Mathematical Foundations of Data Sciences

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

October 7, 2017

Chapter 10

Inverse Problems

The main references for this chapter are [29, 37, 20].

10.1 Inverse Problems Regularization

Increasing the resolution of signals and images requires to solve an ill posed inverse problem. This corresponds to inverting a linear measurement operator that reduces the resolution of the image. This chapter makes use of convex regularization introduced in Chapter **??** to stabilize this inverse problem.

We consider a (usually) continuous linear map $\Phi : \mathcal{S} \to \mathcal{H}$ where $\mathcal{S}$ can be an Hilbert or a more general Banach space. This operator is intended to capture the hardware acquisition process, which maps a high resolution unknown signal $f_0 \in \mathcal{S}$ to a noisy low-resolution obervation

$$y = \Phi f_0 + w \in \mathcal{H}$$

where $w \in \mathcal{H}$ models the acquisition noise. In this section, we do not use a random noise model, and simply assume that $\|w\|_{\mathcal{H}}$ is bounded.

In most applications, $\mathcal{H} = \mathbb{R}^P$ is finite dimensional, because the hardware involved in the acquisition can only record a finite (and often small) number P of observations. Furthermore, in order to implement numerically a recovery process on a computer, it also makes sense to restrict the attention to $\mathcal{S} = \mathbb{R}^N$, where N is number of point on the discretization grid, and is usually very large, $N \gg P$. However, in order to perform a mathematical analysis of the recovery process, and to be able to introduce meaningful models on the unknown f_0, it still makes sense to consider infinite dimensional functional space (especially for the data space $\mathcal{S}$).

The difficulty of this problem is that the direct inversion of Φ is in general impossible or not advisable because Φ^{-1} have a large norm or is even discontinuous. This is further increased by the addition of some measurement noise w, so that the relation $\Phi^{-1}y = f_0 + \Phi^{-1}w$ would leads to an explosion of the noise $\Phi^{-1}w$.

We now gives a few representative examples of forward operators Φ.

Denoising. The case of the identity operator $\Phi = \mathrm{Id}_{\mathcal{S}}$, $\mathcal{S} = \mathcal{H}$ corresponds to the classical denoising problem, already treated in Chapters **??** and **??**.

De-blurring and super-resolution. For a general operator Φ, the recovery of f_0 is more challenging, and this requires to perform both an inversion and a denoising. For many problem, this two goals are in contradiction, since usually inverting the operator increases the noise level. This is for instance the case for the deblurring problem, where Φ is a translation invariant operator, that corresponds to a low pass filtering with some kernel h

$$\Phi f = f \star h. \tag{10.1}$$

One can for instance consider this convolution over $\mathcal{S} = \mathcal{H} = L^2(\mathbb{T}^d)$, see Proposition 3. In practice, this convolution is followed by a sampling on a grid $\Phi f = \{(f \star h)(x_k) \; ; \; 0 \leqslant k < P\}$, see Figure 10.1, middle, for an example of a low resolution image Φf_0. Inverting such operator has important industrial application to upsample the content of digital photos and to compute high definition videos from low definition videos.

Interpolation and inpainting. Inpainting corresponds to interpolating missing pixels in an image. This is modelled by a diagonal operator over the spacial domain

$$(\Phi f)(x) = \begin{cases} 0 & \text{if} \quad x \in \Omega, \\ f(x) & \text{if} \quad x \notin \Omega. \end{cases} \tag{10.2}$$

where $\Omega \subset [0,1]^d$ (continuous model) or $\{0, \ldots, N-1\}$ which is then a set of missing pixels. Figure 10.1, right, shows an example of damaged image Φf_0.

Original f_0 Low resolution Φf_0 Masked Φf_0

Figure 10.1: Example of inverse problem operators.

Medical imaging. Most medical imaging acquisition device only gives indirect access to the signal of interest, and is usually well approximated by such a linear operator Φ. In scanners, the acquisition operator is the Radon transform, which, thanks to the Fourier slice theorem, is equivalent to partial Fourier mesurments along radial lines. Medical resonance imaging (MRI) is also equivalent to partial Fourier measures

$$\Phi f = \left\{ \hat{f}(x) \; ; \; x \in \Omega \right\}. \tag{10.3}$$

Here, Ω is a set of radial line for a scanner, and smooth curves (e.g. spirals) for MRI.

Other indirect application are obtained by electric or magnetic fields measurements of the brain activity (corresponding to MEG/EEG). Denoting $\Omega \subset \mathbb{R}^3$ the region around which measurements are performed (e.g. the head), in a crude approximation of these measurements, one can assume $\Phi f = \{(\psi \star f)(x) \; ; \; x \in \partial\Omega\}$ where $\psi(x)$ is a kernel accounting for the decay of the electric or magnetic field, e.g. $\psi(x) = 1/\|x\|^2$.

10.2 Theoretical Study of Quadratic Regularization

We now give a glimpse on the typical approach to obtain theoretical guarantee on recovery quality in the case of Hilbert space. The goal is not to be exhaustive, but rather to insist on the modelling hypotethese, namely smoothness implies a so called "source condition", and the inherent limitations of quadratic methods (namely slow rates and the impossibility to recover information in $\ker(\Phi)$, i.e. to achieve super-resolution).

10.2.1 Singular Value Decomposition

Finite dimension. Let us start by the simple finite dimensional case $\Phi \in \mathbb{R}^{P\times N}$ so that $\mathcal{S} = \mathbb{R}^N$ and $\mathcal{H} = \mathbb{R}^P$ are Hilbert spaces. In this case, the Singular Value Decomposition (SVD) is the key to analyze the operator very precisely, and to describe linear inversion process.

Proposition 20 (SVD)**.** *There exists* $(U, V) \in \mathbb{R}^{N\times R} \times \mathbb{R}^{P\times R}$*, where* $R = \mathrm{rank}(\Phi) = \dim(\mathrm{Im}(\Phi))$*, with* $U^\top U = V^\top V = \mathrm{Id}_R$*, i.e. having orthogonal columns* $(u_m)_{m=1}^R \subset \mathbb{R}^N, (v_m)_{m=1}^R \subset \mathbb{R}^P$*, and* $(\sigma_m)_{m=1}^R$ *with* $\sigma_m > 0$*, such that*

$$\Phi = U \,\mathrm{diag}_m(\sigma_m) V^\top = \sum_{m=1}^R \sigma_m u_m v_m^\top. \tag{10.4}$$

Proof. We first analyze the problem, and notice that if $\Phi = U\Sigma V^\top$ with $\Sigma = \mathrm{diag}_m(\sigma_m)$, then $\Phi\Phi^\top = U\Sigma^2 U^\top$ and then $V^\top = \Sigma^{-1}U^\top\Phi$. We can use this insight. Since $\Phi\Phi^\top$ is a positive symmetric matrix, we write its eigendecomposition as $\Phi\Phi^\top = U\Sigma^2U^\top$ where $\Sigma = \mathrm{diag}_{m=1}^R(\sigma_m)$ with $\sigma_m > 0$. We then define $V \overset{\text{def.}}{=} \Phi^\top U\Sigma^{-1}$. One then verifies that

$$V^\top V = (\Sigma^{-1}U^\top\Phi)(\Phi^\top U\Sigma^{-1}) = \Sigma^{-1}U^\top(U\Sigma^2U^\top)U\Sigma^{-1} = \mathrm{Id}_P \quad \text{and} \quad U\Sigma V^\top = U\Sigma\Sigma^{-1}U^\top\Phi = \Phi.$$

□

This theorem is still valid with complex matrice, replacing $^\top$ by *. Expression (10.4) describes Φ as a sum of rank-1 matrices $u_m v_m^\top$. One usually order the singular values $(\sigma_m)_m$ in decaying order $\sigma_1 \geqslant \ldots \geqslant \sigma_R$. If these values are different, then the SVD is unique up to ± 1 sign change on the singular vectors.

The left singular vectors U is an orthonormal basis of $\mathrm{Im}(\Phi)$, while the right singular values is an orthonormal basis of $\mathrm{Im}(\Phi^\top) = \ker(\Phi)^\perp$. The decomposition (10.4) is often call the "reduced" SVD because one has only kept the R non-zero singular values. The "full" SVD is obtained by completing U and V to define orthonormal bases of the full spaces $\mathbb{R}^P$ and $\mathbb{R}^N$. Then Σ becomes a rectangular matrix of size $P \times N$.

