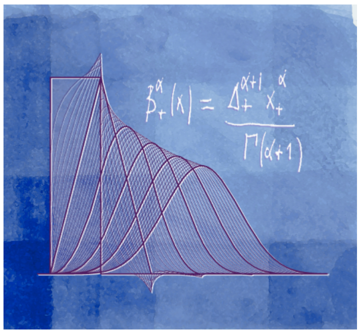
ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Sparse stochastic processes

Personal work:

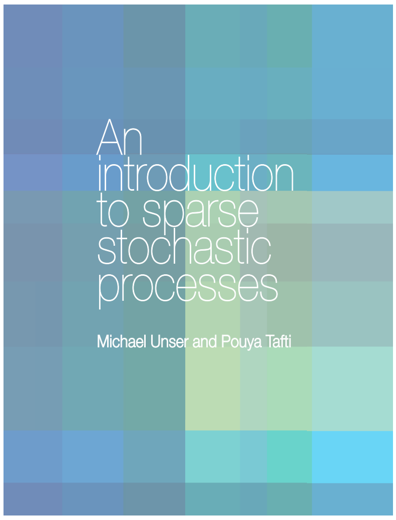
Small research project(s)

Prof. Michael Unser, LIB



EPFL Doctoral School EDEE, Course EE-726, Spring 2017

Table of content



ERC Grant **FUN-SP** (2011-2016)

A Functional Framework for Sparse, Non-Gaussian Signal Processing and Bioimaging

	Space-domain description	Fourier characterization	Generalization
1. Derivative operator	$\mathrm{D} = \frac{\mathrm{d}}{\mathrm{d}x}$	$j\omega$	Diff. operator
2. Discrete derivative	$\mathrm{D}_d s(x) = s(x) - s(x-1)$	$1 - e^{-j\omega}$	Num. analysis
3. Green function	$u(x) = \mathrm{D}^{-1}\{\delta\}(x)$	$\frac{1}{j\omega} + \pi\delta(\omega)$	Distribution theory
4. B-spline basis	$\varphi(x) = \mathrm{rect}(x - \frac{1}{2}) = \mathrm{D}_d u(x)$	$\left(\frac{1-e^{-j\omega}}{j\omega}\right) = \frac{\hat{\mathrm{D}}_d(\omega)}{\hat{\mathrm{D}}(\omega)}$	Spline theory
5. Spline (piecewise constant)	$s(x) = b_1 + \sum_{k\in\mathbb{Z}} a[k]u(x-k)$	$\frac{A(e^{j\omega})-A(1)}{j\omega}$	Spline theory
6. **Wavelet basis** (Haar)	$\psi(x) = \frac{\mathrm{d}}{\mathrm{d}x}\phi(x)$ with $\phi = \varphi * \varphi$	$j\omega\left(\frac{1-e^{-j\omega}}{j\omega}\right)^2$	**This grant**: WP 2.
7. Gaussian process	Brownian motion: $\frac{\mathrm{d}}{\mathrm{d}x}s(x) = w(x)$	$S(\omega) \sim \frac{1}{\lvert\omega\rvert}$	Classical theory
8. **Sparse process**	Compound Poisson process	$S(\omega) \sim \frac{1}{\lvert\omega\rvert}$	**This grant**: WP 2.
9. **Regularization functional**	$\mathrm{TV}(s) = \lVert \mathrm{D}s \rVert_{L_1(\mathbb{R})}$	N/A	**This grant**: WP 2.
10. **Key points/singularities**	$\{x_k : \lvert g(x_k)\rvert > T, g'(x_k) = 0\}$ with $g(x) = \mathrm{D}(\phi * s)(x)$	N/A	**This grant**: WP 2.

