Wavelets, Shearlets and Geometric Frames: Part II

Philipp Grohs[1] and Axel Obermeier[2]

October 22, 2014

[1] ETH Zürich
[2] ETH Zürich, supported by SNF grant 146356

Outline

3. Curvelets, shearlets and parabolic molecules
4. Related systems
5. Applications

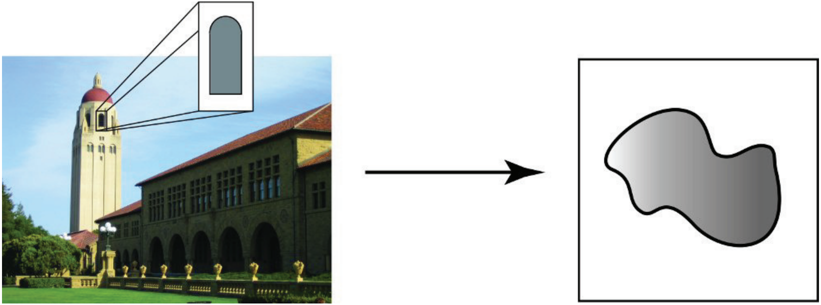
Recall...

Definition (Donoho (2001))

The set $\mathcal{E}$ of cartoon images is given by

$$\mathcal{E} := \{f = f_0 + \chi_B f_1\},$$

where $f_0, f_1 \in C^2([0,1]^2)$ and χ_B is the indicator function of $B \subset [0,1]^2$ with C^2 boundary.



Recall benchmark approximation rate for $\mathcal{E}$ is N^{-1}!

Limitations of Wavelets

Theorem

We have that

$$s^* \left(\mathcal{E}, \mathcal{W}^{2D}(\varphi, \psi, \alpha)\right) = \frac{1}{2}.$$

This is one magnitude short of the optimal rate N^{-1}.

Can we find better dictionaries??

Goal

Construct (tight) frame for $L^2(\mathbb{R}^2)$ such that for all $f \in \mathcal{E}$ we have

$$\|f - f_N\|_2 \lesssim N^{-1},$$

where f_N is the reconstruction from the N biggest frame coefficients.

3. Curvelets, Shearlets and Parabolic Molecules

Curvelets: A first Breakthrough

Main Idea: Wavelets are supported in isotropic quadrilaterals of width $\sim 2^j$. Too many such quadrilaterals are needed to cover the singularity curve.

$\rightsquigarrow$ basis functions supported in an anisotropic rectangle of length $\sim 2^{j/2}$ and width $\sim 2^j$. If we also allow rotations along (say) $2^{j/2}$ equispaced angles at scale j we might have a chance.

Curvelets: A first Breakthrough

Inspired by this idea we seek to construct systems of the form

$$\varphi_{j,l,k}(x) := 2^{3j/4}\psi(D_{2^j}R_{\theta_{j,l}}x - k),$$

where

$$D_a := \mathrm{diag}(a, \sqrt{a}),\ R_\theta = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix}$$

and

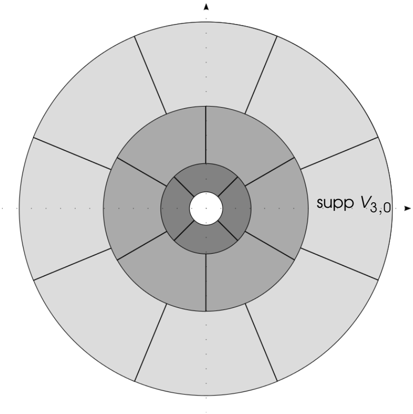
$$\theta_{j,l} \sim 2^{-j/2} l 2\pi, \quad l = -2^{j/2}, \ldots, 2^{j/2}.$$

How to realize such a system?

