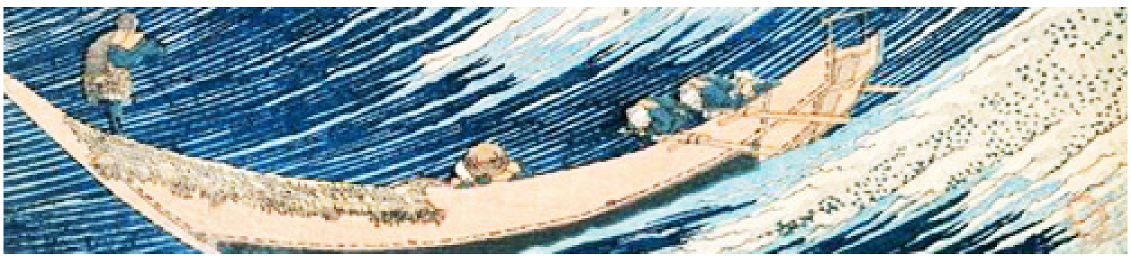
Mathematical Foundations of Data Sciences



Gabriel Peyré
CNRS & DMA
École Normale Supérieure
gabriel.peyre@ens.fr
www.gpeyre.com
www.numerical-tours.com

November 2, 2017

Chapter 1

Shannon Theory

The main reference is [29].

1.1 Analog vs. Discrete Signals

To develop numerical tools and analyze their performances, the mathematical modeling is usually done over a continuous setting. An analog signal is a 1D function $f_0 \in \mathrm{L}^2([0,1])$ where $[0,1]$ denotes the domain of acquisition, which might for instance be time. An analog image is a 2D function $f_0 \in \mathrm{L}^2([0,1]^2)$ where the unit square $[0,1]^2$ is the image domain.

Although these notes are focussed on the processing of sounds and natural images, most of the methods extend to multi-dimensional datasets, which are higher dimensional mappings

$$f_0 : [0,1]^d \to [0,1]^s$$

where d is the dimensionality of the input space ($d = 1$ for sound and $d = 2$ for images) whereas s is the dimensionality of the feature space. For instance, gray scale images corresponds to $(d = 2, s = 1)$, videos to $(d = 3, s = 1)$, color images to $(d = 2, s = 3)$ where one has three channels (R, G, B). One can even consider multi-spectral images where $(d = 2, s \gg 3)$ that is made of a large number of channels for different light wavelengths. Figures 1.1 and 1.2 show examples of such data.

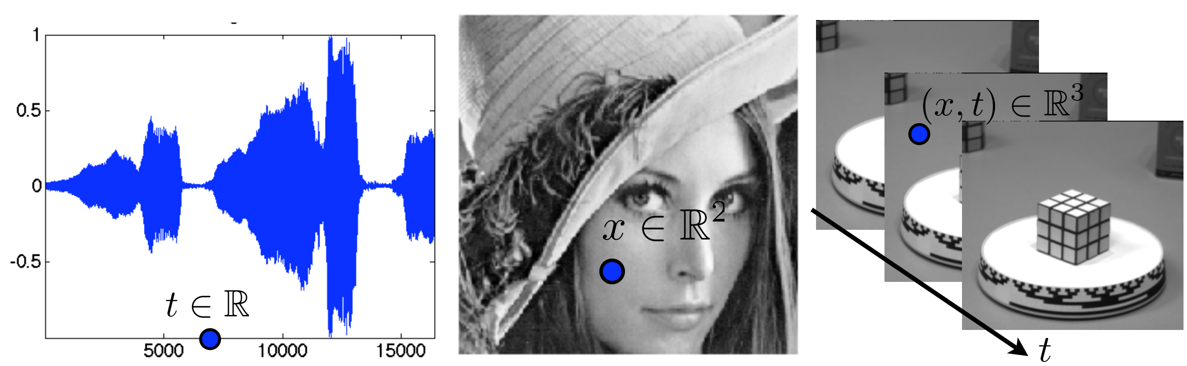


Figure 1.1: Examples of sounds ($d = 1$), image ($d = 2$) and videos ($d = 3$).

