On the Structure of Anisotropic Frames

P. Grohs
ETH Zurich, Seminar for Applied Mathematics

ESI Modern Methods of Time-Frequency Analysis
Motivation
Geometric Multiscale Analysis

Multiscale representation systems which are sensitive to anisotropic features and which provide sparse approximations thereof.

Such constructions are typically based on parabolic scaling and a directional transform such as rotation.
Curvelets
Curvelets
Curvelets
Curvelets
Curvelets
Curvelets
A Zoo of Transforms

- Labate, Lim, Kutyniok and Weiss (2005): Shearlets, where rotation is replaced by shearing, ensuring uniform treatment of continuous and digital transform. Initial construction bandlimited.
All these different systems possess optimally sparse approximation properties for functions exhibiting singularities on lower-dimensional manifolds. This has been shown separately for each system in a list of papers.

**Question:** Is it really necessary to go through these proofs for each single system, or might a higher-level viewpoint be useful?
Goal

Meta-Theorem

All frame systems based on parabolic scaling (specifically curvelets and shearlets) posses the exact same approximation properties, whenever the generating functions are sufficiently smooth, as well as localized in space and frequency.

Concrete results, i.e., questions of type: given $C^N$ basis function with $M$ directional vanishing moments, create frame from rotation/shearing + anisotropic dilation, what are the approximation properties?
Main Ideas

Introduce class of function systems \((m_\lambda)_{\lambda \in \Lambda} \subset L_2(\mathbb{R}^2)\) which includes all known systems based on parabolic scaling.

Show that any two function systems within a class possess equivalent approximation properties.
Previous Work

- Curvelet/Shearlet molecules (Candès-Demanet 2004, Guo-Labate 2008)
  - Unification of curvelet/shearlet-type systems based on rotation/shearing
  - Does not include shearlet/curvelet-based methods
  - Nonquantitative – only basis functions with infinitely many vanishing moments and superpolynomial decay in space \( \sim \) not valid in practice

- Coorbit molecules (Dahlke-Steidl-Teschke 2008-2012, Gröchenig-Piotrowski 2008), see Gabi Steidl’s talk Wed 10:00-11:00.
  - Powerful framework for systems built from group representation
  - Does not include cone adapted shearlets or curvelet-type constructions
  - Abstract concept
Coorbit Molecules

Group $\mathcal{G}$, representation $\pi$, window $g$, “well-spread” point-set $(x_\lambda)_{\lambda \in \Lambda} \subset \mathcal{G}$, “envelope function” $H$. A system $(m_\lambda)_{\lambda \in \Lambda} \subset L^2(\mathbb{R}^d)$ is a system of coorbit molecules if

$$|\langle m_\lambda, \pi(z)g \rangle| \lesssim H(z^{-1}x_\lambda)$$  \hspace{1cm} (1)

Can be defined for (non cone-adapted) shearlets, but:

- What is $H$??
- How to apply to cone-adapted shearlet or curvelet-type systems??
- How to verify (1) for concrete system??

Need more concrete approach in terms of analytical properties of the family $(m_\lambda)_{\lambda \in \Lambda}$ and more general approach in terms of the types of directional transforms allowed.
Parabolic Molecules
Basic Structure

System \((m_\lambda)_{\lambda \in \Lambda}\) where every \(m_\lambda\) is associated with a scale, a direction and a location.
Parametrization

“Phase Space”

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

**Definition**

A *parametrization* consists of an index set \( \Lambda \) and a mapping \( \Phi_\Lambda : \Lambda \to \mathbb{P}, \; \Phi_\Lambda(\lambda) := (s_\lambda, \theta_\lambda, x_\lambda). \)

*Canonical parametrization:*

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

\[ \Phi^0((j, l, k)) := (j, \pi l 2^{-\lfloor j/2 \rfloor}, R_{-\theta_\lambda} D_{2^{-s_\lambda}} k) \]

\[ R_\theta := \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix} \quad D_a := \begin{pmatrix} a & 0 \\ 0 & \sqrt{a} \end{pmatrix}. \]
Canonical Parametrization

Figure: Left: frequency tiling associated with canonical parametrization. Right: Translational grids.
Definition of Parabolic Molecules

Definition

Let $(\Lambda, \Phi^\Lambda)$ be a parametrization. A family $(m_\lambda)_{\lambda \in \Lambda}$ is called a family of parabolic molecules of order $(R, M, N_1, N_2)$ if it can be written as

$$m_\lambda(x) = 2^{3s_\lambda/4} a^{(\lambda)} (D_2^{s_\lambda} R^{\theta_\lambda} (x - x_\lambda))$$

such that

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

for all $|\beta| \leq R$. The implicit constants are uniform over $\lambda \in \Lambda$. 
A Closer Look

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

for all \(||\beta|| \leq R\) means that \(a^{(\lambda)}\) ...