ERC Grant **GlobalBioIm** (2016-2020)

Global integrative framework for Computational Bio-Imaging

Continuous-domain formulation: Simplified picture in 1D

Signal $s \in \mathcal{X} \subseteq \mathcal{S}'(\mathbb{R})$ (native space)

Linear measurement operator $\boldsymbol{\nu} : s \mapsto \mathbf{y} = (y_1, \cdots, y_M) = \boldsymbol{\nu}(s)$ $\qquad \boldsymbol{\nu} : \mathcal{X} \to \mathbb{R}^M$

from continuum to discrete (finite dimensional)

- Tikhonov (or L_2) regularization

$$s_{\mathrm{Tik}} = \arg\min_{s\in\mathcal{H}_{\mathrm{D}}} \underbrace{\lVert \mathbf{y} - \boldsymbol{\nu}(s)\rVert_2^2}_{\text{data consistency}} + \underbrace{\lambda \lVert \mathrm{D}s \rVert_{L_2}^2}_{\text{regularization}}$$

- Total-variation (or sparsity-promoting) regularization

$$s_{\mathrm{TV}} = \arg\min_{s\in\mathrm{BV}(\mathbb{R})} \underbrace{\lVert \mathbf{y} - \boldsymbol{\nu}(s)\rVert_2^2}_{\text{data consistency}} + \underbrace{\lambda \lVert \mathrm{D}s \rVert_{\mathcal{M}}}_{\text{total variation}}$$

- Optimum (Bayesian) statistical reconstrution: MMSE

Measurement model: $\mathbf{y} = \nu(s) + \mathbf{n} \in \mathbb{R}^M$

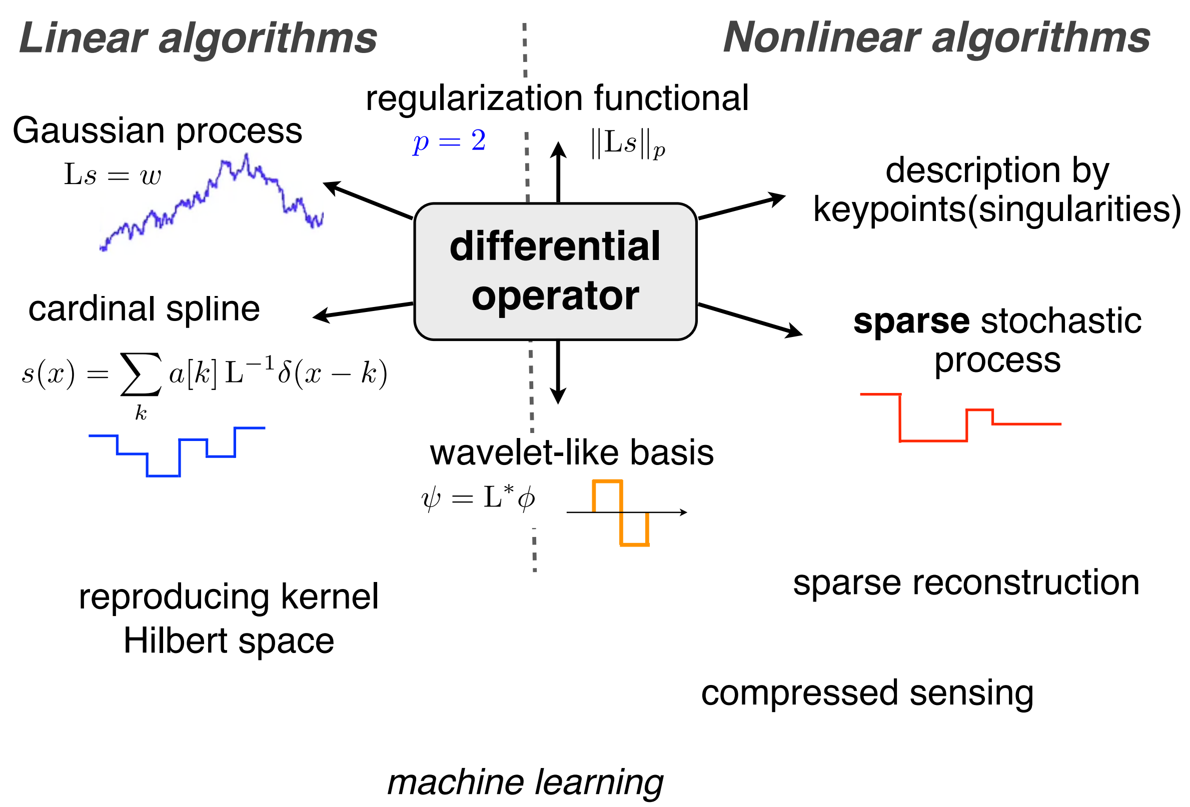
- signal $s \in \mathcal{S}'(\mathbb{R})$: realization of a stochastic process with probability measure $\mathscr{P}_s$
- discrete noise $\mathbf{n}$: i.i.d. zero-mean Gaussian with variance σ^2