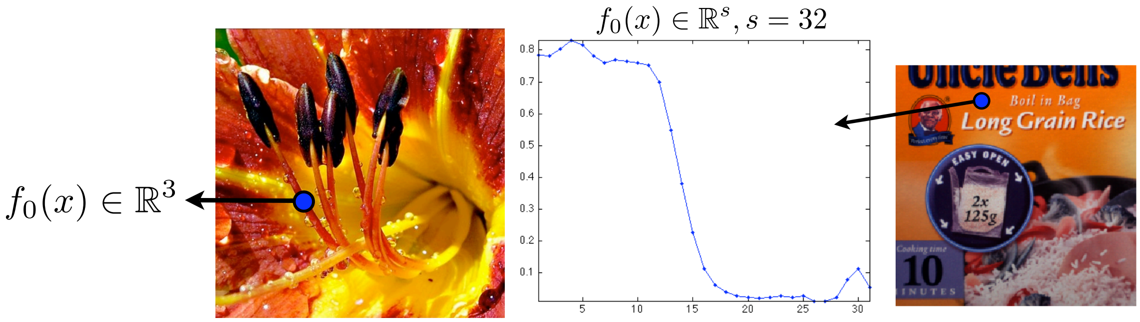


Figure 1.2: Example of color image $s = 3$ and multispectral image ($s = 32$).

1.1.1 Acquisition and Sampling

Signal acquisition is a low dimensional projection of the continuous signal performed by some hardware device. This is for instance the case for a microphone that acquires 1D samples or a digital camera that acquires 2D pixel samples. The sampling operation thus corresponds to mapping from the set of continuous functions to a discrete finite dimensional vector with N entries.

$$f_0 \in \mathrm{L}^2([0,1]^d) \mapsto f \in \mathbb{C}^N$$

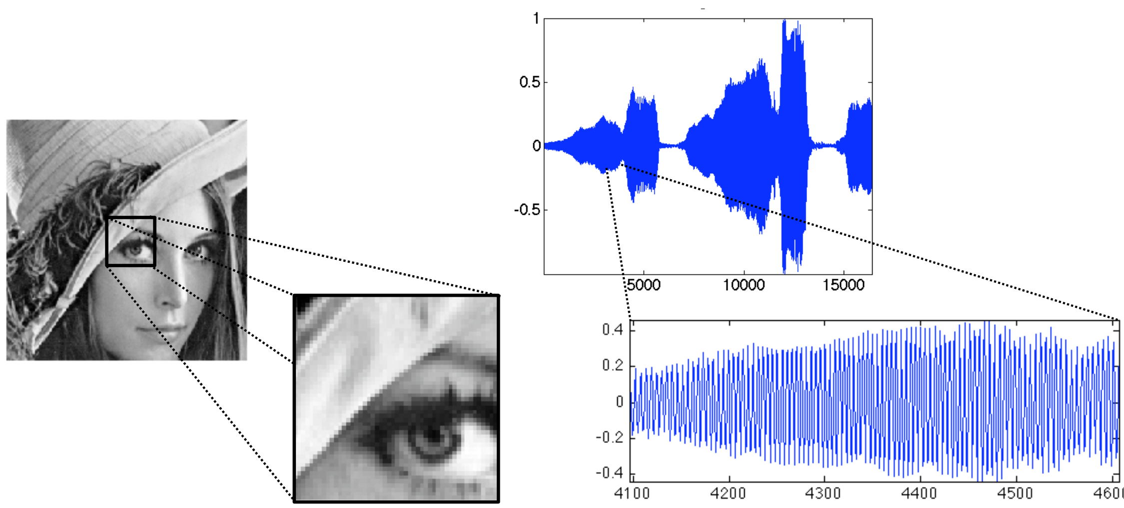


Figure 1.3: Image and sound discretization.

Figure 1.3 shows examples of discretized signals.

1.1.2 Linear Translation Invariant Sampler

A translation invariant sampler performs the acquisition as an inner product between the continuous signal and a constant impulse response h translated at the sample location

$$f[n] = \int_{-S/2}^{S/2} f_0(x)h(n/N - x)\mathrm{d}x = f_0 \star h(n/N). \tag{1.1}$$

The precise shape of $h(x)$ depends on the sampling device, and is usually a smooth low pass function that is maximal around $x = 0$. The size S of the sampler determines the precision of the sampling device, and is usually of the order of $1/N$ to avoid blurring (if S is too large) or aliasing (if S is too small).

Section **??** details how to reverse the sampling operation in the case where the function is smooth.

