p,\delta}$ are the truncated Hermite expansion of the weighted Gaussian, so that they form $(2p+1)$-order accurate filters in the sense that (11.2) holds. Since we are going to use adaptive parameterization where both $\delta = \delta_x$ *and* p increase with N, we now explicitly specify the dependence of φ on both. The corresponding $\varphi_{p,\delta}$-filter reads

$$S_N^{\varphi_{p,\delta}} f(x) = \sum_{|k|\le N} \varphi_{p,\delta}\Big(\frac{|k|}{N}\Big) \widehat{f}(k) e^{ikx}.$$

It can be expressed in terms of the associated mollifier, $\mathcal{S}_N^{\varphi_{p,\delta}}(x)$,

$$S_N^{\varphi_{p,\delta}} f(x) = \mathcal{S}_N^{\varphi_{p,\delta}} * (S_N f)(x), \qquad \mathcal{S}_N^{\varphi_{p,\delta}} := \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \varphi_{p,\delta}\Big(\frac{|k|}{N}\Big) e^{ikx}$$

We observe that, in this case, neither the filter nor its associated mollifier are compactly supported. Relaxing the constraint of having compact support in either physical space — as for $\Phi_p = \rho_2 D_p$ in (10.5) — or the Fourier space — as for $\Phi_p \leftrightarrow \varphi_p$ in (11.7), will enable us to obtain exponential accuracy after appropriate *adaptive* choice of the free parameters,

$$\delta_x := \sqrt{\theta d_x N}, \quad p_N := \theta^2 d_x N, \tag{11.13}$$

where d_x in (10.4) defines the usual analytic neighborhood enclosing x, and with $\theta < 1$ at our disposal. We use $\mathcal{S}^{\varphi}_{p_N,\delta_x}(x)$ to abbreviate the notation of

the corresponding mollifier

$$\mathcal{S}^{\varphi}_{p_N,\delta_x}(x) \equiv \mathcal{S}^{\varphi_{p_N,\delta_x}}_N(x) = \frac{1}{2\pi}\sum_{k=-\infty}^{\infty} \varphi_{p_N,\delta_x}\Big(\frac{|k|}{N}\Big)e^{ikx}.$$

To estimate the error,

$$S^{\varphi_{p_N,\delta_x}}_N f(x) - f(x) = \mathcal{S}^{\varphi}_{p_N,\delta_x} * (S_N f)(x) - f(x),$$

we first need to quantify the exponentially rapid decay of $\mathcal{S}^{\varphi}_{p_N,\delta_x}(x)$ in both physical *and* Fourier space.

We appeal to (2.11). Our choice of $\delta_x = \sqrt{\theta d_x N}$ corresponds to $\beta = \theta d_x$ and the spatial exponential decay in (2.11a) yields

$$|\mathcal{S}^{\varphi}_{p_N,\delta_x}(x)| \lesssim 2^{p_N}\sqrt{\frac{N}{d_x}}\Big(e^{-\eta_1\frac{Nx^2}{\theta d_x}} + e^{-\eta_2\frac{N}{\theta d_x}}\Big), \qquad |x| \le \pi.$$

Since $p_N = \theta^2 d_x N$, we have $2^{p_N} \le e^{\kappa_2\theta^2 d_x N}$ with $\kappa_2 := \log(2)$, and the last inequality confirms the exponential decay of $\Phi_{p_N,\delta_x}(x)$ outside the d_x-neighborhood of the origin. Indeed, by choosing sufficiently small $\theta < 1$,

$$\begin{aligned} |\mathcal{S}^{\varphi}_{p_N,\delta_x}(x)| &\lesssim \sqrt{\frac{N}{d_x}}\left(e^{(\kappa_2\theta^2 - \eta_1\theta)d_x N} + e^{\big(\kappa_2\theta^2 - \frac{\eta_2}{\theta {d_x}^2}\big)d_x N}\right) \\ &\lesssim \sqrt{\frac{N}{d_x}}e^{-\eta d_x N}, \qquad \theta d_x \le |x| < \pi. \end{aligned} \tag{11.14}$$