A typical example is for $\Phi f = f \star h$ over $\mathbb{R}^P = \mathbb{R}^N$, in which case the Fourier transform diagonalizes the convolution, i.e.

$$\Phi = (u_m)_m^* \,\mathrm{diag}(\hat{h}_m)(u_m)_m \tag{10.5}$$

where $(u_m)_n \overset{\text{def.}}{=} \frac{1}{\sqrt{N}} e^{\frac{2\mathrm{i}\pi}{N}nm}$ so that the singular values are $\sigma_m = |\hat{h}_m|$ (removing the zero values) and the singular vectors are $(u_m)_n$ and $(v_m\theta_m)_n$ where $\theta_m \overset{\text{def.}}{=} |\hat{h}_m|/\hat{h}_m$ is a unit complex number.

Computing the SVD of a full matrix $\Phi \in \mathbb{R}^{N\times N}$ has complexity N^3.

Compact operators. One can extend the decomposition to compact operators $\Phi : \mathcal{S} \to \mathcal{H}$ between separable Hilbert space. A compact operator is such that ΦB_1 is pre-compact where $B_1 = \{s \in \mathcal{S} \,;\, \|s\| \leqslant 1\}$ is the unit-ball. This means that for any sequence $(\Phi s_k)_k$ where $s_k \in B_1$ one can extract a converging sub-sequence. Note that in infinite dimension, the identity operator $\Phi : \mathcal{S} \to \mathcal{S}$ is never compact.

Compact operators Φ can be shown to be equivalently defined as those for which an expansion of the form (10.4) holds

$$\Phi = \sum_{m=1}^{+\infty} \sigma_m u_m v_m^\top \tag{10.6}$$

where $(\sigma_m)_m$ is a decaying sequence onverging to 0, $\sigma_m \to 0$. Here in (10.6) convergence holds in the operator norm, which is the algebra norm on linear operator inherited from those of $\mathcal{S}$ and $\mathcal{H}$

$$\|\Phi\|_{\mathcal{L}(\mathcal{S},\mathcal{H})} \overset{\text{def.}}{=} \min_{\|\Phi u\|_{\mathcal{H}}} \|u\|_{\mathcal{S}} \leqslant 1.$$

For Φ having an SVD decomposition (10.6), $\|\Phi\|_{\mathcal{L}(\mathcal{S},\mathcal{H})} = \sigma_1$.

When $\sigma_m = 0$ for $m > R$, Φ has a finite rank $R = \dim(\mathrm{Im}(\Phi))$. As we explain in the sections bellow, when using *linear* recovery methods (such as quadratic regularization), the inverse problem is equivalent to

a finite dimensional problem, since one can restrict its attention to functions in $\ker(\Phi)^\perp$ which as dimension R. Of course, this is not true anymore when one can retrieve function inside $\ker(\Phi)$, which is often referred to as a "super-resolution" effect of non-linear methods. Another definition of compact operator is that they are the limit of finite rank operator. They are thus in some sense the extension of finite dimensional matrices, and are the correct setting to model ill-posed inverse problems. This definition can be extended to linear operator between Banach spaces, but this conclusion does not holds.

Typical example of compact operator are matrix-like operator with a continuous kernel $k(x,y)$ for $(x,y) \in \Omega$ where Ω is a compact sub-set of $\mathbb{R}^d$ (or the torus $\mathbb{T}^d$), i.e.

$$\Phi f = \int_\Omega k(x,y) f(y) \mathrm{d}y$$

where $\mathrm{d}y$ is the Lebesgue measure. An example of such a setting which generalizes (10.5) is when $\Phi f = h \star h$ on $\mathbb{T}^d = (\mathbb{R}/2\pi\mathbb{Z})^d$, which is corresponds to a translation invariant kernel $k(x,y) = h(x-y)$, in which case $u_m(x) = (2\pi)^{-d/2} e^{\mathrm{i}\omega x}$, $\sigma_m = |\hat{f}_m|$.

Pseudo inverse. In the case where $w = 0$, it makes to try to directly solve $\Phi f = y$. The two obstruction for this is that one not necessarily has $y \in \mathrm{Im}(\Phi)$ and even so, there are an infinite number of solutions if $\ker(\Phi) \neq \{0\}$. The usual workaround is to solve this equation in the least square sense

$$f^+ \overset{\text{def.}}{=} \operatorname*{argmin}_{\Phi f = y^+} \|f\|_{\mathcal{S}} \quad \text{where} \quad y^+ = \mathrm{Proj}_{\mathrm{Im}(\Phi)}(y) = \operatorname*{argmin}_{z \in \mathrm{Im}(\Phi)} \|y - z\|_{\mathcal{H}}.$$

The following proposition shows how to compute this least square solution using the SVD and by solving linear systems involving either $\Phi\Phi^*$ or $\Phi^*\Phi$.

Proposition 21. *One has*

$$f^+ = \Phi^+ y \quad \textit{where} \quad \Phi^+ = V \operatorname{diag}_m(1/\sigma_m) U^*. \tag{10.7}$$

In case that $\mathrm{Im}(\Phi) = \mathcal{H}$*, one has* $\Phi^+ = \Phi^*(\Phi\Phi^*)^{-1}$*. In case that* $\ker(\Phi) = \{0\}$*, one has* $\Phi^+ = (\Phi^*\Phi)^{-1}\Phi^*$*.*

Proof. Since U is an ortho-basis of $\mathrm{Im}(\Phi)$, $y^+ = UU^*y$, and thus $\Phi f = y^+$ reads $U\Sigma V^* f = UU^* y$ and hence $V^* f = \Sigma^{-1} U^* y$. Decomposition orthogonally $f = f_0 + r$ where $f_0 \in \ker(\Phi)^\perp$ and $r \in \ker(\Phi)$, one has $f_0 = VV^* f = V\Sigma^{-1}U^* y = \Phi^+ y$ is a constant. Minimizing $\|f\|^2 = \|f_0\|^2 + \|r\|^2$ is thus equivalent to minimizing $\|r\|$ and hence $r = 0$ which is the desired result. If $\mathrm{Im}(\Phi) = \mathcal{H}$, then $R = N$ so that $\Phi\Phi^* = U\Sigma^2 U^*$ is the eigen-decomposition of an invertible and $(\Phi\Phi^*)^{-1} = U\Sigma^{-2}U^*$. One then verifies $\Phi^*(\Phi\Phi^*)^{-1} = V\Sigma U^* U \Sigma^{-2} U^*$ which is the desired result. One deals similarly with the second case. $\square$

For convolution operators $\Phi f = f \star h$, then

$$\Phi^+ y = y \star h^+ \quad \text{where} \quad \hat{h}_m^+ = \begin{cases} \hat{h}_m^{-1} & \text{if} \quad \hat{h}_m \neq 0 \\ 0 & \text{if} \quad \hat{h}_m = 0. \end{cases}.$$

10.2.2 Tikonov Regularization

Regularized inverse. When there is noise, using formula (10.7) is not acceptable, because then

$$\Phi^+ y = \Phi^+ \Phi f_0 + \Phi^+ w = f_0^+ + \Phi^+ w \quad \text{where} \quad f_0^+ \overset{\text{def.}}{=} \mathrm{Proj}_{\ker(\Phi)^\perp},$$

so that the recovery error is $\|\Phi^+ y - f_0^+\| = \|\Phi^+ w\| \geqslant \|w\|/\sigma_R$. The noise is thus amplified by the inverse $1/\sigma_R$ of the smallest amplitude non-zero singular values, which can be very large. In infinite dimension, one typically has $R = +\infty$, so that the inverse is actually not bounded (discontinuous). It is thus mendatory to replace Φ^+ by a regularized approximate inverse, which should have the form

$$\Phi_\lambda^+ = V \operatorname{diag}_m(\mu_\lambda(\sigma_m)) U^* \tag{10.8}$$

where μ_λ, indexed by some parameter $\lambda > 0$, is a regularization of the inverse, that should typically satisfies

$$\mu_\lambda(\sigma) \leqslant C_\lambda < +\infty \quad \text{and} \quad \lim_{\lambda \to 0} \mu_\lambda(\sigma) = \frac{1}{\sigma}$$

Variational regularization. A typical example of such regularized inverse is obtained by considering a penalized least square involving a regularization functional

$$f_\lambda \overset{\text{def.}}{=} \operatorname*{argmin}_{f\in\mathcal{S}} \|y-\Phi f\|_{\mathcal{H}}^2 + \lambda J(f)^2 \tag{10.9}$$

where J is some regularization functional which should at least be continuous on $\mathcal{S}$. The simplest example is the quadratic norm $J = \|\cdot\|_{\mathcal{S}}^2$,

$$f_\lambda \overset{\text{def.}}{=} \operatorname*{argmin}_{f\in\mathcal{S}} \|y-\Phi f\|_{\mathcal{H}}^2 + \lambda \|f\|^2 \tag{10.10}$$

which can be conveniently rewritten in the basis of singular vectors as

$$f_\lambda = \operatorname*{argmin}_{f\in\mathrm{Im}(\Phi^*)} \sum_m (\sigma_m \langle f,\, v_m\rangle - \langle y,\, v_m\rangle)^2 + \lambda \langle f,\, v_m\rangle^2 \tag{10.11}$$

where we have use the fact that necessarily $f_\lambda \in \mathrm{Im}(\Phi^*)$ because of the square penalty. The minimization (10.11) boils down to independent scalar minimization over each coordinate $f_m \overset{\text{def.}}{=} \langle f,\, v_m\rangle$ and the first order condition reads

$$\sigma_m(\sigma_m f_m - y_m) + \lambda f_m = 0 \quad \text{where} \quad y_m \overset{\text{def.}}{=} \langle y,\, v_m\rangle.$$