1.2 Shannon Sampling Theorem

Reminders about Fourier transform. For $f \in L^1(\mathbb{R})$, its Fourier transform is defined as

$$\forall\, \omega \in \mathbb{R}, \quad \hat{f}(\omega) \stackrel{\text{def.}}{=} \int_{\mathbb{R}} f(x) e^{-\mathrm{i}x\omega} \mathrm{d}x. \tag{1.2}$$

One has $\|\hat{f}\|^2 = (2\pi)^{-1}\|f\|^2$, so that $f \mapsto \hat{f}$ can be extended by continuity to $L^2(\mathbb{R})$, which corresponds to computing $\hat{f}$ as a limit when $T \to +\infty$ of $\int_{-T}^{T} f(x)e^{-\mathrm{i}x\omega}\mathrm{d}x$. When $\hat{f} \in L^1(\mathbb{R})$, one can invert the Fourier transform so that

$$f(x) = \int_{\mathbb{R}} \hat{f}(\omega) e^{\mathrm{i}x\omega} \mathrm{d}\omega, \tag{1.3}$$

which shows in particular that f is continuous with vanishing limits at $\pm\infty$.

The Fourier transform $\mathcal{F} : f \mapsto \hat{f}$ exchanges regularity and decay. For instance, if $f \in C^p(\mathbb{R})$ with an integrable Fourier transform, then $\mathcal{F}(f^{(p)})(\omega) = (\mathrm{i}\omega)^{-p}\hat{f}(\omega)$ so that $|\hat{f}(\omega)| = O(1/|\omega|^p)$. Conversely,

$$\int_{\mathbb{R}} (1 + |\omega|)^{-p} |\hat{f}(\omega)| \mathrm{d}\omega < +\infty \quad \Longrightarrow \quad f \in C^p(\mathbb{R}). \tag{1.4}$$

Reminders about Fourier series. We denote $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$ the torus. A function $f \in L^2(\mathbb{T})$ is 2π-periodic, and can be viewed as a function $f \in L^2([0, 1])$ (beware that this means that the boundary points are glued together), and its Fourier coefficients are

$$\forall\, n \in \mathbb{Z}, \quad \hat{f}_n \stackrel{\text{def.}}{=} \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-\mathrm{i}xn} \mathrm{d}x.$$

This formula is equivalent to the computation of an inner-product $\hat{f}_n = \langle f,\, e_n \rangle$ for the inner-product $\langle f,\, g \rangle \stackrel{\text{def.}}{=} \frac{1}{2\pi}\int_{\mathbb{T}} f(x)\bar{g}(x)\mathrm{d}x$. For this inner product, $(e_n)_n$ is orthonormal and is actually an Hilbert basis, meaning that one reconstruct with the following converging series

$$f = \sum_{n \in \mathbb{Z}} \langle f,\, e_n \rangle e_n \tag{1.5}$$

which means $\|f - \sum_{n=-N}^{N} \langle f,\, e_n \rangle e_n\|_{L^2(\mathbb{T})} \to 0$ for $N \to +\infty$. The pointwise convergence of (1.5), and is ensured (and there is normal convergence) when for instance $f \in C^3(\mathbb{T})$.

Poisson formula. The poisson formula connects the Fourier transform and the Fourier series to sampling and periodization operators. For some function $\hat{f}(\omega)$ defined on $\mathbb{R}$, its periodization reads

$$\hat{f}_P(\omega) \stackrel{\text{def.}}{=} \sum_n f(\omega - 2\pi n). \tag{1.6}$$

This formula makes sense if $\hat{f} \in L^1(\mathbb{R})$, and in this case $\|\hat{f}_P\|_{L^1(\mathbb{T})} \leqslant \|\hat{f}\|_1$ The Poisson formula, state in Proposition 1 bellow, corresponds to proving that the following diagram

$$\begin{array}{cccccc} & f(x) & \xrightarrow{\mathcal{F}} & \hat{f}(\omega) & \\ \text{sampling} & \downarrow & & \downarrow & \text{periodization} \\ & (f(n))_n & \xrightarrow{\text{Fourier serie}} & \sum_n f(n)e^{-\mathrm{i}\omega n} & \end{array}$$

is actually commutative.