Curvelets: A first Breakthrough

Start with decomposition of Fourier space into 'parabolic wedges' of aspect ratio 'length $\sim$ width^2' (eg. length $\sim 2^j$, width $\sim 2^{j/2}$) with associated partition-of-unity $\{V_{j,l}\}_{j\in\mathbb{N},\ l\in\{-2^{j/2},\dots,2^{j/2}\}}$, s.t.

$$\sum_{j,l} |V_{j,l}(\xi)|^2 = 1.$$



Curvelets: A first Breakthrough

Build dictionary by modulating the partition functions:

$$\hat{\varphi}_{j,l,k}(\xi) := 2^{-3j/4} \exp(2\pi i R_{\theta_{j,l}}^{-1} D_{2^{-j}} k \xi) V_{j,l}(\xi), \quad j \in \mathbb{N},\ l \in \{-2^{j/2}, \ldots, 2^{j,}$$

where

$$D_a := \operatorname{diag}(a, \sqrt{a}),\ R_\theta = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix}$$

and

$$\theta_{j,l} \sim 2^{-j/2} l 2\pi, \quad l = -2^{j/2}, \ldots, 2^{j/2}.$$

We collect all indices in the index set Λ and get dictionary $\{\varphi_\lambda\}_{\lambda \in \Lambda}$.

Curvelets: A first Breakthrough

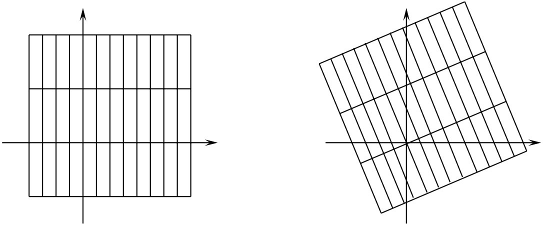
Space-picture:

$$\varphi_{j,l,k} := 2^{-3j/4} T_{U_{j,l}k} \mathcal{F}^{-1} V_{j,l}, \quad j \in \mathbb{N},\ l \in \{-2^{j/2}, \ldots, 2^{j/2}\},\ k \in \mathbb{Z}^2,$$

where

$$T_y f(\cdot) := f(\cdot - y),$$

$$U_{j,l} := R_{\theta_{j,l}}^{-1} D_{2^{-j}},$$



j: scale, l: angle, k: location.

Curvelets: A first Breakthrough

Theorem

The system $\{\varphi_\lambda\}_{\lambda\in\Lambda}$ constitutes a Parseval frame for $L^2(\mathbb{R}^2)$, i.e.

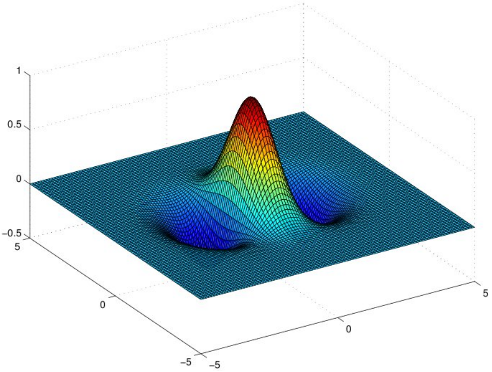
$$\|f\|_{L^2}^2 = \sum_{\lambda\in\Lambda} |\langle f, \varphi_\lambda\rangle|^2$$

and

$$f = \sum_{\lambda\in\Lambda} \langle f, \varphi_\lambda\rangle \varphi_\lambda. \quad (1)$$

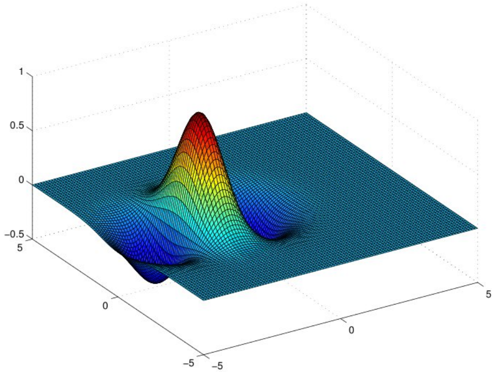
proof

Curvelets: A first Breakthrough



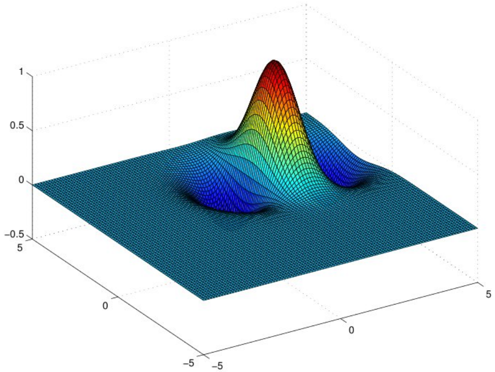
Curvelets are...