One thus has $f_\lambda = \Phi_\lambda^+$ where Φ_λ^+ is in the form (10.8) for the specific choice of function

$$\forall\, \sigma \in \mathbb{R}, \quad \mu_\lambda(\sigma) = \frac{\sigma}{\sigma^2+\lambda}.$$

The question is to understand how to choose λ as a function of the noise level $\|w\|_{\mathcal{H}}$ in order to guarantees that $f_\lambda \to f_0$ and furthermore establish convergence speed. One first needs to ensure at least $f_0 = f_0^+$, which in turns requires that $f_0 \in \mathrm{Im}(\Phi^*) = \ker(\Phi)^\perp$. Indeed, an important drawback of linear recovery methods (such as quadratic regularization) is that necessarily $f_\lambda \in \mathrm{Im}(\Phi^*) = \ker(\Phi^\perp)$ so that no information can be recovered inside $\ker(\Phi)$. Non-linear methods must be used to achieve a "super-resoltution" effect and recover this missing information.

Source condition. In order to ensure convergence speed, one quantify this condition and impose a so-called source condition of order β, which reads

$$f_0 \in \mathrm{Im}((\Phi^*\Phi)^\beta) = \mathrm{Im}(V\,\mathrm{diag}(\sigma_m^{2\beta})V^*).$$

This condition means that there should exists $z \in \mathcal{S}$ such that $f_0 = V\,\mathrm{diag}(\sigma_m^{2\beta})V^* z$, i.e. $z = V\,\mathrm{diag}(\sigma_m^{-2\beta})V^* f_0$. Denoting $\rho \overset{\text{def.}}{=} \|z\|$, we thus in fact impose the following constraint

$$\sum_m \sigma_m^{-2\beta} \langle f,\, v_m\rangle^2 \leqslant \rho^2 < +\infty. \tag{$S_{\beta,\rho}$}$$

This is a Sobolev-type constraint, similar to those imposed in 8.4. A prototypical example is for a low-pass filter $\Phi f = f \star h$ where h as a slow polynomial-like decay of frequency, i.e. $|\hat{h}_m| \sim 1/m^\alpha$ for large m. In this case, since v_m is the Fourier basis, the source condition ($S_{\beta,\rho}$) reads

$$\sum_m \|m\|^{2\alpha\beta} |\hat{f}_m|^2 \leqslant \rho^2 < +\infty. \tag{$S_{\beta,\rho}$}$$

which is a Sobolev ball of radius ρ and differential order $\alpha\beta$.

Sublinear convergence speed. The following theorem shows that this source condition leads to a convergence speed of the regularization.

Theorem 31. *Assuming the source condition* ($S_{\beta,\rho}$) *with* $0 < \beta \leqslant 2$*, then the solution of* (10.10) *for* $\|w\| \leqslant \delta$ *satisfies*

$$\|f_\lambda - f_0\| \leqslant C \rho^{\frac{1}{\beta+1}} \delta^{\frac{\beta}{\beta+1}}$$

for a constant C *which depends only on* β*, and for a choice*

$$\lambda \sim \delta^{\frac{2}{\beta+1}} \rho^{-\frac{2}{\beta+1}}.$$

Proof. Because of the source condition, $f_0 \in \mathrm{Im}(\Phi^*)$. Denoting

$$f_\lambda^0 \stackrel{\text{def.}}{=} \Phi_\lambda^+(\Phi f_0)$$

so that $f_\lambda = f_\lambda^0 + \Phi_\lambda^+ w$, one has for any regularized inverse of the form (10.8)

$$\|f_\lambda - f_0\| \leqslant \|f_\lambda - f_\lambda^0\| + \|f_\lambda^0 - f_0\|. \tag{10.12}$$

The term $\|f_\lambda - f_\lambda^0\|$ is a variance term which account for residual noise, and thus decays when λ increases (more regularization). The term $\|f_\lambda^0 - f_0\|$ is independent of the noise, it is a bias term coming from the approximartion (smoothing) of f_0, and thus increases when λ increases. The choice of an optimal λ thus results in a bias-variance tradeoff between these two terms. Assuming

$$\forall \sigma \geqslant 0, \quad \mu_\lambda(\sigma) \leqslant C_\lambda$$

the variance terme is bounded as

$$\|f_\lambda - f_\lambda^0\|^2 = \|\Phi_\lambda^+ w\|^2 = \sum_m \mu_\lambda(\sigma_m)^2 w_m^2 \leqslant C_\lambda^2 \|w\|_{\mathcal{H}}^2.$$

The bias term is bounded as

$$\|f_\lambda^0 - f_0\|^2 = \sum_m (1 - \mu_\lambda(\sigma_m)\sigma_m)^2 f_{0,m}^2 = \sum_m \left(\frac{1 - \mu_\lambda(\sigma_m)\sigma_m}{\sigma^\beta} \right)^2 \sigma^{2\beta} f_{0,m}^2 \leqslant D_{\lambda,\beta}^2 \rho^2 \tag{10.13}$$

where we assumed

$$\forall \sigma \geqslant 0, \quad \left| \frac{1 - \mu_\lambda(\sigma)\sigma}{\sigma^\beta} \right| \leqslant D_{\lambda,\beta}. \tag{10.14}$$

Putting (10.13) and (10.14) together, one obtains

$$\|f_\lambda - f_0\| \leqslant C_\lambda \delta + D_{\lambda,\beta} \rho. \tag{10.15}$$

In the case of the regularization (10.10), one has $\mu_\lambda(\sigma) = \frac{\sigma}{\sigma^2+\lambda}$, and thus $\frac{1-\mu_\lambda(\sigma)\sigma}{\sigma^\beta} = \frac{\lambda\sigma^\beta}{\sigma^2+\lambda}$. For $\beta \leqslant 2$, one verifies that

$$C_\lambda = \frac{1}{2\sqrt{\lambda}} \quad \text{and} \quad D_{\lambda,\beta} = C_\beta \lambda^{\frac{\beta}{2}},$$

for some constant C_β. Equalizing the contributions of the two terms in (10.15) (a better constant would be reached by finding the best λ) leads to selecting $\frac{\delta}{\sqrt{\lambda}} = \lambda^{\frac{\beta}{2}} \rho$ i.e. $\lambda = (\delta/\rho)^{\frac{2}{\beta+1}}$. With this choice,

$$\|f_\lambda - f_0\| = O(\delta/\sqrt{\lambda}) = O(\delta(\delta/\rho)^{-\frac{1}{\beta+1}}) = O(\delta^{\frac{\beta}{\beta+1}} \rho^{\frac{1}{\beta+1}}).$$

□

This theorem shows that using larger $\beta \leqslant 2$ leads to faster convergence rates as $\|w\|$ drops to zero. The rate (10.12) however suffers from a "saturation" effect, indeed, choosing $\beta > 2$ does not helps (it gives the same rate as $\beta = 2$), and the best possible rate is thus

$$\|f_\lambda - f_0\| = O(\rho^{\frac{1}{3}} \delta^{\frac{2}{3}}).$$

By choosing more alternative regularization functional μ_λ and choosing β large enough, one can show that it is possible to reach rate $\|f_\lambda - f_0\| = O(\delta^{1-\kappa})$ for an arbitrary small $\kappa > 0$, but one cannot reach a linear rate $\|f_\lambda - f_0\| = O(\|w\|)$. Such rates are achievable using non-linear sparse ℓ^1 regularizations as detailed in Chapter 11.

10.3 Quadratic Regularization

After this theoretical study in infinite dimension, we now turn our attention to more practical matters, and focus only on the finite dimensional setting.

Convex regularization. Following (10.9), the ill-posed problem of recovering an approximation of the high resolution image $f_0 \in \mathbb{R}^N$ from noisy measures $y = \Phi f_0 + w \in \mathbb{R}^P$ is regularized by solving a convex optimization problem

$$f^\star \in \operatorname*{argmin}_{f \in \mathbb{R}^N} \ \mathcal{E}(f) \stackrel{\text{def.}}{=} \frac{1}{2}\|y - \Phi f\|^2 + \lambda J(f) \tag{10.16}$$

where $\|y - \Phi f\|^2$ is the data fitting term (here $\|\cdot\|$ is the ℓ^2 norm on $\mathbb{R}^P$) and $J(f)$ is a convex functional on $\mathbb{R}^N$.