Proposition 1 (Poisson formula)**.** *Assume that $\hat{f}$ has compact support and that $|f(x)| \leqslant C(1+|x|)^{-3}$ for some C. Then one has*

$$\forall\, \omega \in \mathbb{R}, \quad \sum_n f(n) e^{-\mathrm{i}\omega n} = \hat{f}_P(\omega). \tag{1.7}$$

Proof. Since $\hat{f}$ is compactly supported, $\hat{f}_P$ is well defined (it involves only a finite sum) and since f has fast decay, using (1.4), $\hat{f}_P$ is C^1. It is thus the sum of its Fourier transform

$$\hat{f}_P(\omega) = \sum_k c_k e^{\mathrm{i}k\omega}, \tag{1.8}$$

where

$$c_k = \frac{1}{2\pi}\int_0^{2\pi} \hat{f}_P(\omega) e^{-\mathrm{i}k\omega} \mathrm{d}\omega = \frac{1}{2\pi}\int_0^{2\pi} \sum_n f(x-2\pi n) e^{-\mathrm{i}k\omega} \mathrm{d}\omega.$$

One has

$$\int_0^{2\pi} \sum_n |f(x-2\pi n) e^{-\mathrm{i}k\omega}| \mathrm{d}\omega = \int_{\mathbb{R}} |f|$$

which is bounded because $\hat{f} \in L^1(\mathbb{R})$ (it has a compact support and is C^1), so one can exchange the sum and integral

$$c_k = \sum_n \frac{1}{2\pi}\int_0^{2\pi} f(x-2\pi n) e^{-\mathrm{i}k\omega} \mathrm{d}\omega = \frac{1}{2\pi}\int_{\mathbb{R}} f(x) e^{-\mathrm{i}k\omega} \mathrm{d}\omega = f(-k)$$

where we used the inverse Fourier transform formula (1.3), which is legit because $\hat{f} \in L^1(\mathbb{R})$. ☐

Shannon theorem. Shannon sampling theorem state a sufficient condition ensuring that the sampling operator $f \mapsto (f(ns))_n$ is invertible for some sampling step size $s > 0$. It require that $\mathrm{supp}(\hat{f}) \subset [-\pi/s, \pi/s]$, which, thanks to formula (1.3), implies that $\hat{f}$ is C^∞ (in fact it is even analytic).

Theorem 1. *If $|f(x)| \leqslant C(1+|x|)^{-3}$ for some C and $\mathrm{supp}(\hat{f}) \subset [-\pi/s, \pi/s]$, then one has*

$$\forall\, x \in \mathbb{R}, \quad f(x) = \sum_n f(ns)\,\mathrm{sinc}(x/s - n) \quad \textit{where} \quad \mathrm{sinc}(u) = \frac{\sin(\pi u)}{\pi u} \tag{1.9}$$

with uniform convergence.

Proof. The change of variable $g = f(s\cdot)$ results in $\hat{g} = s\hat{f}(s\cdot)$ so that we can restrict our attention to $s = 1$. The compact support hypothesis implies $\hat{f}(\omega) = 1_{[-\pi,\pi]}(\omega)\hat{f}_P(\omega)$. Combining the inversion formula (1.3) with Poisson formula (1.8)

$$f(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \hat{f}_P(\omega) e^{\mathrm{i}\omega x} \mathrm{d}\omega = \frac{1}{2\pi}\int_{-\pi}^{\pi} \sum_n f(n) e^{\mathrm{i}\omega(x-n)} \mathrm{d}\omega.$$

Since f has fast decay, $\int_{-\pi}^{\pi} \sum_n |f(n) e^{\mathrm{i}\omega(x-n)}| \mathrm{d}\omega = \sum_n |f(n)| < +\infty$, so that one can exchange summation and integration and obtain

$$f(x) = \sum_n f(n) \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{\mathrm{i}\omega(x-n)} \mathrm{d}\omega = \sum_n f(n)\,\mathrm{sinc}(x-n).$$

☐

1.3 Shannon Source Coding Theorem

We consider an alphabet $(x_1, \ldots, x_K)$ of K symbols, and assume at our disposal some probability distribution over this alphabet, which is just an histogram $p = (p_1, \ldots, p_K) \in \mathbb{R}_+^K$ in the simplex, i.e. $\sum_k p_k = 1$.

The entropy of such an histogram is

$$H(p) \stackrel{\text{def.}}{=} -\sum_k p_k \log_2(p_k)$$

with the convention $O \log_2(0) = 0$.