Curvelets: A first Breakthrough



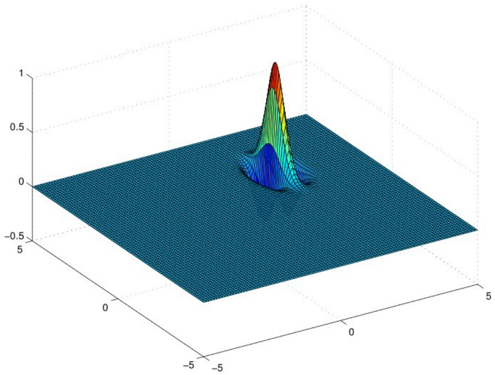
...essentially waveforms whose essential support satisfies 'width $\sim$ length^2'...

Curvelets: A first Breakthrough



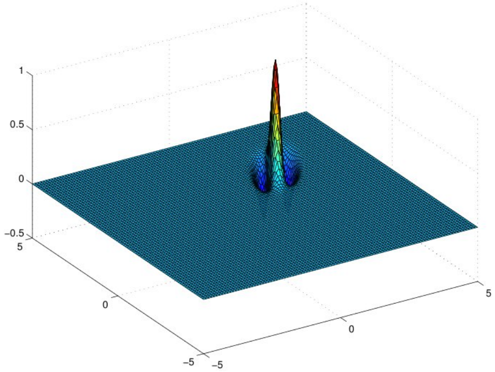
...oscillatory across the shorter edge and low-pass along the longer edge...

Curvelets: A first Breakthrough



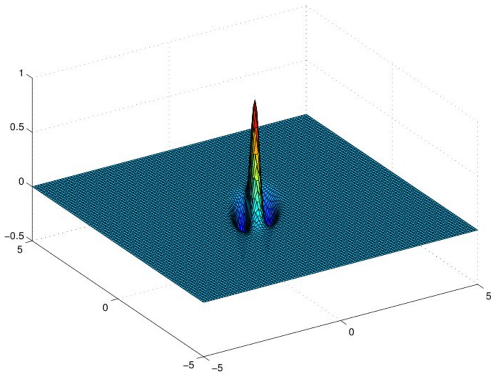
...scaled...

Curvelets: A first Breakthrough



...rotated...

Curvelets: A first Breakthrough



...and translated.

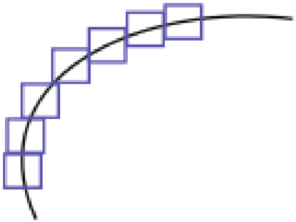
Anisotropic Scaling

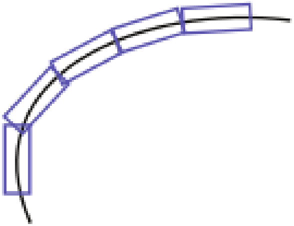
Important: In contrast to 2D-wavelets which have essential support in quadrilaterals

$$2^{-j}[k_1 - a, k_1 + a] \times [k_2 - a, k_2 + a],$$

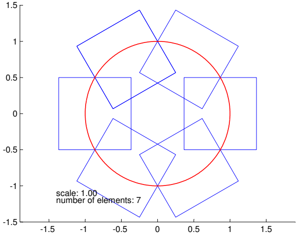
the supports of curvelets obeys a *parabolic scaling law*

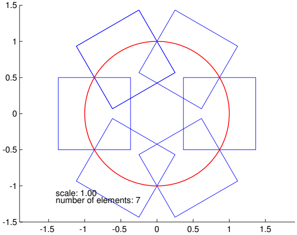
$$\text{length} \sim 2^j \quad \text{and} \quad \text{width} \sim 2^{j/2}.$$





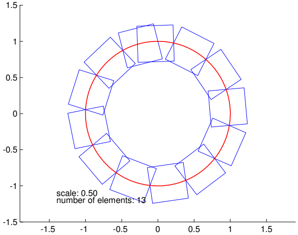
Heuristic: Isotropic Dilation vs. Anisotropic Dilation