Next, we consider decay the Fourier space decay of $\varphi_{p_N,\delta_x}\big(\frac{|k|}{N}\big)$. Appealing to (2.11b) with $\beta = \theta d_x$, we find

$$\left|\varphi_{p_N,\delta_x}\Big(\frac{|k|}{N}\Big)\right| \lesssim c_{p_N,N}e^{-\frac{\theta d_x}{2}|k|}, \qquad c_{p_N,N} = \sum_{j=0}^{p_N}\frac{1}{j!}\left(\frac{\delta_x^2}{2}\right)^j.$$

The pre-factor $c_{p,N}$ has exponential growth of order

$$c_{p_N,N} = \sum_{j=0}^{\theta^2 d_x N}\frac{1}{j!}\left(\frac{\theta d_x N}{2}\right)^j \lesssim e^{\eta_\theta\frac{\theta d_x}{2}N},$$

but with a coefficient η_θ, which can be made sufficiently small by decreasing θ. Consequently, the last two inequalities imply that $|\varphi_{p_N,\delta_x}\big(\frac{|k|}{N}\big)|$ decay exponentially fast for $|k| > N$, i.e.,

$$\left|\varphi_{p_N,\delta_x}\Big(\frac{|k|}{N}\Big)\right| \lesssim e^{-\Big(1-\eta_\theta\frac{N}{|k|}\Big)\frac{\theta d_x}{2}|k|} \lesssim e^{-\eta d_x|k|}, \qquad |k| > N. \tag{11.15}$$

Equipped with (11.14), (11.15) we are ready to prove the following theorem.

Theorem 11.6. (Exponentially accurate filter) (Tanner 2006) Assume that $f(\cdot)$ is piecewise analytic, and let

$$S_N^{\varphi_{p,\delta}} f(x) := \sum_{|k|\le N} \varphi_{p,\delta}\left(\frac{|k|}{N}\right) \widehat{f}(k) e^{ikx}$$

denote the filtered Fourier projection, based on the quadratic exponential filter

$$\varphi_{p,\delta}(\xi) = \varphi_p(\delta\xi) := e^{-\frac{(\delta\xi)^2}{2}} \sum_{j=0}^{p} \frac{1}{j!}\Big(\frac{(\delta\xi)^2}{2}\Big)^j, \tag{11.16a}$$

of degree $p = p_N := \theta^2 d_x N$, with adaptive scaling $\delta = \delta_x := \sqrt{\theta d_x N}$. Here,

$$d_x = \frac{1}{\pi} dist\Big\{x, \big\{c_1, \ldots, c_J\big\}\Big\} [mod\ \pi],$$

defines a πd_x-neighborhood of analyticity around x. Then, for sufficiently small $\theta < 1$, there exists $\eta = \eta_{\theta,f} > 0$ such that the adaptive filter $S_N^{\varphi_{p_N,\delta_N}} f(x)$ recovers $f(x)$ with the following exponential accuracy:

$$|S_N^{\varphi_{p_N,\delta_x}} f(x) - f(x)| \lesssim \sqrt{\frac{N}{d_x}} e^{-\eta d_x N}. \tag{11.16b}$$

The constant $\eta = \eta_{\theta,f} > 0$ is dictated by the specific piecewise analyticity properties of f. The exponential adaptive filter takes the final form

$$S_N^{\varphi_{p_N,\delta_x}} f(x) = \sum_{|k|\le N} \left[\sum_{j=0}^{[\theta^2 d_x N]} \frac{1}{j!}\left(\frac{\theta d_x k^2}{2N}\right)^j\right] e^{-\frac{\theta d_x k^2}{2N}} \widehat{f}(k) e^{ikx}.$$

Proof. We proceed with an error decomposition similar to (11.9):

$$\begin{aligned} S_N^{\varphi_{p_N,\delta_x}} f(x) - f(x) &= \mathcal{S}^{\varphi}_{p_N,\delta_x} * f(x) - f(x) + \mathcal{S}^{\varphi}_{p_N,\delta_x} * S_N f(x) - \mathcal{S}^{\varphi}_{p_N,\delta_x} * f(x) \\ &= \overbrace{\int_{\theta d_x \le |y| \le \pi} \mathcal{S}^{\varphi}_{p_N,\delta_x}(y)\Big[f(x) - f(x-y)\Big]dy}^{\mathcal{I}_1} \\ &+ \overbrace{\int_{|y|\le \theta d_x} \mathcal{S}^{\varphi}_{p_N,\delta_x}(y)\Big[f(x) - f(x-y)\Big]dy}^{\mathcal{I}_2} + \overbrace{\Big(S_N \mathcal{S}^{\varphi}_{p_N,\delta_x} - \mathcal{S}^{\varphi}_{p_N,\delta_x}\Big) * f(x)}^{\mathcal{I}_3 = \text{truncation}}. \end{aligned} \tag{11.17}$$