Lemma 1. *One has*

$$0 \leqslant H(p) \leqslant \log_2(K).$$

Proof. We consider the following constrained optimization problem

$$\min_p \left\{ f(p) \; ; \; g(p) = \sum_k p_k = 1 \right\}$$

where $f = -H$. According to the linked extrema theorem, at an optimum $p^\star$, $\nabla f(p^\star) = \lambda \nabla g(p^\star)$ for some $\lambda \in \mathbb{R}$, so that here $\log(p_k^\star) + 1 = \lambda$, i.e. $p_k^\star = c$ is constant, and since $\sum_k p_k^\star = 1$, one has $p_k^* = 1/K$ and thus $H(p) = \log_2(K)$. $\square$

A code $c_k = c(x_k)$ associate to each symbol x_k a code word $c_k \in \{0,1\}^{\mathbb{N}}$ with a varying length $|c_k| \in \mathbb{N}^*$. We denote the average length associated to this code as

$$L(c) \stackrel{\text{def.}}{=} \sum_k p_k |c_k|.$$

A prefix code $c_k = c(x_k)$ is such that no word c_k is the beginning of another word c'_k. This is equivalent to be able to embed the $(c_k)_k$ as leaves of a binary tree T, with the code being output of a traversal from root to leaves (with a convention that going to a left (resp. right) child output a 0 (resp. a 1). We denote $c = \text{Leaves}(T)$ such prefix property. The following fundamental lemma describes the set of prefix code using an inequality.

Lemma 2 (Kraft inequality)**.** *(i) For a code c, if there exists a tree T such that $c = \text{Leaves}(T)$ then*

$$\sum_k 2^{-|c_k|} \leqslant 1. \tag{1.10}$$

(ii) Conversely, if $(\ell_k)_k$ are such that

$$\sum_k 2^{-\ell_k} \leqslant 1 \tag{1.11}$$

then there exists a code $c = \text{Leaves}(T)$ such that $|c_k| = \ell_k$.

Proof. $\Rightarrow$ We suppose $c = \text{Leaves}(T)$. We denote $m = \max_k |c_k|$ and consider the full binary tree. Bellow each c_k, one has a sub-tree of height $m - |c_k|$. This sub-tree has $2^{m-|c_k|}$ leaves. Since all these sub-trees do not overlap, the total number of leaf do not exceed the total number of leaves 2^m of the full binary tree, hence

$$\sum_k 2^{m-|c_k|} \leqslant 2^m,$$

hence (1.10).

$\Leftarrow$ Conversely, we assume (1.10) holds. Without loss of generality, we assume that $|c_1| \leqslant \ldots \leqslant |c_K|$. We start by putting a sub-tree of height $2^{m-|c_1|}$. Since the second tree is smaller, one can put it immediately aside, and continue this way. Since $\sum_k 2^{m-|c_k|} \leqslant 2^m$, this ensure that we can stack side-by-side all these sub-tree, and this defines a proper sub-tree of the full binary tree. $\square$

We now are ready to state and prove Shannon theory for entropic coding.

Theorem 2. *(i) If $c = Leaves(T)$ for some tree T, then*

$$L(c) \geqslant H(p).$$

(ii) Conversely, there exists a code c with $c = Leaves(T)$ such that

$$L(c) \leqslant H(p) + 1.$$

Proof. First, we consider the following optimization problem

$$\min_{\ell=(\ell_k)_k} \left\{ f(\ell) \stackrel{\text{def.}}{=} \sum_k \ell_k p_k \; ; \; g(\ell) \stackrel{\text{def.}}{=} \sum_k 2^{-\ell_k} \leqslant 1 \right\}. \tag{1.12}$$

We fist show that at an optimal $\ell^\star$, the constraint is saturated, i.g. $g(\ell^\star) = 1$. Indeed, if $g(\ell^\star) = 2^{-u} < 1$, with $u > 0$, we define $\ell'_k \stackrel{\text{def.}}{=} \ell^\star_k - u$, which satisfies $g(\ell') = 1$ and also $f(\ell') = \sum_k (\ell_k - u) p_k < f(\ell^\star)$, which is a contradiction. So we can restrict in (1.12) the constraint to $g(\ell) = 1$ and apply the linked extra theorem, which shows that necessarily, there exists $\lambda \in \mathbb{R}$ with $\nabla f(\ell^\star) = \nabla g(\ell^\star)$, i.e. $(p_k)_k = -\lambda \ln(2)(2^{-\ell^\star_k})_k$. Since $\sum_k p_k = \sum_k 2^{-\ell^\star_k} = 1$, we deduce that $\ell^\star_k = -\log(p_k)$.