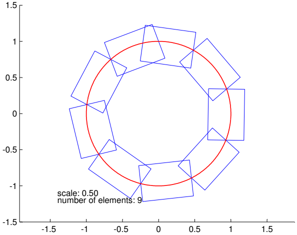




Using elements satisfying parabolic scaling relation length $\sim 2^{-j/2}$, width $\sim 2^{-j}$

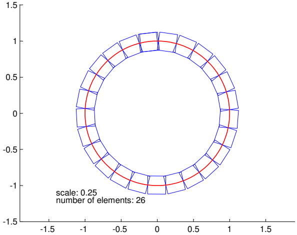
Heuristic: Isotropic Dilation vs. Anisotropic Dilation

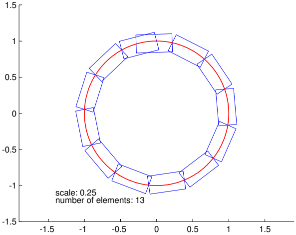




we can cover the singularity curve with $\sim 2^{j/2}$ elements, as opposed to $\sim 2^{j}$ for isotropic methods (such as wavelets)

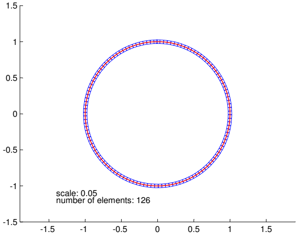
Heuristic: Isotropic Dilation vs. Anisotropic Dilation

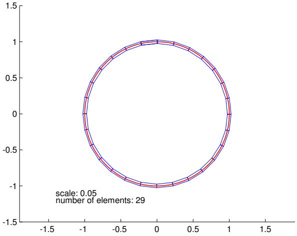




so if we cut at scale J we only need $\sim 2^{J/2}$ coefficients

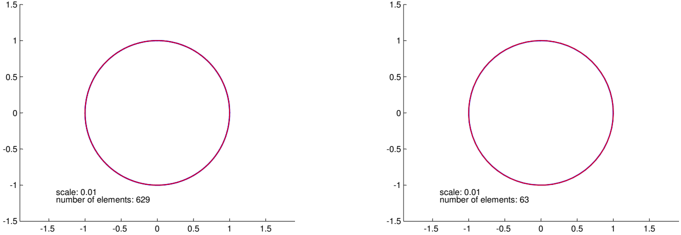
Heuristic: Isotropic Dilation vs. Anisotropic Dilation





cartoon functions are in the Sobolev space $H^{1/2}$, therefore, cutting of at scale J will induce an error of order $2^{-J/2}$

Heuristic: Isotropic Dilation vs. Anisotropic Dilation



and we arrive at the desired rate.

Optimality of Curvelets

Theorem (Candès-Donoho (2004))
Curvelets are optimal for cartoon-images, e.g.,

$$s^*(\mathcal{E}, \mathit{Curvelet}) = s^*(\mathcal{E}) = 1.$$

Actual proof *much* more complicated.

Shearlets

- So far theoretical construction
- Next step: construction of fast algorithms
- ☹ How implement rotations for digital data on a grid?
-  Replace rotation by shearing!

Shearlets

Index set

$$\Lambda^\sigma := \left\{(\varepsilon, j, l, k) \in \mathbb{Z}_2 \times \mathbb{Z}^4 : \varepsilon \in \{0,1\},\ j \geq 0,\ l = -2^{\lfloor \frac{j}{2} \rfloor}, \cdots, 2^{\lfloor \frac{j}{2} \rfloor}\right\},$$

and the shearlet system

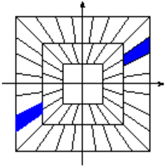
$$\Sigma := \{\sigma_\lambda : \lambda \in \Lambda^\sigma\},$$

with

$$\sigma_{(\varepsilon,0,0,k)}(\cdot) = \varphi(\cdot - k), \quad \sigma_{(\varepsilon,j,l,k)}(\cdot) = 2^{3j/4}\psi^\varepsilon\left(D^\varepsilon_{2^j} S^\varepsilon_{l,j} \cdot - k\right), \quad j \geq 1,$$