The Lagrange multiplier λ weights the importance of these two terms, and is in practice difficult to set. Simulation can be performed on high resolution signal f_0 to calibrate the multiplier by minimizing the super-resolution error $\|f_0 - \tilde{f}\|$, but this is usually difficult to do on real life problems.

In the case where there is no noise, $w = 0$, the Lagrange multiplier λ should be set as small as possible. In the limit where $\lambda \to 0$, the unconstrained optimization problem (10.16) becomes a constrained optimization

$$f^\star = \operatorname*{argmin}_{f \in \mathbb{R}^N} \ \{J(f) \ ; \ \Phi f = y \ \}. \tag{10.17}$$

Quadratic Regularization. The simplest class of prior functional are quadratic, and can be written as

$$J(f) = \frac{1}{2}\|Gf\|^2_{\mathbb{R}^K} = \frac{1}{2}\langle Lf, \, f\rangle_{\mathbb{R}^N} \tag{10.18}$$

where $G \in \mathbb{R}^{K \times N}$ and where $L = G^* G \in \mathbb{R}^{N \times N}$ is a positive semi-definite matrix. The special case (10.10) is recovered when setting $G = L = \mathrm{Id}_N$.

Writing down the first order optimality conditions for (10.16) leads to

$$\nabla \mathcal{E}(f) = \Phi^*(\Phi f - y) + \lambda L f = 0,$$

hence, if

$$\ker(\Phi) \cap \ker(G) = \{0\},$$

then (10.18) has a unique minimizer f_λ, which is obtained by solving a linear system

$$f_\lambda = (\Phi^* \Phi + \lambda L)^{-1} \Phi^* y. \tag{10.19}$$

In the special case where L is diagonalized by the singular basis $(v_m)_m$ of Φ, i.e. $L = V \operatorname{diag}(\alpha_m^2) V^*$, then f_λ reads in this basis

$$\langle f_\lambda, \, v_m\rangle = \frac{\sigma_m}{\sigma_m^2 + \lambda \alpha_m^2} \langle y, \, v_m\rangle. \tag{10.20}$$

Example of convolution. A specific example is for convolution operator

$$\Phi f = h \star f, \tag{10.21}$$

and using $G = \nabla$ be a discretization of the gradient operator, such as for instance using first order finite differences (2.16). This corresponds to the discrete Sobolev prior introduced in Section 9.1.2. Such an operator compute, for a d-dimensional signal $f \in \mathbb{R}^N$ (for instance a 1-D signal for $d = 1$ or an image when $d = 2$), an approximation $\nabla f_n \in \mathbb{R}^d$ of the gradient vector at each sample location n. Thus typically, $\nabla : f \mapsto (\nabla f_n)_n \in \mathbb{R}^{N \times d}$ maps to d-dimensional vector fields. Then $-\nabla^* : \mathbb{R}^{N \times d} \to \mathbb{R}^N$ is a discretized divergence operator. In this case, $\Delta = -GG^*$ is a discretization of the Laplacian, which is itself a convolution operator. One then has

$$\hat{f}_{\lambda,m} = \frac{\hat{h}_m^* \hat{y}_m}{|\hat{h}_m|^2 - \lambda \hat{d}_{2,m}}, \tag{10.22}$$

where $\hat{d}_2$ is the Fourier transform of the filter d_2 corresponding to the Laplacian. For instance, in dimension 1, using first order finite differences, the expression for $\hat{d}_{2,m}$ is given in (2.18).

10.3.1 Solving Linear System

When Φ and L do not share the same singular spaces, using (10.20) is not possible, so that one needs to solve the linear system (10.19), which can be rewritten as

$$Af = b \quad \text{where} \quad A \overset{\text{def.}}{=} \Phi^* \Phi + \lambda L \quad \text{and} \quad b = \Phi^* y.$$

It is possible to solve exactly this linear system with direct methods for moderate N (up to a few thousands), and the numerical complexity for a generic A is $O(N^3)$. Since the involved matrix A is symmetric, the best option is to use Choleski factorization $A = BB^*$ where B is lower-triangular. In favorable cases, this factorization (possibly with some re-re-ordering of the row and columns) can take advantage of some sparsity in A.

For large N, such exact resolution is not an option, and should use approximate iterative solvers, which only access A through matrix-vector multiplication. This is especially advantageous for imaging applications, where such multiplications are in general much faster than a naive $O(N^2)$ explicit computation. If the matrix A is highly sparse, this typically necessitates $O(N)$ operations. In the case where A is symmetric and positive definite (which is the case here), the most well known method is the conjugate gradient methods, which is actually an optimization method solving

$$\min_{f \in \mathbb{R}^N} \mathcal{E}(f) \overset{\text{def.}}{=} \mathcal{Q}(f) \overset{\text{def.}}{=} \langle Af, f \rangle - \langle f, b \rangle \tag{10.23}$$

which is equivalent to the initial minimization (10.16). Instead of doing a naive gradient descent (as studied in Section 10.4.2 bellow), stating from an arbitrary $f^{(0)}$, it compute a new iterate $f^{(\ell+1)}$ from the previous iterates as

$$f^{(\ell+1)} \overset{\text{def.}}{=} \operatorname*{argmin}_f \left\{ \mathcal{E}(f) \; ; \; f \in f^{(\ell)} + \operatorname{Span}(\nabla \mathcal{E}(f^{(0)}), \ldots, \nabla \mathcal{E}(f^{\ell})) \right\}.$$

The crucial and remarkable fact is that this minimization can be computed in closed form at the cost of two matrix-vector product per iteration, for $k \geqslant 1$ (posing initially $d^{(0)} = \nabla \mathcal{E}(f^{(0)}) = Af^{(0)} - b$)

$$f^{(\ell+1)} = f^{(\ell)} - \tau_\ell d^{(\ell)} \quad \text{where} \quad d^{(\ell)} = g_\ell + \frac{\|g^{(\ell)}\|^2}{\|g^{(\ell-1)}\|^2} d^{(\ell-1)} \quad \text{and} \quad \tau_\ell = \frac{\langle g_\ell, d^{(\ell)} \rangle}{\langle Ad^{(\ell)}, d^{(\ell)} \rangle} \tag{10.24}$$

$g^{(\ell)} \overset{\text{def.}}{=} \nabla \mathcal{E}(f^{(\ell)}) = Af^{(\ell)} - b$. It can also be shown that the direction $d^{(\ell)}$ are orthogonal, so that after $\ell = N$ iterations, the conjugate gradient computes the unique solution $f^{(\ell)}$ of the linear system $Af = b$. It is however rarely used this way (as an exact solver), and in practice much less than N iterates are computed. It should also be noted that iterations (10.24) can be caried over for an arbitrary smooth convex function $\mathcal{E}$, and it typically improves over the gradient descent (although in practice quasi-Newton method are often preferred).

10.4 Non-Quadratic Regularization

10.4.1 Total Variation Regularization

A major issue with quadratic regularization such as (10.18) is that they typically leads to blurry recovered data f_λ, which is thus not a good approximation of f_0 when it contains sharp transition such as edges in images. This is can clearly be seen in the convolutive case (10.22), this the restoration operator $\Phi_\lambda^+\Phi$ is a filtering, which tends to smooth sharp part of the data.

This phenomena can also be understood because the restored data f_λ always belongs to $\mathrm{Im}(\Phi^*) = \ker(\Phi)^\perp$, and thus cannot contains "high frequency" details that are lost in the kernel of Φ. To alleviate this shortcoming, and recover missing information in the kernel, it is thus necessarily to consider non-quadratic and in fact non-smooth regularization.

Total variation. The most well know instance of such a non-quadratic and non-smooth regularization is the total variation prior. For smooth function $f : \mathbb{R}^d \mapsto \mathbb{R}$, this amounts to replacing the quadratic Sobolev energy (often called Dirichlet energy)

$$J_{\mathrm{Sob}}(f) \overset{\text{def.}}{=} \frac{1}{2} \int_{\mathbb{R}^d} \|\nabla f\|_{\mathbb{R}^d}^2 \mathrm{d}x,$$

where $\nabla f(x) = (\partial_{x_1} f(x), \ldots, \partial_{x_d} f(x))^\top$ is the gradient, by the (vectorial) L^1 norm of the gradient

$$J_{\mathrm{TV}}(f) \overset{\text{def.}}{=} \int_{\mathbb{R}^d} \|\nabla f\|_{\mathbb{R}^d} \mathrm{d}x.$$

We refer also to Section 9.1.1 about these priors. Simply "removing" the square ^{2} inside the integral might seems light a small change, but in fact it is a game changer.

