α-Molecules: Wavelets, Shearlets, and Beyond

Gitta Kutyniok
(Technische Universität Berlin)

joint with

Philipp Grohs (ETH Zürich)

Sandra Keiper & Martin Schäfer (Technische Universität Berlin)

30th Intl. Colloquium on Group Theoretical Methods in Physics
Ghent University, Belgium, July 14 – 18, 2014
Outline

1. Anisotropic Phenomena
   - Modern Multivariate Data
   - Problem with Wavelets

2. Shearlet Systems
   - Compactly Supported Shearlets
   - Optimal Sparse Approximation
   - Applications with Compressed Sensing

3. $\alpha$-Molecules
   - Curvelets
   - Parabolic Molecules
   - General Framework
   - Approximation Properties

4. Conclusions
Many important multivariate problem classes are governed by *anisotropic features*, which require efficient encoding strategies.

The anisotropic structure can be given...

- ...explicitly.
  - Imaging Science: Edges.
- ...implicitly.
  - Transport Equations: Shock fronts.
Methodology:
Exploit a carefully designed representation system \((\psi_\lambda)_{\lambda} \subseteq \mathcal{H}\):

\[
\mathcal{H} \supseteq \mathcal{C} \ni f \rightarrow (\langle f, \psi_\lambda \rangle)_{\lambda} \rightarrow \sum_{\lambda} \langle f, \psi_\lambda \rangle \psi_\lambda = f
\]

Two Main Goals:

1. Decomposition
2. Efficient representations

Main Desiderata:

- Multiscale representation system.
- Partition of Fourier domain.
- Fast decomposition and reconstruction algorithm.
- Optimally sparse approximation of the considered class.

\(\Rightarrow\) Here: Functions governed by anisotropic features.
What is an Image?
What is an Image?
What is an Image?
Anisotropic/Cartoon Structures

Images:

- Intuitively main structure in images.
- Justified by neurophysiology.

Field et al., 1993
Anisotropic/Cartoon Structures

Images:

- Intuitively main structure in images.
- Justified by neurophysiology.

PDEs:

- Linear transport problem:

\[ u := \mathbf{b} \cdot \nabla u + cu = f \text{ in } D \quad u = g \text{ on } \Gamma_- \]

(Source: Dahmen, K, Lim, Welper, Schwab; 2013)
Fitting Model

Definition (Donoho; 2001):
The set of cartoon-like functions $\mathcal{E}^2(\mathbb{R}^2)$ is defined by

$$\mathcal{E}^2(\mathbb{R}^2) = \{ f \in L^2(\mathbb{R}^2) : f = f_0 + f_1 \cdot \chi_B \},$$

where $B \subset [0, 1]^2$ with $\partial B$ a closed $C^2$-curve, $f_0, f_1 \in C^2_0([0, 1]^2)$. 
Fitting Model

Definition (Donoho; 2001):
The set of cartoon-like functions $\mathcal{E}^2(\mathbb{R}^2)$ is defined by

$$\mathcal{E}^2(\mathbb{R}^2) = \{ f \in L^2(\mathbb{R}^2) : f = f_0 + f_1 \cdot \chi_B \},$$

where $B \subset [0, 1]^2$ with $\partial B$ a closed $C^2$-curve, $f_0, f_1 \in C^2_0([0, 1]^2)$.

Theorem (Donoho; 2001):
Let $(\psi_\lambda)_\lambda \subseteq L^2(\mathbb{R}^2)$. Allowing only polynomial depth search, the optimal asymptotic approximation error of $f \in \mathcal{E}^2(\mathbb{R}^2)$ is

$$\| f - f_N \|_2^2 \asymp N^{-2}, \quad N \to \infty,$$

where $f_N = \sum_{\lambda \in \Lambda_N} c_\lambda \psi_\lambda$. 

\[ \]
Definition (1D): Let $\psi \in L^2(\mathbb{R})$ be a wavelet. Then the associated homogeneous continuous wavelet system is defined by

$$\left\{ a^{-1/2} \psi \left( \frac{x - b}{a} \right) : a > 0, b \in \mathbb{R} \right\}.$$
Definition (1D): Let $\psi \in L^2(\mathbb{R})$ be a wavelet. Then the associated homogeneous continuous wavelet system is defined by

$$\left\{ a^{-1/2} \psi \left( \frac{x - b}{a} \right) : a > 0, b \in \mathbb{R} \right\}.$$

Affine Systems:

- Let $A_1 = \mathbb{R}^+ \ltimes \mathbb{R}$ be the affine group with multiplication given by

  $$(a_1, b_1) \cdot (a_2, b_2) = (a_1 a_2, b_1 + a_1 b_2).$$

- Then

  $$\pi : A_1 \to \mathcal{U}(L^2(\mathbb{R})), \quad \pi(a, b)\psi(x) = a^{-1/2} \psi \left( \frac{x - b}{a} \right)$$

  is a unitary representation of $A_1$. 
Review of 1D Wavelets

Definition (1D): Let $\psi \in L^2(\mathbb{R})$ be a wavelet. Then the associated homogeneous continuous wavelet system is defined by

$$\left\{ a^{-1/2} \psi \left( \frac{x-b}{a} \right) : a > 0, b \in \mathbb{R} \right\}.$$

Affine Systems:

- Let $A_1 = \mathbb{R}^+ \rtimes \mathbb{R}$ be the affine group with multiplication given by

$$(a_1, b_1) \cdot (a_2, b_2) = (a_1 a_2, b_1 + a_1 b_2).$$

- Then

$$\pi : A_1 \to \mathcal{U}(L^2(\mathbb{R})), \quad \pi(a, b)\psi(x) = a^{-1/2} \psi \left( \frac{x-b}{a} \right)$$

is a unitary representation of $A_1$.

