PERGAMON Computers and Mathematics with Applications 41 (2001) 111–133

An International Journal
computers & mathematics
with applications

www.elsevier.nl/locate/camwa

Microlocal Filtering with Multiwavelets

R. Ashino
Division of Mathematical Sciences
Osaka Kyoiku University, Kashiwara, Osaka 582, Japan
ashino@cc.osaka-kyoiku.ac.jp

C. Heil
School of Mathematics
Georgia Institute of Technology, Atlanta, GA 30332-0160, U.S.A.
heil@math.gatech.edu

M. Nagase
Department of Mathematics
Graduate School of Science, Osaka University, Toyonaka, 560, Japan
nagase@math.wani.osaka-u.ac.jp

R. Vaillancourt
Department of Mathematics and Statistics
University of Ottawa, Ottawa, ON, Canada K1N 6N5
remiv@mathstat.uottawa.ca

Dedicated to Professor Yoshimi Saito on the occasion of his 60^{th} birthday

(Received December 1999; accepted January 2000)

Abstract—Hyperfunctions in $\mathbb{R}^n$ are intuitively considered as sums of boundary values of holomorphic functions defined in infinitesimal wedges in $\mathbb{C}^n$. Orthonormal multiwavelets, which are a generalization of orthonormal single wavelets, generate a multiresolution analysis by means of several scaling functions. Microlocal analysis is briefly reviewed and a multiwavelet system adapted to microlocal filtering is proposed. A rough estimate of the microlocal content of functions or signals is obtained from their multiwavelet expansions. A fast algorithm for multiwavelet microlocal filtering is presented and several numerical examples are considered.

Keywords—Microlocal analysis, Filter, Multiwavelet, Analytic representation.

1. INTRODUCTION

The extraordinary development of wavelets in recent years have made them present in a large part of our high-technology world [1]. Wavelets are being incorporated in engineering standards for image and audio signal compression. One of the first standards based on wavelets is "wavelet scalar quantization", adopted by the U.S. Federal Bureau of Investigation (FBI) in 1997 to encode

This work was partially supported through a Grant-in-Aid of Japan, NSERC of Canada and the Centre de recherches mathématiques of the Université de Montréal, and the National Science Foundation.

PII: S0898-1221(00)00261-3

fingerprints. The forthcoming still-image compression standard known as JPEG2000 includes a wavelet option and MPEG-4, the next video compression standard, will be entirely wavelet-based. Wavelet-based high-compression modems will increase the usable bandwidth of electrical signals over a telephone line by a factor of 250 from 4 kHz to 1 MHz. Developments in wavelets have influenced many pure and applied mathematicians and scientists in such disparate fields as numerical analysis, computer vision, human vision, turbulence, statistics, physics, and medicine. This paper is an attempt to use wavelets in the study of hyperfunctions and their microlocal analysis.

Intuitively, hyperfunctions, which were introduced by Sato [2] and extensively developed by the Kyoto School of Mathematics, can be considered as sums of boundary values of holomorphic functions defined in infinitesimal wedges. Hyperfunctions are powerful tools in several applications; for example, vortex sheets in two-dimensional fluid dynamics are a realization of one-dimensional hyperfunctions. Analytic continuation in domains of special forms plays a key role in the theory of hyperfunctions. A simple example of a hyperfunction is the Dirac delta measure $\delta(x)$, which, when applied to a continuous functions $f(x)$, produces the value $f(0)$,

$$\int_{\mathbb{R}} f(x)\,\delta(x)\,dx = f(0).$$

Since, in Schwartz's theory of distributions [3], smooth testing functions of compact support cannot be holomorphic functions, Sato used the Cauchy integral formula to define $\delta(x)$ applied to a holomorphic function $f(z)$ on an open set $D \subset \mathbb{C}$. Assuming that $0 \in D$ and letting $\gamma = \partial D$ denote the boundary of D, we have

$$\frac{1}{2\pi i}\oint_{\gamma} \frac{f(z)}{z}\,dz = f(0).$$

In the limit as the path γ is shifted to $-\gamma_+ + \gamma_-$, as shown in Figure 1, this formula becomes

$$\frac{1}{2\pi i}\oint_{\gamma} \frac{f(z)}{z}\,dz = \int_{-\infty}^{+\infty} \left(-\frac{1}{2\pi i}\right)\left(\frac{1}{x+i0} - \frac{1}{x-i0}\right) f(x)\,dx.$$

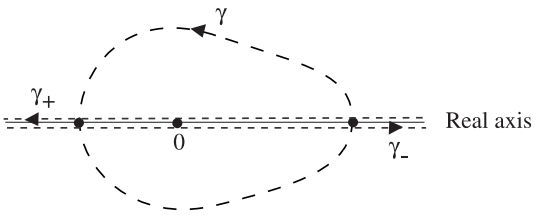


Figure 1. Shifting the path γ to $-\gamma_+ + \gamma_-$.

Thus,

$$\delta(x) = -\frac{1}{2\pi i}\left(\frac{1}{x+i0} - \frac{1}{x-i0}\right)$$

is defined as the limit of two holomorphic functions, one holomorphic in the upper half-plane, and the other holomorphic in the lower half-plane.

The classical construction of an orthonormal wavelet basis begins with a single function, the scaling function, which satisfies a particular functional equation known variously as refinement equation, dilation equation, or two-scale difference equation. This scaling function determines a multiresolution analysis for $L^2(\mathbb{R}^n)$, which in turn determines a wavelet basis [4]. The construction of a multiwavelet basis is analogous, the difference being that several scaling functions are used to generate the multiresolution analysis instead of a single scaling function. The corresponding wavelet basis is then generated by several *multiwavelets*. One advantage of multiwavelets is

that they can incorporate properties that a single wavelet cannot simultaneously possess, such as orthogonality, symmetry, short support, and high approximation order. Such constructions have shown potential in signal processing applications such as image compression or denoising [5,6]. Multiwavelets may well be ideally suited to multichannel signals such as color images (which are two-dimensional, three-channel signals) and stereo audio signals (which are one-dimensional, two-channel signals).

In this paper, we present a particular multiwavelet construction which is suited for microlocal filtering. Microlocal analysis plays an important role in the theory of hyperfunctions, partial differential operators, and many other areas. In this theory, one can define the product of distributions and discuss the partial regularity of multidimensional distributions with respect to any independent variable. Microlocal analysis could be called "local Fourier analysis".

We shall construct multidimensional multiwavelets which have the property that decomposition into this basis reveals directions of analyticity. The resolution of these multiwavelets in any given direction of analyticity can be made as fine as desired, at the cost of increasing the multiplicity of the multiwavelet basis. Each multiwavelet corresponds to one direction of analyticity, and each coefficient of the multiwavelet expansion of a function or signal gives a rough estimate of its microlocal content, or microanalyticity. This is impossible for single wavelets ("uniwavelets"). Since projections defined by means of our multiwavelets give a rough microlocal decomposition, they will be called *microlocal filters*. Microlocal filtering can be done numerically, and we provide some numerical examples. Furthermore, while, for simplicity, we restrict attention to $L^2(\mathbb{R}^n)$, we note that microlocal filtering could be applied to more general function classes. Finally, we present a library of orthonormal multiwavelet bases which may be suitable for implementation via the best basis wavelet packet algorithm.

2. MICROLOCAL ANALYSIS

In this section, we briefly review microlocal analysis for Schwartz distributions based on the theory of hyperfunctions, following [7].

We shall consider cones in the space $\mathbb{R}^n_y$ of imaginary coordinates. We assume that every cone has its vertex at the origin. A subset Γ of $\mathbb{R}^n$ is called a *cone* if for all $t > 0$, we have $t\Gamma \subset \Gamma$. A cone Γ is said to be *proper* if $\Gamma \setminus \{0\}$ is contained in an open half-space with boundary through the origin. If Γ and Γ' are two cones, the notation $\Gamma' \subset\subset \Gamma$ means that $\overline{\Gamma'} \cap \{|y| = 1\}$ is strictly contained in the interior of Γ. Note that the relation $\Gamma' \subset\subset \Gamma$ is not the same as $\Gamma' \subset \Gamma$. For example, if Γ' is a closed cone and Γ is an open cone, the origin is an exceptional point.

If Δ is an open set in $\mathbb{R}^n$, then the subset $\mathbb{R}^n + i\Delta := \{z = x + iy; x \in \mathbb{R}^n, y \in \Delta\}$ in $\mathbb{C}^n$ is called a *tubular domain*. A tubular domain of the form $\mathbb{R}^n + i\Gamma$ where Γ is a cone in $\mathbb{R}^n$ is called a *wedge* with *edge* $\mathbb{R}^n$ and *opening* Γ.

Let Γ be a cone $\mathbb{R}^n_y$. If Δ is an open set in $\mathbb{R}^n_y$ which approaches Γ asymptotically near the origin from the interior of Γ, then the subset $U = \mathbb{R}^n + i\Delta$ of $\mathbb{C}^n$ is called an *infinitesimal wedge* with opening Γ, and is denoted by $\mathbb{R}^n + i\Gamma 0$ (see Figure 2).

The following notation will be used.

NOTATION 1.

- Γ is an open cone in $\mathbb{R}^n$.
- Γ' is a closed cone in $\mathbb{R}^n$.
- $C^0(\mathbb{R}^n)$ is the space of continuous functions on $\mathbb{R}^n$.
- $\mathcal{D}'(\mathbb{R}^n)$ is the space of distributions on $\mathbb{R}^n$.
- $\mathcal{S}'(\mathbb{R}^n)$ is the space of slowly increasing distributions on $\mathbb{R}^n$.
- $\mathcal{O}(U)$ is the set of holomorphic functions in the open set $U \in \mathbb{C}^n$.
- $\mathbb{Z}_+ = \{0, 1, 2, \dots\}$ is the set of natural numbers including zero.
- $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$ with $\alpha_j \in \mathbb{Z}_+$ is a multi-index of nonnegative integers.

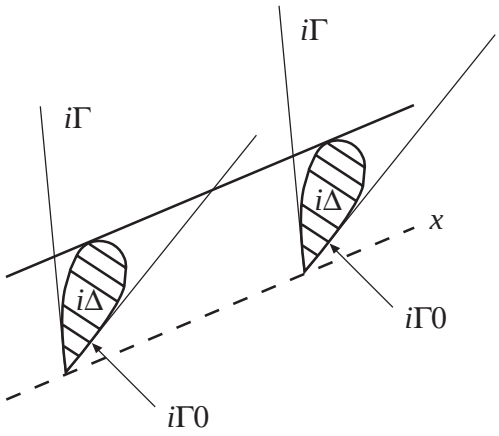


Figure 2. An infinitesimal wedge $\mathbb{R}^n + i\Gamma 0$.