To recall, $\mathcal{S}_N^{\varphi_{p,\delta}}(x)$ abbreviates the mollifier associated with the filter $\varphi_{p,\delta}$:

$$\mathcal{S}_N^{\varphi_{p,\delta}}(x) = \frac{1}{2\pi}\sum_{k=-\infty}^{\infty} \varphi_{p,\delta}\Big(\frac{|k|}{N}\Big)e^{ikx}.$$

We first observe the addition of a truncation error term, $\mathcal{I}_3$, which is due to the fact that $\varphi_{p,\delta}(\xi)$ is no longer compactly supported on $(-1,1)$, i.e., $\mathcal{S}_N^{\varphi_{p,\delta}}(x)$ is no longer a trigonometric polynomial of degree $\le N$. But the truncation error term is exponentially small because $|\varphi_{p,\delta}(|k|/N)|$ are; indeed, by (11.15) we have

$$|\mathcal{I}_3| \lesssim \big\|S_N\mathcal{S}^{\varphi}_{p_N,\delta_x} - \mathcal{S}^{\varphi}_{p_N,\delta_x}\big\|_{L^\infty} \lesssim \sum_{|k|>N} e^{-\eta d_x|k|} \lesssim e^{-\eta d_x N}. \tag{11.18}$$

The first term in the error decomposition can be made exponentially small because of the rapid decay of $\mathcal{S}^{\varphi}_{p_N,\delta_x}(x)$. Indeed, since the support of the first integrand is bounded θd_x from the origin, we find, thanks to (11.14),

$$|\mathcal{I}_1| \lesssim \int_{|y|\ge\theta d_x}^{\pi} |\mathcal{S}^{\varphi}_{p_N,\delta_x}(x)|dy \lesssim \sqrt{\frac{N}{d_x}}e^{-\eta d_x N}. \tag{11.19}$$

We now come to the second term $\mathcal{I}_2$. It can be made small because of the accuracy of $\varphi_{p_N,\delta_x}(\xi)$, which in turn implies that $\mathcal{S}^{\varphi}_{p_N,\delta_x}(x)$ has vanishing moments to order p_N, so that the *local* moments of $\Phi_{p_N,\delta_x}(x)$ equal

$$\overbrace{\int_{|y|\le\theta d_x} y^n\mathcal{S}^{\varphi}_{p_N,\delta_x}(y)dy}^{\text{moments associated with } \mathcal{I}_2} = -\int_{|y|\ge\theta d_x}^{\pi} y^n\mathcal{S}^{\varphi}_{p_N,\delta_x}(y)dy;$$

but the rapid decay of $\mathcal{S}^{\varphi}_{p_N,\delta_x}(y)$ in (11.14) implies

$$\Big|\int_{|y|\ge\theta d_x}^{\pi} y^n\mathcal{S}^{\varphi}_{p_N,\delta_x}(y)dy\Big| \lesssim \sqrt{\frac{N}{d_x}}\pi^{p_N}e^{-\eta d_x N} \lesssim \sqrt{\frac{N}{d_x}}e^{d_xN(\kappa_3\theta^2-\eta)}, \qquad \kappa_3=\log(\pi).$$

The estimate of $\mathcal{I}_2$ now follows the lines of theorem 11.5 using a similar decomposition into two terms, $\mathcal{I}_{21}+\mathcal{I}_{22}$, each of which is exponentially small due to the rapid decay of $\mathcal{S}^{\varphi}_{p_N,\delta_x}(y)$ outside the origin, (11.14). ■

Remark 11.7. (Exponentially accurate mollifier) Observe that the mollifier $\mathcal{S}_N^{\varphi_{p,\delta}}(x)$ associated with the filter $\varphi_{p,\delta_N}(\xi)$ in (11.16a) is exponentially close to $\Phi_{\delta,p}(Ny)$; consult (2.13) and (2.14a). Accordingly, we find the exponentially accurate *mollifier* $\Phi_{\delta,p}(Ny)$: with $\delta=\delta_x := \theta^2 d_x N$ and $p=p_N := \sqrt{\theta d_x N}$

$$\Phi_{\delta_x,p_N}(Ny) = \frac{1}{\sqrt{\theta d_x N}}e^{-\frac{Ny^2}{2\theta d_x}} \times \sum_{j=0}^{[\theta^2 d_x N]}\frac{(-1)^j}{4^j j!}H_{2j}\Big(\frac{\sqrt{N}y}{\sqrt{2\theta d_x}}\Big). \tag{11.20a}$$