$\Rightarrow$ Posterior probability density function: $p\big(s(x_0)|\mathbf{y}\big)$ for any $x_0 \in \mathbb{R}$

$$s_{\mathrm{MMSE}}(x_0|\mathbf{y}) = \mathbb{E}\{s(x_0|\mathbf{y})\} = \int_{\mathbb{R}} s p(s|\mathbf{y}) \mathrm{d}s$$

Remarkable property:

When *s* is a **Gaussian** stochastic process: **perfect equivalence** with Tikhonov/ smoothing spline reconstruction



Personal research project

Choose your **operator** (or class of operators) and go through the whole construction process;
i.e., specify

- Null space
- Green's function
- B-spline
- Stabilized inverse operator(s)
- Reproducing kernel
- Native space
- Wavelet basis
- corresponding stochastic processes
  sparse vs. *Gaussian*
- representer theorem(s)
- optimal (MMSE) interpolator

Symbolic/numerical computation using **Mathematica**

Calculation of Green's functions

```
FourierTransform[HeavisideTheta[x], x, ω, FourierParameters -> {1, -1}]
```

$$-\frac{i}{\omega} + \pi\,\mathrm{DiracDelta}[\omega]$$

```
InverseFourierTransform[1/(I ω), ω, x, FourierParameters -> {1, -1}]
```

$$\frac{\mathrm{Sign}[x]}{2}$$

```
InverseFourierTransform[1/(I ω)^2, ω, x, FourierParameters -> {1, -1}]
```

$$\frac{1}{2}\,x\,\mathrm{Sign}[x]$$

Derivative operator

$$\mathrm{L} = \mathrm{D} = \frac{\mathrm{d}}{\mathrm{d}x}$$

- Null space

$$\mathcal{N}_{\mathrm{D}} = \mathrm{span}\{p_1\}, \quad p_1(x) = 1$$

Biorthogonal system: $\{\phi_1, p_1\} = \{\delta, 1\}$

Projection operator: $\mathrm{Proj}_{\mathcal{N}_{\mathrm{D}}}\{f\} = \langle \phi_1, f\rangle p_1 = f(0)$

hypothesis: $f : \mathbb{R} \to \mathbb{R}$ must be continuous

- Green's function(s)

$u(x) = \mathbb{1}_+(x)$: Heaviside function

$\mathrm{D}\{u\}(x) = \delta(x)$: Dirac impulse

Canonical solution: $\rho_{\mathrm{D}}(x) = \mathcal{F}^{-1}\{\frac{1}{\mathrm{j}\omega}\} = \frac{1}{2}\mathrm{sign}(x)$

Derivative operator: B-spline

Derivative $\quad \mathrm{D}f(x) = \dfrac{\mathrm{d}f(x)}{\mathrm{d}x} \qquad \mathrm{D} \overset{\mathcal{F}}{\longleftrightarrow} \mathrm{j}\omega$