- $|\alpha| = \alpha_1 + \alpha_2 + \cdots + \alpha_n$ is the length of the multi-index α.
- $\partial_x^\alpha = \partial_{x_1}^{\alpha_1} \partial_{x_2}^{\alpha_2} \ldots \partial_{x_n}^{\alpha_n}$ and $D^\alpha = (-i)^{|\alpha|}\, \partial_x^\alpha$.

We now define some classes of functions and distributions.

Definition 1. *The above notation is used.*

(a) *A continuous function $g \in C^0(\mathbb{R}^n)$ is said to be slowly increasing or exponentially decreasing, respectively, on a closed cone Γ' if there exist positive numbers C and M, or C and δ, respectively, such that*

$$|g(x)| \le C\,(1+|x|)^M \qquad or \qquad |g(x)| \le C\,e^{-\delta|x|}, \qquad on\ \Gamma'. \tag{1}$$

(b) *A distribution $f \in \mathcal{D}'(\mathbb{R}^n)$ is said to be slowly increasing or exponentially decreasing, respectively, on an open cone Γ if it can be represented as a finite sum*

$$f(x) = \sum_{|\alpha| \le m} D^\alpha g_\alpha(x), \tag{2}$$

where $g_\alpha \in C^0(\mathbb{R}^n)$ are slowly increasing, or exponentially decreasing functions, respectively, on every closed cone $\Gamma' \subset\subset \Gamma$.

(c) *A holomorphic function $f \in \mathcal{O}(U)$ is said to be slowly increasing in U if it can be represented as a finite sum*

$$f(z) = \sum_{|\alpha| \le m} D^\alpha g_\alpha(z), \tag{3}$$

where each function $g_\alpha \in \mathcal{O}(U)$ is continuous on the closure $\overline{U}$ of U and satisfies the estimate

$$|g_\alpha(z)| \le C_\alpha\,(1+|z|)^M, \qquad on\ \overline{U}. \tag{4}$$

If $U = \mathbb{R}^n + i\Gamma 0$ is an infinitesimal wedge, $f(z)$ is said to be a slowly increasing holomorphic function in the infinitesimal wedge U.

Let $f(z)$ be a slowly increasing holomorphic function in an infinitesimal wedge $\mathbb{R}^n + i\Gamma 0$. Then, by definition, there exists a tubular domain $\mathbb{R}^n + i\Delta$ such that $f(z)$ has representation (3) in $\mathbb{R}^n + i\Delta$, where the slowly increasing holomorphic functions $g_\alpha(z)$ are continuous on $\overline{\mathbb{R}^n + i\Delta}$. Hence, the restriction of $f(z)$ to $\mathbb{R}^n$ yields a slowly increasing distribution $f(x)$ on $\mathbb{R}^n$ that is defined by the right-hand side of (3). Further, for every sequence of points $\{y^{(k)}\}$ lying inside Δ and tending to 0 as $k \to \infty$, we have

$$f(x) = \lim_{\substack{y^{(k)} \to 0 \\ y^{(k)} \in \Delta}} f\left(x + iy^{(k)}\right). \tag{5}$$

For simplicity, we shall write (5) as

$$f(x) = f(x + i\Gamma 0), \tag{6}$$

indicating that $f(x)$ is a "generalized boundary value" of a holomorphic function in an infinitesimal wedge $\mathbb{R}^n + i\Gamma 0$. Such a distribution can be thought of as being analytic with respect to the direction of Γ.

We now turn to the study of the directional analyticity of a distribution.

Definition 2. *A distribution $f(x)$ is said to be analytic with respect to a direction ξ_0 if it can be represented as a finite sum of limits $f_j(x + i\Gamma_j 0)$ of slowly increasing holomorphic functions $f_j(z)$ in $\mathbb{R}^n + i\Gamma_j 0$ such that for every j, we have*

$$\Gamma_j \cap \{y \in \mathbb{R}^n; y \cdot \xi_0 < 0\} \neq \emptyset.$$

Remark 1. The above representation of $f(x)$ by a finite sum of limits $f_j(x + i\Gamma_j 0)$ of slowly increasing holomorphic functions corresponds to an intuitive definition of hyperfunctions. Let Ω be an open set in $\mathbb{R}^n$. If, for $j = 1, \ldots, N$, $F_j(z)$ is a holomorphic function defined on an infinitesimal wedge $\Omega + i\Gamma_j 0$, then intuitively a hyperfunction is a commutative formal sum

$$f(x) = \sum_{j=1}^{N} F_j\,(x + i\Gamma_j 0)$$

of boundary values of holomorphic functions $F_j(z)$ defined in $\Omega + i\Gamma_j 0$ (see [8]). Figure 3 illustrates $\Omega + i\Gamma_j 0$. Note that an infinitesimal wedge of the form $\mathbb{R}^n + i\Gamma_j 0$ is a tubular domain, but, in general, an infinitesimal wedge of the form $\Omega + i\Gamma_j 0$ need not be a tubular domain if $\Omega \neq \mathbb{R}^n$.

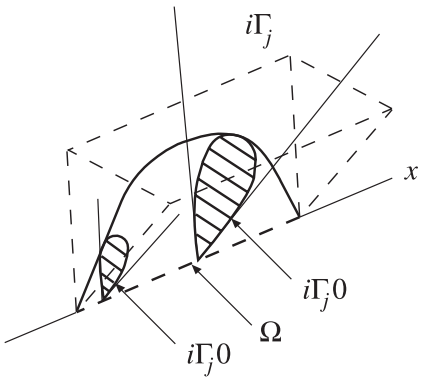


Figure 3. An infinitesimal wedge $\Omega + i\Gamma_j 0$.

In this paper, the Fourier transform $\hat{f}(\xi)$ of $f(x)$ is defined by

$$\hat{f}(\xi) := \int_{\mathbb{R}^n} f(x)\, e^{-ix\xi}\, dx.$$

To characterize the microanalyticity of a slowly increasing distribution $f \in \mathcal{S}'(\mathbb{R}^n)$ by its Fourier transform $\hat{f}$, we introduce the *dual cone*, Γ°, of Γ defined by

$$\Gamma^\circ := \{\xi \in \mathbb{R}^n; y \cdot \xi \geq 0, \text{ for every } y \in \Gamma\}$$

(see Figure 4). If Γ is a cone in $\mathbb{R}^n$ then the dual cone Γ° is a closed convex cone in $\mathbb{R}^n$. Moreover, Γ° is a proper cone. The complement of Γ° is denoted by $(\Gamma^\circ)^c$.

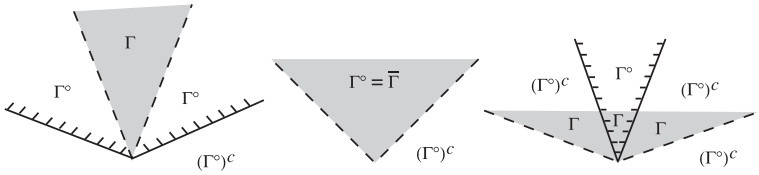


Figure 4. Open cone Γ, dual cone Γ°, and complement $(\Gamma^\circ)^c$ of dual cone.

The following two lemmas are standard (see [7]).

LEMMA 1. *Let Γ be an open convex cone. A slowly increasing distribution $f(x) \in \mathcal{S}'(\mathbb{R}^n)$ can be represented as the limit $f(x + i\Gamma 0)$ of a slowly increasing holomorphic function $f(z)$ in the infinitesimal wedge $\mathbb{R}^n + i\Gamma 0$ if and only if the Fourier transform $\hat{f}$ of f is exponentially decreasing in the open cone $(\Gamma^\circ)^c$, the complement of Γ°, that is, $\hat{f}$ is exponentially decreasing on every closed proper subcone $\Gamma' \subset\subset (\Gamma^\circ)^c$.*

The convex hull of an open cone Γ is denoted by $\mathrm{co}\Gamma$. It can be shown that $\Gamma^\circ = (\mathrm{co}\Gamma)^\circ$.

LEMMA 2. BOCHNER. *Let Γ be an open connected cone. Every slowly increasing holomorphic function in the infinitesimal wedge $\mathbb{R}^n + i\Gamma 0$ can be extended to a slowly increasing holomorphic function in the infinitesimal wedge $\mathbb{R}^n + i(\mathrm{co}\Gamma)0$.*

Hereafter, we shall always assume that the opening Γ of an infinitesimal wedge is convex. The larger the opening Γ, the more regular a slowly increasing distribution $f(x + i\Gamma 0)$ will be. The largest opening Γ is the whole space, in which case $f(x + i\Gamma 0)$ is analytic. The next largest possible openings Γ are half-spaces.

Let a slowly increasing distribution $f(x)$ be analytic with respect to a direction ξ_0. Then, by Definition 2, $f(x)$ can be represented as a finite sum of limits $f_j(x + i\Gamma_j 0)$ of slowly increasing holomorphic functions in $\mathbb{R}^n + i\Gamma_j 0$ such that $\Gamma_j \cap \{y \in \mathbb{R}^n; y \cdot \xi_0 < 0\} \neq \emptyset$ for each j. By Lemma 1, each Fourier transform $\hat{f}_j(\xi)$ is exponentially decreasing in the open cone $(\Gamma_j^\circ)^c$. Since $\xi_0 \notin \Gamma_j^\circ$, there exists an open cone Γ containing ξ_0 such that the Fourier transform $\hat{f}(\xi) = \sum \hat{f}_j(\xi)$ is exponentially decreasing in Γ.

It is desirable to localize the directional decay of a function in ξ-space (Fourier space), because local nonsmoothness of a function f in x-space corresponds to slow decay of the Fourier transform $\hat{f}$ along some directions at infinity. Each such direction corresponds to a point on the unit sphere $\mathbb{S}^{n-1}$ in ξ-space. Therefore, we shall use the coordinates $(x, \xi) \in \mathbb{R}^n \times \mathbb{S}^{n-1}$ to represent a point $x \in \mathbb{R}^n$ together with a direction $\xi \in \mathbb{S}^{n-1}$.

