



Linear Filtering

Spectral Theory

Nonlinear singular vectors

1-homogeneous functions

Existence questions

Error estimates

Multiscale Methods

Variational methods

Gradient flow

Inverse scale space flow

Chapter 2

Multiscale Methods

Nonlinear Multiscale Methods for Image and Signal Analysis
SS 2015

Michael Moeller
Computer Vision
Department of Computer Science
TU München



Linear Filtering

Spectral Theory

- Nonlinear singular vectors
- 1-homogeneous functions
- Existence questions
- Error estimates

Multiscale Methods

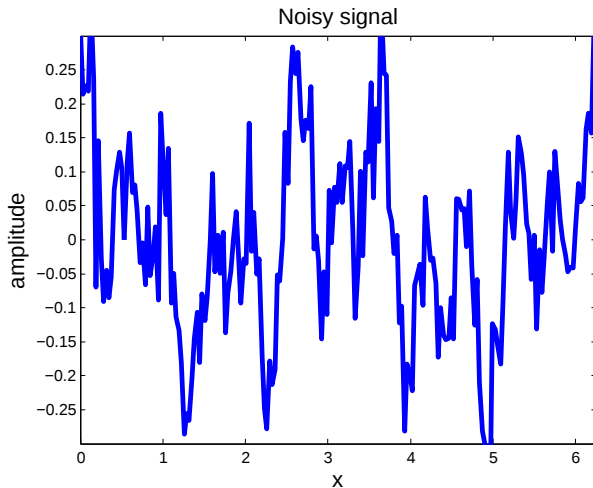
- Variational methods
- Gradient flow
- Inverse scale space flow

Linear image and signal filtering

Linear signal denoising

Multiscale Methods

Michael Moeller



Linear Filtering

Spectral Theory

- Nonlinear singular vectors
- 1-homogeneous functions
- Existence questions
- Error estimates

Multiscale Methods

- Variational methods
- Gradient flow
- Inverse scale space flow

Linear signal denoising

Multiscale Methods

Michael Moeller



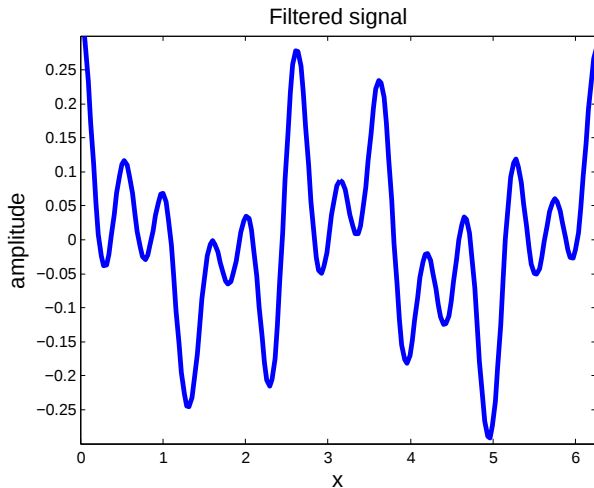
Linear Filtering

Spectral Theory

- Nonlinear singular vectors
- 1-homogeneous functions
- Existence questions
- Error estimates

Multiscale Methods

- Variational methods
- Gradient flow
- Inverse scale space flow



Linear image inpainting



Linear Filtering

Spectral Theory

- Nonlinear singular vectors
- 1-homogeneous functions
- Existence questions
- Error estimates

Multiscale Methods

- Variational methods
- Gradient flow
- Inverse scale space flow



Linear image inpainting

Multiscale Methods

Michael Moeller



Linear Filtering

Spectral Theory

- Nonlinear singular vectors
- 1-homogeneous functions
- Existence questions
- Error estimates

Multiscale Methods

- Variational methods
- Gradient flow
- Inverse scale space flow



Linear image deblurring



Multiscale Methods

Michael Moeller



Linear Filtering

Spectral Theory

- Nonlinear singular vectors
- 1-homogeneous functions
- Existence questions
- Error estimates

Multiscale Methods

- Variational methods
- Gradient flow
- Inverse scale space flow

Linear image deblurring



Multiscale Methods

Michael Moeller



Linear Filtering

Spectral Theory

- Nonlinear singular vectors
- 1-homogeneous functions
- Existence questions
- Error estimates

Multiscale Methods

- Variational methods
- Gradient flow
- Inverse scale space flow

Linear image denoising



Multiscale Methods

Michael Moeller



Linear Filtering

Spectral Theory

- Nonlinear singular vectors
- 1-homogeneous functions
- Existence questions
- Error estimates

Multiscale Methods

- Variational methods
- Gradient flow
- Inverse scale space flow

Linear image denoising



Multiscale Methods

Michael Moeller



Linear Filtering

Spectral Theory

- Nonlinear singular vectors
- 1-homogeneous functions
- Existence questions
- Error estimates

Multiscale Methods

- Variational methods
- Gradient flow
- Inverse scale space flow

Linear image sharpening



Multiscale Methods

Michael Moeller



Linear Filtering

Spectral Theory

- Nonlinear singular vectors
- 1-homogeneous functions
- Existence questions
- Error estimates

Multiscale Methods

- Variational methods
- Gradient flow
- Inverse scale space flow

Linear image sharpening



Multiscale Methods

Michael Moeller



Linear Filtering

Spectral Theory

- Nonlinear singular vectors
- 1-homogeneous functions
- Existence questions
- Error estimates

Multiscale Methods

- Variational methods
- Gradient flow
- Inverse scale space flow



How can we understand the behavior of linear filters?

Consider for instance the simple linear sharpening

$$\hat{u} = \text{imfilter}(f, k) = k * f$$

with a kernel

$$k = \text{fspecial}('unsharp') = \begin{bmatrix} -0.1667 & -0.6667 & -0.1667 \\ -0.6667 & 4.3333 & -0.6667 \\ -0.1667 & -0.6667 & -0.1667 \end{bmatrix}.$$

Remember the Convolution Theorem:

$$\hat{u} = k * f \Rightarrow \mathcal{F}(\hat{u}) = \mathcal{F}(k)\mathcal{F}(f)$$

Linear image and signal filtering



Linear Filtering

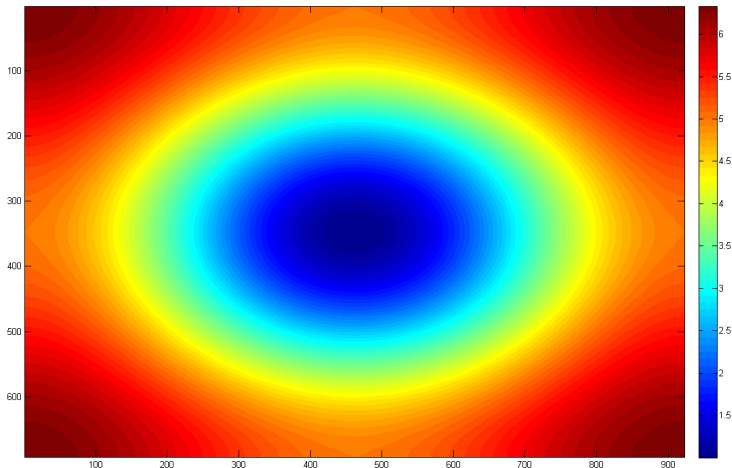
Spectral Theory

- Nonlinear singular vectors
- 1-homogeneous functions
- Existence questions
- Error estimates

Multiscale Methods

- Variational methods
- Gradient flow
- Inverse scale space flow

Absolute values of $\mathcal{F}(k)$.