$$D^0_a = D_a,\ D^1_a := \operatorname{diag}(\sqrt{a}, a),\ S_{l,j} := \begin{pmatrix} 1 & l2^{-\lfloor j/2 \rfloor} \\ 0 & 1 \end{pmatrix},$$

$$S^1_{l,j} = (S^0_{l,j})^\top.$$



$\operatorname{supp} \mathcal{F}\varphi \subset [-2,2]^2$,
$\operatorname{supp} \mathcal{F}\psi^0 \subset ([-4,-4] \cup [1,4]) \times [-2,2]$,
$\operatorname{supp} \mathcal{F}\psi^1 \subset [-2,2] \times ([-4,-1] \cup [1,4])$

Shearlets

Theorem (Guo-Labate (2008))

With a bandlimited shearlet frame Σ we have

$$s^*(\mathcal{E}, \Sigma) = s^*(\mathcal{E}) = 1.$$

Shearlets

- Shearing better adapted to data sampled on digital grid
- Fast algorithms exist
- Also compactly supported shearlet frames available (Kutyniok et. al. (2012))
- Software and publications available at `www.shearlet.org`
- More general concept: *parabolic molecules* encompass all known constructions and yield simple proofs of optimal cartoon approximation (G-Kutyniok (2014)).

Literature

- Candès, Dohoho. New tight frames of curvelets and optimal representations of objects with piecewise C^2 singularities. *Communications in Pure and Applied Mathematics* (2004).
- Kutyniok, Labate (Eds.). Shearlets: Multiscale analysis for multivariate data. Birkhäuser/Springer (2012).
- Grohs, Kutyniok. Parabolic Molecules. *Foundations of Computational Mathematics* (2014).

4. Related Systems

Different Anisotropies

Consider dilation matrix $D_a^\beta := \mathrm{diag}(a, a^\beta)$. Build frames based on the principle

$$\varphi_{j,l,k}(x) := 2^{j(1+\beta)/2}\psi(D_{2^j}^\beta R_{\theta_{j,l}} x - k),$$

where

$$\theta_{j,l} \sim 2^{-j(1-\beta)} l 2\pi, \quad l = -2^{j/2}, \dots, 2^{j(1-\beta)}.$$

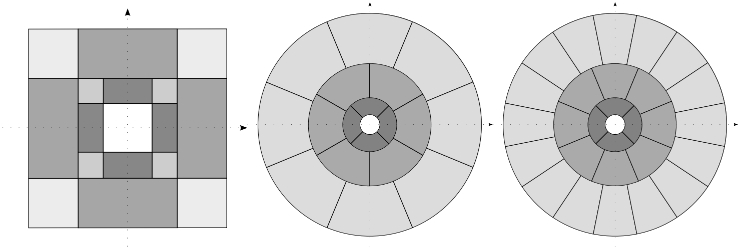
$\beta = 1$ Wavelets

$\beta = 1/2$ Curvelets, shearlets, parabolic molecules

$\beta = 0$ Ridgelets

$\rightsquigarrow$ β-molecules (G-Keiper-Kutyniok-Schaefer (2014))

Fourier Partitionings



Left to right: Fourier partitioning associated to wavelets, curvelets, ridgelets.

Approximation Results

New signal class of generalized cartoon-images

$$\mathcal{E}^\alpha := \{f = f_0 + \chi_B f_1\},$$

where $f_0, f_1 \in C^\alpha([0,1]^2)$ and χ_B is the indicator function of $B \subset [0,1]^2$ with C^2 boundary.

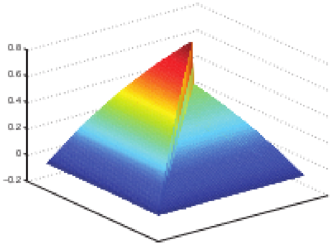
Theorem (G-Keiper-Kutyniok-Schaefer (2014))

Frames of β-molecules are optimal for $\mathcal{E}^\alpha$ if $\beta = \alpha^{-1}$ and $\beta \in [1/2, 1]$.