DEFINITION 3. *A distribution $f(x) \in \mathcal{D}'(\mathbb{R}^n)$ is said to be analytic at $x_0 \in \mathbb{R}^n$ if there exists an open neighborhood $V \subset \mathbb{R}^n$ of x_0 such that the restriction $f|_V$ of f on V is analytic in V. The set of all points $x \in \mathbb{R}^n$ where f is not analytic is called the singular support of f and is denoted by* $\mathrm{sing\,supp} f$.

DEFINITION 4. *A distribution $f(x)$ is said to be microanalytic or microlocal analytic at $(x_0, \xi_0) \in \mathbb{R}^n \times \mathbb{S}^{n-1}$ if there exists a distribution $g(x)$ which is analytic with respect to the direction ξ_0 such that $f(x) - g(x)$ is analytic in a neighborhood of x_0. The set of all points $(x, \xi) \in \mathbb{R}^n \times \mathbb{S}^{n-1}$ where f is not microanalytic is called the singular spectrum of f and is denoted by* $\mathrm{S.S.}f$.

REMARK 2. Let π be the natural projection from $\mathbb{R}^n \times \mathbb{S}^{n-1}$ to $\mathbb{R}^n$. Then

$$\pi(\mathrm{S.S.}f) = \mathrm{sing\,supp} f.$$

Hence, an analytic function is a function which is microanalytic at every point (x, ξ).

The following lemma shows that the singular spectrum is an invariant concept under analytic local coordinate transformations (see [7]).

LEMMA 3. *Let $y = F(x)$ be an analytic local coordinate transformation with Jacobian matrix $dF(x) = (\frac{\partial F_j}{\partial x_k})$. Then*

$$\left(x, \frac{dF(x)^\top \eta}{\|dF(x)^\top \eta\|}\right) \in \mathrm{S.S.}f(x) \implies (y, \eta) \in \mathrm{S.S.}f\left(F^{-1}(y)\right)$$

and

$$\left(x, \frac{dF(x)^\top \eta}{\|dF(x)^\top \eta\|}\right) \notin \mathrm{S.S.}f(x) \implies (y, \eta) \notin \mathrm{S.S.}f\left(F^{-1}(y)\right),$$

where $dF(x)^\top$ denotes the transpose matrix of $dF(x)$.

3. MULTIWAVELETS

We define multiwavelets in this section, referring to [9] for detail. The following notation will be used.

NOTATION 2.

- Given $f \in L^2(\mathbb{R}^n)$, $f_{jk}(x)$ denotes the scaled and shifted functions

$$f_{jk}(x) = 2^{nj/2} f\left(2^j x - k\right), \qquad j \in \mathbb{Z}, \quad k \in \mathbb{Z}^n.$$

- Given $f_1, \ldots, f_d \in L^2(\mathbb{R}^n)$, F denotes the vector-valued function $F = (f_1, \ldots, f_d) \in L^2(\mathbb{R}^n)^d$ and F_{jk} denotes the scaled and shifted vector-valued functions

$$F_{jk} = ((f_1)_{jk}, \ldots, (f_d)_{jk}), \qquad j \in \mathbb{Z}, \quad k \in \mathbb{Z}^n.$$

- $E = \{0,1\}^n \setminus \{(0, \ldots, 0)\}$ is the set of $2^n - 1$ vertices of the n-dimensional unit cube less the origin.
- $D = \{1, \ldots, d\}$ for a positive integer d.
- $\langle u, v \rangle = \int u(x)\bar{v}(x)\, dx$ denotes the inner product on $L^2(\mathbb{R}^n)$.

DEFINITION 5. *Let the $2^n - 1$ functions $\Psi_\varepsilon := (\psi_{\varepsilon 1}, \ldots, \psi_{\varepsilon d}) \in L^2(\mathbb{R}^n)^d$, $\varepsilon \in E$, be such that the system*

$$\left\{(\psi_{\varepsilon\delta})_{jk}(x) := 2^{nj/2}\, \psi_{\varepsilon\delta}\left(2^j x - k\right)\right\}_{\varepsilon \in E,\, \delta \in D,\, j \in \mathbb{Z},\, k \in \mathbb{Z}^n} \tag{7}$$

forms an orthonormal basis for $L^2(\mathbb{R}^n)$. Then $\{\Psi_\varepsilon\}_{\varepsilon \in E}$ is called a family of $2^n - 1$ multiwavelets, or multiwavelet functions, the functions $(\psi_{\varepsilon\delta})_{jk}$ are called multiwavelets, (7) is called an orthonormal multiwavelet basis, and any $f \in L^2(\mathbb{R}^n)$ admits the multiwavelet expansion

$$f = \sum_{\substack{\varepsilon \in E,\, \delta \in D \\ j \in \mathbb{Z},\, k \in \mathbb{Z}^n}} \langle f, (\psi_{\varepsilon\delta})_{jk} \rangle\, (\psi_{\varepsilon\delta})_{jk}. \tag{8}$$

4. ONE-DIMENSIONAL MICROLOCAL FILTERING

In the one-dimensional case, i.e., $n = 1$, the definition of hyperfunctions is simple. Let Ω be an open subset of $\mathbb{R}$. An open set $U \subset \mathbb{C}$ is called a *complex neighborhood* of Ω if Ω is a relatively closed subset of U. The directions of analyticity are $\mathbb{S}^0 = \{\pm 1\}$. For a complex neighborhood U of Ω, corresponding to the directions ± 1, define

$$U_+ := U \cap \{\mathrm{Im} z > 0\}, \qquad U_- := U \cap \{\mathrm{Im} z < 0\}.$$

We can take U sufficiently small so that $U \setminus \Omega = U_+ \cup U_-$ (see Figure 5).

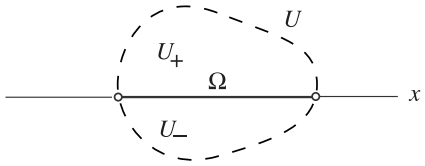


Figure 5. A complex neighborhood U of Ω.

A hyperfunction $f(x)$ on Ω is defined by the difference of two holomorphic functions

$$f(x) = F_+(x + i0) - F_-(x - i0), \qquad F_\pm \in \mathcal{O}(U_\pm).$$

For simplicity, we write $F_\pm \in \mathcal{O}(U_\pm)$ to mean $F_+ \in \mathcal{O}(U_+)$ and $F_- \in \mathcal{O}(U_-)$; the same convection applies throughout the rest of this paper. The pair (F_+, F_-) is called a *defining function*. Such a pair is not unique, since, for each holomorphic function $G \in \mathcal{O}(U)$, the defining function $(F_+ + G, F_- + G)$ determines the same hyperfunction f.

Our goal in this one-dimensional setting is to construct a multiwavelet function $\Psi := (\psi_+, \psi_-)$, corresponding to the directions ± 1 of analyticity, such that the functions $(\psi_\pm)_{jk}$ are localized in both x and ξ and the pair of projections

$$\mathcal{P}_\pm f := \sum_{j\in\mathbb{Z},\, k\in\mathbb{Z}} \langle f, (\psi_\pm)_{jk}\rangle\, (\psi_\pm)_{jk}, \qquad f \in L^2(\mathbb{R}), \tag{9}$$

yield a defining function $(\mathcal{P}_+ f, \mathcal{P}_- f)$ of f. In this case, each term $\langle f, (\psi_\pm)_{jk}\rangle\, (\psi_\pm)_{jk}$ may serve as a comparison function g in the sense of Definition 4 to measure the microanalyticity of f. Moreover, even when dealing with a function $\tilde{f} \notin L^2(\mathbb{R})$, if there exists a real analytic function h, say, a polynomial, such that $\tilde{f} - h \in L^2(\mathbb{R})$, then we can measure the microanalyticity of $\tilde{f}$ by measuring the microanalyticity of $\tilde{f} - h$.

Our construction of multiwavelet functions is closely related to the classical Hardy spaces $H^2(\mathbb{R}_\pm)$ defined by

$$H^2(\mathbb{R}_+) := \left\{ f \in L^2(\mathbb{R});\ \hat{f}(\xi) = 0,\ \text{for a.a. } \xi \le 0 \right\},$$
$$H^2(\mathbb{R}_-) := \left\{ f \in L^2(\mathbb{R});\ \hat{f}(\xi) = 0,\ \text{for a.a. } \xi \ge 0 \right\}.$$

Hereafter, we shall use the convention that $[a, b]$ denotes the interval $[b, a]$ when $b < a$.

THEOREM 1. *Consider the multiwavelet function $\Psi := (\psi_+, \psi_-)$ where $\psi_\pm$ are defined by $\hat{\psi}_\pm = \chi_{[\pm 2\pi, \pm 4\pi]}$ (see Figure 6) and let $\mathcal{P}_\pm$ be the projections defined by (9). Then for each $f \in L^2(\mathbb{R})$, $\mathcal{P}_\pm f$ can be continued analytically to the upper and lower half-planes $H_+ := \{z \in \mathbb{C}; \mathrm{Im} z > 0\}$ and $H_- := \{z \in \mathbb{C}; \mathrm{Im} z < 0\}$, respectively.*

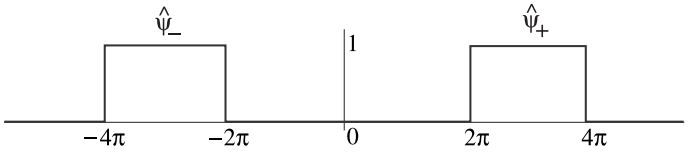


Figure 6. Functions $\hat{\psi}_\pm = \chi_{[\pm 2\pi, \pm 4\pi]}$.

PROOF. It is well known that every element of $H^2(\mathbb{R}_\pm)$ can be continued analytically to $H_\pm$. Since $L^2(\mathbb{R}) = H^2(\mathbb{R}_+) \oplus H^2(\mathbb{R}_-)$, where $\oplus$ denotes the orthogonal sum, it is enough to show that $\psi_\pm$ is a uniwavelet function for $H^2(\mathbb{R}_\pm)$. But this is immediate since $\{(\psi_\pm)_{jk}\}_{j\in\mathbb{Z},\, k\in\mathbb{Z}}$ forms an orthonormal basis for $H^2(\mathbb{R}_\pm)$ (see, for example, [10; 11, Section 5; 12, Section 7.6]). Hence, $\{(\psi_+)_{jk}, (\psi_-)_{jk}\}_{j\in\mathbb{Z},\, k\in\mathbb{Z}}$ forms an orthonormal basis for $L^2(\mathbb{R})$, which means that Ψ is a multiwavelet function for $L^2(\mathbb{R})$. ∎

REMARK 3. The multiwavelet system defined in Theorem 1 is associated with a multiresolution analysis. In particular, if $\varphi_\pm$ is defined by $\hat{\varphi}_\pm = \chi_{[0, \pm 2\pi]}$, then $\Phi := (\varphi_+, \varphi_-)$ is a multiscaling function for the above multiwavelets. We refer to [9] for the connection between multiscaling functions, multiresolution analysis, and multiwavelets.