Indeed, while $J_{\mathrm{Sob}}(1_\Omega) = +\infty$ where 1_Ω is the indicator a set Ω with finite perimeter $|\Omega| < +\infty$, one can show that $J_{\mathrm{TV}}(1_\Omega) = |\Omega|$, if one interpret ∇f as a distribution Df (actually a vectorial Radon measure) and $\int_{\mathbb{R}^d} \|\nabla f\|_{\mathbb{R}^d} \mathrm{d}x$ is replaced by the total mass $|Df|(\Omega)$ of this distribution $m = Df$

$$|m|(\Omega) = \sup \left\{ \int_{\mathbb{R}^d} \langle h(x), \mathrm{d}m(x) \rangle \ ; \ h \in \mathcal{C}(\mathbb{R}^d \mapsto \mathbb{R}^d), \forall x, \|h(x)\| \leqslant 1 \right\}.$$

The total variation of a function such that Df has a bounded total mass (a so-called bounded variation function) is hence defined as

$$J_{\mathrm{TV}}(f) \overset{\text{def.}}{=} \sup \left\{ \int_{\mathbb{R}^d} f(x) \operatorname{div}(h)(x) \mathrm{d}x \ ; \ h \in \mathcal{C}_c^1(\mathbb{R}^d; \mathbb{R}^d), \|h\|_\infty \leqslant 1 \right\}.$$

Generalizing the fact that $J_{\mathrm{TV}}(1_\Omega) = |\Omega|$, the functional co-area formula reads

$$J_{\mathrm{TV}}(f) = \int_{\mathbb{R}} \mathcal{H}_{d-1}(L_t(f)) \mathrm{d}t \quad \text{where} \quad L_t(f) = \{x \ ; \ f(x) = t\}$$

and where $\mathcal{H}_{d-1}$ is the Hausforf measures of dimension $d-1$, for instance, for $d = 2$ if L has finite perimeter $|L|$, then $\mathcal{H}_{d-1}(L) = |L|$ is the perimeter of L.

Discretized Total variation. For discretized data $f \in \mathbb{R}^N$, one can define a discretized TV semi-norm as detailed in Section 9.1.2, and it reads, generalizing (9.6) to any dimension

$$J_{\mathrm{TV}}(f) = \sum_n \|\nabla f_n\|_{\mathbb{R}^d}$$

where $\nabla f_n \in \mathbb{R}^d$ is a finite difference gradient at location indexed by n.

The discrete total variation prior $J_{\mathrm{TV}}(f)$ defined in (9.6) is a convex but non differentiable function of f, since a term of the form $\|\nabla f_n\|$ is non differentiable if $\nabla f_n = 0$. We defer to chapters 12 and 13 the study of advanced non-smooth convex optimization technics that allows to handle this kind of functionals.

In order to be able to use simple gradient descent methods, one needs to smooth the TV functional. The general machinery proceeds by replacing the non-smooth ℓ^2 Euclidean norm $\|\cdot\|$ by a smoothed version, for instance

$$\forall\, u \in \mathbb{R}^d, \quad \|u\|_\varepsilon \overset{\text{def.}}{=} \sqrt{\varepsilon^2 + \|u\|}.$$

This leads to the definition of a smoothed approximate TV functional, already introduced in (9.12),

$$J_{\mathrm{TV}}^\varepsilon(f) \overset{\text{def.}}{=} \sum_n \|\nabla f_n\|_\varepsilon$$

One has the following asymptotics for $\varepsilon \to \{0, +\infty\}$

$$\|u\|_\varepsilon \overset{\varepsilon \to 0}{\longrightarrow} \|u\| \quad \text{and} \quad \|u\|_\varepsilon = \varepsilon + \frac{1}{2\varepsilon}\|u\|^2 + O(1/\varepsilon^2)$$

which suggest that $J_{\mathrm{TV}}^\varepsilon$ interpolates between J_{TV} and J_{Sob}.

The resulting inverse regularization problem (10.16) thus reads

$$f_\lambda \overset{\text{def.}}{=} \operatorname*{argmin}_{f \in \mathbb{R}^N} \mathcal{E}(f) = \frac{1}{2}\|y - \Phi f\|^2 + \lambda J_{\mathrm{TV}}^\varepsilon(f) \tag{10.25}$$

It is a strictly convex problem (because $\|\cdot\|_\varepsilon$ is strictly convex for $\varepsilon > 0$) so that its solution f_λ is unique.

10.4.2 Gradient Descent Method

The optimization program (10.25) is a example of smooth unconstrained convex optimization of the form

$$\min_{f \in \mathbb{R}^N} \mathcal{E}(f) \tag{10.26}$$

where $\mathcal{E} : \mathbb{R}^N \to \mathbb{R}$ is a $\mathcal{C}^1$ function. Recall that the gradient $\nabla \mathcal{E} : \mathbb{R}^N \mapsto \mathbb{R}^N$ of this functional (not to be confound with the discretized gradient $\nabla f \in \mathbb{R}^N$ of f) is defined by the following first order relation

$$\mathcal{E}(f + r) = \mathcal{E}(f) + \langle f,\, r\rangle_{\mathbb{R}^N} + O(\|r\|_{\mathbb{R}^N}^2)$$

where we used $O(\|r\|_{\mathbb{R}^N}^2)$ in place of $o(\|r\|_{\mathbb{R}^N})$ (for differentiable function) because we assume here $\mathcal{E}$ is of class $\mathcal{C}^1$ (i.e. the gradient is continuous).

For such a function, the gradient descent algorithm is defined as

$$f^{(\ell+1)} = f^{(\ell)} - \tau_\ell \nabla \mathcal{E}(f^{(\ell)}),$$

where the step size $\tau_\ell > 0$ should be small enough to guarantee convergence, but large enough for this algorithm to be fast.

To allow for not-too-small steps,

One needs to quantify the smoothness of $\mathcal{E}$. This is enforced by requiring that the gradient is L-Lipschitz, i.e.

$$\forall\, (f, g) \in (\mathbb{R}^N)^2, \quad \|\nabla \mathcal{E}(f) - \nabla \mathcal{E}(g)\| \leqslant L\|f - g\|. \tag{$\mathcal{R}_L$}$$

In order to obtain fast convergence of the iterates themselve, it is needed that the function has enough "curvature" (i.e. is not too flat), which corresponds to imposing that $\mathcal{E}$ is M-strongly convex

$$\forall\, (f, g), \in (\mathbb{R}^N)^2, \quad \langle \nabla \mathcal{E}(f) - \nabla \mathcal{E}(g),\, f - g\rangle \geqslant M\|f - g\|^2. \tag{$\mathcal{S}_L$}$$

The following proposition express these conditions as constraints on the hessian for $\mathcal{C}^2$ functions.

Proposition 22. *If $\mathcal{E}$ is of class $\mathcal{C}^2$, conditions ($\mathcal{R}_L$) and ($\mathcal{S}_L$) are equivalent to*

$$\forall f, \quad M\mathrm{Id}_{N\times N} \preceq \partial^2\mathcal{E}(f) \preceq L\mathrm{Id}_{N\times N} \tag{10.27}$$

where $\partial^2\mathcal{E}(f) \in \mathbb{R}^{N\times N}$ is the Hessian of $\mathcal{E}$, and where $\preceq$ is the natural order on symmetric matrices, i.e.

$$A \preceq B \quad \Longleftrightarrow \quad \forall u \in \mathbb{R}^N, \quad \langle Au,\, u\rangle \leqslant \langle Bu,\, u\rangle.$$

Condition (10.27) thus reads that the singular values of $\partial^2\mathcal{E}(f)$ should be contained in the interval $[M, L]$. The upper bound is also equivalent to $\|\partial^2\mathcal{E}(f)\|_{\mathrm{op}} \leqslant L$ where $\|\cdot\|_{\mathrm{op}}$ is the operator norm, i.e. the largest singular value. In the special case of a quadratic function $\mathcal{Q}$ of the form (10.23), $\partial^2\mathcal{E}(f) = A$ is constant, so that $[M, L]$ can be chosen to be the range of the singular values of A.

The following theorem ensure the convergence of the gradient descent with a linear speed.

Theorem 32. *If f satisfy conditions ($\mathcal{R}_L$) and ($\mathcal{S}_L$), assuming there exists $(\tau_{\min}, \tau_{\max})$ such that*

$$0 < \tau_{\min} \leqslant \tau_\ell \leqslant \tau_{\max} < \frac{2M}{L} \tag{10.28}$$

then there exists $0 \leqslant \rho < 1$ such that

$$\|f^{(\ell)} - f^\star\| \leqslant \rho^\ell \|f^{(0)} - f^\star\| \tag{10.29}$$

where $f^\star$ is the unique solution to (10.26).