(i) If $c = \text{Leave}(T)$, the by Kraft inequality (1.10), necessarily $\ell_k = |c_k|$ satisfy the constraints of (1.12), and thus $H(p) = f(\ell^\star) \leqslant f(\ell) = L(\ell)$.

(ii) We define $\ell_k \stackrel{\text{def.}}{=} \lceil -\log_2(p_k) \rceil \in \mathbb{N}^*$. Then $\sum_k 2^{-\ell_k} \leqslant \sum_k 2^{\log_2(p_k)} = 1$, so that these lengths satisfy (1.11). Thanks to Proposition 2 (ii), there thus exists a prefix code c with $|c_k| = \lceil -\log_2(p_k) \rceil$. Furthermore

$$L(c) = \sum_k p_k \lceil -\log_2(p_k) \rceil \leqslant \sum_k p_k(-\log_2(p_k) + 1) = H(p) + 1.$$

□

Bibliography

[1] P. Alliez and C. Gotsman. Recent advances in compression of 3d meshes. In N. A. Dodgson, M. S. Floater, and M. A. Sabin, editors, *Advances in multiresolution for geometric modelling*, pages 3–26. Springer Verlag, 2005.

[2] P. Alliez, G. Ucelli, C. Gotsman, and M. Attene. Recent advances in remeshing of surfaces. In *AIM@SHAPE repport*. 2005.

[3] Amir Beck. *Introduction to Nonlinear Optimization: Theory, Algorithms, and Applications with MATLAB*. SIAM, 2014.

[4] Stephen Boyd, Neal Parikh, Eric Chu, Borja Peleato, and Jonathan Eckstein. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends® in Machine Learning*, 3(1):1–122, 2011.

[5] Stephen Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[6] E. Candès and D. Donoho. New tight frames of curvelets and optimal representations of objects with piecewise C^2 singularities. *Commun. on Pure and Appl. Math.*, 57(2):219–266, 2004.

[7] E. J. Candès. The restricted isometry property and its implications for compressed sensing. *Compte Rendus de l'Académie des Sciences*, Serie I(346):589–592, 2006.

[8] E. J. Candès, L. Demanet, D. L. Donoho, and L. Ying. Fast discrete curvelet transforms. *SIAM Multiscale Modeling and Simulation*, 5:861–899, 2005.

[9] A. Chambolle. An algorithm for total variation minimization and applications. *J. Math. Imaging Vis.*, 20:89–97, 2004.

[10] Antonin Chambolle, Vicent Caselles, Daniel Cremers, Matteo Novaga, and Thomas Pock. An introduction to total variation for image analysis. *Theoretical foundations and numerical methods for sparse recovery*, 9(263-340):227, 2010.

[11] Antonin Chambolle and Thomas Pock. An introduction to continuous optimization for imaging. *Acta Numerica*, 25:161–319, 2016.

[12] S.S. Chen, D.L. Donoho, and M.A. Saunders. Atomic decomposition by basis pursuit. *SIAM Journal on Scientific Computing*, 20(1):33–61, 1999.

[13] F. R. K. Chung. Spectral graph theory. *Regional Conference Series in Mathematics, American Mathematical Society*, 92:1–212, 1997.

[14] Philippe G Ciarlet. Introduction à l'analyse numérique matricielle et à l'optimisation. 1982.

[15] P. L. Combettes and V. R. Wajs. Signal recovery by proximal forward-backward splitting. *SIAM Multiscale Modeling and Simulation*, 4(4), 2005.

[16] P. Schroeder et al. D. Zorin. Subdivision surfaces in character animation. In *Course notes at SIGGRAPH 2000*, July 2000.

[17] I. Daubechies, M. Defrise, and C. De Mol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Commun. on Pure and Appl. Math.*, 57:1413–1541, 2004.

[18] I. Daubechies and W. Sweldens. Factoring wavelet transforms into lifting steps. *J. Fourier Anal. Appl.*, 4(3):245–267, 1998.

[19] D. Donoho and I. Johnstone. Ideal spatial adaptation via wavelet shrinkage. *Biometrika*, 81:425–455, Dec 1994.

[20] Heinz Werner Engl, Martin Hanke, and Andreas Neubauer. *Regularization of inverse problems*, volume 375. Springer Science & Business Media, 1996.