Finite difference operator

$$\mathrm{D_d}f(x) = f(x) - f(x-1) \qquad \mathrm{D_d} \overset{\mathcal{F}}{\longleftrightarrow} 1 - \mathrm{e}^{-\mathrm{j}\omega}$$

$$= (\beta_+^0 * \mathrm{D}f)(x)$$

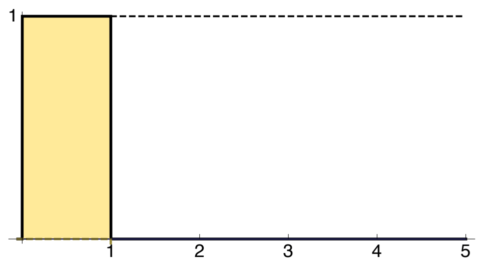
B-spline of degree 0

$$\beta_+^0(x) = \mathrm{D_d}\mathrm{D}^{-1}\delta(x) = \mathrm{D_d}\mathbb{1}_+(x)$$

$$\updownarrow$$

$$\hat{\beta}_+^0(\omega) = \frac{1 - \mathrm{e}^{-\mathrm{j}\omega}}{\mathrm{j}\omega}$$

$$\beta_+^0(x) = \mathbb{1}_+(x) - \mathbb{1}_+(x-1)$$



Derivative operator: Native space (Sobolev)

Lebesgue space of finite-energy functions:

$$L_2(\mathbb{R}) = \{f : \mathbb{R} \to \mathbb{R} \quad \text{s.t.} \quad \|f\|_{L_2} \triangleq \left(\int_{\mathbb{R}} |f(x)|^2 \mathrm{d}x\right)^{1/2} < \infty\}$$

- Tikhonov regularization

$$\mathcal{H}_{\mathrm{D}} = \{f : \mathbb{R} \to \mathbb{R} \quad \text{s.t.} \quad \|\mathrm{D}f\|_{L_2}^2 < \infty\}$$

Hilbert space equipped with inner product: $\langle f, g\rangle_{\mathrm{D}} = \langle \mathrm{D}f, \mathrm{D}g\rangle + f(0)g(0)$

Generalization: $\langle f, g\rangle_{\mathrm{D}} = \langle \mathrm{D}f, \mathrm{D}g\rangle + \langle \phi_1, f\rangle\langle \phi_1, g\rangle$

subject to $\langle \phi_1, 1\rangle = 1$ (biorthogonality)

Stabilized inverse operator

Motivation: Unstable differential equation

$$\mathrm{D}s = w \qquad \Rightarrow \qquad s(x) = \mathrm{D}^{-1}\{w\}(x) = (u * w)(x) = \int_{-\infty}^{x} w(\tau)\mathrm{d}\tau \text{ ???}$$

Problem: D^{-1} (integrator) is unstable on $\mathcal{H}_{\mathrm{D}}$ (Native space)

- Introduction of suitable boundary condition(s)

$$\mathrm{D}s = w \quad \text{s.t.} \quad \langle \phi_1, s\rangle = s(0) = 0$$

Solution: $s(x) = \mathrm{D}_{\phi_1}^{-1}\{w\}(x) = \int_{\mathbb{R}} g_{\phi_1}(x, y)w(y)\mathrm{d}y \quad = \int_{0}^{x} w(\tau)\mathrm{d}\tau$

Schwartz kernel (generalized impulse response) of inverse operator :

$$g_{\phi_1}(x, y) = \rho_{\mathrm{D}}(x - y) - p_1(x)q_1(y)$$

with $q_1(y) = \langle \phi_1, \rho_{\mathrm{D}}(\cdot - y)\rangle = \rho_{\mathrm{D}}(-y)$

(see Theorem XXX, RKHS notes)

Synthesis of (non-uniform) splines

Take w to be a series of Dirac impulses (innovation):

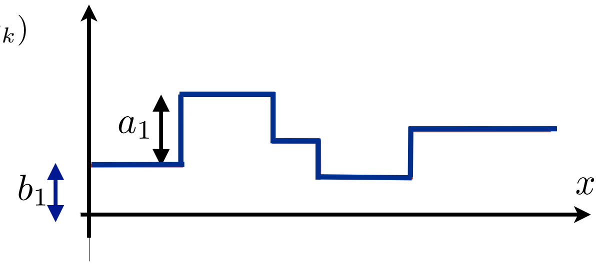
$$w = \sum_k a_k \delta(\cdot - x_k) \qquad \text{(simplifiying assumption } x_k > 0\text{)}$$