- Discretization:

$$A_1 \to \{(2^{-j}, 2^{-j}m) : j, m \in \mathbb{Z}\}.$$
Review of 2-D Wavelets

Definition (1D): Let \( \phi \in L^2(\mathbb{R}) \) be a scaling function and \( \psi \in L^2(\mathbb{R}) \) be a wavelet. Then the inhomogeneous discrete wavelet system is defined by

\[
\{ \phi(x - m) : m \in \mathbb{Z} \} \cup \{ 2^{j/2} \psi(2^j x - m) : j \geq 0, m \in \mathbb{Z} \}.
\]
Review of 2-D Wavelets

Definition (1D): Let $\phi \in L^2(\mathbb{R})$ be a scaling function and $\psi \in L^2(\mathbb{R})$ be a wavelet. Then the inhomogeneous discrete wavelet system is defined by

$$\{\phi(x - m) : m \in \mathbb{Z}\} \cup \{2^{j/2} \psi(2^j x - m) : j \geq 0, m \in \mathbb{Z}\}.$$ 

Definition (2D): An inhomogeneous discrete wavelet system is defined by

$$\{\phi^{(1)}(x - m) : m \in \mathbb{Z}^2\} \cup \{2^j \psi^{(i)}(2^j x - m) : j \geq 0, m \in \mathbb{Z}^2, i = 1, 2, 3\},$$

where

$$\psi^{(1)}(x) = \phi(x_1)\psi(x_2),$$

$$\phi^{(1)}(x) = \phi(x_1)\phi(x_2)$$ and

$$\psi^{(2)}(x) = \psi(x_1)\phi(x_2),$$

$$\psi^{(3)}(x) = \psi(x_1)\psi(x_2).$$

Theorem: Discrete wavelets provide optimally sparse approximations for functions $f \in L^2(\mathbb{R}^2)$, which are $C^2$ apart from point singularities:

$$\|f - f_N\|_2^2 \lesssim N^{-1}, \quad N \to \infty.$$
Wavelet Decomposition: JPEG2000
Beyond Wavelets...

**Observation:**
- For $f \in E^2(\mathbb{R}^2)$, wavelets only achieve $\|f - f_N\|_2^2 \asymp N^{-1}$, $N \to \infty$.
- Wavelets **cannot** approximate curvilinear singularities optimally sparse.
- Reason: **Isotropic** structure of wavelets:

$$2^j \psi(\begin{pmatrix} 2^j & 0 \\ 0 & 2^j \end{pmatrix} x - m)$$

**Intuitive explanation:**
Main Goal

Design a representation system which...

- fits into the framework of affine systems,
- provides optimally sparse approximation of cartoons,
- allows for compactly supported analyzing elements,
- is associated with fast decomposition algorithms,
- treats the continuum and digital ‘world’ uniformly.
Main Goal

Design a representation system which...

- fits into the framework of affine systems,
- provides optimally sparse approximation of cartoons,
- allows for compactly supported analyzing elements,
- is associated with fast decomposition algorithms,
- treats the continuum and digital ‘world’ uniformly.

Non-exhaustive list of approaches:

- Ridgelets (Candès and Donoho; 1999)
- Curvelets (Candès and Donoho; 2002)
- Contourlets (Do and Vetterli; 2002)
- Bandlets (LePennec and Mallat; 2003)
- Shearlets (K and Labate; 2006)
What is a Shearlet?
Parabolic scaling (‘width $\approx$ length\(^2\)’):

$$A_a = \begin{pmatrix} a & 0 \\ 0 & a^{1/2} \end{pmatrix}, \quad a > 0.$$ 

Historical remark:

Scaling and Orientation

Parabolic scaling (‘width $\approx$ length$^2$’):

$$A_a = \begin{pmatrix} a & 0 \\ 0 & a^{1/2} \end{pmatrix}, \quad a > 0.$$ 

Historical remark:


Orientation via shearing:

$$S_s = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix}, \quad s \in \mathbb{R}.$$ 

Advantage:

- Shearing $S_k$, $k \in \mathbb{Z}$, leaves the digital grid $\mathbb{Z}^2$ invariant.
- Uniform theory for the continuum and digital situation.
Shearlet Systems

**Full Affine Group of Motions:**

- Let $A_n = GL_n \ltimes \mathbb{R}^n$ with multiplication given by

  $$(M_1, t_1) \cdot (M_2, t_2) = (M_1 M_2, t_1 + M_1 t_2).$$

- Then

  $\pi : A_n \to \mathcal{U}(L^2(\mathbb{R}^n)), \quad \pi(M, t)\psi(x) = |\det(M)|^{-1/2} \psi(M^{-1}(x - t))$

  is a unitary representation of $A_n$. 
Shearlet Systems

Full Affine Group of Motions:
- Let $A_n = GL_n \rtimes \mathbb{R}^n$ with multiplication given by

\[
(M_1, t_1) \cdot (M_2, t_2) = (M_1 M_2, t_1 + M_1 t_2).
\]

- Then

$$\pi : A_n \to U(L^2(\mathbb{R}^n)), \quad \pi(M, t)\psi(x) = |\det(M)|^{-1/2} \psi(M^{-1}(x - t))$$

is a unitary representation of $A_n$.

Shearlet Group:
- Let $S = (\mathbb{R}^+ \times \mathbb{R}) \rtimes \mathbb{R}^2$ with multiplication given by

\[
(a, s, t) \cdot (a', s', t') = (aa', s + s' \sqrt{a}, t + S_s A_a t').
\]

- Then

$$\sigma : S \to U(L^2(\mathbb{R}^2)), \quad \sigma(a, s, t)\psi(x) = a^{-3/4} \psi(A_a^{-1} S_s^{-1}(x - t))$$

is a unitary representation of $S$. 
Shearlet Systems

Shearlet Group:

- Let $S = (\mathbb{R}^+ \times \mathbb{R}) \ltimes \mathbb{R}^2$, and consider

\[ \sigma : S \to \mathcal{U}(L^2(\mathbb{R}^2)), \quad \sigma(a, s, t)\psi(x) = a^{-\frac{3}{4}}\psi(A_a^{-1}S_s^{-1}(x - t)) \]