REMARK 4. The multiwavelets $\{(\psi_+)_{jk}, (\psi_-)_{jk}\}_{j\in\mathbb{Z},\, k\in\mathbb{Z}}$ of Theorem 1 provide us with the best possible resolution in ξ-space, because the directions of analyticity are $\mathbb{S}^0 = \{\pm 1\}$. In x-space,

however, $(\psi_\pm)_{jk}$ are multiples of cardinal sine functions; thus, the resolution is rough, but the coefficients of the multiwavelet expansion can be calculated numerically.

Next, we consider the possibility of obtaining a finer resolution in x-space while keeping the best resolution in ξ-space. In one sense, this is impossible. In fact, if we keep the best resolution in ξ-space, $\psi_\pm$ must be uniwavelet functions of $H^2(\mathbb{R}_\pm)$. However, it is known that there exist no smooth (in ξ-space) wavelets in $H^2(\mathbb{R}_\pm)$. More precisely, there are no wavelet functions $\psi_+ \in H^2(\mathbb{R}_+)$ satisfying the following two conditions:

$$\left|\hat{\psi}_+\right| \text{ is continuous on } \mathbb{R}$$

and

$$\left|\hat{\psi}_+(\xi)\right| = O\left((1+|\xi|)^{-\alpha-1/2}\right), \qquad \text{as } |\xi| \to \infty \text{ for some } \alpha > 0.$$

See, for example, [12, Section 7.6; 13]

One solution to this problem is to use a multiwavelet *frame*. A collection of functions $\{(\psi_+)_{jk}, (\psi_-)_{jk}\}_{j\in\mathbb{Z},\, k\in\mathbb{Z}}$ is called a *tight frame* for $L^2(\mathbb{R})$ with frame bound A if

$$A\,\|f\|^2 = \sum_{\substack{\delta\in\{\pm\}\\ j\in\mathbb{Z},\, k\in\mathbb{Z}}} |\langle f, (\psi_\delta)_{jk}\rangle|^2, \qquad \text{for all } f \in L^2(\mathbb{R}).$$

In this case, each $f \in L^2(\mathbb{R})$ can be expanded in a series of the form

$$f = \frac{1}{A} \sum_{\substack{\delta\in\{\pm\}\\ j\in\mathbb{Z},\, k\in\mathbb{Z}}} \langle f, (\psi_\delta)_{jk}\rangle\, (\psi_\delta)_{jk}. \tag{10}$$

We refer to [4] or [11] for the derivation of basic properties of frames. One disadvantage of frames is that the representation in (10) need not be unique in general. However, for many applications, including microlocal filtering, uniqueness is not essential.

Smooth frames for $H^2(\mathbb{R}_\pm)$ were first constructed in [10]. The following construction is from [12, Section 8.4]. Given $\varepsilon > 0$, let s_ε be a C^∞ function such that $s_{2\varepsilon}(x) = s_\varepsilon(x/2)$, $s_\varepsilon(x) = 0$ for $x < \varepsilon$, and $s_\varepsilon^2(x) + c_\varepsilon^2(x) = 1$, where $c_\varepsilon(x) = s_\varepsilon(-x)$. For $0 < \varepsilon \le (1/3)\pi$, define a *bell function* associated with the interval $[\pi, 2\pi]$ by

$$b_\varepsilon(x) := s_\varepsilon(x-\pi)\, c_{2\varepsilon}(x-2\pi),$$

and define ψ^ε by

$$\hat{\psi}^\varepsilon(\xi) := b_\varepsilon\left(\frac{\xi}{2}\right), \qquad \xi \in \mathbb{R}. \tag{11}$$

The following lemma is stated without proof in [12].

LEMMA 4. *Given $0 < \varepsilon \le (1/3)\pi$, let ψ^ε be defined by (11). Then $\{(\psi^\varepsilon)_{j,k}\}_{j,k\in\mathbb{Z}}$ forms a tight frame for $H^2(\mathbb{R}_+)$ with frame bound 1. Moreover, ψ^ε is in the Schwartz class $\mathcal{S}(\mathbb{R})$ of C^∞ functions of fast descent.*

Using these smooth frame uniwavelets for $H^2(\mathbb{R}_+)$, we can construct frame multiwavelets for $L^2(\mathbb{R})$ having finer resolution in x-space.

THEOREM 2. *Let $b_\varepsilon^\pm$ be bell functions associated with the interval $[\pi, 2\pi]$. Then, $b_\varepsilon^+(\xi/2)$ is a bell function associated with the interval $[2\pi, 4\pi]$ and $b_\varepsilon^-(-\xi/2)$ is a bell function associated with the interval $[-4\pi, -2\pi]$. Define $\psi_\pm(x)$ by $\hat{\psi}_\pm(\xi) = b_\varepsilon^\pm(\pm\xi/2)$, respectively (see Figure 7). Then $\{(\psi_+)_{jk}, (\psi_-)_{jk}\}_{j,k\in\mathbb{Z}}$ forms a tight frame for $L^2(\mathbb{R})$ with frame bound 1, and furthermore,*

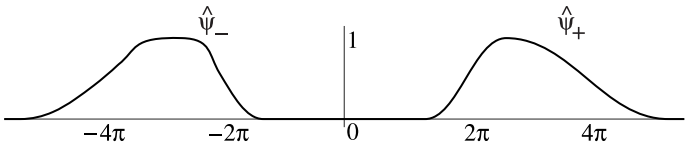


Figure 7. Functions $\hat{\psi}_\pm(\xi) = b_\varepsilon^\pm(\pm\xi/2)$.

$(\psi_\pm)_{jk} \in \mathcal{S}(\mathbb{R})$. Let $\mathcal{P}_\pm$ be projections to $H^2(\mathbb{R}_\pm)$ defined by (9). Then, for each $f \in L^2(\mathbb{R})$, $\mathcal{P}_\pm f$ can be continued analytically to $H_\pm$.

5. MULTIDIMENSIONAL MICROLOCAL FILTERING

We now generalize Theorem 1 to the multidimensional case and provide a library of orthonormal multiwavelet bases from which the best microlocal filtering might be selected by means of a best basis algorithm.

NOTATION 3. The following notation will be used.

- $H = \{\pm 1\}^n$ is a parametrization of the 2^n n-dimensional orthants in $\mathbb{R}^n$. For example, in $\mathbb{R}^2$, $(+1,+1)$, $(-1,+1)$, $(-1,-1)$, and $(+1,-1)$ correspond to the first, second, third, and fourth quadrants, respectively.
- For $\eta = (\eta_1, \ldots, \eta_n) \in H$, denote by Q_η the unit cube $\prod_{k=1}^n [0, \eta_k]$, where $[0,-1]$ stands for the interval $[-1, 0]$.
- For $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n) \in E$ as defined in Notation 2 and $\eta = (\eta_1, \ldots, \eta_n) \in H$, denote the element-wise product by

$$\varepsilon . * \eta := (\varepsilon_1 \eta_1, \ldots, \varepsilon_n \eta_n).$$

- For $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n) \in E$, $\eta = (\eta_1 \ldots, \eta_n) \in H$, and $j \in \mathbb{Z}_+$, define the cube

$$2^j \left(Q_\eta + \varepsilon . * \eta\right) := \left\{ \left(2^j(x_1 + \varepsilon_1\eta_1), \ldots, 2^j(x_n + \varepsilon_n\eta_n)\right) ; (x_1, \ldots, x_n) \in Q_\eta \right\}.$$

  Then let $\mathcal{Q}_{j,\varepsilon,\eta}$ be the collection of unit cubes that cover $2^j\left(Q_\eta + \varepsilon . * \eta\right)$ with overlaps of measure zero, i.e.,

$$\mathcal{Q}_{j,\varepsilon,\eta} := \left\{ \prod_{k=1}^n [\eta_k(\ell_k - 1), \eta_k \ell_k] + 2^j(\varepsilon . * \eta) \; ; \; 1 \le \ell_1, \ldots, \ell_n \le 2^j, \;\; \ell_1, \ldots, \ell_n \in \mathbb{N} \right\},$$

  where $[-(\ell_k - 1), -\ell_k]$ stands for the interval $[-\ell_k, -(\ell_k - 1)]$.
- Given an indexing set K and a collection $\{Q_k\}_{k\in K}$ of subsets of $\mathbb{R}^n$, define

$$\mathcal{Q} := \{Q_k\}_{k\in K} \qquad \text{and} \qquad \iota(\mathcal{Q}) := \bigcup_{k\in K} Q_k.$$

- Define

$$2\pi\mathcal{Q}_{j,\varepsilon,\eta} := \{2\pi Q \; ; \; Q \in \mathcal{Q}_{j,\varepsilon,\eta}\}.$$

- Let $\mathbb{Z}_+^{E\times H}$ denote the set of all functions from $E \times H$ to $\mathbb{Z}_+$.
- For a nonnegative integer $N \in \mathbb{Z}_+$, let $\mathbb{Z}_N := \{0, 1, \ldots, N\}$ and denote the set of all functions from $E \times H$ to $\mathbb{Z}_N$ by $\mathbb{Z}_N^{E\times H}$.

THEOREM 3. *Let $j \in \mathbb{Z}_+$, $\varepsilon \in E$, and $\eta \in H$. For $Q \in \mathcal{Q}_{j,\varepsilon,\eta}$, define ψ_Q by*

$$\hat{\psi}_Q = \chi_{2\pi Q},$$

where $\chi_{2\pi Q}$ is the characteristic function of the cube $2\pi Q$. For $\rho \in \mathbb{Z}_+^{E\times H}$, let

$$\mathcal{Q}_\rho := \bigcup_{(\varepsilon,\eta)\in E\times H} 2\pi \mathcal{Q}_{\rho(\varepsilon,\eta),\varepsilon,\eta}.$$

Then $\Psi := (\psi_Q)_{Q\in\mathcal{Q}_\rho}$ is a multiwavelet function.