Middle is 1 and corresponds to the lowest frequency



Linear Filtering

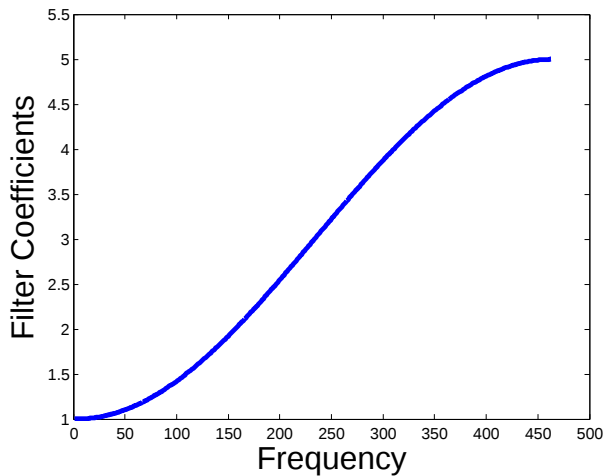
Spectral Theory

- Nonlinear singular vectors
- 1-homogeneous functions
- Existence questions
- Error estimates

Multiscale Methods

- Variational methods
- Gradient flow
- Inverse scale space flow

Or in 1d:





This representation is very intuitive for us, since we have an understanding of frequencies and can look at filters.

But what does it mean mathematically?

What does $\mathcal{F}(\hat{u}) = \mathcal{F}(k)\mathcal{F}(f)$ do?

Pointwise (or componentwise) multiplication
 $\rightarrow \mathcal{F}(k)$ is diagonal!



Let us go back to linear algebra:

Consider

$$\hat{u} = Af$$

for some symmetric positive definite matrix $A \in \mathbb{R}^{n \times n}$.

Note that any linear operator can be written in this form!

There exists an orthonormal basis $\{v_1, \dots, v_n\}$ of eigenvectors of A with eigenvalues $\{\lambda_1, \dots, \lambda_n\}$:

$$Av_j = \lambda_j v_j$$

Linear Filtering

Spectral Theory

- Nonlinear singular vectors
- 1-homogeneous functions
- Existence questions
- Error estimates

Multiscale Methods

- Variational methods
- Gradient flow
- Inverse scale space flow



We write

$$f = \sum_i a_i v_i.$$

Now

$$\hat{u} = Af = \sum_i a_i Av_i = \sum_i \lambda_i a_i v_i.$$

Let us represent $\hat{u}$ in the eigenbasis of A and denote its coefficients by b_i . Then

$$\begin{pmatrix} b_1 \\ \cdot \\ \cdot \\ b_n \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & 0 & \lambda_n \end{pmatrix} \begin{pmatrix} a_1 \\ \cdot \\ \cdot \\ a_n \end{pmatrix}$$

→ We have diagonalized A and you know this since > 3 years.



Linear Filtering

Spectral Theory

Nonlinear singular vectors
1-homogeneous functions
Existence questions
Error estimates

Multiscale Methods

Variational methods
Gradient flow
Inverse scale space flow

$$\begin{pmatrix} b_1 \\ \cdot \\ \cdot \\ b_n \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & 0 & \lambda_n \end{pmatrix} \begin{pmatrix} a_1 \\ \cdot \\ \cdot \\ a_n \end{pmatrix}$$

Engineering interpretation:

- λ_i is the filter coefficients for the i -th *frequency*.
- $\lambda_i > 1$ means boosting the *frequency*, $\lambda_i < 1$ means damping the frequency.
- The interpretation of the *frequency* is given by the eigenvector v_i .
- Any convolution diagonalizes under sin/cos, which yields a classical frequency.
- Other linear operators lead to other meanings of frequencies.



Variational methods can be linear, too...

$$\hat{u} = \arg \min_u \frac{1}{2} \|u - f\|_2^2 + \alpha \|\nabla u\|_2^2. \quad (1)$$

Optimality at

$$0 = \hat{u} - f - \alpha \Delta \hat{u},$$

or

$$\hat{u} = (I - \alpha \Delta)^{-1} f.$$

- Depends linearly on f .
- Also diagonalizes via FFT.
- Variational method (1) is nothing but a special frequency filter...
- ... and does not work very well.

Linear Filtering

Spectral Theory

Nonlinear singular vectors
1-homogeneous functions
Existence questions
Error estimates

Multiscale Methods

Variational methods
Gradient flow
Inverse scale space flow



Now consider

$$\hat{u} = \arg \min_u \frac{1}{2} \|u - f\|_2^2 + \alpha \|\nabla u\|_1, \quad (2)$$

which is highly nonlinear.

Absolutely no eigenvector theory!

Or is there?

Linear Filtering

Spectral Theory

- Nonlinear singular vectors
- 1-homogeneous functions
- Existence questions
- Error estimates

Multiscale Methods

- Variational methods
- Gradient flow
- Inverse scale space flow



Linear Filtering

Spectral Theory

- Nonlinear singular vectors
- 1-homogeneous functions
- Existence questions
- Error estimates

Multiscale Methods

- Variational methods
- Gradient flow
- Inverse scale space flow

Nonlinear Spectral Theory¹

¹Largely based on: M. Benning and M. Burger, *Ground States and Singular Vectors of Convex Variational Regularization Methods*, 2013



Let us start with the (general) previous observation that

$$\hat{u} = \arg \min_u \frac{1}{2} \|u - f\|_2^2 + \frac{\alpha}{2} \|Ku\|_2^2,$$

leads to

$$(I + \alpha K'K)\hat{u} = f$$

such that the singular vectors v of the above problem are the eigenvectors of the symmetric, positive semi-definite matrix $K'K$, i.e. there exist $v_\lambda \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$ such that

$$\lambda v_\lambda = K'K v_\lambda$$

or

$$\lambda v_\lambda \in \partial J(v_\lambda)$$

for $J(u) = \frac{1}{2} \|Ku\|^2$.



The variational model

$$\hat{u} = \arg \min_u \frac{1}{2} \|u - f\|_2^2 + \frac{\alpha}{2} \|Ku\|_2^2,$$

leads to the singular vector description that there exist $v_\lambda \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$ such that

$$\lambda v_\lambda \in \partial J(v_\lambda).$$

The latter makes sense for any convex regularization!

Can we study general J , e.g. TV regularization?



Notation

From now on we denote the set of all proper, convex, lower semi-continuous functions $J : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{\infty\}$ by $\Gamma_0(\mathbb{R}^n)$.

Definition: One-homogeneous

We call $J \in \Gamma_0(\mathbb{R}^n)$ (absolutely) 1-homogeneous, if

$$J(\lambda u) = |\lambda| J(u)$$

holds for all $\lambda \in \mathbb{R}$.

Example: $J(u) = \|Ku\|$ is one-homogeneous for any norm.

[Linear Filtering](#)[Spectral Theory](#)

Nonlinear singular vectors

1-homogeneous functions

Existence questions

Error estimates

[Multiscale Methods](#)

Variational methods

Gradient flow

Inverse scale space flow



Triangle inequality of 1-homogeneous functions

Let $J \in \Gamma_0(\mathbb{R}^n)$ be 1-hom. Then J meets the triangle inequality.

Domain of 1-homogeneous functions

Let $J \in \Gamma_0(\mathbb{R}^n)$ be 1-homogeneous. Then $\text{dom}(J)$ is a linear subspace.