- Discretization:

\[ S \to \{(2^{-j}, -2^{-j/2}k, S_{-2^{-j/2}k}A_{2^{-j}}m) : j, k \in \mathbb{Z}, m \in \mathbb{Z}^2\} \]

Remark: Discretization by coorbit theory (Feichtinger, Gröchenig; 1990).
Shearlet Systems

Shearlet Group:

- Let $S = (\mathbb{R}^+ \times \mathbb{R}) \rtimes \mathbb{R}^2$, and consider

$$\sigma : S \to \mathcal{U}(L^2(\mathbb{R}^2)), \quad \sigma(a, s, t) \psi(x) = a^{-\frac{3}{4}} \psi(A_a^{-1} S_s^{-1}(x - t))$$

- Discretization:

$$S \to \{(2^{-j}, -2^{-j/2} k, S_{-2^{-j/2} k} A_{2^{-j}} m) : j, k \in \mathbb{Z}, m \in \mathbb{Z}^2\}.$$

Remark: Discretization by coorbit theory (Feichtinger, Gröchenig; 1990).

Definition (K, Labate; 2006):

For $\psi \in L^2(\mathbb{R}^2)$, the associated shearlet system is defined by

$$\{2^{\frac{3j}{4}} \psi(S_k A_{2^j} \cdot - m) : j, k \in \mathbb{Z}, m \in \mathbb{Z}^2\}.$$
Shearlet Systems

Shearlet Group:
- Let $S = (\mathbb{R}^+ \times \mathbb{R}) \ltimes \mathbb{R}^2$, and consider
  \[
  \sigma : S \to \mathcal{U}(L^2(\mathbb{R}^2)), \quad \sigma(a, s, t)\psi(x) = a^{-3/4}\psi(A^{-1}_{a} S^{-1}_{s}(x - t))
  \]
- Discretization:
  \[
  S \to \{(2^{-j}, -2^{-j/2}k, S_{-2^{-j/2}k} A_{2^{-j}}m) : j, k \in \mathbb{Z}, m \in \mathbb{Z}^2\}.
  \]

Remark: Discretization by coorbit theory (Feichtinger, Gröchenig; 1990).

Definition (K, Labate; 2006):
For $\psi \in L^2(\mathbb{R}^2)$, the associated shearlet system is defined by
\[
\{2^{3j/4}\psi(S_k A_{2^j} \cdot -m) : j, k \in \mathbb{Z}, m \in \mathbb{Z}^2\}.
\]

∽ Problem: Non-uniform treatment of directions.
Example of Classical (Band-Limited) Shearlet

Let $\psi \in L^2(\mathbb{R}^2)$ be defined by

$$\hat{\psi}(\xi) = \hat{\psi}(\xi_1, \xi_2) = \hat{\psi}_1(\xi_1) \hat{\psi}_2\left(\frac{\xi_2}{\xi_1}\right),$$

where

- $\psi_1$ wavelet, $\text{supp}(\hat{\psi}_1) \subseteq [-2, -\frac{1}{2}] \cup [\frac{1}{2}, 2]$ and $\hat{\psi}_1 \in C^\infty(\mathbb{R})$.
- $\text{supp}(\hat{\psi}_2) \subseteq [-1, 1]$ and $\hat{\psi}_2 \in C^\infty(\mathbb{R})$. 
Example of Classical (Band-Limited) Shearlet

Let $\psi \in L^2(\mathbb{R}^2)$ be defined by

$$\hat{\psi}(\xi) = \hat{\psi}(\xi_1, \xi_2) = \hat{\psi}_1(\xi_1) \hat{\psi}_2(\frac{\xi_2}{\xi_1}),$$

where

- $\psi_1$ wavelet, $\text{supp}(\hat{\psi}_1) \subseteq [-2, -\frac{1}{2}] \cup [\frac{1}{2}, 2]$ and $\hat{\psi}_1 \in C^\infty(\mathbb{R})$.
- $\text{supp}(\hat{\psi}_2) \subseteq [-1, 1]$ and $\hat{\psi}_2 \in C^\infty(\mathbb{R})$.

Induced tiling of frequency domain:
Example of Classical (Band-Limited) Shearlet

Let $\psi \in L^2(\mathbb{R}^2)$ be defined by

$$\hat{\psi}(\xi) = \hat{\psi}(\xi_1, \xi_2) = \hat{\psi}_1(\xi_1) \hat{\psi}_2\left(\frac{\xi_2}{\xi_1}\right),$$

where

- $\psi_1$ wavelet, $\text{supp}(\hat{\psi}_1) \subseteq [-2, -\frac{1}{2}] \cup [\frac{1}{2}, 2]$ and $\hat{\psi}_1 \in C^\infty(\mathbb{R})$.
- $\text{supp}(\hat{\psi}_2) \subseteq [-1, 1]$ and $\hat{\psi}_2 \in C^\infty(\mathbb{R})$.

**Induced tiling of frequency domain:**

![Diagram showing the induced tiling of frequency domain](image-url)
Definition (K, Labate; 2006):
The (cone-adapted) discrete shearlet system $\mathcal{SH}(\phi, \psi, \tilde{\psi})$ generated by
$\phi \in L^2(\mathbb{R}^2)$ and $\psi, \tilde{\psi} \in L^2(\mathbb{R}^2)$ is the union of

$$\{\phi(\cdot - m) : m \in \mathbb{Z}^2\},$$

$$\{2^{3j/4} \psi(S_k A_{2j} \cdot - m) : j \geq 0, |k| \leq \left\lfloor 2^{j/2} \right\rfloor, m \in \mathbb{Z}^2\},$$

$$\{2^{3j/4} \tilde{\psi}(\tilde{S}_k \tilde{A}_{2j} \cdot - m) : j \geq 0, |k| \leq \left\lfloor 2^{j/2} \right\rfloor, m \in \mathbb{Z}^2\}.$$
(Cone-adapted) Discrete Shearlet Systems