Proof. First, we claim that $\bigcup_{j\in\mathbb{Z}} 2^j\mathcal{Q}_\rho$ is a nonoverlapping covering of $\mathbb{R}^n$. Consider the case $\rho \equiv 0$. We have

$$\mathcal{Q}_0 = \bigcup_{(\varepsilon,\eta)\in E\times H} 2\pi \mathcal{Q}_{0,\varepsilon,\eta} = \bigcup_{(\varepsilon,\eta)\in E\times H} 2\pi \left(Q_\eta + \varepsilon .* \eta\right).$$

Since $\bigcup_{(\varepsilon,\eta)\in E\times H} \left(Q_\eta + \varepsilon .* \eta\right)$ is a nonoverlapping covering of $\prod_{k=1}^n [-2,2]\backslash\prod_{k=1}^n (-1,1)$ and

$$\bigcup_{j\in\mathbb{Z}} 2^j \left(\prod_{k=1}^n [-2,2]\Big\backslash \prod_{k=1}^n (-1,1)\right) = \mathbb{R}^n,$$

then $\bigcup_{j\in\mathbb{Z}} 2^j\mathcal{Q}_0$ is a nonoverlapping covering of $\mathbb{R}^n$. For an arbitrary fixed element $(\varepsilon,\eta)\in E\times H$, since the set $\bigcup_{j\in\mathbb{Z}} 2^j 2\pi \left(Q_\eta + \varepsilon .* \eta\right)$ is invariant under multiplication by 2^{j_0} for any $j_0\in\mathbb{Z}$, we have

$$\begin{aligned}\bigcup_{j\in\mathbb{Z}} 2^j\mathcal{Q}_0 &= \bigcup_{(\varepsilon,\eta)\in E\times H} \left[\bigcup_{j\in\mathbb{Z}} 2\pi \left(Q_\eta + \varepsilon .* \eta\right)\right] \\ &= \bigcup_{(\varepsilon,\eta)\in E\times H} \left[\bigcup_{j\in\mathbb{Z}} 2\pi 2^{\rho(\varepsilon,\eta)} \left(Q_\eta + \varepsilon .* \eta\right)\right] \\ &= \bigcup_{j\in\mathbb{Z}} 2^j\mathcal{Q}_\rho.\end{aligned}$$

Next, consider a cube $2\pi Q\in\mathcal{Q}_\rho$. Because $\hat{\psi}_Q = \chi_{2\pi Q}$, the collection

$$\left\{e^{ik\xi}\hat{\psi}_Q(\xi)\right\}_{k\in\mathbb{Z}^n}$$

forms an orthonormal basis for the space of $L^2(\mathbb{R}^n_\xi)$ functions supported on $2\pi Q$. By Plancherel's formula,

$$\{\psi_Q(x-k)\}_{k\in\mathbb{Z}^n}$$

is an orthonormal basis for the space of $L^2(\mathbb{R}^n_x)$ functions whose Fourier transforms are supported in $2\pi Q$. Taking the union of these orthonormal bases for $2\pi Q\in\mathcal{Q}_\rho$, we obtain an orthonormal basis for the space of functions whose Fourier transforms are supported in $\iota(\mathcal{Q}_\rho)$ defined in Notation 3. The dilates of those functions form an orthonormal basis for functions whose Fourier transforms are supported in dilates of $\iota(\mathcal{Q}_\rho)$, and combining these in all possible ways, we obtain the following multiwavelet orthonormal basis for $L^2(\mathbb{R}^n)$:

$$\left\{2^{nj/2}\psi_Q\left(2^j x - k\right)\right\}_{Q\in\mathcal{Q}_\rho, j\in\mathbb{Z}, k\in\mathbb{Z}^n}. \qquad \blacksquare$$

Corollary 1. *For $\eta\in H$, define φ_η by $\hat{\varphi}_\eta = \chi_{2\pi Q_\eta}$. Then $\Phi := (\varphi_\eta)_{\eta\in H}$ is a multiscaling function for the multiwavelet system of Theorem 3.*

Remark 5. Since each ψ_Q is the product of a complex exponential and cardinal sine functions in the variables x_k, $k=1,\ldots,n$, multiwavelet coefficients can be calculated numerically.

The points $(\varepsilon,\eta)\in E\times H$ can be thought of as rough directions of analyticity. By choosing ρ so that $\rho(\varepsilon,\eta)$ is large, that is, by taking the set $2^{\rho(\varepsilon,\eta)}\left(Q_\eta + \varepsilon .* \eta\right)$ to be a large cube, the

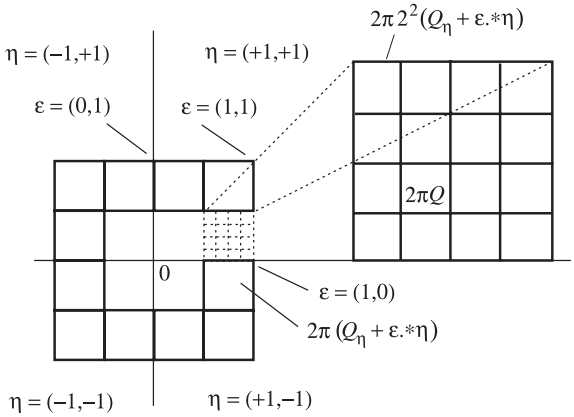


Figure 8. Two-dimensional example of $\mathcal{Q}_\rho$.

Fourier transform of each function ψ_Q for $Q \in \mathcal{Q}_{\rho(\varepsilon,\eta),\varepsilon,\eta}$ has support contained in a cube $2\pi Q \in 2\pi\mathcal{Q}_{\rho(\varepsilon,\eta),\varepsilon,\eta}$ which subtends a very small angle as viewed from the origin (see Figure 8).

By Lemma 1, for a given function f, the cubes Q for which ψ_Q have large wavelet coefficients indicate fairly well the directions along which the Fourier transform of f is concentrated, that is, the directions of analyticity of f. The cost for a good angular resolution in ξ-space is many multiwavelets. Even though these multiwavelets have rough localization in x-space, they still can be used as a tool for microlocal filtering.

DEFINITION 6. *Let $N \in \mathbb{Z}_+$. The family of orthonormal multiwavelet bases*

$$\mathcal{MFB}_N := \left\{ \left\{ 2^{nj/2}\psi_Q\left(2^j x - k\right) \right\}_{Q\in\mathcal{Q}_\rho, j\in\mathbb{Z}, k\in\mathbb{Z}^n} \rho \in \mathbb{Z}_N^{E\times H} \right\}$$

is called a library of microlocal filtering bases of level N.

Once we have a library of microlocal filtering bases $\mathcal{MFB}_N$ for a given function f, we can find the best microlocal filtering basis for f in $\mathcal{MFB}_N$ by an entropy functional criterion as in [14].

6. FAST MICROLOCAL FILTERING ALGORITHM

To implement the multiwavelet transform of f, we need the scaling coefficients at high resolution. Recall that in the uniwavelet case, at very high resolution, the scaling functions are usually close to the delta function; hence, the samples of the function f are often used as scaling coefficients. However, for multiwavelets we need expansion coefficients for d scaling functions. Simply using nearby samples as scaling coefficients may be a bad choice. Data samples need to be preprocessed (*prefiltered*) to produce reasonable values for the expansion coefficients of scaling functions at the highest scale.

Our design of prefilter is the following. Let φ_η, $\eta \in H$, be the scaling functions defined by $\hat{\varphi}_\eta = \chi_{2\pi Q_\eta}$ in Corollary 1. Assume that $f \in \overline{\mathrm{Span}}\{(\varphi_\eta)_{j_0 k}\}_{\eta\in H,\, k\in\mathbb{Z}^n}$ for large j_0, that is, $\mathrm{supp}\hat{f} \subset 2\pi 2^{j_0}[-1,1]^n$. Then,

$$f(x) = \sum_{\eta\in H,\, k\in\mathbb{Z}^n} \langle f, (\varphi_\eta)_{j_0 k}\rangle\, (\varphi_\eta)_{j_0 k}(x).$$

By Plancherel's formula,

$$\begin{aligned}
\langle f, (\varphi_\eta)_{j_0 k}\rangle &= (2\pi)^{-n} \left\langle \hat{f}, \widehat{(\varphi_\eta)_{j_0 k}} \right\rangle \\
&= (2\pi)^{-n} \left\langle \hat{f}, 2^{-nj_0/2} e^{-ik\xi/2^{j_0}}\, \widehat{\varphi_\eta}\left(\frac{\xi}{2^{j_0}}\right) \right\rangle \\
&= 2^{-nj_0/2}\,(2\pi)^{-n} \int_{\mathbb{R}^n} e^{ik\xi/2^{j_0}}\, \hat{f}(\xi)\chi_{2\pi 2^{j_0} Q_\eta}(\xi)\, d\xi \\
&= 2^{-nj_0/2} \mathcal{F}^{-1}\left[\hat{f}(\xi)\chi_{2\pi 2^{j_0} Q_\eta}(\xi)\right]\left(\frac{k}{2^{j_0}}\right).
\end{aligned}$$

Hence,

$$f(x) = \sum_{\eta \in H,\, k \in \mathbb{Z}^n} \mathcal{F}^{-1}\left[2^{-nj_0}\chi_{2\pi 2^{j_0}Q_\eta}(\xi)\hat{f}(\xi)\right]\left(\frac{k}{2^{j_0}}\right) 2^{nj_0/2}(\varphi_\eta)_{j_0 k}(x).$$

Denote

$$f_\eta(x) = \sum_{k \in \mathbb{Z}^n} \mathcal{F}^{-1}\left[2^{-nj_0}\chi_{2\pi 2^{j_0}Q_\eta}\hat{f}\right](x), \qquad \eta \in H.$$

Since $2^{nj_0/2}(\varphi_\eta)_{j_0 k}$ are close to delta functions for sufficiently large j_0, it follows that

$$f(x) \approx \sum_{\eta \in H} f_\eta(x).$$

Therefore, $f_\eta(x)$ can be regarded as prefiltered data for each scaling function φ_η. In this case, the prefilters are

$$P_\eta = \mathcal{F}^{-1} \circ 2^{-nj_0}\chi_{2\pi 2^{j_0}Q_\eta} \circ \mathcal{F}, \qquad \eta \in H,$$

for sufficiently large j_0. Here $2^{-nj_0}\chi_{2\pi 2^{j_0}Q_\eta}$ denotes the multiplication operator by the function $2^{-nj_0}\chi_{2\pi 2^{j_0}Q_\eta}$.