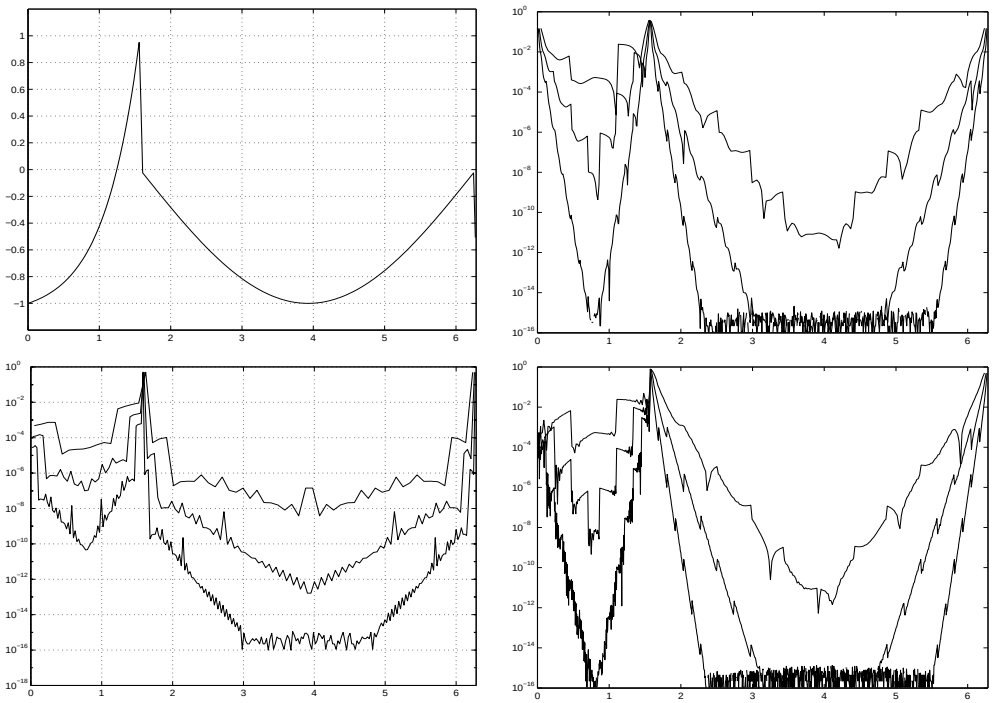

Figure 11.2: *Top left* Function $f(x)$ in (10.12). *Top right*: Log-error in its reconstruction from $S_N f$, $N = 32, 64, 128$ using the optimal filter (11.16b). *Bottom left*: Log-error in reconstruction of f from $I_N f$, $N = 32, 64, 128$ using the 4th-order normalized adaptive mollifier (10.21c). *Bottom right*: The same using the optimal pseudo-spectral mollifier (11.20b)

It is particularly useful to implement in the discrete case, where we end up with the exponentially accurate discrete mollifier (Tanner 2006, Theorem 4.2)

$$\Big| h \sum_{\nu=0}^{2N} \Phi_{\delta_x, p_N}\big(N(x - y_\nu)\big) f(y_\nu) - f(x) \Big| \lesssim \sqrt{\frac{N}{d_x}} e^{-\eta d_x N}. \tag{11.20b}$$

Figure 11.2 from (Tanner 2006), illustrates the superior convergence rate of the optimal filter (11.16b) (with $\theta \sim 1/4$) and its associated mollifier (11.20b).

Acknowledgments. A large portion of the material covered in this review is based on the work of the author with various collaborators, and it is my pleasure to acknowledge here Shlomo Engelberg, Anne Gelb, David Gottlieb, Jared Tanner and Jing Zou. I am indebted in particular to David Gottlieb, for years of friendship and collaboration which began with our first paper,

(Gottlieb and Tadmor 1985), the forerunner of a large body of subsequent work. Research was supported by NSF grant DMS04-07704 and by ONR grant N00014-91-J-1076.