... possesses \(M\) (almost) vanishing moments in the \(x_1\)-direction

... is of smoothness \(N_1 + N_2\) in \(x_2\) and \(N_1\) in \(x_1\)

... decays like \(\langle x \rangle^{-R}\) in space.
Frequency Localization

Figure: Left: The weight function
\[
\min \left( 1, 2^{-s \lambda} + |\xi_1| + 2^{-s \lambda}/|\xi_2| \right)^M \langle |\xi| \rangle^{-N_1} \langle \xi_2 \rangle^{-N_2} \text{ for } s \lambda = 3, M = 3, N_1 = N_2 = 2. 
\]

Right: Approximate Frequency support of a corresponding molecule \( \hat{m}_\lambda \) with \( \theta_\lambda = \pi/4 \).
Second Generation Curvelets

Pick window functions $W(r), V(t)$ which are both real, nonnegative, $C^\infty$ s.t.

$$\text{supp } W \subseteq \left[\frac{1}{2}, 2\right], \quad \text{supp } V \subseteq [-1, 1]$$

and

$$\sum_{j \in \mathbb{Z}} W(2^{-j}r)^2 = 1, \quad \sum_{l \in \mathbb{Z}} V(t - l)^2 = 1.$$

Define in polar coordinates

$$\hat{\gamma}(j, 0, 0)(r, \omega) := 2^{-3j/4} W(2^{-j}r) V\left(2^{\lfloor j/2 \rfloor} \omega\right),$$

$$\gamma(j, l, k)(\cdot) := \gamma(j, 0, 0) \left(R_{\theta(j, l, k)} \left(\cdot - x(j, l, k)\right)\right),$$

With appropriate modifications for $j = 0$ define the second-generation curvelet frame

$$\Gamma^0 := \{\gamma_\lambda : \lambda \in \Lambda^0\}.$$
No big surprise:

Lemma

\( \Gamma^0 \) constitutes a system of parabolic molecules of arbitrary order.
Similar Systems

- Hart Smith’s parabolic frame
- Borup and Nielsen’s construction
- Curvelet molecules

All based on rotation. How about shearing?
Shearlets

Index set

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

and the shearlet system

\[ \Sigma := \left\{ \sigma_\lambda : \lambda \in \Lambda^\sigma \right\}, \]

with

\[ \sigma_{(\varepsilon,0,0,k)}(\cdot) = \varphi(\cdot - k), \quad \sigma_{(\varepsilon,j,l,k)}(\cdot) = 2^{3j/4} \psi_{j,l,k}^\varepsilon \left( D_{2j}^{\varepsilon} S_{l,j} \cdot - k \right), \quad j \geq 1, \]

where \( D_a^0 = D_a, \ D_a^1 := \text{diag}(\sqrt{a}, a), \ S_{l,j} := \begin{pmatrix} 1 & l2^{-\left\lfloor j/2 \right\rfloor} \\ 0 & 1 \end{pmatrix} \) and

\[ S_{l,j}^1 = \left( S_{l,j}^0 \right)^\top. \]
Shearlets

**Definition**

We call $\Sigma$ a system of *shearlet molecules* of order $(R, M, N_1, N_2)$ if the functions $\varphi, \psi_0^{j,l,k}, \psi_1^{j,l,k}$ satisfy

$$
|\partial^\beta \hat{\psi}_{j,l,k}(\xi_1, \xi_2)| \lesssim \min \left( 1, 2^{-s\lambda} + |\xi_{1+\varepsilon}| + 2^{s\lambda}/2 |\xi_{2-\varepsilon}| \right)^M \langle |\xi| \rangle^{-N_1} \langle \xi_{2-\varepsilon} \rangle^{-N_2}
$$

for every $\beta \in \mathbb{N}^2$ with $|\beta| \leq R$.