[21] M. Figueiredo and R. Nowak. An EM Algorithm for Wavelet-Based Image Restoration. *IEEE Trans. Image Proc.*, 12(8):906–916, 2003.

[22] M. S. Floater and K. Hormann. Surface parameterization: a tutorial and survey. In N. A. Dodgson, M. S. Floater, and M. A. Sabin, editors, *Advances in multiresolution for geometric modelling*, pages 157–186. Springer Verlag, 2005.

[23] Simon Foucart and Holger Rauhut. *A mathematical introduction to compressive sensing*, volume 1. Birkhäuser Basel, 2013.

[24] I. Guskov, W. Sweldens, and P. Schröder. Multiresolution signal processing for meshes. In Alyn Rockwood, editor, *Proceedings of the Conference on Computer Graphics (Siggraph99)*, pages 325–334. ACM Press, August8–13 1999.

[25] A. Khodakovsky, P. Schröder, and W. Sweldens. Progressive geometry compression. In *Proceedings of the Computer Graphics Conference 2000 (SIGGRAPH-00)*, pages 271–278, New York, July 23–28 2000. ACMPress.

[26] L. Kobbelt. $\sqrt{3}$ subdivision. In Sheila Hoffmeyer, editor, *Proc. of SIGGRAPH'00*, pages 103–112, New York, July 23–28 2000. ACMPress.

[27] M. Lounsbery, T. D. DeRose, and J. Warren. Multiresolution analysis for surfaces of arbitrary topological type. *ACM Trans. Graph.*, 16(1):34–73, 1997.

[28] S. Mallat. *A Wavelet Tour of Signal Processing, 3rd edition.* Academic Press, San Diego, 2009.

[29] Stephane Mallat. *A wavelet tour of signal processing: the sparse way.* Academic press, 2008.

[30] D. Mumford and J. Shah. Optimal approximation by piecewise smooth functions and associated variational problems. *Commun. on Pure and Appl. Math.*, 42:577–685, 1989.

[31] Neal Parikh, Stephen Boyd, et al. Proximal algorithms. *Foundations and Trends® in Optimization*, 1(3):127–239, 2014.

[32] Gabriel Peyré. *L'algèbre discrète de la transformée de Fourier.* Ellipses, 2004.

[33] Gabriel Peyré and Marco Cuturi. Computational optimal transport. 2017.

[34] J. Portilla, V. Strela, M.J. Wainwright, and Simoncelli E.P. Image denoising using scale mixtures of Gaussians in the wavelet domain. *IEEE Trans. Image Proc.*, 12(11):1338–1351, November 2003.

[35] E. Praun and H. Hoppe. Spherical parametrization and remeshing. *ACM Transactions on Graphics*, 22(3):340–349, July 2003.

[36] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Phys. D*, 60(1-4):259–268, 1992.

[37] Filippo Santambrogio. Optimal transport for applied mathematicians. *Birkäuser, NY*, 2015.

[38] Otmar Scherzer, Markus Grasmair, Harald Grossauer, Markus Haltmeier, Frank Lenzen, and L Sirovich. *Variational methods in imaging*. Springer, 2009.

[39] P. Schröder and W. Sweldens. Spherical Wavelets: Efficiently Representing Functions on the Sphere. In *Proc. of SIGGRAPH 95*, pages 161–172, 1995.

[40] P. Schröder and W. Sweldens. Spherical wavelets: Texture processing. In P. Hanrahan and W. Purgathofer, editors, *Rendering Techniques '95*. Springer Verlag, Wien, New York, August 1995.

[41] C. E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27(3):379–423, 1948.

[42] A. Sheffer, E. Praun, and K. Rose. Mesh parameterization methods and their applications. *Found. Trends. Comput. Graph. Vis.*, 2(2):105–171, 2006.

[43] Jean-Luc Starck, Fionn Murtagh, and Jalal Fadili. *Sparse image and signal processing: Wavelets and related geometric multiscale analysis*. Cambridge university press, 2015.

[44] W. Sweldens. The lifting scheme: A custom-design construction of biorthogonal wavelets. *Applied and Computation Harmonic Analysis*, 3(2):186–200, 1996.

[45] W. Sweldens. The lifting scheme: A construction of second generation wavelets. *SIAM J. Math. Anal.*, 29(2):511–546, 1997.