Definition (K, Labate; 2006): The (cone-adapted) discrete shearlet system \( \mathcal{SH}(\phi, \psi, \tilde{\psi}) \) generated by \( \phi \in L^2(\mathbb{R}^2) \) and \( \psi, \tilde{\psi} \in L^2(\mathbb{R}^2) \) is the union of

\[
\begin{align*}
\{ \phi(\cdot - m) : m \in \mathbb{Z}^2 \}, \\
\{ 2^{3j/4} \psi(S_k A_{2j} \cdot - m) : j \geq 0, |k| \leq \lceil 2^{j/2} \rceil, m \in \mathbb{Z}^2 \}, \\
\{ 2^{3j/4} \tilde{\psi}(\tilde{S}_k \tilde{A}_{2j} \cdot - m) : j \geq 0, |k| \leq \lceil 2^{j/2} \rceil, m \in \mathbb{Z}^2 \}.
\end{align*}
\]

Theorem (K, Labate, Lim, Weiss; 2006): For \( \psi, \tilde{\psi} \) classical shearlets, \( \mathcal{SH}(\phi, \psi, \tilde{\psi}) \) is a Parseval frame for \( L^2(\mathbb{R}^2) \):

\[
A \| f \|^2 \leq \sum_{\sigma \in \mathcal{SH}(\phi, \psi, \tilde{\psi})} |\langle f, \sigma \rangle|^2 \leq B \| f \|^2 \quad \text{for all } f \in L^2(\mathbb{R}^2)
\]

holds for \( A = B = 1 \).
Compactly Supported Shearlets

Theorem (Kittipoom, K, Lim; 2012): Let $\phi, \psi, \tilde{\psi} \in L^2(\mathbb{R}^2)$ be compactly supported, and let $\hat{\psi}, \hat{\tilde{\psi}}$ satisfy certain decay condition. Then $\mathcal{SH}(\phi, \psi, \tilde{\psi})$ forms a shearlet frame with controllable frame bounds.

Remark: Exemplary class with $B/A \approx 4.$
Compactly Supported Shearlets

Theorem (Kittipoom, K, Lim; 2012):
Let $\phi, \psi, \tilde{\psi} \in L^2(\mathbb{R}^2)$ be compactly supported, and let $\hat{\psi}, \hat{\tilde{\psi}}$ satisfy certain decay condition. Then $\mathcal{SH}(\phi, \psi, \tilde{\psi})$ forms a shearlet frame with controllable frame bounds.

Remark: Exemplary class with $B/A \approx 4$.

Theorem (K, Lim; 2011):
Let $\phi, \psi, \tilde{\psi} \in L^2(\mathbb{R}^2)$ be compactly supported, and let $\hat{\psi}, \hat{\tilde{\psi}}$ satisfy certain decay condition. Then $\mathcal{SH}(\phi, \psi, \tilde{\psi})$ provides an optimally sparse approximation of $f \in \mathcal{E}^2(\mathbb{R}^2)$, i.e.,

$$\|f - f_N\|_2^2 \leq C \cdot N^{-2} \cdot (\log N)^3, \quad N \to \infty.$$
Recent Approaches to Fast Shearlet Transforms

www.ShearLab.org:
- Separable Shearlet Transform \((\text{Lim}; 2009)\)
- Digital Shearlet Transform \((K, \text{Shahram, Zhuang}; 2011)\)
- 2D&3D (parallelized) Shearlet Transform \((K, \text{Lim, Reisenhofer}; 2013)\)

Additional Code:
- Filter-based implementation \((\text{Easley, Labate, Lim}; 2009)\)

Theoretical Approaches:
- Adaptive Directional Subdivision Schemes \((K, \text{Sauer}; 2009)\)
- Shearlet Unitary Extension Principle \((\text{Han, K, Shen}; 2011)\)
- Gabor Shearlets \((\text{Bodmann, K, Zhuang}; 2013)\)
Applications...
Problem from Neurobiology: Alzheimer Research

- Detection of characteristics of Alzheimer.
- Separation of spines and dendrites.
Numerical Results of Separation

(Source: Brandt, K, Lim, Südermann; 2010)
Separation using Compressed Sensing

General Approach:
Let
- $x = x_1^0 + x_2^0 \in \mathcal{H}$ be a signal.
- $\Phi_i$ be ONBs such that $c_i^0$ with $x_i^0 = \Phi_i c_i^0$ is sparse, $i = 1, 2$.
- Consider

$$x = x_1^0 + x_2^0 = \Phi_1 c_1^0 + \Phi_2 c_2^0 = [\Phi_1 | \Phi_2] \begin{bmatrix} c_1^0 \\ c_2^0 \end{bmatrix}.$$  

$l_1$ Minimization Problem (Elad, Starck, Querre, Donoho; 2005):

$$(\hat{c}_1, \hat{c}_2) = \text{argmin}(\|c_1\|_1 + \|c_2\|_1) \text{ subject to } x = \Phi_1 c_1 + \Phi_2 c_2.$$
Separation using Compressed Sensing

General Approach:

Let

- $x = x_1^0 + x_2^0 \in \mathcal{H}$ be a signal.
- $\Phi_i$ be ONBs such that $c_i^0$ with $x_i^0 = \Phi_i c_i^0$ is sparse, $i = 1, 2$.
- Consider

$$x = x_1^0 + x_2^0 = \Phi_1 c_1^0 + \Phi_2 c_2^0 = [\Phi_1 \mid \Phi_2] \begin{bmatrix} c_1^0 \\ c_2^0 \end{bmatrix}.$$ 

$l_1$ Minimization Problem (Elad, Starck, Querre, Donoho; 2005):

$$(\hat{c}_1, \hat{c}_2) = \arg\min (\|c_1\|_1 + \|c_2\|_1) \text{ subject to } x = \Phi_1 c_1 + \Phi_2 c_2.$$ 

Theorems (Donoho, K; 2013) (K; 2014):

- Wavelets: Optimally sparse approximations of points.
- Shearlets: Optimally sparse approximations of curves.