Convention: Domain of 1-homogeneous functions

Without restriction of generality (for variational problems), we will assume that any 1-homogeneous $J \in \Gamma_0(\mathbb{R}^n)$ maps from $\mathbb{R}^n$ to $\mathbb{R}$.

Remark: Using the above convention we conclude that such a J is continuous and defines a semi-norm on $\mathbb{R}^n$.

[Linear Filtering](#)[Spectral Theory](#)[Nonlinear singular vectors](#)[1-homogeneous functions](#)[Existence questions](#)[Error estimates](#)[Multiscale Methods](#)[Variational methods](#)[Gradient flow](#)[Inverse scale space flow](#)

Kernel of 1-homogeneous functions

Let $J \in \Gamma_0(\mathbb{R}^n)$ be 1-homogeneous. Then

$$\ker(J) = \{u \in \mathbb{R}^n \mid J(u) = 0\}$$

is a linear subspace.

Remark: J defines a norm on $\ker(J)^\perp$.

Nonnegativity of 1-homogeneous functions

Let $J \in \Gamma_0(\mathbb{R}^n)$ be 1-hom. Then $J(0) = 0$ and $J(u) \geq 0$ for all u .

[Linear Filtering](#)[Spectral Theory](#)[Nonlinear singular vectors](#)[1-homogeneous functions](#)[Existence questions](#)[Error estimates](#)[Multiscale Methods](#)[Variational methods](#)[Gradient flow](#)[Inverse scale space flow](#)

Subdifferential of 1-homogeneous functions

Let $J \in \Gamma_0(\mathbb{R}^n)$ be 1-hom. and subdifferentiable at u . Then

$$\partial J(u) = \{p \in \mathbb{R}^n \mid J(u) = \langle p, u \rangle, J(v) \geq \langle p, v \rangle \forall v \in \mathbb{R}^n\}$$

0-homogeneous subdifferential

Let $J \in \Gamma_0(\mathbb{R}^n)$ be 1-hom. and subdifferentiable at u . Then

$$\partial J(au) = \partial J(u)$$

holds for all $a > 0$.



Linear Filtering

Spectral Theory

Nonlinear singular vectors

1-homogeneous functions

Existence questions

Error estimates

Multiscale Methods

Variational methods

Gradient flow

Inverse scale space flow

Let us return to

$$\lambda u \in \partial J(u).$$

What about normalization?

- Linear case: If v_λ meets $\lambda v_\lambda \in \partial J(v_\lambda)$, then $v = av_\lambda$ meets $\lambda v \in \partial J(v)$, too.
- One-homogeneous: If v_λ meets $\lambda v_\lambda \in \partial J(v_\lambda)$, then $v = av_\lambda$ meets $\frac{\lambda}{a} v \in \partial J(v)$.

[Linear Filtering](#)[Spectral Theory](#)[Nonlinear singular vectors](#)[1-homogeneous functions](#)[Existence questions](#)[Error estimates](#)[Multiscale Methods](#)[Variational methods](#)[Gradient flow](#)[Inverse scale space flow](#)

Definition: Generalized singular vector

For $J \in \Gamma_0(\mathbb{R}^n)$ we call a v_λ with $\|v_\lambda\|_2 = 1$ a singular vector of J with singular value $\lambda \in \mathbb{R}$ if

$$\lambda v_\lambda \in \partial J(v_\lambda).$$

Observations for J being one-homogeneous:

- If there exists a v_λ with

$$\lambda v_\lambda \in \partial J(v_\lambda)$$

then $\tilde{v}_\lambda = \frac{v_\lambda}{\|v_\lambda\|_2}$ is a singular vector to J .

- For a singular value λ it holds that $\lambda = J(v_\lambda) \geq 0$.
- Smaller singular values correspond to smaller “frequencies”.



Linear Filtering

Spectral Theory

Nonlinear singular vectors

1-homogeneous functions

Existence questions

Error estimates

Multiscale Methods

Variational methods

Gradient flow

Inverse scale space flow

What about the existence of singular vectors?

[Linear Filtering](#)[Spectral Theory](#)

Nonlinear singular vectors

1-homogeneous functions

[Existence questions](#)

Error estimates

[Multiscale Methods](#)

Variational methods

Gradient flow

Inverse scale space flow

Definition: Ground States

Let $J \in \Gamma_0(\mathbb{R}^n)$ be 1-homogeneous. A *ground state* of J is defined by

$$u_0 = \arg \min_{\substack{u \in \ker(J)^\perp \\ \|u\|_2=1}} J(u).$$

Ground states exist

Let $J \in \Gamma_0(\mathbb{R}^n)$ be 1-homogeneous, and let $\ker(J) + \text{dom}(J)^\perp \neq \mathbb{R}^n$. Then a ground state exists.

Proof: Board.

[Linear Filtering](#)[Spectral Theory](#)

Nonlinear singular vectors

1-homogeneous functions

[Existence questions](#)

Error estimates

[Multiscale Methods](#)

Variational methods

Gradient flow

Inverse scale space flow

Ground states are singular vectors

Let $J \in \Gamma_0(\mathbb{R}^n)$ be one-homogeneous with ground state u_0 .
Then u_0 is a singular vector with the singular value $\lambda_0 = J(u_0)$.

Proof: Board

Remark: A ground state is a singular vector with the smallest possible singular values: $\lambda \geq \lambda_0$ for all singular values λ .

[Linear Filtering](#)[Spectral Theory](#)

Nonlinear singular vectors

1-homogeneous functions

[Existence questions](#)

Error estimates

[Multiscale Methods](#)

Variational methods

Gradient flow

Inverse scale space flow

Examples:

- Ground states for $J(u) = \|u\|_1$.
- Singular vectors for $J(u) = \|u\|_1$.

Conclusion: Ground states do not need to be unique.

- Singular vectors for 1d-TV: $J(u) = \|D_x u\|_1$.

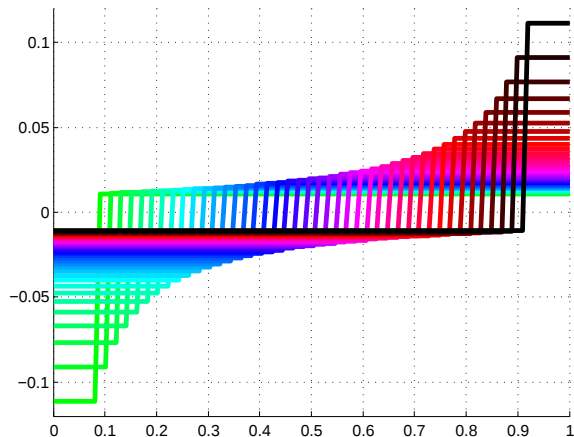
$$u_1 = 1/c$$

$$u_i = \begin{cases} u_{i-1} + 1/c & \text{for } i \leq c \\ u_{i-1} - 1/c & \text{for } c < i \leq n \end{cases}$$

Then $(D_x)^T u \in \mathbb{R}^n$ is a singular vector.