Denote

$$V_j^\eta = \overline{\mathrm{Span}}\left\{(\varphi_\eta)_{jk}\right\}_{k \in \mathbb{Z}^n}, \qquad \eta \in H, \qquad j \in \mathbb{Z}.$$

For each $\eta \in H$, $\{V_j^\eta\}_{j \in \mathbb{Z}}$ is an increasing sequence of subspaces which defines a (uniwavelet) multiresolution analysis of $V^\eta \subset L^2(\mathbb{R}^n)$. We can apply a fast uniwavelet transform and truncate filter coefficients at each resolution level of $\{V_j^\eta\}_{j \in \mathbb{Z}}$. Finally, since $\{V^\eta\}_{\eta \in H}$ is an orthogonal decomposition of $L^2(\mathbb{R}^n)$, we need only sum up all the filter coefficients with respect to $\eta \in H$.

To apply this algorithm to images, we put $n = 2$ and $j_0 = 0$ and use the two-dimensional fast Fourier transform for prefiltering. Figure 9 is an example of microlocal prefiltering of images. In this figure, if brightness is scaled from one to zero, white is one and black is zero.

In this example, the prefiltered image by $P_{(+,-)}$ has maximum energy among the four prefiltered images by $P_{(\pm,\pm)}$, because it is the brightest. The following tableau lists the energy of the four prefiltered images.

Energy $= 1.0\mathrm{e}+06^*$

$$\begin{array}{cc} 2.2059 & 2.0735 \\ 2.5749 & 8.5187. \end{array}$$

Experimentally, we observed that singular parts, or details, of images contain less energy than regular parts, or approximations, of images. Hence, we suspect that the prefiltered image by $P_{(+,-)}$ contains regular parts of the original image and the other three prefiltered images contain singular parts of the original image. Let us look at the three filtered images by $P_{(+,-)} + P_{(-,-)}$, $P_{(+,-)} + P_{(+,+)}$, and $P_{(+,-)} + P_{(-,+)}$.

In the first case, the support of the Fourier transform of the filtered image by $P_{(+,-)} + P_{(-,-)}$ is contained in the half-space $\{(\xi, \eta) \in \mathbb{R}^2 \ ; \ \eta \leq 0\}$. Hence, Lemma 1 implies that there exist an open cone Γ_1 containing $(\xi, \eta) = (0, -1)$ and a holomorphic function $f_1(z)$ in the infinitesimal wedge $\mathbb{R}^2 + i\Gamma_1 0$ such that the filtered image by $P_{(+,-)} + P_{(-,-)}$ is represented as the limit $f_1(x + i\Gamma_1 0)$.

Similarly, in the second case, the support of the Fourier transform of the filtered image by $P_{(+,-)} + P_{(+,+)}$ is contained in the half-space $\{(\xi, \eta) \in \mathbb{R}^2; \xi \geq 0\}$. Hence, by Lemma 1, there exist an open cone Γ_2 containing $(\xi, \eta) = (1, 0)$ and a holomorphic function $f_2(z)$ in the infinitesimal wedge $\mathbb{R}^2 + i\Gamma_2 0$ such that the filtered image by $P_{(+,-)} + P_{(+,+)}$ is represented as the limit $f_2(x + i\Gamma_1 0)$. Hence, these two filtered images by $P_{(+,-)} + P_{(-,-)}$ and $P_{(+,-)} + P_{(+,+)}$ are, in a sense, "approximations" of the original.

In the third case, however, the support of the Fourier transform of the filtered image by $P_{(+,-)} + P_{(-,+)}$ is not restricted to half spaces. Hence, we may assume the Fourier transform

Figure 9. Microlocal prefiltering of an image. White is one and black is zero.

of the filtered image by $P_{(+,-)} + P_{(-,+)}$ cannot have exponential decay on any half-space (this is an assumption in dealing with images). Then the filtered image by $P_{(+,-)} + P_{(-,+)}$ cannot be represented as a boundary value of a single holomorphic function in an infinitesimal wedge. However, it can be represented as a sum of boundary values of several holomorphic functions in infinitesimal wedges. This means that the filtered image by $P_{(+,-)} + P_{(-,+)}$, in a sense, may be a "detail" only. These "regularities" and "singularities" can be seen in Figure 9.

7. NUMERICAL EXAMPLES

We present commonly used examples of hyperfunctions in $\mathbb{R}^2$ and illustrate their microanalyticity numerically.

Each example is formulated as a MATLAB image A with mn pixels in the form of an $m \times n$ matrix with nonnegative real elements. The two-dimensional discrete fast Fourier transform of A produces a matrix B of the same dimensions as A. Then B is prefiltered by splitting it into four parts of dimensions $(m/2) \times (n/2)$,

$$B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix},$$

and each part is filtered by three disjoint masks. Each of these twelve parts is appropriately padded with zeros to the original size $m \times n$ and its discrete fast inverse Fourier transform is analyzed for the directions of microlocal analyticity of the given function. The twelve filter masks are shown in Figure 10, where black is one and white is zero. Each of the center prefilter masks is the sum of the three outside adjacent masks. The masks are arranged in a 4×4 matrix and an image whose Fourier transform has been prefiltered or filtered by the (i, j) mask will be referred to as its $(i, j)^{\text{th}}$ part.

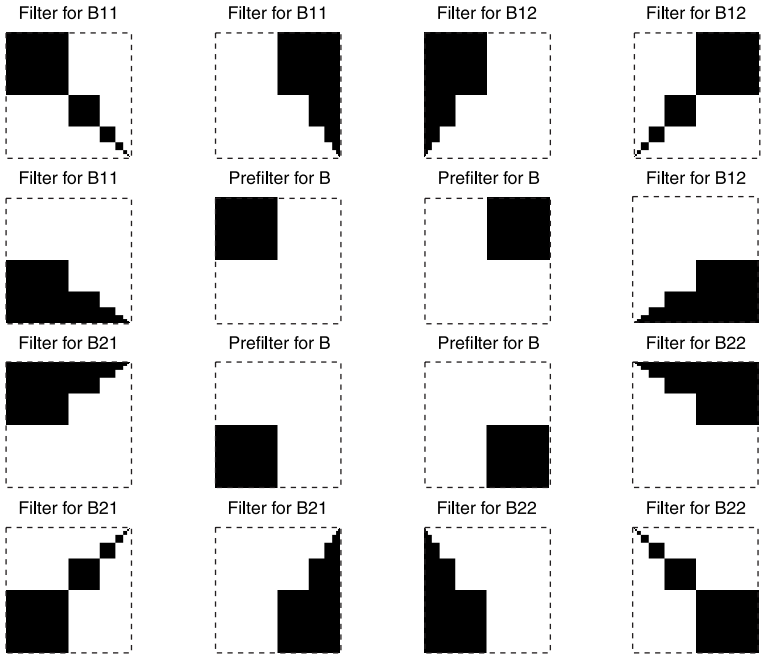


Figure 10. Each of the four center prefilter masks is the sum of the three adjoining filter masks padded with zeros. Black is one and white is zero. The dashed frames are used to delimit the masks.

7.1. A Hyperfunction with Microanalytic Direction in its Singular Support

Consider the one-variable hyperfunction

$$f(x_1) = (x_1 + i0)^{\lambda},$$

with defining functions

$$F_+(z_1) = z_1^{\lambda}, \qquad F_-(z_1) = 0.$$

The function $f(x_1)$ is analytic with respect to the direction $\xi_1 = -1$ and has singular spectrum

$$\text{S.S. } f(x_1) = \{x_1 = 0\} \times \{\xi_1 = +1\}.$$

Let us transform this one-variable hyperfunction $f(x_1)$ to the two-variable hyperfunction $f(y_1 - y_2)$ in $\mathbb{R}^2$. Define an analytic transformation $y = F(x)$ from $\mathbb{R}^2$ to $\mathbb{R}^2$ by

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = A \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Since this transformation is linear, its Jacobian matrix, $dF(x)$, is equal to A,

$$dF(x) = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

Hence, $y = F(x)$ is a bijection on $\mathbb{R}^2$ and its inverse, $x = F^{-1}(y)$, is

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = A^{-1} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}.$$

The singular spectrum of $f(x_1)$ as a two-variable function is

$$\text{S.S.}\, f(x_1) = \left\{ (x_1, x_2, \xi_1, \xi_2) \in \mathbb{R}^2 \times \mathbb{S}^1;\ x_1 = 0,\, \xi_1 = +1 \right\}.$$

Since

$$F(\{x_1 = 0\}) = \{y_1 - y_2 = 0\}$$

and

$$\xi = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{dF(x)^\top \eta}{\|dF(x)^\top \eta\|} = \frac{A^\top \eta}{\|A^\top \eta\|},$$

that is,

$$\eta = \left(A^\top\right)^{-1} \begin{bmatrix} c \\ 0 \end{bmatrix}, \qquad c = \left\|A^\top \eta\right\| \implies \eta = \begin{bmatrix} c \\ -c \end{bmatrix}, \qquad c = \frac{1}{\sqrt{2}},$$

then Lemma 3 implies that the function $f(y_1 - y_2)$ has singular spectrum

$$\text{S.S.}\, f(y_1 - y_2) = \left\{ (y_1, y_2, \eta_1, \eta_2) \in \mathbb{R}^2 \times \mathbb{S}^1;\ y_1 = y_2,\, \eta_1 = \frac{1}{\sqrt{2}},\, \eta_2 = -\frac{1}{\sqrt{2}} \right\}.$$

See also [15, p. 222]. For numerical convenience, we shall consider this function on a 256×256 matrix A with singularity along the secondary diagonal

$$a(m - s + 1, r) = \left[\frac{r - s - i\epsilon}{(r - s)^2 + 2\epsilon^2} \right]^\lambda, \qquad r = 1, 2, \ldots, m, \quad s = 1, 2, \ldots, m,$$

with $\lambda = 1/2$ and $\epsilon = 10^{-3}$. Figure 11 shows the result of the numerical microlocal analysis of f by the twelve masks of Figure 10.