- Formal solution of differential equation

$$\mathrm{D}s = w \quad \text{s.t.} \quad \langle \phi_1, s \rangle = s(0) = b_1$$

Solution: $s(x) = b_1 + \mathrm{D}_{\phi_1}^{-1}\{w\}(x) = b_1 + \int_{\mathbb{R}} g_{\phi_1}(x,y) w(y) \mathrm{d}y = b_1 + \int_0^x w(\tau) \mathrm{d}\tau$

$$\Rightarrow \qquad s(x) = b_1 + \sum_k a_k u(x - x_k)$$



Synthesis of Gaussian process

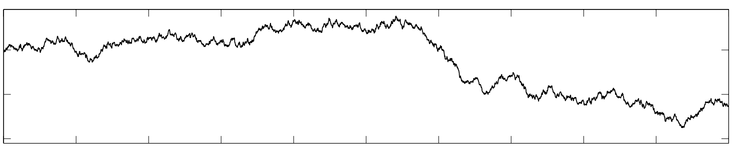
Take w to be a Gaussian white noise process

- Solution of stochastic differential equation

$$\mathrm{D}s = w \quad \text{s.t.} \quad \langle \phi_1, s \rangle = s(0) = 0$$

Solution: $s(x) = \mathrm{D}_{\phi_1}^{-1}\{w\}(x) = \int_{\mathbb{R}} g_{\phi_1}(x,y) w(y) \mathrm{d}y = \int_0^x w(\tau) \mathrm{d}\tau$

Yields Brownian motion (Wiener in 1923)



Autocorrelation / reproducing kernel

Brownian motion: $s = \mathrm{D}_{\phi_1}^{-1}\{w\}$ where w is a white Gaussian noise

with $\phi_1 = \delta$

- Autocorrelation function

$$a_s(x,y) = \mathbb{E}\{s(x)s(y)\} = \mathbb{E}\left\{\mathrm{D}_{\phi_1}^{-1}\{w\}(x)\mathrm{D}_{\phi_1}^{-1}\{w\}(y)\right\}$$

- Functional equivalence

$a_s : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is the reproducing kernel of the Hilbert space (RKHS)

$$\begin{aligned}\mathcal{H}_{\mathrm{D},\phi_1} &= \left\{f \in \mathcal{H}_{\mathrm{D}} : \langle \phi_1, f\rangle = 0\right\} \\ &= \{f = \mathrm{D}_{\phi_1}^{-1} w : w \in L_2(\mathbb{R})\}\end{aligned}$$

$$\Rightarrow \quad a_s(x,y) = h_{\mathrm{D}}(x,y) = \tfrac{1}{2}\big(|x| + |y| - |x-y|\big)$$

L_2 representer theorem (Schoenberg 1964, Kimeldorf-Wahba 1971)

- Variational interpolation problem (smoothing splines)

(P2) $$\arg\min_{f \in \mathcal{H}_{\mathrm{D}}} \left(\sum_{m=1}^{M} |y_m - f(x_m)|^2 + \lambda \|\mathrm{D}f\|_{L_2}^2\right)$$

Convex loss function: $F : \mathbb{R}^M \times \mathbb{R}^M \to \mathbb{R}$

Sample values: $\boldsymbol{f} = \big(f(x_1), \dots, f(x_M)\big)$

(P2') $$\arg\min_{f \in \mathcal{H}} \left(F(\boldsymbol{y}, \boldsymbol{f}) + \lambda \|\mathrm{D}f\|_{L_2}^2\right)$$

(Schölkopf-Smola 2001)

Green's function of $\mathrm{D}^*\mathrm{D} = -\mathrm{D}^2$: $\rho_{\mathrm{D}^*\mathrm{D}}(x) = \mathcal{F}^{-1}\left\{\frac{1}{\omega^2}\right\}(x) = -\frac{1}{2}|x|$

Representer theorem (theory of RKHS)

The generic parametric form of the solution of (P2') is

$$f(x) = b_1 + \sum_{m=1}^{M} a_m |x - x_m|$$

Supports the theory of SVM, kernel methods, etc.