Line Singularities

Signal class

$$\mathcal{L}^\alpha = \{f \in C^\alpha, \text{ apart from line discontinuities}\}.$$



Theorem (Candès (1999), G-Keiper-Kutyniok-Schaefer (2014))

0-molecules (aka ridgelets) are optimal for the signal class $\mathcal{L}^\alpha$, e.g.

$$s^*(\mathcal{L}^\alpha, \textit{ridgelets}) = s^*(\mathcal{L}^\alpha) = \alpha/2.$$

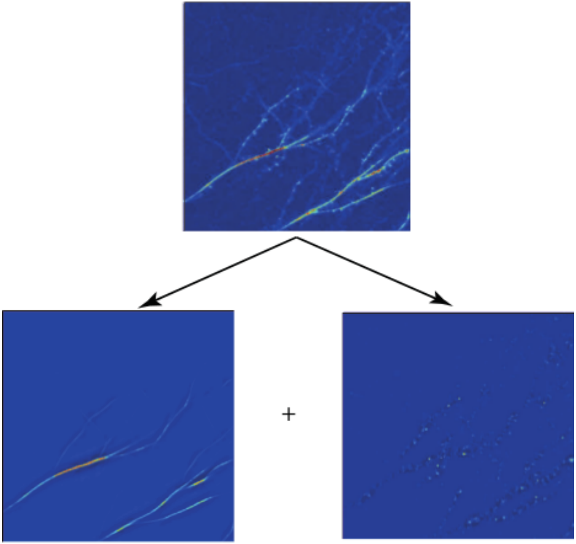
Literature

- Candès. Ridgelets and the representation of mutilated Sobolev functions. *SIAM Journal on Mathematical Analysis* (2001).
- Grohs, Keiper, Kutyniok and Schaefer. Cartoon Approximation with α-Curvelets. preprint (2014), available from `www.math.ethz.ch/~pgrohs/research`.
- Grohs, Keiper, Kutyniok and Schaefer. α-Molecules. preprint (2014), available from `www.math.ethz.ch/~pgrohs/research`.

5. Some Applications

Morphological Component Analysis

Goal: Separate Signal into curvelike and pointlike components.



Source: G. Kutyniok

Morphological Component Analysis

Wavelets are optimal for point-like features, curvelets/shearlets/parabolic molecules are optimal for curve-like features

For wavelet dictionary $\mathcal{W} = (\psi_{j,k})$ and curvelet frame $\Gamma = (\gamma_{j,k,l})$ consider combined dictionary $\mathcal{W} \cup \Gamma$ and, given

$$f = f_{curv} + f_{point}$$

seek the sparsest representation

$$f = \underbrace{\sum_{jk} c_{jk} \psi_{j,k}}_{\hat{f}_{point}} + \underbrace{\sum_{j,k,l} d_{j,k,l} \gamma_{j,k,l}}_{\hat{f}_{curv}} .$$

- Actual algorithm uses bandpass filter on $f = \sum_i P_i f$ and solves, for each frequency part

$$(\hat{f}^i_{point}, \hat{f}^i_{curv}) := \operatorname*{arg\,min}_{f^j_{point} + f^i_{curv} = P_i f} \|(\langle f^i_{point}, \psi_{j,k}\rangle)_{j,k}\|_{\ell^1} + \|(\langle f^i_{curv}, \gamma_{j,k,l}\rangle)_{j,k,l}\|_{\ell^1}.$$

- Under certain conditions it holds that

$$\lim_{i\to\infty} \frac{\|\hat{f}^i_{point} - P_i f_{point}\| + \|\hat{f}^i_{curv} - P_i f_{curv}\|}{\|P_i f_{point}\| + \|P_i f_{curv}\|} = 0.$$

- Donoho, Kutyniok. Microlocal Analysis of the Geometric Separation Problem. *Communications in Pure and Applied Mathematics* (2012).
- Software available at `www.shearlab.org`
- Link between Harmonic Analysis and Compressed Sensing

Solving Transport PDEs

Equations of the form

$$s \cdot \nabla u(x,s) + \kappa u(x,s) = f(x,s) + Q(u)(x,s), \quad (x,s) \in \Omega \times \mathbb{S}^{d-1},$$

where $\Omega \subset \mathbb{R}^d$, κ absorption coefficient, f source term and $Q(u)$ scattering operator (for instance $Q(u)(x,s) = \int_{\mathbb{S}^{d-1}} K(s,t)u(x,t)dt$) + inflow BCs.