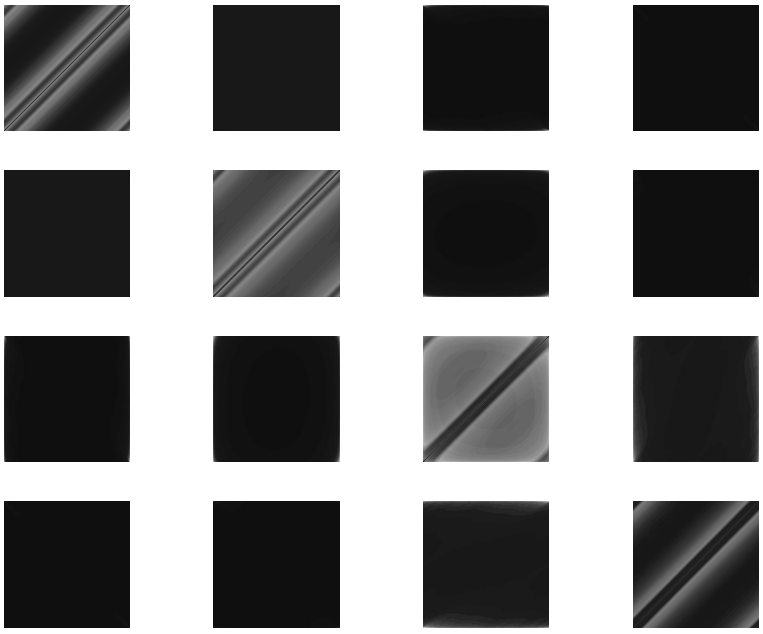


Figure 11. The microlocal content of the twelve partitions of the original figure. White is one and black is zero in a gray scale.

The energy in each of the twelve parts, measured by their Frobenius norm, is shown in the following tableau.

Energy $= 1.0\text{e} + 06^*$

$$\begin{array}{cccc} 9.4796 & 0.4414 & 0.1644 & 0.0661 \\ 0.4414 & & & 0.0727 \\ 0.1644 & & & 0.5355 \\ 0.0661 & 0.0727 & 0.5355 & 8.7079. \end{array}$$

The difference between the sum of the twelve parts and the original figure, in the maximum norm, is

$$2.7102\text{e} - 10.$$

It is seen from Figure 11 that the top left and the bottom right filters pick up almost all the energy in the Fourier transformed image. Hence, the direction along the main diagonal, $r = s$, is in the singular spectrum of f and the direction along the secondary diagonal, $m - r + 1 = s$, is a direction of microanalyticity.

7.2. An Image with Support of Its Fourier Transform in the First Quadrant

In this example, the Fourier transform B of image A is the function

$$\hat{f}(\xi_1, \xi_2) = \begin{cases} \xi_1\xi_2, & \text{for } \xi_1 \geq 0, \quad \xi_2 \geq 0, \\ 0, & \text{otherwise.} \end{cases}$$

The inverse Fourier-Laplace transform of $\hat{f}(\xi_1, \xi_2)$ is

$$f(x_1 + iy_1, x_2 + iy_2) = \frac{1}{4\pi^2} \frac{1}{(x_1 + iy_1)^2} \frac{1}{(x_2 + iy_2)^2}.$$

Prefiltering and filtering of the Fourier transform of A is done by the prefilters and filters of Figure 10. The energy in each of the twelve filtered parts, measured by their Frobenius norm, is shown in the following tableau.

Energy $= 1.0\text{e} + 05^*$

$$\begin{array}{cccc} 0.0000 & 0.0000 & 2.3459 & 6.1648 \\ 0.0000 & & & 2.3159 \\ 0.0000 & & & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000. \end{array}$$

The difference between the sum of the twelve parts and the original figure, in the Frobenius norm, is

$$8.3745\text{e} - 12.$$

In this example, it was convenient to take for image A the inverse Fourier transform of image B. These are shown in Figure 12, where, in a gray scale, black is one and white is zero. The pixels in the 8×8 central square of image A have been set to 164 to enhance the remaining pixels. Due to boundary discontinuities of finite-size images, it would have been difficult to start from A to produce an image B with zero energy outside the first quadrant in Fourier space as shown in the above tableau.

7.3. A Hyperfunction without Microanalytic Direction in the Singular Support

Consider the one-variable distribution

$$f(x_1) = \delta(x_1),$$

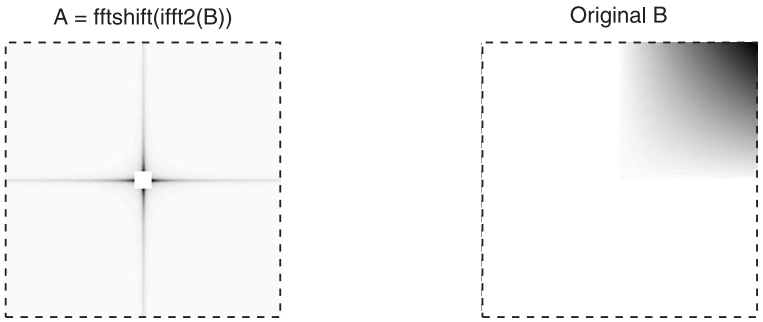


Figure 12. Image A built from image B in the Fourier space. Black is one and white is zero. The dashed frames are used to delimit the figures.

where the Dirac mass $\delta(x)$ has defining functions

$$F_{\pm}(z_1) = -\frac{1}{2\pi i}\,\frac{1}{z_1}.$$

The function $f(x_1)$ has singular spectrum

$$\text{S.S.}\, f(x_1) = \{x_1 = 0\} \times \{\xi_1 = \pm 1\}$$

in opposite directions, and hence, has no direction of microanalyticity. Let us transform this one-variable distribution to a two-variable distribution $f(y_1 - y_2)$ in $\mathbb{R}^2$. Such a distribution is called a line impulse in the image processing literature [16, p. 4]. By Lemma 3, the function $f(y_1 - y_2)$ has singular spectrum

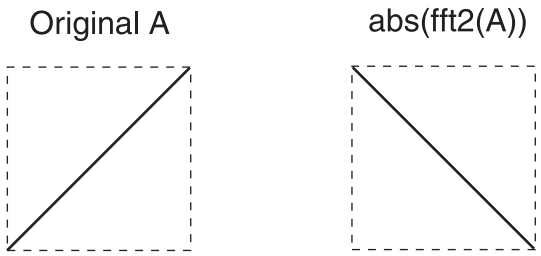
$$\text{S.S.}\, f(y_1 - y_2) = \left\{(y_1, y_2, \eta_1, \eta_2) \in \mathbb{R}^2 \times \mathbb{S}^1;\ y_1 = y_2,\ (\eta_1, \eta_2) = \pm\left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)\right\}.$$

See also [15, p. 222].

In this example a line impulse is simulated numerically by a 256×256 matrix A with elements on the secondary diagonal equal to 100,

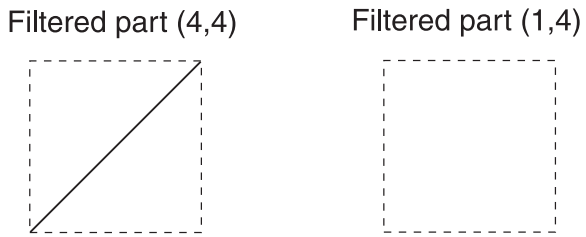
$$a(257 - r, r) = 100, \qquad r = 1, 2, \ldots, 256,$$

and all the other elements are set to zero, as shown in Figure 13a. In this figure, black is one and white is zero in a gray scale. The absolute value of the fast Fourier transform of the original figure is shown in Figure 13b. Prefiltering and filtering by the filters of Figure 10 recover the line structure of the original image. The filtered part $(4, 4)$, shown in Figure 13c, contains the secondary diagonal line at about half the intensity of the original line. The filtered parts $(1, 1)$, $(2, 2)$, and $(3, 3)$ (not shown) are similar to part $(4, 4)$. The remaining 12 filtered parts are white (part $(1, 4)$ is shown in Figure 13d).



(a) (b)

Figure 13. Filtering of an image containing a line impulse along the secondary diagonal. Black is one and white is zero. The dashed frames are used to delimit the figures.



(c) (d)

Figure 13. (cont.)

The energy in each of the twelve parts, measured by their Frobenius norm, is shown in the following tableau.

Energy = 1.0e + 05*

2.8862	0.0853	0.0000	0.0000
0.0853			0.0000
0.0000			0.0853
0.0000	0.0000	0.0853	2.8862.

The difference between the sum of the twelve parts and the original figure, in the Frobenius norm, is

$$8.3745\text{e} - 12.$$

It is seen from the tableau that the prefilters for the matrices B_{11} and B_{22} pick up all the energy in the Fourier transformed image. Since most of this energy is picked up by the top left and bottom right filters, it is seen that the opposite directions along the secondary diagonal are in the singular spectrum of f.

7.4. Denoising

In this example, a vertical line impulse of height 100 in column 60 of a 256×256 matrix with remaining elements equal to zero, is superimposed with a random noise of maximum height 100 over all the pixels of the matrix. One hundred minus the original image is shown in the left part of Figure 14. In this figure, white is one and black is zero in a gray scale.

Prefiltering and filtering by the filters of Figure 10 recover the line structure of the original image as shown in the right part of Figure 14. The energy in each of the twelve filtered parts, measured by their Frobenius norm, is shown in the following tableau.

Energy = 1.0e + 06*

0.5475	0.5459	0.5486	0.5469
0.5425			0.5524
2.9939			3.1255
0.5636	0.5656	1.3740	1.3729.

The filtered images in positions $(3, 1)$ and $(3, 4)$ contain most of the energy.

The difference between the sum of the twelve parts and the original figure, in the Frobenius norm, is

$$6.4367\text{e}{-11}.$$

The difference between the sum of parts $(3, 1)$ and $(3, 4)$ and the original figure, in the Frobenius norm, is

$$1.1990\text{e} + 04.$$

It is seen that the filters in positions $(3, 1)$ and $(3, 4)$ of Figure 13 recover the vertical line by eliminating much of the noise. One hundred minus their sum is shown in the right part of Figure 14.

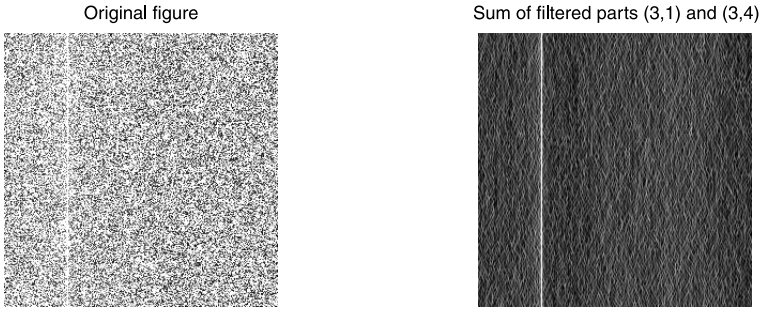


Figure 14. Original figure with noise and image from the sum of filtered parts $(3,1)$ and $(3,4)$. White is one and black is zero in a gray scale.