REFERENCES

S. Abarbanel, D. Gottlieb and E. Tadmor (1986) 'Spectral methods for discontinuous problems', in "Numerical Methods for Fluid Dynamics II", Proceedings of the 1985 Conference on Numerical Methods for Fluid Dynamics (K. W. Morton and M. J. Baines, eds.), Clarendon Press, Oxford, pp. 129-153.

N.S. Banerjee and J. Geer (1997) Exponential approximations using Fourier series partial sums, ICASE Report No. 97-56, NASA Langley Research Center, 1997.

N. Bary (1964) 'Treatise of Trigonometric Series', The Macmillan Company, New York.

R. Bauer (1995) 'Band filters for determining shock locations, Ph.D. thesis, Applied mathematics, Brown University.

J. Bourgain (1988) 'A remark on the uncertainty principle for Hilbertian basis', J. Func. Anal. 79 , 136-143.

J. P. Boyd (1995) 'A Lag-Averaged Generalization of Euler's Method for Accelerating Series', Appl. Math. Comput., 143–166.

J. P. Boyd (1996) 'The Erfc-Log Filter and the Asymptotic of the Euler and Vandeven Sequence Accelerations', Proceedings of the Third International Conference on Spectral and High Order Methods, 267–276.

J. P. Boyd (2005) 'Trouble with Gegenbauer reconstruction for defeating Gibbs' phenomenon: Runge phenomenon in the diagonal limit of Gegenbauer polynomial approximations', J. Comput. Physics. —bf 204, 253–264.

O. Bruno, Y. Han and M. Pohlman (2006), 'Accurate, high-order representation of complex three-dimensional surfaces via Fourier-Continuation analysis', preprint.

E. Candes and F. Guo (2002) 'New multiscale transform, minimum total variation synthesis:applications to edge-preserving image reconstruction', Signal Processing 82, 1519-1543.

E. Candes and J. Romberg (2006) 'Quantitative robust uncertainty principles and optimally sparse decompositions', Found. Comput. Math. **6** , 227-254

E. Candes, J. Romberg, and T. Tao (2006a) 'Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information', IEEE Transactions on Information Theory **52**, 489–509.

E. Candes, J. Romberg, and T. Tao (2006b) 'Stable signal recovery from incomplete and inaccurate measurements', Communications on Pure and Applied Mathematics **59**, 1208–1223.

J. Canny (1986) 'A computational approach to edge detection', IEEE Trans. Pattern Analysis and Machine Intelligence, 8:679–714.

H.S. Carslaw (1952) 'Introduction to the Theory of Fourier's Series and Integrals', Dover.

J. Clark (1989) 'Authenticating edges produced by zero-crossing algorithms', IEEE Trans. Pattern Analysis and Machine Learning **11(2)**, 43–57.