$\Rightarrow$ Asymptotic optimal separation of points and curves can be proven!
Numerical Results of Separation

Wavelet Expansion

Shearlet Expansion

(Source: Brandt, K, Lim, Sündermann; 2010)
Numerical Results of Inpainting

(Source: Lim; 2014)
Inpainting using Compressed Sensing

General Approach:
Let

- $x^0 \in \mathcal{H}$ be a signal.
- $\Phi$ be an ONB such that $c^0$ with $x^0 = \Phi c^0$ is sparse.
- $\mathcal{H} = \mathcal{H}_K \oplus \mathcal{H}_M$ with orthogonal projections $P_K$ and $P_M$.

$l_1$ Minimization Problem (Elad, Starck, Querre, Donoho; 2005):

$$\hat{c} = \arg\min_c \|c\|_1 \text{ s.t. } P_K x^0 = P_K \Phi c \implies \hat{x} = \Phi \hat{c}$$
Inpainting using Compressed Sensing

General Approach:
Let
- $x^0 \in \mathcal{H}$ be a signal.
- $\Phi$ be an ONB such that $c^0$ with $x^0 = \Phi c^0$ is sparse.
- $\mathcal{H} = \mathcal{H}_K \oplus \mathcal{H}_M$ with orthogonal projections $P_K$ and $P_M$.

$L_1$ Minimization Problem (Elad, Starck, Querre, Donoho; 2005):

$$\hat{c} = \text{argmin}_c \| c \|_1 \text{ s.t. } P_K x^0 = P_K \Phi c \implies \hat{x} = \Phi \hat{c}$$

Theorems (King, K, Zhuang; 2014)(Genzel, K; 2014):
- Shearlets: Optimally sparse approximations of images.
- Asymptotic optimal inpainting can be proven!
Numerical Results of Inpainting

(Source: Lim; 2014)
Towards a General Framework…
Multiscale representation systems designed by Applied Harmonic Analysis concepts and Group Theoretical Methods have established themselves as a standard tool in applied mathematics and various application areas.

Examples:
- Wavelets.
- Curvelets.
- Shearlets.
- Ridgelets.
- ...
Applied Harmonic Analysis

Multiscale representation systems designed by Applied Harmonic Analysis concepts and Group Theoretical Methods have established themselves as a standard tool in applied mathematics and various application areas.

Examples:
- Wavelets.
- Curvelets.
- Shearlets.
- Ridgelets.
- ...

Key Property:
*Fast Algorithms combined with Sparse Approximation Properties!*
Multiscale representation systems designed by *Applied Harmonic Analysis* concepts and *Group Theoretical Methods* have established themselves as a standard tool in applied mathematics and various application areas.

**Examples:**
- Wavelets.
- Curvelets.
- Shearlets.
- Ridgelets.
- ...

**Key Property:**
*Fast Algorithms combined with Sparse Approximation Properties!*

**Question:**
*Does there exist a General Framework?*
Curvelets

Definition (Candès, Donoho; 2002): Let

- \( W \in C^\infty(\mathbb{R}) \) be a wavelet with \( \text{supp}(W) \subseteq (\frac{1}{2}, 2) \),
- \( V \in C^\infty(\mathbb{R}) \) be a ‘bump function’ with \( \text{supp}(V) \subseteq (-1, 1) \).

Then the curvelet system \( \{\gamma(j,l,k)\}_{(j,l,k)} \) is defined by

\[
\hat{\gamma}(j,0,0)(r,\omega) := 2^{-3j/4} W(2^{-j} r) V(2^{j/2} \omega)
\]
and

\[
\gamma(j,l,k)(\cdot) := \gamma(j,0,0)(R_{\theta(j,l,k)}(\cdot - x(j,l,k))).
\]
Definition (Candès, Donoho; 2002): Let
- \( W \in C^\infty(\mathbb{R}) \) be a wavelet with \( \text{supp}(W) \subseteq (\frac{1}{2}, 2) \),
- \( V \in C^\infty(\mathbb{R}) \) be a ‘bump function’ with \( \text{supp}(V) \subseteq (-1, 1) \).

Then the curvelet system \( (\gamma_{j,l,k})_{(j,l,k)} \) is defined by

\[
\hat{\gamma}_{j,0,0}(r, \omega) := 2^{-3j/4} W (2^{-j} r) V(2^{j/2} \omega)
\]

and

\[
\gamma_{j,l,k}(\cdot) := \gamma_{j,0,0}(R_{\theta_{(j,l,k)}}(\cdot - x_{(j,l,k)})).
\]

\(~\sim~\text{Problem: Rotation based + no affine system!}\)
Curvelets

**Definition (Candès, Donoho; 2002):** Let

- \( W \in C^\infty(\mathbb{R}) \) be a wavelet with \( \text{supp}(W) \subseteq (\frac{1}{2}, 2) \),
- \( V \in C^\infty(\mathbb{R}) \) be a ‘bump function’ with \( \text{supp}(V) \subseteq (-1, 1) \).

Then the **curvelet system** \( (\gamma_{(j,l,k)})_{(j,l,k)} \) is defined by

\[
\hat{\gamma}_{(j,0,0)}(r, \omega) := 2^{-3j/4} W(2^{-j} r) V(2^{j/2} \omega)
\]

and

\[
\gamma_{(j,l,k)}(\cdot) := \gamma_{(j,0,0)}(R_{\theta(j,l,k)}(\cdot - x_{(j,l,k)})).
\]

\( \Rightarrow \) **Problem:** Rotation based + no affine system!

**Theorem (Candès, Donoho; 2002):** The curvelet system forms a Parseval frame for \( L^2(\mathbb{R}^2) \) and provides optimally sparse approximations of \( f \in \mathcal{E}^2(\mathbb{R}^2) \), i.e.,

\[
\|f - f_N\|_2^2 \leq C \cdot N^{-2} \cdot (\log N)^3, \quad N \to \infty.
\]
Introduce a Framework which...