7.5. Detection of Smooth and Singular Parts of an Image

In this example, a diagonal line impulse of height 100 is added to an exponentially decreasing radial function on a 256×256 matrix, as shown in Figure 15, where black is one and white is zero in a gray scale.

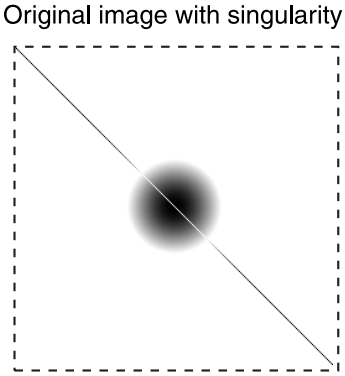


Figure 15. Original smooth image with a line impulse singularity. Black is one and white is zero. The dashed frame delimits the image.

Prefiltering and filtering by the filters of Figure 10 separate the smooth and singular structures of the original image as shown in Figure 16.

The energy in each of the twelve parts, measured by their Frobenius norm, is shown in the following tableau.

Energy $= 1.0\text{e}+06^*$

1.3039	1.8358	2.4944	3.4984
1.8358			4.3728
2.4944			5.6681
3.4984	4.3728	5.6681	4.1156.

The difference between the sum of the twelve parts and the original figure, in the Frobenius norm, is

$$2.4778\text{e}{-10}.$$

The difference between the sum of parts $(3,4)$, $(4,3)$, and $(4,4)$ and the original figure, in the Frobenius norm, is

$$4.2847\text{e}+04.$$

Since much of the energy is contained in the parts $(3,4)$, $(4,3)$, and $(4,4)$ of the figure, one expects that their sum, which is part $(3,3)$ of Figure 16, will recover the regular part although with some distortion away from the origin due to mixing of frequencies in the Fourier space (a phenomenon called *aliasing*) [16, p. 204]. The singular line impulse is recovered in parts $(1,4)$

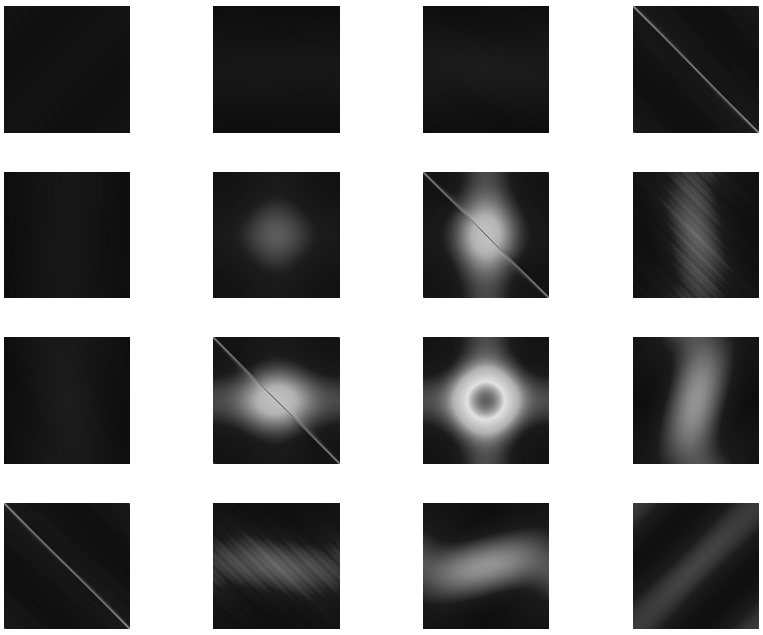


Figure 16. Filtering of a smooth function with a line impulse. White is one and black is zero in a gray scale.

and $(4,1)$ of Figure 16. Thus, the regular and singular parts of an image can be separated by means of the twelve filters.

7.6. Image Compression in the Fourier Domain

In this example, a horizontal line impulse of height 100 is added to a 256×256 zero matrix as shown in Figure 17a. In this figure, black is one and white is zero in a gray scale. The dashed frames are used to delimit the images. The absolute value of the fast Fourier transform of the original figure is shown in Figure 17b. Prefiltering and filtering by the filters of Figure 10 recover the line structure of the original image. The filtered part $(1,3)$ contains the horizontal line at about half the intensity of the original line. The filtered parts $(2,3)$, $(3,3)$, and $(4,3)$ (not shown) are similar to part $(1,3)$. The remaining 12 filtered parts are white (part $(1,4)$ is shown in Figure 17d).

With no compression, the energy in each of the twelve parts, measured by their Frobenius norm, is shown in the following tableau.

Energy $= 1.0\text{e}+05$*

$$\begin{array}{cccc} 0 & 0 & 2.8905 & 0.0855 \\ 0 & & & 0.0855 \\ 0 & & & 0.0855 \\ 0 & 0 & 2.8905 & 0.0855. \end{array}$$

A first compression omits the zero-energy left half-part of the Fourier transform of the original image. This compression does not reduce the quality of the reconstructed image.

A further compression is obtained by taking only parts $(1,3)$, $(1,4)$, and $(2,4)$ of the Fourier transform of the original image. A final compression is obtained by taking only part $(1,3)$ of the Fourier transform of the original image.

The difference between the sum of the 12, 6, 3, and 1 parts which are retained and the original figure, in the Frobenius norm, is,

$$\text{diff}\,12 = 2.9865\text{e}-12, \quad \text{diff}\,6 = 2.9865\text{e}-12, \quad \text{diff}\,3 = 1.1370\text{e}+03, \quad \text{diff}\,1 = 1.1390\text{e}+03,$$

respectively, In the last two cases, compression slightly reduces the quality of the reconstructed image.

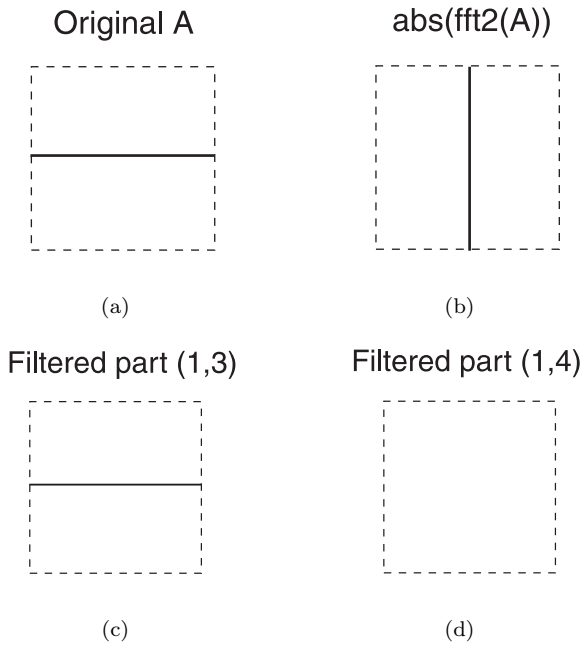


Figure 17. Filtering and compression of an image containing a horizontal line impulse. Black is one and white is zero. The dashed frames are used to delimit the figures.

From one compression to the next, half the number of 256×256 inverse fast Fourier transforms need to be done. In MATLAB 5.3, sparse matrices cannot be used with the fast Fourier transform.

8. CONCLUSION

Hyperfunctions in $\mathbb{R}^n$ have been presented as sums of boundary values of holomorphic functions defined in infinitesimal wedges in $\mathbb{C}^n$. Microlocal analysis has been briefly reviewed and a multiwavelet system adapted to microlocal filtering is proposed. A rough estimate of the microlocal content of functions or signals is obtained from their multiwavelet expansions and a fast algorithm for multiwavelet microlocal filtering is presented. The numerical filtering of a natural image and several simple geometric figures have been prefiltered and filtered in the Fourier space to analyze their microanalytic properties.

REFERENCES

1. B. Burke Hubbard, *The World According to Wavelets: The Story of a Mathematical Technique in the Making*, 2nd Edition, A K Peters, Wellesley, MA, (1998).
2. M. Sato, Theory of hyperfunctions I, *J. Fac. Sci. Univ. Tokyo, Sec. I* **8** (1), 139–193, (1959).
3. L. Schwartz, *Théorie des Distributions*, 3rd Edition, Hermann, Paris, (1966).
4. I. Daubechies, *Ten Lectures on Wavelets*, CBMS-NSF Regional Conference Series in Applied Mathematics, Volume 61, SIAM, Philadelphia, PA, (1992).
5. V. Strela, P.N. Heller, G. Strang, P. Topiwala and C. Heil, The application of multiwavelet filter banks to image processing, *IEEE Trans. Image Processing* **8** (4), 548–563, (1999).
6. X.-G. Xia, J.S. Geronimo, D.P. Hardin and B.W. Suter, Design of prefilters for discrete multiwavelet transforms, *IEEE Trans. Signal Processing* **44** (1), 25–35, (January 1996).
7. A. Kaneko, *Linear Partial Differential Equations with Constant Coefficients,* (In Japanese), Iwanami, Tokyo, (1992).
8. A. Kaneko, *Introduction to Hyperfunctions*, Kluwer Academic, Dordrecht, (1988).
9. R. Ashino, M. Nagase and R. Vaillancourt, A construction of multiwavelets, *Computers Math. Applic.* **32** (3), 23–37, (1996).
10. I. Daubechies, A. Grossmann and Y. Meyer, Painless nonorthogonal expansions, *J. Math. Phys.* **27** (5), 1271–1283, (1986).
11. C. Heil and D. Walnut, Continuous and discrete wavelet transforms, *SIAM Review* **31** (4), 628–666, (1989).
12. E. Hernández and G. Weiss, *A First Course on Wavelets*, CRC Press, Boca Raton, FL, (1996).
13. E. Hernández, X. Wang and G. Weiss, Smoothing minimally supported frequency wavelets. Part I, *J. Fourier Anal. Appl.* **2** (4), 329–340, (1996).

14. R. Coifman and M.V. Wickerhauser, Entropy-based algorithms for best basis selection, *IEEE Trans. Inform. Theory* **38** (2), 713–718, (March 1992).
15. M. Morimoto, *An Introduction to Sato's Hyperfunctions*, Transl. by M. Morimoto, Translations of Mathematical Monographs, Volume 129, American Mathematical Society, Providence, RI, (1993).
16. J.S. Lim, *Two-Dimensional Signal and Image Processing*, Prentice Hall P T R, Englewood Cliffs, NJ, (1990).