Proof. Since $\nabla\mathcal{E}(f^\star) = 0$, one has

$$f^{(\ell+1)} - f^\star = (f^{(\ell)} - f^\star) - \tau_\ell(\nabla\mathcal{E}(f^{(\ell)}) - \nabla\mathcal{E}(f^\star)).$$

Hence, using strong convexity and Lipschitz gradient

$$\begin{aligned}\|f^{(\ell+1)} - f^\star\|^2 &= \|f^{(\ell)} - f^\star\|^2 - 2\tau_\ell\langle f^{(\ell)} - f^\star,\, \nabla\mathcal{E}(f^{(\ell)}) - \nabla\mathcal{E}(f^\star)\rangle + \tau_\ell^2\|\nabla\mathcal{E}(f^{(\ell)}) - \nabla\mathcal{E}(f^\star)\|^2 \\ &\leqslant P(\tau_\ell)\|f^{(\ell)} - f^\star\|^2 \quad \text{where} \quad P(\tau) = 1 - 2M\tau + L^2\tau^2.\end{aligned}$$

Figure **??** shows visually the shape of the second order polynomial P, which shows that condition (10.34) on τ_ℓ implies

$$P(\tau_\ell)^{\frac{1}{2}} \leqslant \rho \overset{\text{def.}}{=} \max(P(\tau_{\min}), P(\tau_{\max}))^{\frac{1}{2}} < 1,$$

which shows the desired result. □

The error decay rate (10.32), although it is geometrical $O(\rho^\ell)$ is called a "linear rate" in the optimization litterature. It is a "global" rate because it hold for all ℓ (and not only for large enough ℓ). The best (smallest) rate ρ is obtained when choosing

$$\tau_\ell = \frac{M}{L^2} \quad \Longrightarrow \quad \rho = 1 - \frac{M^2}{L^2}. \tag{10.30}$$

In the case of a quadratic functional of the form (10.23), one can sharpen the convergence proof because the iterates are computed in closed form using matrix multiplication

$$f^{(\ell)} - f^\star = (\mathrm{Id}_N - \tau_\ell A)(f^{(0)} - f^\star)$$

which immediately leads to the following proposition.

Proposition 23. *For $\mathcal{E}(f) = \langle A,\, f\rangle - \langle b,\, f\rangle$ with the singular values of A upper-bounded by L, assuming there exists $(\tau_{\min}, \tau_{\max})$ such that*

$$0 < \tau_{\min} \leqslant \tau_\ell \leqslant \tilde{\tau}_{\max} < \frac{2}{L} \tag{10.31}$$

then there exists $0 \leqslant \tilde{\rho} < 1$ *such that*

$$\|f^{(\ell)} - f^\star\| \leqslant \tilde{\rho}^\ell \|f^{(0)} - f^\star\|. \tag{10.32}$$

If the singular values are lower bounded by M*, then the best rate* $\tilde{\rho}$ *is obtained for*

$$\tau_\ell = \frac{1}{L+M} \quad \Longrightarrow \quad \tilde{\rho} \overset{\text{def.}}{=} \frac{L-M}{L+M}. \tag{10.33}$$

The maximum allowable step size $\tilde{\tau}_{\max}$ in (10.34) is much larger than $\tau_{\max}$ given in (10.34), and the optimal rate (10.33) is also much better (smaller) than the one in (10.30). In particular, if $\varepsilon \overset{\text{def.}}{=} M/L \ll 1$ (which is the typical setup for ill-posed problems), then

$$\rho \sim 1 - \varepsilon^2 \quad \text{and} \quad \tilde{\rho} \sim 1 - 2\varepsilon.$$

These two results are however complementary. Indeed, if the gradient descent converges, then ultimately $f^{(\ell)}$ is close to $f^\star$, so that one can approximate up to second order $\mathcal{E}(f) \approx \mathcal{E}(f^\star) + \langle Af, f\rangle - \langle f, b\rangle$ with $A = \partial^2 \mathcal{E}(f^\star)$ and $b = -\nabla \mathcal{E}(f^\star)$. So that the "local" rate, the one obtained after a large enough of iterations, is actually driven by $\tilde{\rho}$ and not ρ. It is thus important to distinguish between the global rate and the local rate. In practice, descent algorithm typically have two phase: a first "slow" phase govern by the global rate, and a second "fast" phase governed by the local rate. Unfortunately, the optimal step sizes τ_ℓ are in general different for the two phase, so that optimal adaptation of step size is a difficult problems. This is why more advanced users typically use various line search strategies (to find the optimal step size at each iteration) or use second order information using quasi-Newton technics (BFGS).

The convergence result of Proposition 23 does not requires strong convexity, while Theorem 32 does. In the general non-strongly convex case, it is still possible to prove convergence, but the rate is only sub-linear, and is only on the value of $\mathcal{E}$, not on the iterate $f^{(\ell)}$ themselves. Note that in this case, the solution of the minimization problem is not necessarily unique. The proof is more technical.

Theorem 33. *If* f *satisfy conditions* $(\mathcal{R}_L)$*, assuming there exists* $(\tau_{\min}, \tau_{\max})$ *such that*

$$0 < \tau_{\min} \leqslant \tau_\ell \leqslant \tau_{\max} < \frac{2}{L}, \tag{10.34}$$

then $f^{(\ell)}$ *converges to a solution* $f^\star$ *of* (10.26) *and there exists* $C > 0$ *such that*

$$\mathcal{E}(f^{(\ell)}) - \mathcal{E}(f^\star) \leqslant \frac{C}{\ell}.$$

10.4.3 Examples of Gradient Computation

Note that the gradient of a quadratic function $\mathcal{Q}(f)$ of the form (10.23) reads

$$\nabla \mathcal{G}(f) = Af - b.$$

In particular, one retrieves that the first order optimality condition $\nabla \mathcal{G}(f) = 0$ is equivalent to the linear system $Af = b$.

For the quadratic fidelity term $\mathcal{G}(f) = \frac{1}{2}\|\Phi f - y\|^2$, one thus obtains

$$\nabla \mathcal{G}(f) = \Phi^*(\Phi y - y).$$

In the special case of the regularized TV problem (10.25), the gradient of $\mathcal{E}$ reads

$$\nabla \mathcal{E}(f) = \Phi^*(\Phi y - y) + \lambda \nabla J_{\mathrm{TV}}^\varepsilon(f).$$

Recall the chain rule for differential reads $\partial(\mathcal{G}_1 \circ \mathcal{G}_2) = \partial \mathcal{G}_1 \circ \partial \mathcal{G}_2$, but that gradient vectors are actually transposed of differentials, so that for $\mathcal{E} = \mathcal{F} \circ \mathcal{H}$ where $\mathcal{F} : \mathbb{R}^P \to \mathbb{R}$ and $\mathcal{H} : \mathbb{R}^N \to \mathbb{R}^P$, one has

$$\nabla \mathcal{E}(f) = [\partial \mathcal{H}(f)]^* \left(\nabla \mathcal{F}(\mathcal{H} f)\right).$$

Since $J_{\mathrm{TV}}^{\varepsilon} = \|\cdot\|_{1,\varepsilon} \circ \nabla$, one thus has

$$\nabla J_{\mathrm{TV}}^{\varepsilon} = \nabla^{\star} \circ (\partial \|\cdot\|_{1,\varepsilon}) \quad \text{where} \quad \|u\|_{1,\varepsilon} = \sum_n \|u_n\|_{\varepsilon}$$

so that

$$J_{\mathrm{TV}}^{\varepsilon}(f) = -\operatorname{div}(\mathcal{N}^{\varepsilon}(\nabla f)),$$

where $\mathcal{N}^{\varepsilon}(u) = (u_n/\|u_n\|_{\varepsilon})_n$ is the smoothed-normalization operator of vector fields (the differential of $\|\cdot\|_{1,\varepsilon}$), and where $\operatorname{div} = -\nabla^*$ is minus the adjoint of the gradient.

Since $\operatorname{div} = -\nabla^*$, their Lipschitz constants are equal $\|\operatorname{div}\|_{\mathrm{op}} = \|\nabla\|_{\mathrm{op}}$, and is for instance equal to $\sqrt{2d}$ for the discretized gradient operator. Computation shows that the Hessian of $\|\cdot\|_{\varepsilon}$ is bounded by $1/\varepsilon$, so that for the smoothed-TV functional, the Lipschitz constant of the gradient is upper-bounded by

$$L = \frac{\|\nabla\|^2}{\varepsilon} + \|\Phi\|_{\mathrm{op}}^2.$$

Furthermore, this functional is strongly convex because $\|\cdot\|_{\varepsilon}$ is ε-strongly convex, and the Hessian is lower bounded by

$$M = \varepsilon + \sigma_{\min}(\Phi)^2$$

where $\sigma_{\min}(\Phi)$ is the smallest singular value of Φ. For ill-posed problems, typically $\sigma_{\min}(\Phi) = 0$ or is very small, so that both L and M degrades (tends respectively to 0 and $+\infty$) as $\varepsilon \to 0$, so that gradient descent becomes prohibitive for small ε, and it is thus required to use dedicated non-smooth optimization methods detailed in the following chapters. On the good news side, note however that in many case, using a small but non-zero value for ε often leads to better a visually more pleasing results, since it introduce a small blurring which diminishes the artifacts (and in particular the so-called "stair-casing" effect) of TV regularization.