Other examples of operators

See chapters 5 and 6

- One-dimensional operators

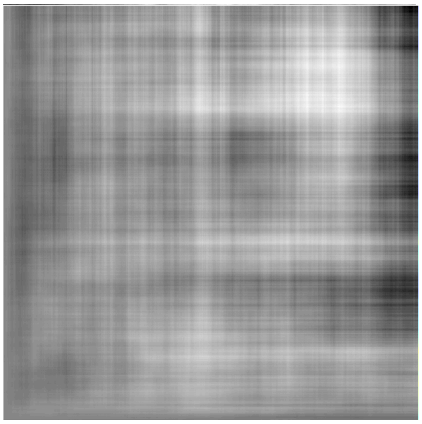
  - Pure derivatives: $\mathrm{D}^m \quad \xleftrightarrow{\mathcal{F}} \quad (\mathrm{j}\omega)^n$

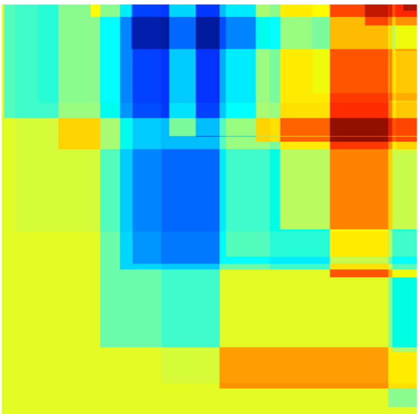
  - Ordinary differential operators: $\mathrm{L} = \mathrm{P}_{\boldsymbol{\alpha}} = \mathrm{P}_{\alpha_1} \cdots \mathrm{P}_{\alpha_N}$
    with $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_N) \in \mathbb{C}^N$ and $\mathrm{P}_{\alpha_n} = \mathrm{D} - \alpha_n \mathrm{Id}$

  - Fractional derivatives: $\partial_\tau^\gamma \quad \xleftrightarrow{\mathcal{F}} \quad (\mathrm{j}\omega)^{\frac{\gamma}{2}+\tau}(-\mathrm{j}\omega)^{\frac{\gamma}{2}-\tau}$

- Multi-dimensional operators

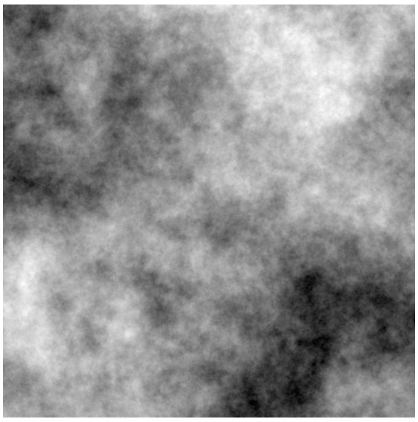
  - Separable operators: $\hat{L}(\omega_1, \dots, \omega_d) = \prod_{i=1}^d \hat{P}_{\boldsymbol{\alpha}}(\omega_i)$
  - Fractional Laplacian: $(-\Delta)^{\frac{\gamma}{2}} \quad \xleftrightarrow{\mathcal{F}} \quad \|\boldsymbol{\omega}\|^\gamma$





Gaussian

$$\mathrm{L} = \mathrm{D}_{r_1}\mathrm{D}_{r_2} \quad \xleftrightarrow{\mathcal{F}} \quad (\mathrm{j}\omega_1)(\mathrm{j}\omega_2)$$



$$\mathrm{L} = (-\Delta)^{1/2} \quad \xleftrightarrow{\mathcal{F}} \quad \|\boldsymbol{\omega}\|$$

Sparse encounters
(Anza-Borrego desert circa 1995)

Derivative operator: Native space (BV)