Generalized singular vectors

Example: 1d TV



Linear Filtering

Spectral Theory

Nonlinear singular vectors

1-homogeneous functions

Existence questions

Error estimates

Multiscale Methods

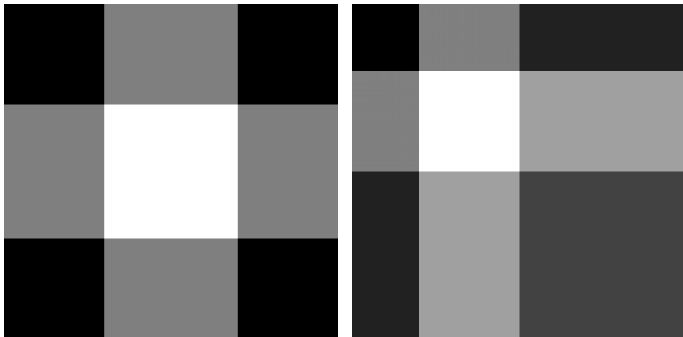
Variational methods

Gradient flow

Inverse scale space flow



Example: 2d TV



Linear Filtering

Spectral Theory

Nonlinear singular vectors

1-homogeneous functions

Existence questions

Error estimates

Multiscale Methods

Variational methods

Gradient flow

Inverse scale space flow



Usual behavior in the linear case: For symmetric positive semi-definite matrices, there exists an orthonormal basis of singular vectors.

$$\lambda_1 \langle v_1, v_2 \rangle = \langle Av_1, v_2 \rangle = \langle v_1, Av_2 \rangle = \lambda_2 \langle v_1, v_2 \rangle$$

How about the generalized case?

→ Singular vectors to different singular values are NOT necessarily orthonormal! E.g. ℓ^1 .



Can we always find an orthonormal basis consisting of singular vectors?

Linear case, $J(u) = \frac{1}{2} \|Ku\|_2^2$: Rayleigh Principle

Given $\{u_1, \dots, u_n\}$, compute

$$u_{n+1} = \arg \min_{\substack{u \in \ker(J)^\perp \\ \|u\|_2 = 1 \\ u \in \text{span}(u_1, \dots, u_n)^\perp}} J(u).$$

One can show: The u_n are singular vectors of K and form an orthonormal system. After completing with an orthonormal system that spans $\ker(K)$, one obtains an orthonormal basis.

Does the same strategy work in the nonlinear case?

[Linear Filtering](#)[Spectral Theory](#)

Nonlinear singular vectors

1-homogeneous functions

[Existence questions](#)

Error estimates

[Multiscale Methods](#)

Variational methods

Gradient flow

Inverse scale space flow

Consider

$$J(u) = \|Ku\|_1 \quad \text{with} \quad K = \begin{pmatrix} 1 & -0.1 \\ 0 & 1 \end{pmatrix}$$

Easy to verify: $u_0 = (1, 0)^T$ is a ground state.

Orthogonal: $u_1 = (0, 1)^T$, but for $v = (1, 10)^T$ we have

$$\langle J(u_1)u_1, v \rangle = 1.1 \cdot 10 = 11$$

but $J(v) = 10$.

$\Rightarrow u_1 = (0, 1)^T$ cannot be a singular vector.

This does not mean there is no basis of singular vectors!



Linear Filtering

Spectral Theory

Nonlinear singular vectors

1-homogeneous functions

Existence questions

Error estimates

Multiscale Methods

Variational methods

Gradient flow

Inverse scale space flow

Board computation:

In what way does the proof for showing that the ground state is a singular vector fail in the case of the Rayleigh principle?

A basis of singular vectors?

Again, consider

$$J(u) = \|Ku\|_1 \quad \text{with} \quad K = \begin{pmatrix} 1 & -0.1 \\ 0 & 1 \end{pmatrix}$$

Then for $q = (-1, 1)^T$ we find

$$K^T q = (-1, 1.1), KK^T q = (-1.11, 1.1)$$

such that

$$\langle K^T q, K^T q \rangle = \|K(K^T q)\|_1 = J(K^T q).$$

Moreover,

$$\langle K^T q, z \rangle = \langle q, Kz \rangle \leq \|Kz\|_1 = J(z).$$

Thus, $K^T q \in \partial J(K^T q)$, and after normalization, $K^T q$ is a singular vector.



A basis of singular vectors?

Again, consider

$$J(u) = \|Ku\|_1 \quad \text{with} \quad K = \begin{pmatrix} 1 & -0.1 \\ 0 & 1 \end{pmatrix}$$

Then for $q_2 = (1, 1)^T$ we find

$$K^T q_2 = (1, 0.9), \quad KK^T q_2 = (0.91, 0.9)$$

such that

$$\langle K^T q_2, K^T q_2 \rangle = \|K(K^T q_2)\|_1 = J(K^T q_2).$$

Moreover,

$$\langle K^T q_2, z \rangle = \langle q_2, Kz \rangle \leq \|Kz\|_1 = J(z).$$

Thus, $K^T q_2 \in \partial J(K^T q_2)$, and after normalization, $K^T q_2$ is a singular vector which is orthogonal to q_1 !



[Linear Filtering](#)[Spectral Theory](#)

Nonlinear singular vectors

1-homogeneous functions

[Existence questions](#)

Error estimates

[Multiscale Methods](#)

Variational methods

Gradient flow

Inverse scale space flow

Open problem 1: Existence of a basis

For any 1-homogeneous $J \in \Gamma_0(\mathbb{R}^n)$ there exists a basis of generalized singular vectors?

Remark: There are some particular cases for which the above provably holds, as we will see in the next chapter.

Open problem 2: Numerical computation of singular vectors²

The problem of finding elements that satisfy

$$\lambda v_\lambda \in \partial J(v_\lambda), \quad \|v_\lambda\|_2 = 1$$

is highly degenerate. How can we find such v_λ in practice?

²Also see Hein and Buehler, *An inverse power method for nonlinear eigenproblems with applications in 1-spectral clustering and sparse PCA*, 2010.



Linear Filtering

Spectral Theory

Nonlinear singular vectors

1-homogeneous functions

Existence questions

Error estimates

Multiscale Methods

Variational methods

Gradient flow

Inverse scale space flow

Orthogonality to kernel

For any 1-homogeneous $J \in \Gamma_0(\mathbb{R}^n)$, if $v \in \text{kern}(J)$ and $p \in \partial J(0)$ then $p \perp v$. In particular, all singular vectors with nonzero singular values are in $\text{kern}(J)^\perp$.

Conclusion: Singular vectors with singular value 0 are orthogonal to those that have non-zero singular values.



Linear Filtering

Spectral Theory

Nonlinear singular vectors

1-homogeneous functions

Existence questions

Error estimates

Multiscale Methods

Variational methods

Gradient flow

Inverse scale space flow

Error Estimates

What is the theory of nonlinear singular vectors good for?



Error estimates – clean data

Let $J \in \Gamma_0(\mathbb{R}^n)$ be 1-homogeneous, and let v_λ be a singular vector with singular value λ . Then

$$u(t) = \arg \min_u \frac{1}{2} \|u - \gamma v_\lambda\|^2 + tJ(u)$$

is explicitly given by

$$u(t) = \begin{cases} (\gamma - t\lambda)v_\lambda & \text{if } t\lambda \leq \gamma, \\ 0 & \text{else.} \end{cases}$$

Proof: Board.