Essentially same definition as curvelets, only with rotation replaced by shearing.
Shearlet Molecules are Parabolic Molecules

Define *shearlet parametrization*

\[ \Phi^\sigma(\lambda) = (s_\lambda, \theta_\lambda, x_\lambda) := \left( j, \varepsilon \pi / 2 + \arctan(-12^{-\lfloor j/2 \rfloor}), (S_1^\varepsilon)^{-1} D_2^{\varepsilon-j} k \right), \]

\[ \lambda \in \Lambda^\sigma. \]

**Lemma (G-Kutyniok (2012))**

*Assume that the shearlet system \( \Sigma \) constitutes a system of shearlet molecules of order \((R, M, N_1, N_2)\). Then \( \Sigma \) constitutes a system of parabolic molecules of order \((R, M, N_1, N_2)\), associated to the parametrization \((\Lambda^\sigma, \Phi^\sigma)\).*
Shearlets

This implies that the following systems constitute systems of parabolic molecules:

- bandlimited shearlets
- compactly supported shearlets
- shearlet molecules as introduced by Guo and Labate (2008).
Theorem (G-Kutyniok (2012))

Known systems to date based on parabolic scaling are parabolic molecules (in particular curvelets and bandlimited and compactly supported shearlets).
Almost Orthogonality
Index Distance Function

Definition

For two points \((s_\lambda, \theta_\lambda, x_\lambda), (s_\mu, \theta_\mu, x_\mu) \in \mathbb{P}\) define the \textit{parabolic distance}

\[
\omega(\lambda, \mu) := 2 |s_\lambda - s_\mu| \left( 1 + 2^{s_{\lambda_0}} d(\lambda, \mu) \right),
\]

and

\[
d(\lambda, \mu) := |\theta_\lambda - \theta_\mu|^2 + |x_\lambda - x_\mu|^2 + |\langle e_\lambda, x_\lambda - x_\mu \rangle|.
\]

where \(\lambda_0 = \text{argmin}(s_\lambda, s_\mu)\) and \(e_\lambda = (\cos(\theta_\lambda), \sin(\theta_\lambda))^\top\).

Almost Orthogonality

Theorem (G-Kutyniok (2012))

Let \((m_\lambda)_{\lambda \in \Lambda}, (p_\mu)_{\mu \in M}\) be two systems of parabolic molecules of order \((R, M, N_1, N_2)\) with

\[
R \geq 2N, \quad M > 4N - \frac{5}{4}, \quad N_1 \geq 2N + \frac{3}{4}, \quad N_2 \geq 2N.
\]

Then

\[
|\langle m_\lambda, p_\mu \rangle| \lesssim \omega \left((s_\lambda, \theta_\lambda, x_\lambda), (s_\mu, \theta_\mu, x_\mu)\right)^{-N}.
\]


Proof by “hard analysis”.
Applications
1. Sparsity Equivalence
**Sparsity Equivalence**

**Definition**

Let \((m_\lambda)_{\lambda \in \Lambda}\) and \((p_\mu)_{\mu \in M}\) be frame systems, and let \(0 < p \leq 1\). Then \((m_\lambda)_{\lambda \in \Lambda}\) and \((p_\mu)_{\mu \in M}\) are sparsity equivalent in \(\ell_p\), if

\[
\left\| \left( \langle m_\lambda, p_\mu \rangle \right)_{\lambda \in \Lambda, \mu \in \Lambda^0} \right\|_{\ell_p \to \ell_p} < \infty.
\]

**Theorem (G-Kutyniok (2012))**

Assume that \((m_\lambda)_{\lambda \in \Lambda}\) is a system of parabolic molecules associated with \(\Lambda\) of order \((R, M, N_1, N_2)\) such that

\[
R \geq 2 \frac{2}{p}, \quad M > 4 \frac{2}{p} - \frac{5}{4}, \quad N_1 \geq 2 \frac{2}{p} + \frac{3}{4}, \quad N_2 \geq 2 \frac{2}{p}.
\]

Then \((m_\lambda)_{\lambda \in \Lambda}\) is sparsity equivalent to \(\Gamma^0\) in \(\ell_p\).
Cartoon Approximation

**Theorem (G-Kutyniok (2012))**

Assume that \((m_\lambda)_{\lambda \in \Lambda}\) is a system of parabolic molecules of order \((R, M, N_1, N_2)\) such that

(i) \((m_\lambda)_{\lambda \in \Lambda}\) constitutes a frame for \(L^2(\mathbb{R}^2)\),

(ii) \(\Lambda\) is \(k\)-admissible for all \(k > 2\),

(iii) it holds that

\[
R \geq 6, \quad M > 12 - \frac{5}{4}, \quad N_1 \geq 6 + \frac{3}{4}, \quad N_2 \geq 6.
\]