Lebesgue space of integrable functions:

$$L_1(\mathbb{R}) = \{f : \mathbb{R} \to \mathbb{R} \quad \text{s. t.} \quad \|f\|_{L_1} \triangleq \int_{\mathbb{R}} |f(x)|\mathrm{d}x < \infty\}$$

Space of signed and bounded Borel measures:

$$\mathcal{M}(\mathbb{R}) = \left\{w : \|w\|_{\mathcal{M}} \triangleq \sup_{\|\varphi\|_\infty \le 1} \langle w, \varphi\rangle < \infty\right\}$$

$\mathcal{M}(\mathbb{R})$ as the weak extension of $L_1(\mathbb{R})$: $w \in L_1(\mathbb{R}) \Rightarrow \|w\|_{\mathcal{M}} = \|f\|_{L_1}$

$L_1(\mathbb{R}) \subseteq \mathcal{M}(\mathbb{R})$; $\delta(\cdot - x_0) \notin L_1(\mathbb{R})$, while $\delta(\cdot - x_0) \in \mathcal{M}(\mathbb{R})$ with $\|\delta(\cdot - x_0)\|_{\mathcal{M}} = 1$

- Total-variation regularization

$$\mathcal{M}_{\mathrm{D}}(\mathbb{R}) = \{f : \mathbb{R} \to \mathbb{R} \quad \text{s. t.} \quad \|\mathrm{D}f\|_{\mathcal{M}} < \infty\}$$

Banach space equipped with norm: $\|f\|_{\mathrm{D}} = \|\mathrm{D}f\|_{\mathcal{M}} + |f(0)|^2$

TV representer theorem

Linear measurement operator $\boldsymbol{\nu} : \mathcal{M}_{\mathrm{D}} \to \mathbb{R}^M$

$$s \mapsto \boldsymbol{\nu}(s) = (\langle \nu_1, s\rangle, \dots, \langle \nu_M, s\rangle)$$

- General linear inverse problem with TV regularization

(P1) $$\arg \min_{f \in \mathcal{M}_{\mathrm{D}}} \left(\sum_{m=1}^{M} |y_m - \langle \nu_m, f\rangle|^2 + \lambda \|\mathrm{D}f\|_{\mathcal{M}} \right)$$

Convex loss function: $F : \mathbb{R}^M \times \mathbb{R}^M \to \mathbb{R}$

(P1') $$\arg \min_{f \in \mathcal{M}_{\mathrm{D}}} \left(F(\boldsymbol{y}, \boldsymbol{\nu}(f)) + \lambda \|\mathrm{D}f\|_{\mathcal{M}}\right)$$

Representer theorem for total variation regularization

The extreme points of (P1') are necessarily non-uniform splines of the form

$$f(x) = b_1 + \sum_{k=1}^{K} a_k u(x - x_k) \text{ with } K < M \text{ data-depend knots } (x_k).$$

(Unser-Fageot-ward, *SIAM Review*, 2017)