[Linear Filtering](#)[Spectral Theory](#)

Nonlinear singular vectors

1-homogeneous functions

Existence questions

[Error estimates](#)[Multiscale Methods](#)

Variational methods

Gradient flow

Inverse scale space flow

[Linear Filtering](#)[Spectral Theory](#)

Nonlinear singular vectors

1-homogeneous functions

Existence questions

[Error estimates](#)[Multiscale Methods](#)

Variational methods

Gradient flow

Inverse scale space flow

Error estimates – noisy data

Let $J \in \Gamma_0(\mathbb{R}^n)$ be 1-homogeneous, and let v_λ be a singular vector with singular value λ . Let $f = \gamma v_\lambda + n$, and let there exist positive constants μ, η with $\frac{\mu}{\eta} < \gamma$ such that

$$\mu v_\lambda + \eta n \in \partial J(v_\lambda).$$

Then

$$u_\alpha = \arg \min_u \frac{1}{2} \|u - f\|^2 + \alpha J(u)$$

is explicitly given by

$$u_\alpha = \left(\gamma - \alpha \lambda + \frac{\lambda - \mu}{\eta} \right) v_\lambda,$$

if $\alpha \in [1/\eta, 1/\eta + \gamma/\lambda[$.

Example: ℓ^1 regularization – Matlab. Polyhedral discussion.

[Linear Filtering](#)[Spectral Theory](#)

Nonlinear singular vectors

1-homogeneous functions

Existence questions

[Error estimates](#)[Multiscale Methods](#)

Variational methods

Gradient flow

Inverse scale space flow

Remark: Generalizations

All concepts we discussed in this section can be generalized to problems of the form

$$\min_u \frac{1}{2} \|Au - f\|^2 + \alpha J(u)$$

with adapted definitions of ground states and singular vectors. See Benning, Burger, *Ground states and singular vectors of convex variational regularization methods*.

Open problem 3: Arbitrary data fidelity terms

Does the theory extend to even more general variational problems like

$$\min_u H_f(u) + \alpha J(u)?$$



Linear Filtering

Spectral Theory

Nonlinear singular vectors

1-homogeneous functions

Existence questions

Error estimates

Multiscale Methods

Variational methods

Gradient flow

Inverse scale space flow

Nonlinear Multiscale Methods



Linear Filtering

Spectral Theory

Nonlinear singular vectors

1-homogeneous functions

Existence questions

Error estimates

Multiscale Methods

Variational methods

Gradient flow

Inverse scale space flow

Can we use variational methods like

$$u(t) = \arg \min_u \frac{1}{2} \|u - f\|^2 + tJ(u)$$

to define a multiscale decomposition of f ?

(Throughout this chapter J is 1-homogeneous.)



For

$$u(t) = \arg \min_u \frac{1}{2} \|u - f\|^2 + tJ(u)$$

we have already seen: If $\frac{f}{\gamma}$ is a singular vector to singular value λ , then

$$u(t) = \begin{cases} (1 - \frac{\lambda t}{\gamma})f & \text{if } t\lambda \leq \gamma \\ 0 & \text{else.} \end{cases}$$

How can we define a transformation that consists of a single peak if the data is a singular vector?

- $u(t)$ is piecewise linear in time.
- $\partial_t u(t)$ is piecewise constant in time.
- $\partial_{tt} u(t)$ consists of Dirac deltas (=is a measure).

Linear Filtering

Spectral Theory

Nonlinear singular vectors

1-homogeneous functions

Existence questions

Error estimates

Multiscale Methods

Variational methods

Gradient flow

Inverse scale space flow

Multiscale Decompositions by Variational Methods

Let us consider the example of $\frac{f}{\gamma}$ being a singular vector again.

We know

$$u(t) = \begin{cases} (1 - \frac{\lambda t}{\gamma})f & \text{if } t\lambda \leq \gamma \\ 0 & \text{else.} \end{cases}$$

Thus,

$$\partial_t u(t) = \begin{cases} -\frac{\lambda}{\gamma}f & \text{if } t\lambda \leq \gamma \\ 0 & \text{else.} \end{cases}$$

And therefore

$$\partial_{tt} u(t) = \frac{\lambda}{\gamma} \delta(\lambda t - \gamma) f = \frac{1}{t} \delta(\lambda t - \gamma) \cdot f$$

We find that in this case

$$f = \int_0^\infty t \partial_{tt} u(t) dt.$$



Linear Filtering

Spectral Theory

Nonlinear singular vectors

1-homogeneous functions

Existence questions

Error estimates

Multiscale Methods

Variational methods

Gradient flow

Inverse scale space flow



Linear Filtering

Spectral Theory

Nonlinear singular vectors

1-homogeneous functions

Existence questions

Error estimates

Multiscale Methods

Variational methods

Gradient flow

Inverse scale space flow

The next lectures

- Today: Finish variational methods, do gradient flows.
- 30.06: No lecture!
- Monday 06.07. 14:00, Room 02.05.014 Hands-on code!
- Tuesday 07.07. 16:00, Room 02.05.014 Hands-on code!
- 07.07: Gradient flows + inverse scale space flows.
- 14.07: Inverse scale space flows, discussion + summary of the course.
- 28/29.07: Exams (Part 1)
- 17/18.09: Exams (Part 2)



Wavelength decomposition

We define

$$\phi(t) = t\partial_{tt}u(t)$$

and call it the *wavelength decomposition* of f with respect to J .

Questions:

- Why is it really a *decomposition*?
- Why do we call it *wavelength*?

Linear Filtering

Spectral Theory

Nonlinear singular vectors

1-homogeneous functions

Existence questions

Error estimates

Multiscale Methods

Variational methods

Gradient flow

Inverse scale space flow



Spectral decomposition

It holds that

$$f = \text{proj}_{\text{kern}(\mathcal{J})}(f) + \int_0^\infty \phi(t) dt.$$

We should verify that our investigations are reasonable:

Continuity

For any 1-homogeneous $J \in \Gamma_0(\mathbb{R}^n)$ and any $f \in \mathbb{R}^n$ the solution

$$u(t) = \arg \min_u \frac{1}{2} \|u - f\|_2^2 + tJ(u), \quad t \geq 0$$

is Lipschitz continuous in t .

<i>Function</i>	<i>Weak Derivative</i>
Smooth (C^1) Lipschitz Absolutely Continuous Bounded Variation	Continuous (C^0) Bounded (L^∞) Integrable (L^1) Distributional derivative is Radon measure

https://www.math.ucdavis.edu/~hunter/m218a_09/ch3A.pdf



Linear Filtering

Spectral Theory

Nonlinear singular vectors

1-homogeneous functions

Existence questions

Error estimates

Multiscale Methods

Variational methods

Gradient flow

Inverse scale space flow

[Linear Filtering](#)[Spectral Theory](#)

Nonlinear singular vectors

1-homogeneous functions

Existence questions

Error estimates

[Multiscale Methods](#)[Variational methods](#)

Gradient flow

Inverse scale space flow

Interpretation of ϕ

Our definition of $\phi(t) = t \partial_{tt} u(t)$ can be understood in the way that each component ϕ_i is as an element of the space $(W^{1,1})^*$ dual to the Sobolev space $W^{1,1}$. It acts via

$$w \mapsto - \int_0^T u'_i(t) (w(t) + t w'(t)) dt$$

where T is the finite extinction time.

[Linear Filtering](#)[Spectral Theory](#)

Nonlinear singular vectors

1-homogeneous functions

Existence questions

Error estimates

[Multiscale Methods](#)[Variational methods](#)

Gradient flow

Inverse scale space flow

Finite time extinction

For any 1-homogeneous $J \in \Gamma_0(\mathbb{R}^n)$ and any $f \in \mathbb{R}^n$ there exist a $T \in \mathbb{R}$ such that

$$u(t) = \arg \min_u \frac{1}{2} \|u - f\|_2^2 + tJ(u)$$

meets $u(t) = \text{proj}_{\text{kern}(J)}(f)$ for all $t \geq T$.