Then the frame \((m_\lambda)_{\lambda \in \Lambda}\) possesses an almost best \(N\)-term approximation rate of order \(N^{-1+\varepsilon}\), \(\varepsilon > 0\) arbitrary for cartoon images.
2. Function Spaces
Function Spaces

With the curvelet frame

\[ \Gamma^0 := \{ \gamma_{j,l,k} : (j, l, k) \in \Lambda^0 \} \]

introduced above, following Borup-Nielsen (2007), define for \( p, q, \alpha > 1 \) the function spaces \( G^\alpha_{p,q} \) given by the norm

\[
\| f \|_{G^\alpha_{p,q}} := \left( \sum_{j \geq 0, l} \left( 2^{\alpha j} \left( \sum_k \left| \langle f, \gamma_{j,l,k} \rangle \right|^p \right)^{1/p} \right)^q \right)^{1/q}.
\]  \( (2) \)
Equivalence Result

Theorem (G-Kutyniok (2012))

Let $\Sigma = \{\sigma_\lambda : \lambda \in \Lambda\}$ be a frame for $L^2(\mathbb{R}^2)$ with dual frame $\tilde{\Sigma} = \{\tilde{\sigma}_\lambda : \lambda \in \Lambda\}$. Assume further that $\Sigma$, $\tilde{\Sigma}$ are both parabolic molecules of arbitrary order. Then

$$\|f\|_{G^\alpha_{p,q}} \sim \left(\sum_{j \geq 0} \left(2^{\alpha j} \left(\sum_{\lambda \in \Lambda_j} |\langle f, \sigma_\lambda \rangle|^p\right)^{1/p}\right)^q\right)^{1/q} \sim \left(\sum_{j \geq 0} \left(2^{\alpha j} \left(\sum_{\lambda \in \Lambda_j} |\langle f, \tilde{\sigma}_\lambda \rangle|^p\right)^{1/p}\right)^q\right)^{1/q}.$$

Can be made quantitative.
Localization

In general it is hard to establish that dual frame consists of parabolic molecules, but it is possible to show localization results for dual frames:

**Theorem (G (2012, ACHA))**

Let $\Sigma = \{\sigma_\lambda : \lambda \in \Lambda\}$ be a frame for $L_2(\mathbb{R}^2)$ with dual frame $\tilde{\Sigma} = \{\tilde{\sigma}_\lambda : \lambda \in \Lambda\}$. Assume that

$$|\langle \sigma_\lambda, \sigma_\mu \rangle| \leq C \omega(\lambda, \mu)^{-N}.$$ 

Then there exists $N^+$ depending (in an explicit way) on $N$, $C$ such that

$$|\langle \tilde{\sigma}_\lambda, \tilde{\sigma}_\mu \rangle| \lesssim \omega(\lambda, \mu)^{-N^+}.$$
Summary

- Thorough understanding of ingredients of representation systems needed for sparse approximation
- General framework where results can directly be transferred without giving proofs for every single system
- Guideline in designing new representation systems
- Further work (jointly with G. Kutyniok, E. King): Continuous parameters, further properties (separation, Radon inversion, ...), higher dimensions, ...
- How can group-property be relaxed in coorbit theory?


The End

Questions?
Definition

\( \Lambda \) is \( k \)-admissible if

\[
\sup_{\mu \in \Lambda^0} \sum_{\lambda \in \Lambda} \omega(\lambda, \mu)^{-k} < \infty.
\]
$G$ locally compact group.

$\pi : G \to U(\mathcal{H})$, the unitary transforms of a Hilbert space $\mathcal{H}$, irreducible and square-integrable: $\exists \psi \in \mathcal{H}$ s.t.

$$\int_{G} |\langle \pi(z)\psi, \psi \rangle|^2 \, d\mu_L(z) < \infty.$$ 

Window

$g \neq 0 \in \mathcal{B}_w := \{ h \in \mathcal{H} : \langle h, \pi(z)h \rangle \in \mathcal{W}^L(L_\infty, L_1, w(G)) \}$. Points $(x_\lambda)_{\lambda \in \Lambda}$ are $U$-dense ($\bigcup_{\lambda \in \Lambda} x_\lambda U = G$) and relatively separated.