R. DeVore and B. Lucier (1992) 'Wavelets', Acta Numer. **1**, 1–56.

R. DeVore and Lorentz (1993) 'Constructive Approximation', Springer.
D. Donoho (1998) 'Orthonormal ridgelets and linear singularities', www-stat.stanford.edu/~donoho/Reports/1998/ridge-lin-sing.pdf
D. Donoho (2004) 'Compressed sensing', www-stat.stanford.edu/~donoho/Reports/2004/CompressedSensing091604.pdf
D. Donoho and X. Huo (2001) 'Uncertainty principles and ideal atomic decomposition', IEEE Trans. Inform. Theory 47, 2845-2862.
D. Donoho, M. Elad, and V. Temlyakov (2006) 'Stable recovery of sparse overcomplete representations in the presence of noise', IEEE Trans. Information Theory, **52**, 6–18,
D. Donoho, J. Tanner (2005) 'Sparse nonnegative solutions of underdetermined linear equations by linear programming', Proc. Nat. Academy Sciences **102(27)** 9446–9451.
H. Dym and H. McKean (1972) 'Fourier Series and Integrals', Academic Press.
K.S. Eckhoff (1995) 'Accurate reconstructions of functions of finite regularity from truncated series expansions', *Math. Comp.* **64**, 671–690.
K.S. Eckhoff (1998) On a high order numerical method for functions with singularities, *Math. Comp.* 67, 1063-1087 (1998).
K.S. Eckhoff and C. E. Wasberg (1995) 'On the numerical approximation of derivatives by a modified Fourier collocation method', Rep. No. 99, Department of Mathematics, University of Bergen, Norway.
S. Engelberg and E. Tadmor (2007) 'Recovery of edges from noisy spectral data — a new perspective', preprint.
N. Firoozye and V. Sverak (1996) 'Measure filters: an extension of Weiner's theorem', Indiana Univ. Math. J. 45, 695-707.
G. Folland (1992), 'Fourier Analysis and its Applications', Brooks/Cole Publ.
B. Fornberg (1996) 'A Practical Guide to Pseudospectral Methods', Cambridge University Press.
A. Gelb (2004) 'Parameter optimization and reduction of round-off error for the Gegenbauer reconstruction method', J. Sci. Comput., 20(3) 433-459.
A. Gelb and S. Gottlieb (2007) 'The resolution of the Gibbs phenomenon for Fourier spectral methods', Preprint.
A. Gelb and E. Tadmor (1999) 'Detection of edges in spectral data', Appl. Comp. Harmonic Anal., 7 101–135.
A. Gelb and E. Tadmor (2000a) 'Detection of Edges in Spectral Data II. Nonlinear Enhancement', SIAM Journal of Numerical Analysis, 38, 1389–1408.
A. Gelb and E. Tadmor (2000b) 'Enhanced spectral viscosity approximations for conservation laws', Applied Numerical Mathematics 33, 3–21.
A. Gelb and E. Tadmor (2002) 'Spectral reconstruction of one- and two-dimensional piecewise smooth functions from their discrete Data', Mathematical Modeling and Numerical Analysis 36, 155–175.
A. Gelb and E. Tadmor (2006) 'Adaptive edge detectors for piecewise smooth data based on the minmod limiter', Journal of Scientific Computing 28(2-3) 279–306.
A. Gelb and J. Tanner (2006) 'Robust reprojection methods for the resolution of the Gibbs phenomenon', Appl. Comp. Harmonic Anal., 20:1 3–25. .
J. W. Gibbs (1899) 'Fourier Series', Nature 59, 606.

B. I. Golubov (1972) 'Determination of the jump of a function of bounded p-variation by its Fourier series', Math. Notes 12, 444–449.
D. Gottlieb and J. Hesthaven (2001) 'Spectral methods for hyperbolic problems', Journal of Computational Applied Mathematics **128** 83-131.
D. Gottlieb and S. Orszag (1977) 'Numerical Analysis of Spectral Methods: theory and applications', vol. **26** of CBMS-NSF, SIAM.
D. Gottlieb, C.-W. Shu, A. Solomonoff and H. Vandeven (1992) 'On the Gibbs phenomenon I: recovering exponential accuracy from the Fourier partial sum of a non-periodic analytic function', Journal of Computational and Applied Mathematics **43** pp. 81-98.
D. Gottlieb and C.-W. Shu (1995a) 'On the Gibbs phenomenon IV: recovering exponential accuracy in a sub-interval from a Gegenbauer partial sum of a piecewise analytic function', Mathematics of Computation **64**, pp. 1081-1095.
D. Gottlieb and C.-W. Shu (1995b) 'On the Gibbs phenomenon V: recovering exponential accuracy from collocation point values of a piecewise analytic function', Numerische Mathematik **71** pp. 511-526.
D. Gottlieb and C.-W. Shu (1997) 'The Gibbs phenomenon and its resolution', SIAM Review, 39 pp.644-668.
D. Gottlieb and C.-W. Shu (1998)'General theory for the resolution of the Gibbs phenomenon', Accademia Nazionale Dei Lincey, ATTI Dei Convegni Lincey **147** pp. 39-48.
D. Gottlieb and E. Tadmor (1985) 'Recovering pointwise values of discontinuous data within spectral accuracy',in "Progress and Supercomputing in Computational Fluid Dynamics", Proceedings of a 1984 U.S.-Israel Workshop, Progress in Scientific Computing, Vol. 6 (E. M. Murman and S. S. Abarbanel, eds.), Birkhäuser, Boston, pp. 357-375.
L. Grafakos (2004) 'Classical and Modern Fourier Analysis', Pearson Education.
A. Harten (1983) 'High resolution schemes for hyperbolic conservation laws', J. Comput. Phys. 49 , 357–393.
E. Hewitt and R. Hewitt (1979) 'The Gibbs-Wilbraham phenomenon: an episode in Fourier analysis', Archive for History of the Exact Sceinces 21, 129–160.
D. K. Hoffman and D. J. Kouri (2000) 'Hierarchy of local minimum solutions of Heisenberg's uncertainty principle', Phys. Rev. Lett. 85 5263-5267.
Y. Katznelson (1976) 'An Introduction to Harmonic Analysis', Dover.
T. W. Kröner (1995) 'Fourier Analysis', Cambridge.
G. Kvernadze (1998) 'Determination of the jump of a bounded function by its Fourier series', J. Approx. Theory 92, 167–190.
Y. Maday. S. Ould-Kaber and E. Tadmor (1993) 'Legendre pseudospectral viscosity method for nonlinear conservation laws' SIAM J. Numer. Anal. 30, 321-342.
A. Majda, J. McDonough and S. Osher (1978) 'The Fourier method for nonsmooth initial data', Math. Comput. 30, 1041–1081.
S. Mallat (1989) 'Multiresolution approximations and wavelets orthonormal bases of $L^2(R)$', Trans. Amer. Math. Soc., 315, pp. 69–87.
D. Marr and E. Hildreth (1980) 'Theory of edge detection', Proc. Roy. Soc. Lond., B207:187 - 217.
H.N. Mhaskar and J. Prestin (2000) 'On the detection of singularities of a periodic function', Adv. Comput. Math. **12**, 95–131.