Conclusion: $\partial_t u(t) = 0$ for all $t \geq T$.

We can now proof the spectral decomposition theorem.

Remark: For $\text{proj}_{\text{kern}(J)}(f) = 0$ we have

$$f = \int_0^\infty \phi(t) dt.$$



Linear Filtering

Spectral Theory

Nonlinear singular vectors

1-homogeneous functions

Existence questions

Error estimates

Multiscale Methods

Variational methods

Gradient flow

Inverse scale space flow

How does it look for data f that is not a singular vector?

Show videos!



How should we measure the spectral response, i.e. the energy concentrated at a particular frequency/time?

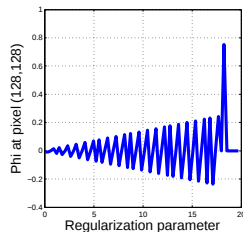
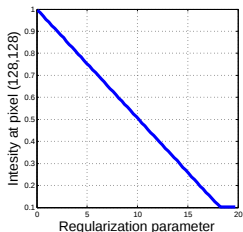
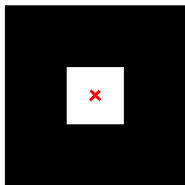
First option: $S(t) = \|\phi(t)\|_1$.

E.g. for f being a singular vector, we have $S(t) = \delta(\lambda t - \gamma) \|f\|_1$



General advice: Be careful in the computation of $\phi(t)$!

Differentiation is an ill-posed problem!



Linear Filtering

Spectral Theory

Nonlinear singular vectors

1-homogeneous functions

Existence questions

Error estimates

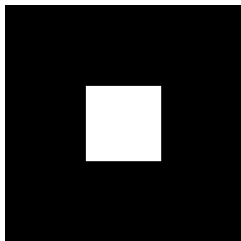
Multiscale Methods

Variational methods

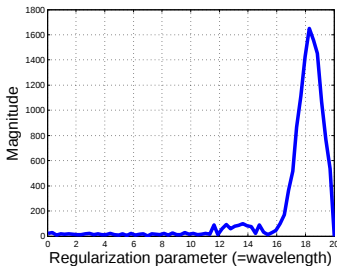
Gradient flow

Inverse scale space flow

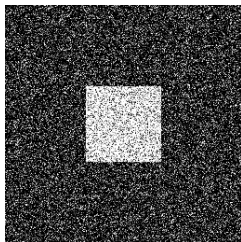
Multiscale Decompositions by Variational Methods



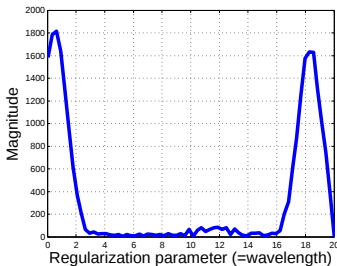
Input image



TV spectrum



Input image



TV spectrum



Linear Filtering

Spectral Theory

- Nonlinear singular vectors
- 1-homogeneous functions
- Existence questions
- Error estimates

Multiscale Methods

Variational methods

- Gradient flow
- Inverse scale space flow

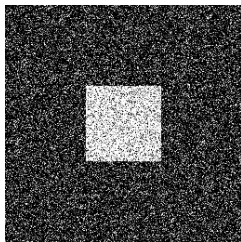


Spectral filtering

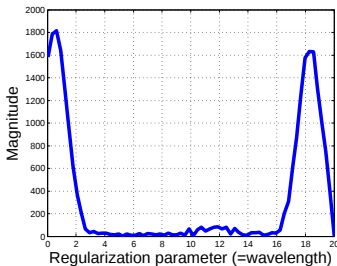
For a function $w : [0, \infty[\rightarrow \mathbb{R}$ we define the spectral filtering as

$$u_w = \int_0^\infty w(t) \phi(t) dt.$$

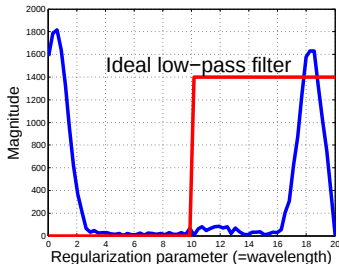
Multiscale Decompositions by Variational Methods



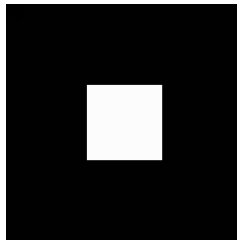
Input image



TV spectrum



Ideal filter



Filtered Image

Linear Filtering

Spectral Theory

- Nonlinear singular vectors
- 1-homogeneous functions
- Existence questions
- Error estimates

Multiscale Methods

Variational methods

- Gradient flow
- Inverse scale space flow



Linear Filtering

Spectral Theory

Nonlinear singular vectors

1-homogeneous functions

Existence questions

Error estimates

Multiscale Methods

Variational methods

Gradient flow

Inverse scale space flow

You know everything you need for implementing nonlinear multiscale methods based on variational regularization.

Your task: Implement it!

[Linear Filtering](#)[Spectral Theory](#)

Nonlinear singular vectors

1-homogeneous functions

Existence questions

Error estimates

[Multiscale Methods](#)

Variational methods

[Gradient flow](#)

Inverse scale space flow

Can we use gradient flows like

$$\partial_t u(t) \in -\partial J(u(t)), \quad u(0) = f,$$

to define a multiscale decomposition of f ?

[Linear Filtering](#)[Spectral Theory](#)

Nonlinear singular vectors

1-homogeneous functions

Existence questions

Error estimates

[Multiscale Methods](#)

Variational methods

[Gradient flow](#)

Inverse scale space flow

First step: Let $f/\gamma \in \partial J(f)$. Then the solution to

$$\partial_t u(t) \in -\partial J(u(t)), \quad u(0) = f,$$

is given by

$$u(t) = \begin{cases} (1 - \frac{\lambda t}{\gamma})f & \text{if } t\lambda \leq \gamma \\ 0 & \text{else.} \end{cases}$$

Exactly the same as for the variational method!

We can keep all definitions!

Details on the existence of solutions to the flow:

e.g. Aubin, Cellina: *Differential Inclusions*. p. 147, Theorem 1.

[Linear Filtering](#)[Spectral Theory](#)

Nonlinear singular vectors

1-homogeneous functions

Existence questions

Error estimates

[Multiscale Methods](#)

Variational methods

[Gradient flow](#)

Inverse scale space flow

Gradient flow multiscale decomposition

Wavelength representation:

$$\phi(t) = t \partial_{tt} u(t)$$

Reconstruction (assuming finite time extinction):

$$f = \int_0^\infty \phi(t) dt + \text{proj}_{\text{kern}(J)}(f)$$

Spectral response:

$$S(t) = \|\phi(t)\|_1 \quad \text{or} \quad S(t) = t \sqrt{\partial_{tt} J(u(t))}$$

Filtering:

$$\hat{u} = \int_0^\infty w(t) \phi(t) dt + \text{proj}_{\text{kern}(J)}(f)$$



Linear Filtering

Spectral Theory

- Nonlinear singular vectors
- 1-homogeneous functions
- Existence questions
- Error estimates

Multiscale Methods

Variational methods

Gradient flow

Inverse scale space flow

Open problem 4: Finite time extinction

Under what assumptions does the gradient flow yield finite time extinction?

Open problem 5: Relation to variational method

Under what assumptions (on J or f) is the gradient flow equivalent to the variational method?