10.5 Examples of Inverse Problems

We detail here some inverse problem in imaging that can be solved using quadratic regularization or non-linear TV.

10.5.1 Deconvolution

The blurring operator (10.1) is diagonal over Fourier, so that quadratic regularization are easily solved using Fast Fourier Transforms when considering periodic boundary conditions. We refer to (10.21) and the correspond explanations. TV regularization in contrast cannot be solved with fast Fourier technics, and is thus much slower.

10.5.2 Inpainting

For the inpainting problem, the operator defined in (10.3) is diagonal in space

$$\Phi = \operatorname{diag}_m(\delta_{\Omega^c}[m]),$$

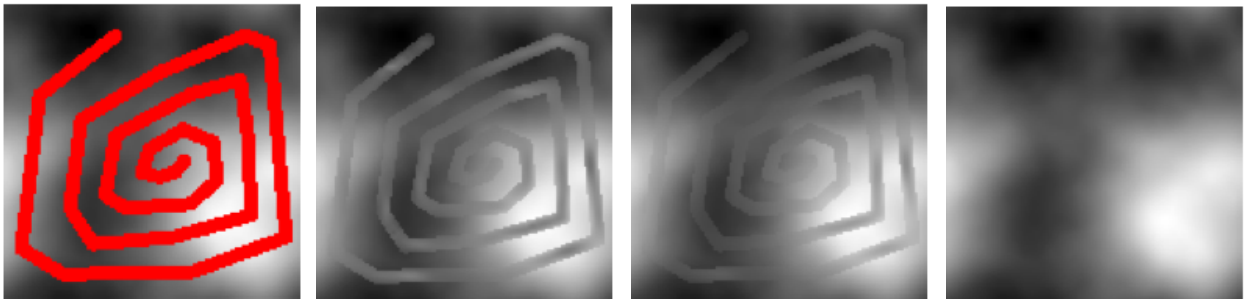
and is an orthogonal projector $\Phi^* = \Phi$.

In the noiseless case, to constrain the solution to lie in the affine space $\{f \in \mathbb{R}^N \;;\; y = \Phi f\}$, we use the orthogonal projector

$$\forall x, \quad P_y(f)(x) = \begin{cases} f(x) & \text{if} \quad x \in \Omega, \\ y(x) & \text{if} \quad x \notin \Omega. \end{cases}$$

In the noiseless case, the recovery (10.17) is solved using a projected gradient descent. For the Sobolev energy, the algorithm iterates

$$f^{(\ell+1)} = P_y(f^{(\ell)} + \tau \Delta f^{(\ell)}).$$



$k = 1$ $k = 10$ $k = 20$ $k = 100$

Figure 10.2: Sobolev projected gradient descent algorithm.

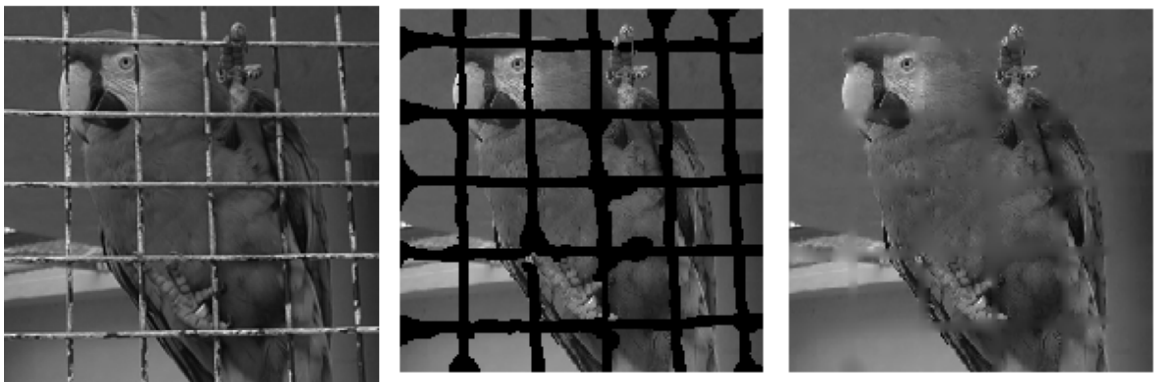


Image f_0 Observation $y = \Phi f_0$ Sobolev $f^\star$

Figure 10.3: Inpainting the parrot cage.

which converges if $\tau < 2/\|\Delta\| = 1/4$. Figure 10.2 shows some iteration of this algorithm, which progressively interpolate within the missing area.

Figure 10.3 shows an example of Sobolev inpainting to achieve a special effect.

For the smoothed TV prior, the gradient descent reads

$$f^{(\ell+1)} = P_y\left(f^{(\ell)} + \tau \operatorname{div}\left(\frac{\nabla f^{(\ell)}}{\sqrt{\varepsilon^2 + \|\nabla f^{(\ell)}\|^2}}\right)\right)$$

which converges if $\tau < \varepsilon/4$.

Figure 10.4 compare the Sobolev inpainting and the TV inpainting for a small value of ε. The SNR is not improved by the total variation, but the result looks visually slightly better.

10.5.3 Tomography Inversion

In medical imaging, a scanner device compute projection of the human body along rays $\Delta_{t,\theta}$ defined

$$x \cdot \tau_\theta = x_1 \cos\theta + x_2 \sin\theta = t$$

where we restrict ourself to 2D projection to simplify the exposition.

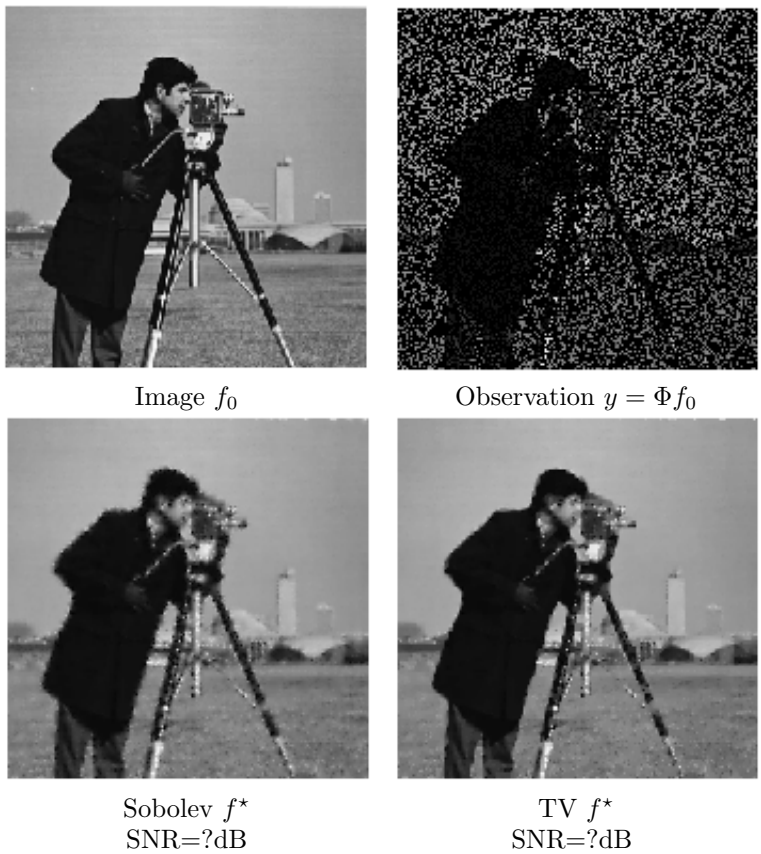


Figure 10.4: Inpainting with Sobolev and TV regularization.

The scanning process computes a Radon transform, which compute the integral of the function to acquires along rays

$$\forall\, \theta \in [0,\pi), \forall\, t \in \mathbb{R}, \quad p_\theta(t) = \int_{\Delta_{t,\theta}} f(x)\, ds = \iint f(x)\, \delta(x \cdot \tau_\theta - t)\, dx$$

see Figure (10.5)

The Fourier slice theorem relates the Fourier transform of the scanned data to the 1D Fourier transform of the data along rays

$$\forall\, \theta \in [0,\pi)\ ,\ \forall\, \xi \in \mathbb{R} \quad \hat{p}_\theta(\xi) = \hat{f}(\xi \cos\theta, \xi \sin\theta). \tag{10.35}$$

This shows that the pseudo inverse of the Radon transform is computed easily over the Fourier domain using inverse 2D Fourier transform

$$f(x) = \frac{1}{2\pi} \int_0^\pi p_\theta \star h(x \cdot \tau_\theta)\, d\theta$$

with $\hat{h}(\xi) = |\xi|$.