Envelope $H \in \mathcal{W}^R(L_\infty, L_1, w(G))$. Return to talk.
$w$ submultiplicative weight: $w(xy) \leq w(x)w(y)$. Then

$$\mathcal{W}^L(L_\infty, L_1, w(G)) := \left\{ f \in L_\infty,_{loc} : \sup_{y \in xQ} |f(y)| \in L_1, w(G) \right\},$$

where $Q \subset G$ is a relatively compact neighborhood of the identity.
$w$ submultiplicative weight: $w(xy) \leq w(x)w(y)$. Then

$$\mathcal{W}^R(L_\infty, L_1, w(G)) := \left\{ f \in L_\infty,\text{loc} : \sup_{y \in Qx} |f(y)| \in L_1, w(G) \right\},$$

where $Q \subset G$ is a relatively compact neighborhood of the identity.
(x_\lambda)_{\lambda \in \Lambda} \text{ separated if } x_\lambda Q \cap x_{\lambda'} Q = \emptyset \text{ for all } \lambda \neq \lambda' \text{ and some compact neighborhood } Q \subset G \text{ of the identity. Relatively separated if finite union of separated sets.}
Banach Algebra Viewpoint

Definition

Define for $N \in \mathbb{N}$ the Banach Algebra

$$\mathcal{B}_N := \{ A : l_2(\Lambda) \rightarrow l_2(\Lambda) : |a_{\lambda,\lambda'}| \lesssim \omega(\lambda, \lambda')^{-N} \text{ for all } \lambda, \lambda' \in \Lambda \}$$

with norm

$$\|A\|_{\mathcal{B}_N} := \inf \{ C_0 : |a_{\lambda,\lambda'}| \leq C_0 \omega(\lambda, \lambda')^{-N} \text{ for all } \lambda, \lambda' \in \Lambda \}.$$
Moore-Penrose Pseudoinverse

**Definition**

Assume that $\mathbf{A}$ is a symmetric matrix with a spectral gap, i.e.

$$\sigma_2(\mathbf{A}) \subset 0 \cup [A, B] \quad 0 < A < B.$$  

Its Moore-Penrose pseudoinverse $\mathbf{A}^+$ is defined by

$$\mathbf{A}^+|_{\ker\mathbf{A}} = 0 \quad \mathbf{A}^+|_{\mathrm{ran}\mathbf{A}} \mathbf{A}|_{\mathrm{ran}\mathbf{A}} = \mathbf{I}|_{\mathrm{ran}\mathbf{A}}.$$  

**Lemma**

Consider a frame $\Sigma = (\sigma_\lambda)_{\lambda \in \Lambda}$ for a Hilbert space $\mathcal{H}$. Then the Gram matrix $(\Sigma, \Sigma)_{\mathcal{H}}$ is symmetric and possesses a spectral gap ($A, B$ being the frame constants). Its Moore-Penrose pseudoinverse satisfies

$$\mathbf{A}^+ = \left(\tilde{\Sigma}, \tilde{\Sigma}\right).$$
Tool I: Landweber Iteration

Lemma

The matrix $A^+$ can be computed via a Landweber-type iteration by the formula

$$A^+ = \beta \sum_{k \in \mathbb{N}} \left( I - \beta A^2 \right)^k A,$$

where $\beta = \frac{2}{A^2 + B^2}$. Furthermore, we have that

$$\sigma_2 \left( \left( I - \beta A^2 \right)^k A \right) \subseteq [-Br^k, Br^k],$$

with

$$r := \frac{B^2 - A^2}{A^2 + B^2} < 1.$$
**Tool II: Boundedness of Multiplication**

**Lemma**

Assume that $A \in \mathcal{B}_{N+L}$ with $L$ large enough (but completely explicit). Then we have for all $B \in \mathcal{B}_N$ such that $AB$ is symmetric the estimate

$$\|AB\|_{\mathcal{B}_N} \leq C \|A\|_{\mathcal{B}_{N+L}} \|B\|_{\mathcal{B}_N}.$$

Argument applies to very general classes of $\omega$ (also used e.g. for wavelets)

Proof is very complicated...
Main Result

Theorem

Assume $A \in \mathcal{B}_{N+L}$ as before. Then $A^+ \in \mathcal{B}_{N^+}$, where

$$N^+ = N \left( 1 - \frac{\log \left( 1 + \beta \|A\|_{\mathcal{B}_{N+L}}^2 C^2 \right)}{\log(r)} \right)^{-1}$$
Proof

We use the Landweber-type iteration and write

\[ A^+ = \beta \sum_{k \in \mathbb{N}} A^{(k)}, \]

where

\[ A^{(k)} := (I - \beta A^2)^k A. \]

Use two different estimates for the entries

\[ a_{\lambda, \lambda'}^{(k)}, \]

of \( A. \)
First Estimate

We claim that

$$\|A^{(k)}\|_{B_N} \leq (1 + \beta \|A\|_{B_{N+L}}^2 C^2)^k \|A\|_{B_N}$$

Proof by induction.