- ...covers all systems known to sparsify cartoons.
- ...enables easy transfer of (sparsity) results between systems.
- ...allows categorization of systems with respect to sparsity behaviors.
- ...is general enough to allow construction of novel systems.
Framework including Shearlets and Curvelets

Introduce a Framework which...

- ...covers all systems known to sparsify cartoons.
- ...enables easy transfer of (sparsity) results between systems.
- ...allows categorization of systems with respect to sparsity behaviors.
- ...is general enough to allow construction of novel systems.

Crucial Ingredient: Parabolic scaling, i.e., a scaling matrix of the type

\[
A_{2^j} = \begin{pmatrix} 2^j & 0 \\ 0 & 2^{j/2} \end{pmatrix}, \quad j \in \mathbb{Z},
\]

since from

\[
E(x_2) \approx \frac{1}{2} \kappa x_2^2 \quad \text{and} \quad E(\ell) = w
\]

follows

\[
w \approx \frac{\kappa}{2} \ell^2 \quad (\text{‘width } \approx \text{ length}^2\text{’}).
\]
Parametrization

Parameter space:

\[ \mathbb{P} := \mathbb{R}_+ \times \mathbb{T} \times \mathbb{R}^2, \]

where \((s, \theta, x) \in \mathbb{P}\) describes scale \(2^s\), orientation \(\theta\), and location \(x\).

Definition: A **parametrization** is a pair \((\Lambda, \Phi_\Lambda)\), where \(\Lambda\) is a discrete index set and \(\Phi_\Lambda\) is a mapping

\[
\Phi_\Lambda : \left\{ \begin{array}{c}
\Lambda \\
\lambda
\end{array} \right\} \rightarrow \mathbb{P},
\]

\[
\lambda \mapsto (s_\lambda, \theta_\lambda, x_\lambda).
\]
Parametrization

Parameter space:

\[ \mathbb{P} := \mathbb{R}_+ \times \mathbb{T} \times \mathbb{R}^2, \]

where \((s, \theta, x) \in \mathbb{P}\) describes scale \(2^s\), orientation \(\theta\), and location \(x\).

Definition: A parametrization is a pair \((\Lambda, \Phi_\Lambda)\), where \(\Lambda\) is a discrete index set and \(\Phi_\Lambda\) is a mapping

\[
\Phi_\Lambda : \begin{cases}
\Lambda & \mapsto \mathbb{P}, \\
\lambda & \mapsto (s_\lambda, \theta_\lambda, x_\lambda). 
\end{cases}
\]

Example: The canonical parametrization \((\Lambda^0, \Phi^0(\lambda))\) is defined by

\[
\Lambda^0 := \left\{ (j, \ell, k) \in \mathbb{Z}^4 : j \geq 0, \ \ell = -2^{\lfloor j/2 \rfloor - 1}, \ldots, 2^{\lfloor j/2 \rfloor - 1} \right\},
\]

and

\[
\Phi^0(j, \ell, k) = (s_\lambda, \theta_\lambda, x_\lambda) = (j, \ell 2^{-\lfloor j/2 \rfloor} \pi, R_{-\theta_\lambda} A_{2^{-s_\lambda}} k).
\]
**Parabolic Molecules**

**Definition (Grohs, K; 2014):**
Let $(\Lambda, \Phi_{\Lambda})$ be a parametrization. Then $(m_\lambda)_{\lambda \in \Lambda}$ is a system of parabolic molecules of order $(L, M, N_1, N_2) \in (\mathbb{Z}_+ \cup \{\infty\})^2 \times \mathbb{Z}_+^2$, if, for all $\lambda \in \Lambda$,

$$m_\lambda(x) = 2^{3s_\lambda/4} g^{(\lambda)}(A_{2^{s_\lambda}} R_{\theta_\lambda}(x - x_\lambda)), \quad \Phi_{\Lambda}(\lambda) = (s_\lambda, \theta_\lambda, x_\lambda),$$

such that, for all $|\beta| \leq L$,

$$\left| \partial^\beta \hat{g}^{(\lambda)}(\xi) \right| \lesssim \min \left(1, 2^{-s_\lambda} + |\xi_1| + 2^{-s_\lambda/2}|\xi_2| \right)^M \langle |\xi| \rangle^{-N_1} \langle |\xi_2| \rangle^{-N_2}.$$  

**Control Parameters:**
- $L$: Spatial localization.
- $M$: Number of directional (almost) vanishing moments.
- $N_1, N_2$: Smoothness of $m_\lambda$.  

Parabolic Molecules

Definition (Grohs, K; 2014):
Let \((\Lambda, \Phi_{\Lambda})\) be a parametrization. Then \((m_{\lambda})_{\lambda \in \Lambda}\) is a system of parabolic molecules of order \((L, M, N_1, N_2) \in (\mathbb{Z}_+ \cup \{\infty\})^2 \times \mathbb{Z}_+^2\), if, for all \(\lambda \in \Lambda\),

\[
m_{\lambda}(x) = 2^{3s_{\lambda}/4} \hat{g}^{(\lambda)}(A_{2^{s_{\lambda}}} R_{\theta_{\lambda}} (x - x_{\lambda})), \quad \Phi_{\Lambda}(\lambda) = (s_{\lambda}, \theta_{\lambda}, x_{\lambda}),
\]

such that, for all \(|\beta| \leq L\),

\[
\left| \partial^{\beta} \hat{g}^{(\lambda)}(\xi) \right| \lesssim \min \left(1, 2^{-s_{\lambda}} + |\xi_1| + 2^{-s_{\lambda}/2} |\xi_2| \right)^M \langle |\xi| \rangle^{-N_1} \langle |\xi_2| \rangle^{-N_2}.
\]

Illustration:
Special Cases

This framework includes...