Imaging devices only capture a limited number of equispaced rays at orientations $\{\theta_k = \pi/k\}_{0 \leqslant k < K}$. This defines a tomography operator which corresponds to a partial Radon transform

$$Rf = (p_{\theta_k})_{0 \leqslant k < K}.$$

Relation (10.35) shows that knowing Rf is equivalent to knowing the Fourier transform of f along rays,

$$\{\hat{f}(\xi \cos(\theta_k), \xi \sin(\theta_k))\ \}_k.$$

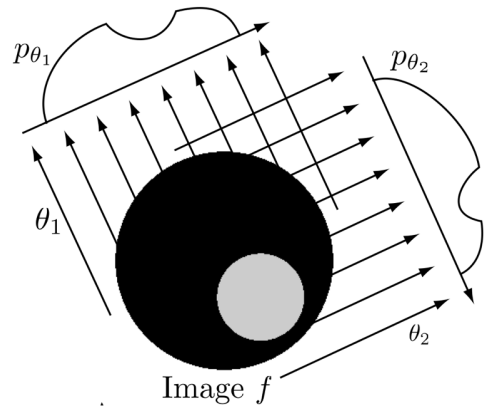


Figure 10.5: Principle of tomography acquisition.

We thus simply the acquisition process over the discrete domain and model it as computing directly samples of the Fourier transform

$$\Phi f = (\hat{f}[\omega])_{\omega \in \Omega} \in \mathbb{R}^P$$

where Ω is a discrete set of radial lines in the Fourier plane, see Figure 10.6, right.

In this discrete setting, recovering from Tomography measures $y = Rf_0$ is equivalent in this setup to inpaint missing Fourier frequencies, and we consider partial noisy Fourier measures

$$\forall\, \omega \in \Omega, \quad y[\omega] = \hat{f}[\omega] + w[\omega]$$

where $w[\omega]$ is some measurement noise, assumed here to be Gaussian white noise for simplicity.

The peuso-inverse $f^+ = R^+ y$ defined in (10.7) of this partial Fourier measurements reads

$$\hat{f}^+[\omega] = \begin{cases} y[\omega] & \text{if} \quad \omega \in \Omega, \\ 0 & \text{if} \quad \omega \notin \Omega. \end{cases}$$

Figure 10.7 shows examples of pseudo inverse reconstruction for increasing size of Ω. This reconstruction exhibit serious artifact because of bad handling of Fourier frequencies (zero padding of missing frequencies).

The total variation regularization (**??**) reads

$$f^\star \in \operatorname*{argmin}_f \frac{1}{2} \sum_{\omega \in \Omega} |y[\omega] - \hat{f}[\omega]|^2 + \lambda \|f\|_{\mathrm{TV}}.$$

It is especially suitable for medical imaging where organ of the body are of relatively constant gray value, thus resembling to the cartoon image model introduced in Section 6.2.4. Figure 10.8 compares this total variation recovery to the pseudo-inverse for a synthetic cartoon image. This shows the hability of the total variation to recover sharp features when inpainting Fourier measures. This should be contrasted with the difficulties that faces TV regularization to inpaint over the spacial domain, as shown in Figure 11.9.

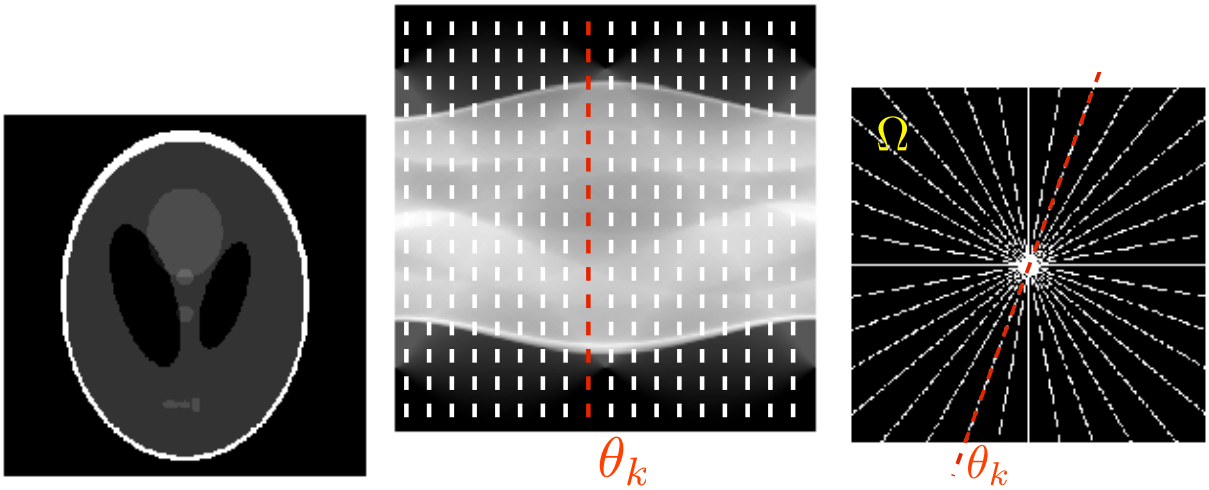


Image f | Radon sub-sampling | Fourier domain

Figure 10.6: Partial Fourier measures.

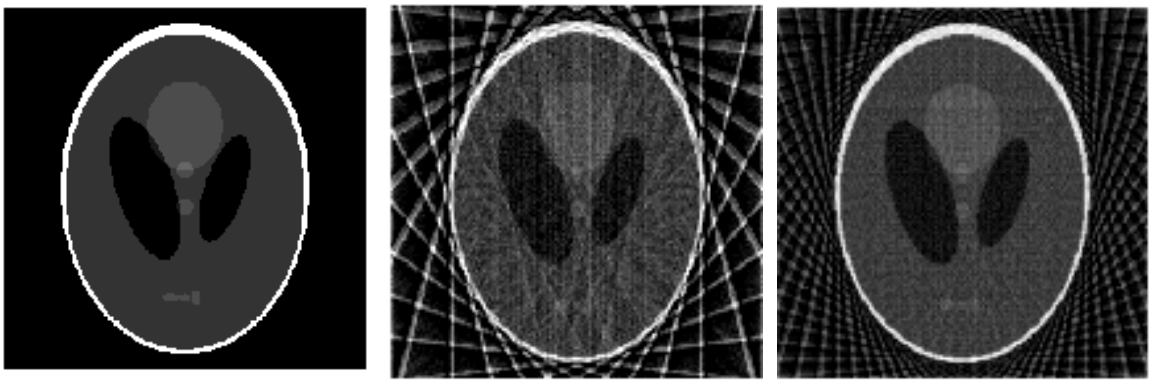


Image f_0 | 13 projections | 32 projections.

Figure 10.7: Pseudo inverse reconstruction from partial Radon projections.

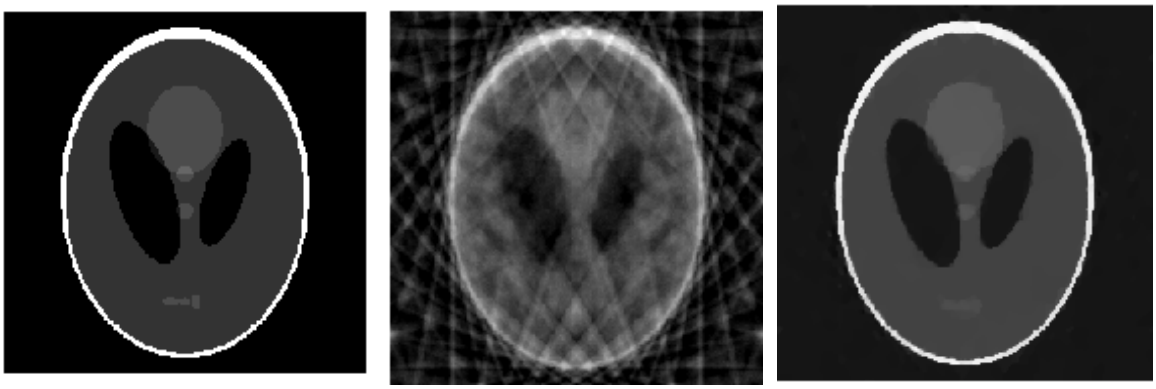


Image f_0 | Pseudo-inverse | TV

Figure 10.8: Total variation tomography inversion.

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] Y. Nesterov. Smooth minimization of non-smooth functions. *Math. Program.*, 103(1, Ser. A):127–152, 2005.

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

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

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

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

[39] 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.

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

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

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

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

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