Obviously true for $k = 0$. Assume $k - 1$. Then

$$A^{(k)} = (1 - \beta A^2) A^{(k-1)}.$$

By our multiplication theorem we have

$$\|AA^{(k-1)}\|_{B_N} \leq C \|A\|_{B_{N+L}} \|A^{(k-1)}\|_{B_N}$$

and, using the same argument

$$\|\beta A^2 A^{(k-1)}\|_{B_N} \leq \beta C^2 \|A\|_{B_{N+L}}^2 \|A^{(k-1)}\|_{B_N}.$$

Using the triangle inequality yields the claim.
Second Estimate

We claim that

\[ |a^{(k)}_{\lambda, \lambda'}| \lesssim r^k. \]

To see this, write

\[ |a^{(k)}_{\lambda, \lambda'}| = |(Ae_\lambda, e_{\lambda'})| \]

use Cauchy-Schwartz and spectral properties to obtain

\[ \cdots \leq \|A^{(k)}e_\lambda\| \|e'_{\lambda}\| \leq \|A^{(k)}\| \leq Br^k. \]
Summing Up

We have

\[ a_{\lambda,\lambda'}^+ = \sum_{k=0}^{\infty} a^{(k)}_{\lambda,\lambda'} . \]

Use above two estimates to get for any \( k_0 \)

\[ |a_{\lambda,\lambda'}^+| \lesssim \sum_{k=0}^{k_0} (1 + \beta \|A\|_{B_{N+L}} C^2)^k \omega(\lambda, \lambda')^{-N} + \sum_{k=k_0+1}^{\infty} r^k . \]

Geometric summation gives

\[ |a_{\lambda,\lambda'}^+| \lesssim (1 + \beta \|A\|_{B_{N+L}} C^2)^{k_0} \omega(\lambda, \lambda')^{-N} + r^{k_0} . \]
Final Step

Put $D := 1 + \beta \|A\|_{B_{N+L}} C^2$.

Want to find $k_0, N^+$ such that both $D^{k_0} \omega(\lambda, \lambda')^{-N}$ and $r^{k_0}$ can be bounded by $\omega(\lambda, \lambda')^{-N^+}$.

Estimate for $r^{k_0}$ yields

$$k_0 = -N^+ \frac{\log(\omega(\lambda, \lambda'))}{\log(r)}$$

Estimate for $D^{k_0} \omega(\lambda, \lambda')$ yields

$$k_0 = (N - N^+) \frac{\log(\omega(\lambda, \lambda'))}{\log(D)}.$$ 

Eliminating $k_0$ yields

$$N^+ = N(1 - \frac{\log(D)}{\log(r)})^{-1}$$

This proves the theorem.
Frame Result

Using the fact that the Gramian of the dual frame is the Moore-Penrose pseudoinverse of the Gramian of the primal frame we have shown the following:

**Theorem (G (2012))**

Let \( \Sigma = \{ \sigma_\lambda : \lambda \in \Lambda \} \) be a frame for \( L_2(\mathbb{R}^2) \) with dual frame \( \tilde{\Sigma} = \{ \tilde{\sigma}_\lambda : \lambda \in \Lambda \} \). Assume that

\[
|\langle \sigma_\lambda, \sigma_\mu \rangle| \leq C \omega(\lambda, \mu)^{-N}.
\]

Then there exists \( N^+ \) depending (in an explicit way) on \( N, C \) such that

\[
|\langle \tilde{\sigma}_\lambda, \tilde{\sigma}_\mu \rangle| \lesssim \omega(\lambda, \mu)^{-N^+}.
\]
Remarks

- $N^+$ depends on $N$, $\|A\|_{\mathcal{B}_{N+L}}$ and the spectrum of $A$.
- Dependence on $\|A\|_{\mathcal{B}_{N+L}}$ is annoying.
- Similar results exist for wavelet Riesz bases. Our results are a generalization which applies also to anisotropic frames.
- Localization is essential in the study of function spaces but also in numerical purposes. For instance, localization ensures fast matrix-vector multiplication and is also essential in the study of adaptive methods to solve PDEs.