- Parabolic Frame (Smith; 1998)
- Second Generation Curvelets (Candès and Donoho; 2002)
- Curvelet Molecules (Candès and Demanet; 2002)
- Bandlimited Shearlets (K and Labate; 2006)
- Frame Decompositions (Borup and Nielsen; 2007)
- Shearlet Molecules (Guo and Labate; 2008)
- Compactly Supported Shearlets (Kittipoom, K, and Lim; 2012)
- ...
What about Wavelets, Ridgelets,…?
Extension of Framework

Main Idea:

- Introduction of a parameter $\alpha \in [0, 1]$ to measure the amount of anisotropy.
- For $a > 0$, define

$$A_{\alpha, a} = \begin{pmatrix} a & 0 \\ 0 & a^\alpha \end{pmatrix}.$$ 

Illustration:

At $\alpha = 0$, Ridgelets; at $\alpha = \frac{1}{2}$, Curvelets/Shearlets; at $\alpha = 1$, Wavelets.
$\alpha$-Molecules

Definition (Grohs, Keiper, K, Schäfer; 2014):
Let $\alpha \in [0, 1]$, and let $(\Lambda, \Phi_\Lambda)$ be a parametrization. Then $(m_\lambda)_{\lambda \in \Lambda}$ is a system of $\alpha$-molecules of order $(L, M, N_1, N_2) \in (\mathbb{Z}_+ \cup \{\infty\})^2 \times \mathbb{Z}_+$, if, for all $\lambda \in \Lambda$,

$$m_\lambda(x) = s_\lambda^{(1+\alpha)/2} g^{(\lambda)}(A_\alpha, s_\lambda R_{\theta_\lambda} (x - x_\lambda)), \quad \Phi_\Lambda(\lambda) = (s_\lambda, \theta_\lambda, x_\lambda),$$

such that, for all $|\beta| \leq L$,

$$\left| \partial^\beta \hat{g}^{(\lambda)}(\xi) \right| \lesssim \min \left(1, s_\lambda^{-1} + |\xi_1| + s_\lambda^{-(1-\alpha)} |\xi_2| \right)^M \langle |\xi| \rangle^{-N_1} \langle \xi_2 \rangle^{-N_2}.$$
**α-Molecules**

**Definition (Grohs, Keiper, K, Schäfer; 2014):**

Let $\alpha \in [0, 1]$, and let $(\Lambda, \Phi_\Lambda)$ be a parametrization. Then $(m_\lambda)_{\lambda \in \Lambda}$ is a system of $\alpha$-molecules of order $(L, M, N_1, N_2) \in (\mathbb{Z}_+ \cup \{\infty\})^2 \times \mathbb{Z}_+$, if, for all $\lambda \in \Lambda$,

$$m_\lambda(x) = s^{(1+\alpha)/2}_\lambda g(\lambda) \left(A_{\alpha, s_\lambda} R_{\theta_\lambda} (x - x_\lambda)\right), \quad \Phi_\Lambda(\lambda) = (s_\lambda, \theta_\lambda, x_\lambda),$$

such that, for all $|\beta| \leq L$,

$$\left| \partial^\beta \hat{g}^{(\lambda)}(\xi) \right| \lesssim \min \left(1, s_\lambda^{-1} + |\xi_1| + s_\lambda^{-(1-\alpha)} |\xi_2| \right)^M \langle |\xi| \rangle^{-N_1} \langle \xi_2 \rangle^{-N_2}.$$

**Examples:**

- Wavelets ($\alpha = 1$)
- Ridgelets ($\alpha = 0$)
- Shearlets, parabolic molecules in general ($\alpha = \frac{1}{2}$)
- $\alpha$-Curvelets ($\alpha \in [0, 1]$)
**Metric Properties of Parametrizations**

**Definition:**
Let $\alpha \in [0, 1]$, and let $(\Lambda, \Phi_\Lambda)$ and $(\Delta, \Phi_\Delta)$ be parametrizations. For $\lambda \in \Lambda$ and $\mu \in \Delta$, we define the index distance by

$$
\omega_\alpha (\lambda, \mu) := \omega_\alpha (\Phi_\Lambda(\lambda), \Phi_\Delta(\mu)) := \max \left\{ \frac{s_\lambda}{s_\mu}, \frac{s_\mu}{s_\lambda} \right\} (1 + d_\alpha (\lambda, \mu)),
$$

with $d_\alpha (\lambda, \mu)$ defined by

$$
s_0^{2(1-\alpha)} |\theta_\lambda - \theta_\mu|^2 + s_0^{2\alpha} |x_\lambda - x_\mu|^2 + \frac{s_0^2}{1 + s_0^{2(1-\alpha)} |\theta_\lambda - \theta_\mu|^2} |\langle e_\lambda, x_\lambda - x_\mu \rangle|^2.
$$

where $s_0 = \min\{s_\lambda, s_\mu\}$ and $e_\lambda = (\cos(\theta_\lambda), -\sin(\theta_\lambda))^T$.

**Remark:** $d_{\frac{1}{2}}$ is the Hart Smith’s phase space metric on $\mathbb{T} \times \mathbb{R}^2$. 
Main Result: Decay of Cross-Grammian

Theorem (Grohs, Keiper, K, Schäfer; 2014):
Let $\alpha \in [0, 1]$, $N > 0$, and let $(m_\lambda)_{\lambda \in \Lambda}$, $(p_\mu)_{\mu \in \Delta}$ be systems of $\alpha$-molecules of order $(L, M, N_1, N_2)$ with

$$L \geq 2N, \quad M > 3N - \frac{3 - \alpha}{2}, \quad N_1 \geq N + \frac{1 + \alpha}{2}, \quad N_2 \geq 2N.$$

Then, for all $\lambda \in \Lambda$ and $\mu \in \Delta$,

$$|\langle m_\lambda, p_\mu \rangle| \lesssim \omega_\alpha (\lambda, \mu)^{-N}.$$
...towards Sparse Approximation Properties!
Strategy