Stationary distribution of a phase-space density u whose evolution is governed by free transport, absorption, external sources and interaction with the surrounding medium via a scattering operator. Examples include radiative transfer (simulation of dense gas at very high temperatures) or socio-economic processes.

Kinetic Transport Equations

Difficulties in the numerical solution:

1. 'Curse of dimensionality': Problem is $2d-1$-dimensional
2. Line singularities transported along rays
3. The equation is not H^1-elliptic – wavelet and FE discretizations do not lead to well-conditioned linear systems
4. Anisotropy – anisotropic meshes cannot be used since they need to be combined for different directions.

Goal: Adaptive approximation schemes which operate in optimal computational complexity (accuracy vs. number of flops).

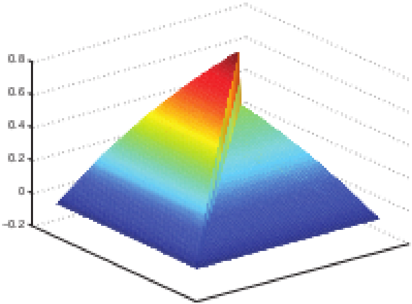
Question: What is the right discretization for such equations?

Solving Transport PDEs

Solution of

$$s \cdot \nabla u + \kappa u = f$$

may be singular along lines $\rightsquigarrow$ use ridgelets for discretization in space!



Solving Transport PDEs

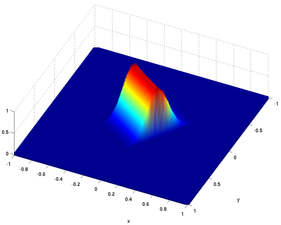
Theorem (G,Obermeier (2014))

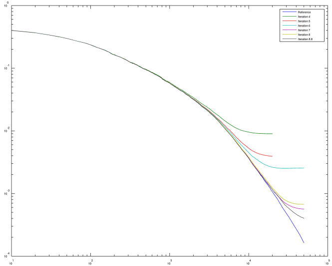
Let u be a solution of $s \cdot \nabla u + \kappa u = f$ *which is* C^n *apart from a line discontinuity in direction s. Then there exists a computable ridgelet-based algorithm SOLVE which computes in N flops an approximation* $u_N \in H^{1,s} := \{v \in L^2 : \, s \cdot \nabla v \in L^2\}$ *with the approximation rate*

$$\|u - u_N\|_{H^{1,s}} \lesssim N^{-(n-1)/2}$$

This rate is optimal.

SOLVING TRANSPORT PDES

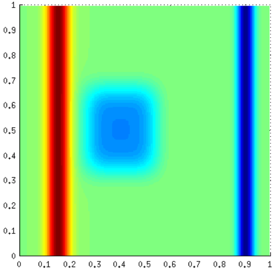




Numerical approximation computed by SOLVE converges exponentially, even if line discontinuities are present in the solution!
Compare ridgelet error $\sim \exp(-\gamma N^{\delta})$ vs. error at least $\sim N^{-1/2}$ with conventional discretizations (wavelets, FEM)!

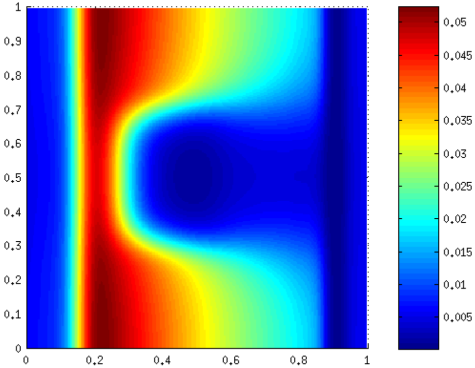
Solving Transport PDEs

Back to full kinetic equation with scattering
$Q(u)(x,s) = \int_{\mathbb{S}^1} u(x,s)ds$



red: source term, blue absorption (scattering around obstacle) **Quantity of interest:** Incident radiation $\int_{\mathbb{S}^1} u(x,s)ds$.