S. Pilipovic and N. Teofanov (2002) 'Wilson bases and ultramodulation spaces', Math. Nachr., 242, 179-196.

N. Srinivasa and K. Rajgopal (1992) "Detection of edges from projections', IEEE Trans. Medical imaging, 11(1), 76-80.

E. Stein (1993) 'Harmonic Analysis: real-variable methods, orthogonality and oscillatory integrals', vol **32** of Princeton Mathematical Series, Princeton Univ. Press.

G. Szego (1958) 'Orthogonal Polynomials', American Mathematical Society, Providence, RI, 1958.

E. Tadmor (1986) 'The exponential accuracy of Fourier and Chebyshev differencing methods', SIAM Journal on Numerical Analysis 23, 1–10.

E. Tadmor (1989) 'Convergence of spectral methods for nonlinear conservation laws', SIAM J. Numer. Anal. 26, 30-44.

E. Tadmor (1994) 'Spectral methods for hyperbolic problems', Lecture notes delivered at Ecole des Ondes, INRIA - Rocquencourt January 24-28, www.cscamm.umd.edu/people/faculty/tadmor/pub/spectral-approximations/Tadmor.INRIA-94.pdf.

E. Tadmor (1998) 'Approximate solutions of nonlinear conservation laws' in "Advanced Numerical Approximation of Nonlinear Hyperbolic Equations" Lecture notes in Mathematics **1697**, 1997 C.I.M.E. course in Cetraro, Italy (A. Quarteroni ed.) Springer Verlag, 1-149.

E. Tadmor and J. Tanner (2002) 'Adaptive Mollifiers - High Resolution Recovery of Piecewise Smooth Data from its Spectral Information', J. Foundations of Comp. Math. 2, 155-189.

E. Tadmor and J. Tanner (2003) 'An adaptive order Godunov type central scheme', "Hyperbolic Problems: Theory, Numerics, Applications", Proceedings of the 9^{th} International Conference in Pasadena, March 2002 (T. Hou & E. Tadmor, eds.), Springer, pp. 871–880.

E. Tadmor and J. Tanner (2005) 'Adaptive filters for piecewise smooth spectral data', IMA Journal of Numerical Analysis 25(4) 635-647.

E. Tadmor and J. Zou (2007) 'Novel edge detection methods for incomplete and noisy spectral data', preprint.

J. Tanner (2006) 'Optimal filter and mollifier for piecewise smooth spectral data', Math. Comp.

T. Tao (2005) 'An uncertainty principle for cyclic groups of prime order', Math. Res. Lett. 12, 121-127.

A. Torchinsky (1986) 'Real-Variable Methods in Harmonic Analysis', Academic Press.

F. Ulupinar and G. Medioni (1988) 'Refining edges detected by the LoG operator', Comput. vision, pattern recognition proc. CVPT'88, 202–207.

H. Vandeven (1991) 'Family of Spectral Filters for Discontinuous Problems', Journal of Scientific Computing, 6(2), 159-192.

A. Zygmund (1959) 'Trigonometric Series', Cambridge University Press.