α-Curvelets

Shearlets

Wavelets

Shearlet Molecules

Ridgelets

\[ \alpha = 0 \]

\[ \alpha = \frac{1}{2} \]

\[ \alpha = 1 \]
Sufficient Condition for Sparsity Equivalence

Definition:
Let $\alpha \in [0, 1]$ and $k > 0$. Two parametrizations $(\Lambda, \Phi_\Lambda)$ and $(\Delta, \Phi_\Delta)$ are $(\alpha, k)$-consistent, if

$$\sup_{\lambda \in \Lambda} \sum_{\mu \in \Delta} \omega_\alpha (\lambda, \mu)^{-k} < \infty \quad \text{and} \quad \sup_{\mu \in \Delta} \sum_{\lambda \in \Lambda} \omega_\alpha (\lambda, \mu)^{-k} < \infty.$$
Sufficient Condition for Sparsity Equivalence

Definition:
Let \( \alpha \in [0, 1] \) and \( k > 0 \). Two parametrizations \((\Lambda, \Phi_\Lambda)\) and \((\Delta, \Phi_\Delta)\) are \((\alpha, k)\)-consistent, if

\[
\sup_{\lambda \in \Lambda} \sum_{\mu \in \Delta} \omega_\alpha(\lambda, \mu)^{-k} < \infty \quad \text{and} \quad \sup_{\mu \in \Delta} \sum_{\lambda \in \Lambda} \omega_\alpha(\lambda, \mu)^{-k} < \infty.
\]

Theorem (Grohs, Keiper, K, Schäfer; 2014):
Let \( 0 < p \leq 1 \), and let \((m_\lambda)_{\lambda \in \Lambda}\) and \((p_\mu)_{\mu \in \Delta}\) be frames of \( \alpha \)-molecules of order \((L, M, N_1, N_2)\) with \((\alpha, k)\)-consistent parametrizations \((\Lambda, \Phi_\Lambda)\) and \((\Delta, \Phi_\Delta)\) for some \( k > 0 \). If

\[
L \geq 2 \frac{k}{p}, \quad M > 3 \frac{k}{p} - \frac{3 - \alpha}{2}, \quad N_1 \geq \frac{k}{p} + \frac{1 + \alpha}{2}, \quad N_2 \geq 2 \frac{k}{p},
\]

then

\[
\left\| \left( (m_\lambda, p_\mu) \right)_{\lambda \in \Lambda, \mu \in \Delta} \right\|_{\ell^p_{L} \to \ell^p_{M}} < \infty.
\]
Definition (K, Lemvig, Lim; 2012), (Keiper; 2012):
The set of cartoon-like functions $\mathcal{E}^\beta(\mathbb{R}^2)$, $\beta \in (1, 2]$ is defined by

$$
\mathcal{E}^\beta(\mathbb{R}^2) = \{ f \in L^2(\mathbb{R}^2) : f = f_0 + f_1 \cdot \chi_B \},
$$

where $B \subset [0, 1]^2$ with $\partial B$ a closed $C^\beta$-curve, $f_0, f_1 \in C^\beta_0([0, 1]^2)$. 
Generalized Image Model

Definition (K, Lemvig, Lim; 2012), (Keiper; 2012):
The set of cartoon-like functions $E^\beta(\mathbb{R}^2)$, $\beta \in (1, 2]$ is defined by

$$E^\beta(\mathbb{R}^2) = \{ f \in L^2(\mathbb{R}^2) : f = f_0 + f_1 \cdot \chi_B \},$$

where $B \subset [0, 1]^2$ with $\partial B$ a closed $C^\beta$-curve, $f_0, f_1 \in C^\beta_0([0, 1]^2)$.

Theorem (Grohs, Keiper, K, Schäfer; 2014):
Let $\alpha \in [\frac{1}{2}, 1)$, $\beta = \alpha^{-1}$. The Parseval frame of $\alpha$-curvelets provides an optimally sparse approximation of $f \in E^\beta(\mathbb{R}^2)$, i.e.,

$$\| f - f_N \|_2^2 \leq C \cdot N^{-\beta} \cdot (\log N)^{\beta+1}, \quad N \to \infty.$$
Sparse Approximation with $\alpha$-Molecules

Theorem (Grohs, Keiper, K, Schäfer; 2014):
Let $\alpha \in [\frac{1}{2}, 1)$, $\beta = \alpha^{-1}$, and let $(m_{\lambda})_{\lambda \in \Lambda}$ be a system of $\alpha$-molecules of order $(L, M, N_1, N_2)$ such that

(i) $(m_{\lambda})_{\lambda \in \Lambda}$ constitutes a frame for $L^2(\mathbb{R}^2)$,
(ii) $(\Lambda, \Phi_\Lambda)$ is $(\alpha, k)$-consistent with the parametrization of $\alpha$-curvelets for all $k > 0$,
(iii) it holds that

$$L \geq k(1+\beta), \quad M \geq \frac{3k}{2}(1+\beta)+\frac{\alpha - 3}{2}, \quad N_1 \geq \frac{k}{2}(1+\beta)+\frac{1+\alpha}{2}, \quad N_2 \geq k(1+\beta).$$

Then, for any $\varepsilon > 0$ and for any $f \in \mathcal{E}^{\beta}(\mathbb{R}^2)$, $(m_{\lambda})_{\lambda \in \Lambda}$ satisfies

$$\|f - f_N\|_2^2 \leq C \cdot N^{-\beta + \varepsilon}, \quad N \to \infty.$$
Let’s conclude...
Applied Harmonic Analysis provides various representation systems such as wavelets, ridgelets, curvelets, and shearlets.

Several systems are based on unitary representations of locally compact groups.

Shearlets are based on parabolic scaling and provide optimally sparse approximations for curvilinear features.

α-Molecules even include, for instance, wavelets and ridgelets.

Sparse approximation results can be derived in a unified manner.
THANK YOU!

References available at:

www.math.tu-berlin.de/~kutyniok

Code available at:

www.ShearLab.org

Related Books:

- Y. Eldar and G. Kutyniok
  Compressed Sensing: Theory and Applications

- G. Kutyniok and D. Labate
  Shearlets: Multiscale Analysis for Multivariate Data