Solving Transport PDEs



Solution computed using ridgelets in space, together with *sparse collocation scheme*, breaking curse of dimensionality.

Solving Transport PDEs

- First provably convergent adaptive solver for transport PDEs
- Etter, Grohs, Obermeier. FFRT-A Fast Finite Ridgelet Transform for Radiative Transport. *SIAM Journal on Multiscale Modeling and Simulation* (2015).
  Grohs, Obermeier. Optimal Adaptive Ridgelet Schemes for Linear Transport Equations. (2014), available from
  `www.math.ethz.ch/~pgrohs/research`.
- Software package available at
  `www.math.ethz.ch/~pgrohs/research/FFRT`.
- Link between Harmonic Analysis and Numerics

Classification of Singularities

Separate edges, corners and smooth regions

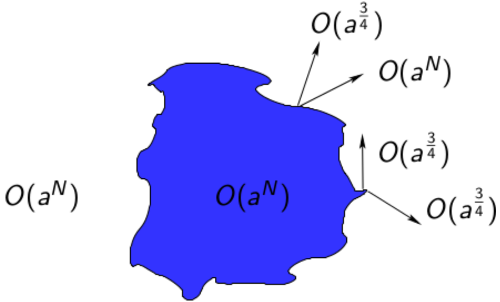
Source: D. Labate

Shearlet coefficients have different decay rates for different types of singularities.

Classification of Singularities

Theorem (Guo-Labate (2008))

With $a = 2^{-j}$, the scale we have



Classification of Singularities

- Very competitive results
- Extension to 3D exists
- Guo, Labate. Analysis and identification of multidimensional singularities using the continuous shearlet transform, from: `Shearlets: Multiscale Analysis for Multivariate Data', Birkhäuser/Springer (2012).
- Software available from `http://www.math.uh.edu/~dlabate/software.html`
- Link between Harmonic Analysis and Geometry

Further Applications

- Time propagation of wave equations (Candès-Demanet (2006))
- Edge detection (Easley-Guo-Labate (2008))
- High quality denoising (Easley-Labate-Colonna (2009))
- Inpainting with theoretical guarantees (Genzel-Kutyniok (2015))
- Fast motion deblurring (G-Kereta-Wiesmann (2014))

SUMMARY

- Partitioning Fourier plane into anisotropic wedges yields dictionaries with built-in directionality
- This strategy yields families of representation systems capable of solving problems for which conventional systems fall short
- Fourier partitioning yields fast algorithms via FFT
- Harmonic Analysis is a treasure trove for designing dictionaries, customized to specific `data architectures'

End of Part II

Thank You!

Appendix: Proofs and Additional Material

Proof Sketch: We have

$$\|f\|_2^2 = \sum_{j,l} \int_{\text{supp } \Phi_{j,l}} |\Phi_{j,l}(\xi)\mathcal{F}f(\xi)|^2 d\xi.$$

Since $\{2^{-3j/4}\exp(2\pi i U_{j,l}k\xi)\}_{k\in\mathbb{Z}^2}$ is an ONB of $L^2(\text{supp } \Phi_{j,l})$,

$$\int |\Phi_{j,l}(\xi)\mathcal{F}f(\xi)|^2 d\xi = \sum_{k\in\mathbb{Z}^2} |\int_{\mathbb{R}^2} 2^{-3j/4}\Phi_{j,l}(\xi)\mathcal{F}f(\xi)\exp(2\pi i U_{j,l}k\xi)d\xi|^2.$$

By Parseval we have

$$\int_{\mathbb{R}^2} \Phi_{j,l}(\xi)\mathcal{F}f(\xi)\exp(2\pi i U_{j,l}k\xi)d\xi = \int_{\mathbb{R}^2} T_{U_{j,l}k}\mathcal{F}^{-1}\Phi_{j,l}(x)f(x)dx.$$

Putting together we get

$$\|f\|_2^2 = \sum_{\lambda\in\Lambda} |\langle f, \varphi_\lambda\rangle|^2.$$

▸ return to talk