For your numerical implementation:

$$\partial_t u(t) = -p(t), \quad p(t) \in \partial J(u(t)).$$

Discrete:

$$\frac{u(t^{k+1}) - u(t^k)}{\tau} = -p(t^{k+1}), \quad p(t^{k+1}) \in \partial J(u(t^{k+1})).$$

Or

$$0 = u(t^{k+1}) - u(t^k) + \tau p(t^{k+1}),$$

i.e.

$$u(t^{k+1}) = \arg \min_u \frac{1}{2} \|u - u(t^k)\|^2 + \tau J(u).$$

Implement it!



Linear Filtering

Spectral Theory

Nonlinear singular vectors

1-homogeneous functions

Existence questions

Error estimates

Multiscale Methods

Variational methods

Gradient flow

Inverse scale space flow

[Linear Filtering](#)[Spectral Theory](#)

Nonlinear singular vectors

1-homogeneous functions

Existence questions

Error estimates

[Multiscale Methods](#)

Variational methods

Gradient flow

Inverse scale space flow

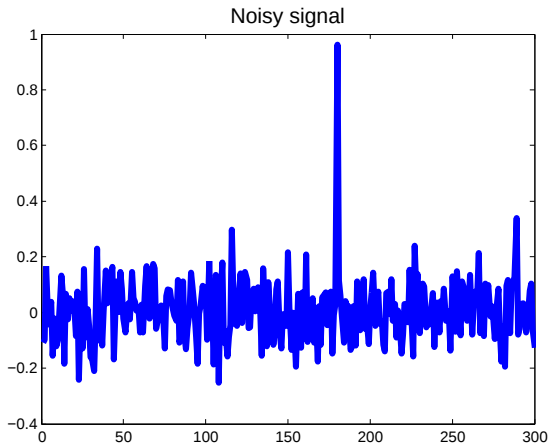
Can we use an inverse scale space flows like

$$\partial_t p(t) = f - u(t), \quad p(t) \in \partial J(u(t)), \quad p(0) = 0,$$

to define a multiscale decomposition of f ?

What is an inverse scale space flow?

Simple problem: Denoise a sparse signal (convex + variational)



Linear Filtering

Spectral Theory

Nonlinear singular vectors

1-homogeneous functions

Existence questions

Error estimates

Multiscale Methods

Variational methods

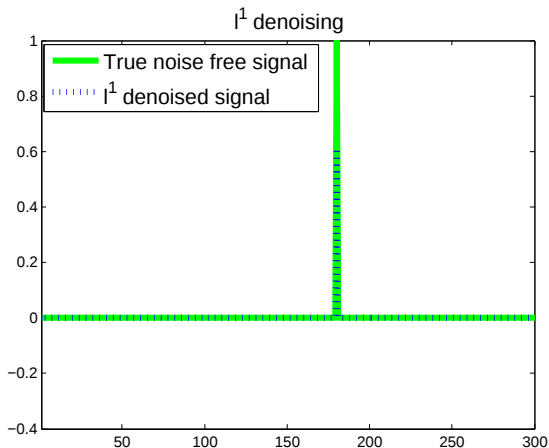
Gradient flow

Inverse scale space flow

Systematic errors

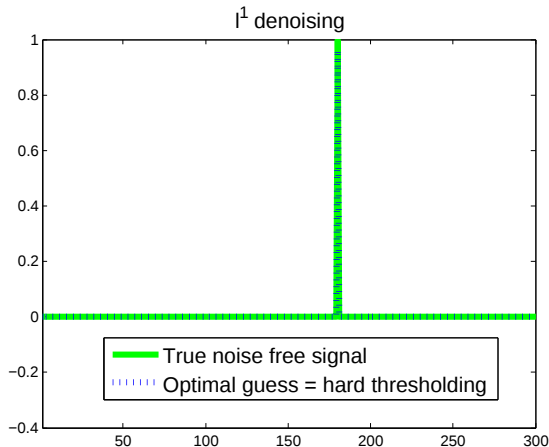
Based on what we have seen:

$$\hat{u} = \arg \min_u \frac{1}{2} \|u - f\|_2^2 + \alpha \|u\|_1$$



Systematic errors

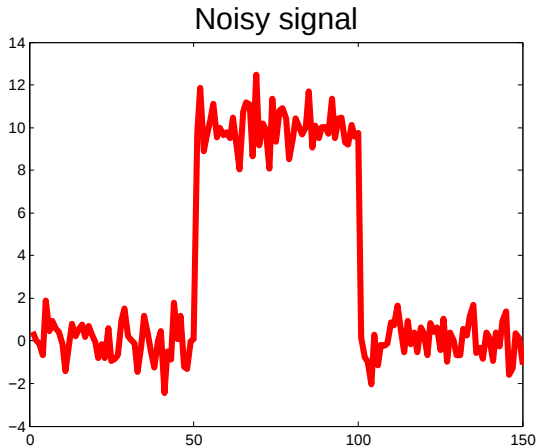
This would be a much better solution...



Noise with zero mean:

→ intensity of nonzero components should not be decreased!





Linear Filtering

Spectral Theory

Nonlinear singular vectors

1-homogeneous functions

Existence questions

Error estimates

Multiscale Methods

Variational methods

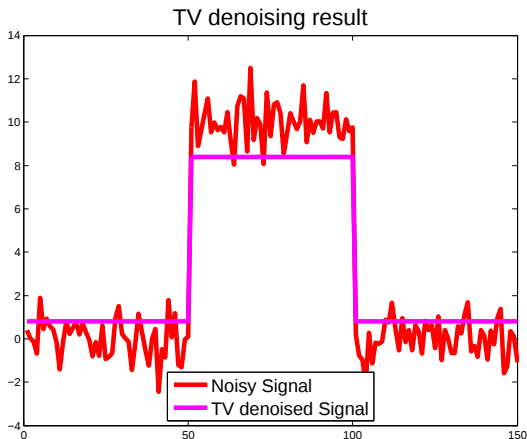
Gradient flow

Inverse scale space flow

Systematic errors

Assume Gaussian noise and piecewise constant signal:

$$\hat{u} = \arg \min_u \frac{1}{2} \|u - f\|^2 + \alpha \|\partial_x u\|_1$$



[Linear Filtering](#)[Spectral Theory](#)

Nonlinear singular vectors

1-homogeneous functions

Existence questions

Error estimates

[Multiscale Methods](#)

Variational methods

Gradient flow

Inverse scale space flow

How can we avoid the systematic error?

Osher et al. 2005: Bregman iteration

$$u^{k+1} = \arg \min_u \frac{1}{2} \|u - f\|_2^2 + \alpha(J(u) - \langle p^k, u \rangle)$$

for $p^k \in \partial J(u^k)$, $p^0 = 0$.

Why does it work? Example ℓ^1 .