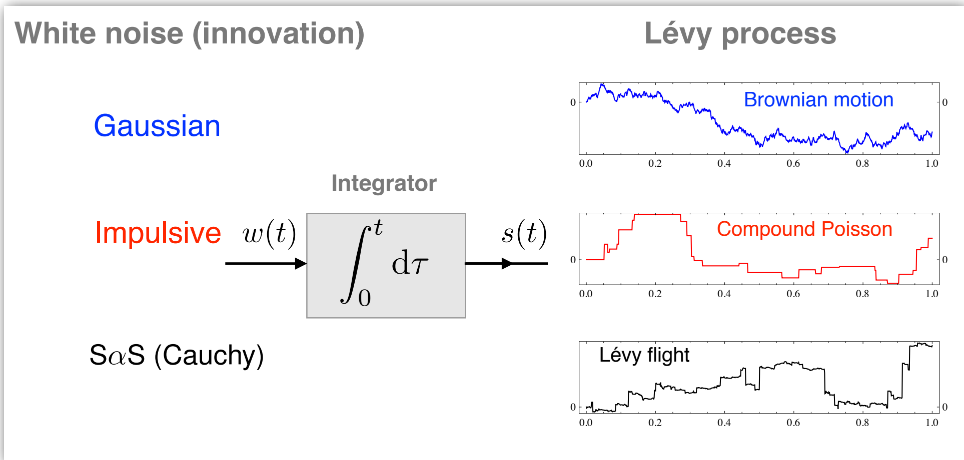
Synthesis of Sparse process

Take w to be a white noise Lévy noise

- Solution of stochastic differential equation

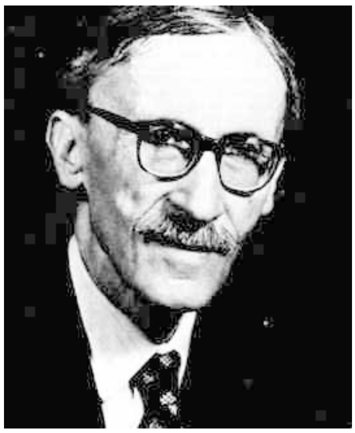
$$\mathrm{D}s = w \quad \text{s.t.} \quad \langle \phi_1, s\rangle = s(0) = 0$$

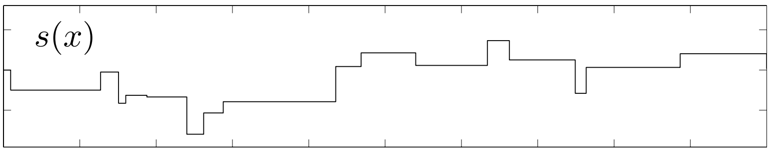
Solution: $$s(x) = \mathrm{D}_{\phi_1}^{-1}\{w\}(x) = \int_{\mathbb{R}} g_{\phi_1}(x,y)w(y)\mathrm{d}y = \int_0^x w(\tau)\mathrm{d}\tau$$



Lévy processes

Constructed by Paul Lévy in the 1930's





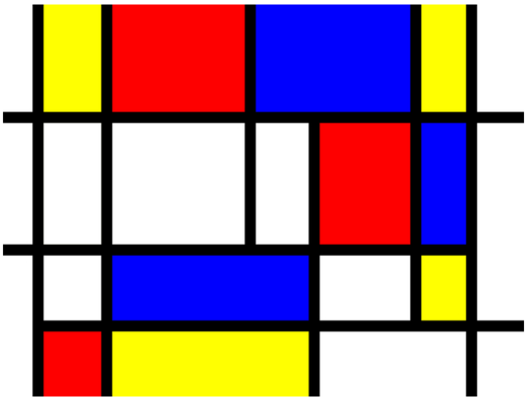
- Non-Gaussian generalization of Wiener process
- Non-stationary
- Self-similar: "$1/\omega$" spectral decay
- Independent increments

  $u[k] = s(k) - s(k-1)$: i.i.d. infinitely divisible (heavy tailed)

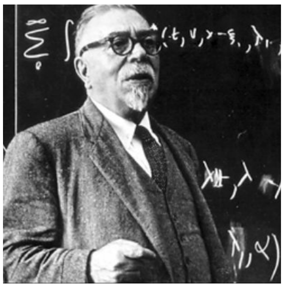
Example: compound Poisson process (piecewise-constant, with random jumps)

☒ **Archetype of a "sparse" random signal**

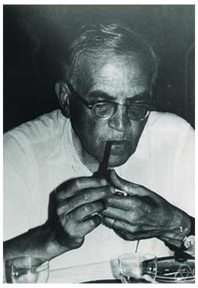
Gaussian vs. Sparse



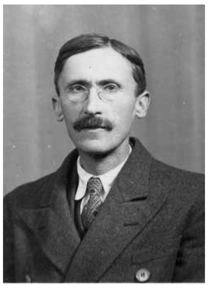
Fourier analysis   Splines   Wavelet analysis



Norbert Wiener



Isaac Schoenberg



Paul Lévy