Linear Filtering

Spectral Theory

Nonlinear singular vectors

1-homogeneous functions

Existence questions

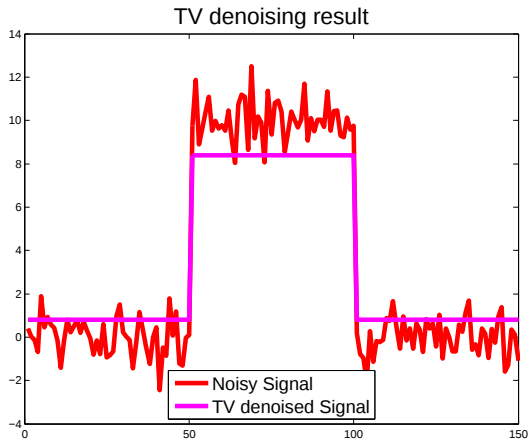
Error estimates

Multiscale Methods

Variational methods

Gradient flow

Inverse scale space flow



Bregman iteration



Linear Filtering

Spectral Theory

Nonlinear singular vectors

1-homogeneous functions

Existence questions

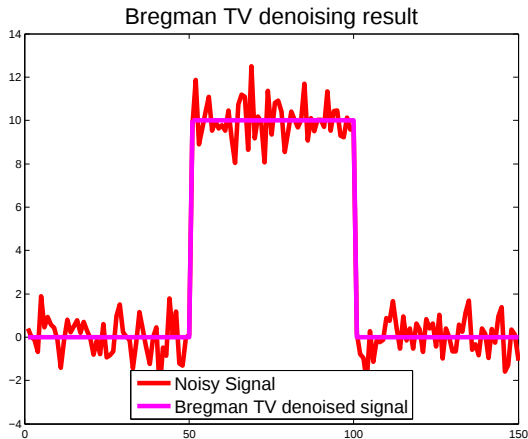
Error estimates

Multiscale Methods

Variational methods

Gradient flow

Inverse scale space flow





Iterative procedure:

$$u^{k+1} = \arg \min_u \frac{1}{2} \|u - f\|_2^2 + \alpha (J(u) - \langle p^k, u \rangle)$$

Optimality condition

$$\begin{aligned} 0 &= u^{k+1} - f + \alpha p^{k+1} - \alpha p^k \\ \Rightarrow \frac{p^{k+1} - p^k}{\frac{1}{\alpha}} &= f - u^{k+1} \end{aligned}$$

Left hand side looks like a derivative. Take $\alpha \rightarrow \infty$.

$$\partial_s p(s) = f - u(s), \quad p(s) \in \partial J(u(s))$$

Inverse Scale Space Flow!

Linear Filtering

Spectral Theory

Nonlinear singular vectors

1-homogeneous functions

Existence questions

Error estimates

Multiscale Methods

Variational methods

Gradient flow

Inverse scale space flow

[Linear Filtering](#)[Spectral Theory](#)

Nonlinear singular vectors

1-homogeneous functions

Existence questions

Error estimates

[Multiscale Methods](#)

Variational methods

Gradient flow

Inverse scale space flow

Some intuition on the behavior of

$$\partial_s p(s) = f - u(s), \quad p(s) \in \partial J(u(s)).$$

Remember $p(t) \in \partial J(u(t))$ means $u(t) \in \partial J^*(p(t))$. Define

$$H(p) = -\langle p, f \rangle + J^*(p).$$

Convex functional!

Inverse scale space flow is

$$\partial_s p(s) \in -\partial H(p(s)).$$

Same structure as for gradient flow (except 1-homogeneity).

[Linear Filtering](#)[Spectral Theory](#)

Nonlinear singular vectors

1-homogeneous functions

Existence questions

Error estimates

[Multiscale Methods](#)

Variational methods

Gradient flow

[Inverse scale space flow](#)

Properties of the ISS flow

Short computations on the board show:

- Existence of a unique solution.³
- $H(p(s))$ is monotonically decreasing.
- We find that $\|u(s) - f\|$ decreases monotonically,
- Integration of the first estimate yields $\|\partial_s p(s)\|^2 \in \mathcal{O}(\frac{1}{s})$,
- which is the same as $\|u(s) - f\|^2 \in \mathcal{O}(\frac{1}{s})$.

³Aubin, Cellina: *Differential Inclusions*. p. 147, Theorem 1.

[Linear Filtering](#)[Spectral Theory](#)

Nonlinear singular vectors

1-homogeneous functions

Existence questions

Error estimates

[Multiscale Methods](#)

Variational methods

Gradient flow

Inverse scale space flow

Behavior of

$$\partial_s p(s) = f - u(s), \quad p(t) \in \partial J(u(s)), \quad p(0) = 0$$

for $\frac{f}{\gamma}$ being a singular vector.

We find

$$p(s) = \begin{cases} s f & \text{if } s \leq \frac{\lambda}{\gamma} \\ \frac{\lambda}{\gamma} f & \text{else.} \end{cases} \quad u(s) = \begin{cases} 0 & \text{if } s \leq \frac{\lambda}{\gamma} \\ f & \text{else.} \end{cases}$$

We observe

- $u(s)$ is piecewise constant in time.
- Only one derivative is needed to obtain a peak!



We define $\psi(s) = \partial_s u(s)$ to be the **frequency coefficients in the inverse scale space decomposition**.

Why

- *frequency?*

The larger λ , the longer it takes for $u(s) = f$, i.e., the peak occurs at larger s .

- *decomposition?*

Starting with $p(0) = 0$, $u(0) = \text{proj}_{\text{ker}(J)}(f)$, we find

$$\int_0^\infty \psi(s) \, ds = f - \text{proj}_{\text{ker}(J)}(f)$$

Frequency vs. Wavelength Representation



Can we express the frequencies $\psi(s)$ by a wavelengths $\phi(s)$?

Filtering the wavelength representation with $w(t)$ should be the same as filtering the frequency representation with $w(1/t)$.

By using $t = 1/s$ we find

$$\int_0^\infty \psi(s) w(s) ds = \int_0^\infty \frac{1}{s^2} \psi(1/s) w(1/s) ds,$$

such that we define $\phi(t) = \frac{1}{t^2} \psi(1/t)$.

We now have a set of tools for nonlinear multiscale analysis with inverse scale space flows!

[Linear Filtering](#)[Spectral Theory](#)

Nonlinear singular vectors

1-homogeneous functions

Existence questions

Error estimates

[Multiscale Methods](#)

Variational methods

Gradient flow

Inverse scale space flow

Inverse scale space flow multiscale decomposition

Frequency representation:

$$\psi(s) = \partial_s u(s)$$

Wavelength representation:

$$\phi(t) = 1/t^2 \psi(1/t)$$

Reconstruction:

$$f = \int_0^\infty \psi(s) ds + \text{proj}_{\text{kern}(J)}(f)$$

Filtering:

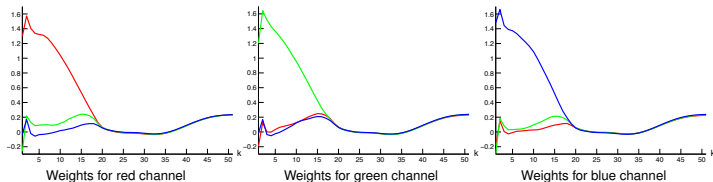
$$\hat{u} = \int_0^\infty w(s) \psi(s) ds + \text{proj}_{\text{kern}(J)}(f)$$



Filter learning

What is the optimal filter for a certain task?

Color image denoising:

[Linear Filtering](#)[Spectral Theory](#)

Nonlinear singular vectors

1-homogeneous functions

Existence questions

Error estimates

[Multiscale Methods](#)

Variational methods

Gradient flow

[Inverse scale space flow](#)

Example Application



Noisy

Multiscale Methods

Michael Moeller



Linear Filtering

Spectral Theory

Nonlinear singular vectors

1-homogeneous functions

Existence questions

Error estimates

Multiscale Methods

Variational methods

Gradient flow

Inverse scale space flow

Example Application



Best Bregman Iteration



Example Application



Learned Optimal Filters

