Konrad-Zuse-Zentrum
für Informationstechnik Berlin

Takustraße 7
D-14195 Berlin-Dahlem
Germany

INGRID DAUBECHIES, GERD TESCHKE, LUMINITA VESE

Iteratively Solving Linear Inverse Problems under General Convex Constraints

ZIB-Report 06-36 (July 2006)

Iteratively Solving Linear Inverse Problems under General Convex Constraints

I. Daubechies*, G. Teschke† and L. Vese‡

3rd August 2006

Abstract

We consider linear inverse problems where the solution is assumed to fulfill some general homogeneous convex constraint. We develop an algorithm that amounts to a projected Landweber iteration and that provides and iterative approach to the solution of this inverse problem. For relatively moderate assumptions on the constraint we can always prove weak convergence of the iterative scheme. In certain cases, i.e. for special families of convex constraints, weak convergence implies norm convergence. The presented approach covers a wide range of problems, e.g. Besov– or BV–restoration for which we present also numerical experiments in the context of image processing.

Keywords. Linear inverse problems, Landweber iteration, Besov- and BV restoration, Generalized shrinkage

1 Scope of the Problem

In a wide range of practical applications one has to solve the problem that the features of interest cannot be observed directly, but have to be interfered from other quantities. Very often a linear model describing the relationship between the features and the measured quantities works quite nicely, i.e. in such situations the task to solve is

$$Tf = h \ .$$

Typically, the observed data are not exactly equal to the image $h = Tf$, but rather a distortion of h, i.e.

$$g = h + e = Tf + e \ .$$

To find an estimate for f from an observation g, one might minimize

$$\|g - Tf\|^2.$$

This kind of problem lies in the class of ill-posed inverse problems if the generalized inverse $T^+ : g \mapsto f^+$ is unbounded. In such cases, the generalized inverse must

*Princeton University, PACM, Washington Road, Princeton, NJ 08544-1000, USA
†Konrad–Zuse–Institute Berlin, Takustr. 7, D-14195 Berlin-Dahlem, Germany
‡Department of Mathematics, UCLA, 405 Hilgard Avenue, Los Angeles, CA 90095-1555, USA

be replaced by bounded approximants, so that numerically stable solutions can be defined.

A typical procedure to avoid these instabilities or to regularize the inverse problem is to modify the functional to be minimized, so that it not only incorporates the discrepancy, but also some a priori knowledge one may have about the solution. More precisely, we consider then a functional of the form

$$\|g - Tf\|_{\mathcal{H}}^2 + 2\alpha J(f), \tag{1.1}$$

where $J(f) < \infty$, or even $J(f) < 1$ is the mathematical translation of the a priori knowledge.

In this paper, we shall consider two different choices of $J(f)$, both adapted to the case where the inverse problem consists in deblurring and denoising a 2-dim. image, as in [DT05], which was in turn, inspired by [DDD04] and [VO02]. Both approaches are natural sequels to [DT05]. In the first approach, we consider $J(f)$ of the same type as in [DT05], but we put it in a more general framework, where $J(f)$ can be any positive, convex, one–homogeneous functional. An extensive discussion of such functionals, in much greater generality than what we present here, is given in [CW05]. In order to be self contained, and to avoid introducing the full complexity of [CW05], we present here a sketch of a simpler version that suffices for our case. In the second approach, $J(f)$ is the same as in [OSV02], but the numerical solution in [OSV02] of a 4-th order nonlinear PDE is replaced by an iterative approach similar to [DDD04] and [DT05].

The paper is organized as follows: in Section 2, we present the framework for our first approach (generalizing [DT05]), which is then discussed in section 3, in several stages: a "warm up" for the case when T^*T has a bounded inverse (in this case, the problem is well–posed), then the ill–posed problem (where 0 is in the spectrum of T^*T), with a convergence analysis. In Section 4, we explain our second approach, an alternative way of minimizing the functional in [OSV02]. Finally, in Section 5, we give numerical results for both.

2 Penalization by a positive, convex, one–homogeneous functional

2.1 Preliminaries

In this Section and the next, we assume that the functional to minimize takes the form (1.1), where J is a positive, convex and one–homogeneous functional. In this case, the variational problem can be recast as follows: Consider J^*, the Fenchel transform or so–called dual functional of J, see [RW98]. Since J is positive and one–homogeneous, there exists a convex set C such that J^* is equal to the indicator function χ_C over C. In Hilbert space, we have total duality between convex sets and positive and one–homogeneous functionals, i.e. $J = (\chi_C)^*$, or

$$(\chi_C)^*(f) = \sup_{h \in C} \langle f, h \rangle = J(f) \ ;$$

see, e.g., [AA03, Cha, CW05]. (Note: [CW05] gives a much more general and complete discussion; we restrict ourselves here to a simple situation, and only sketch

the arguments. For a complete, detailed discussion, we refer the reader to [CW05].) We thus end up with the following reformulation of our problem: given some closed convex set $C \subset \mathcal{H}$ (on which we may still impose extra conditions, below), we wish to minimize

$$\mathcal{F}_C(f) = \|g - Tf\|^2 + 2\alpha \sup_{h \in C} \langle f, h \rangle \ , \tag{2.1}$$

where we assume T to be a bounded operator from $\mathcal{H}$ to itself, with $\|T\| < 1$. We shall consider two particular cases in more detail.

Example 1. As in [DDD04], a particular orthonormal basis $\{\phi_\lambda\}_{\lambda \in \Lambda}$ in $\mathcal{H}$ is preselected, the prior is defined as

$$J(f) = \sum_{\lambda \in \Lambda} |\langle f, \phi_\lambda \rangle| \ .$$

This can, of course, be viewed as a special case of (2.1), since in this case

$$C = \{h \in L_{2,per}([0,1]^2); \ |\langle h, \phi_\lambda \rangle| \leq 1, \quad \forall \lambda\}. \ .$$

Similarly, the case with the prior

$$|f|_w = \sum_{\lambda \in \Lambda} w_\lambda |\langle f, \phi_\lambda \rangle| \ , \ \text{with} \ \inf_{\lambda \in \Lambda} w_\lambda > 0 \ ,$$

fits also into the framework of (2.1), C now defined by

$$C = \{h \in L_{2,per}([0,1]^2); |\langle h, \phi_\lambda \rangle| \leq w_\lambda^{-1}, \quad \forall \lambda\} \ .$$

When $T \neq I$ and the problem is ill-posed, the resulting minimization scheme amounts to Landweber iteration with thresholding applied in each step.

Example 2. In the BV regularization framework, [ROF92], [RO94], one considers functionals of the form

$$\|g - Tf\|^2_{L_2(\Omega)} + 2\alpha \int_\Omega |\nabla f| \tag{2.2}$$

and minimizes over all possible $f \in L_2(\Omega)$. Expressing this functional by means of a convex set C one discovers that C is the L_2-closure of

$$C = \{h \in L_{2,per}([0,1]^2); \ h = \text{div } v, \ \text{where } v \text{ is a 2d field that satisfies}$$

$$\|v\|_{\ell_\infty} = \sup_{(x,y) \in [0,1]^2} (|v_1(x,y)|^2 + |v_2(x,y)|^2)^{1/2} \leq 1\} \ ,$$

i.e. we may again write

$$\sup_{h \in C} \langle f, h \rangle = \int_\Omega |\nabla f| = |f|_{BV} \ ;$$

for details on the structure of C we refer the reader to [Mey02]. It turns out that results on iterative strategies developed for Example 1 carry over to the BV case and that much of the analysis elaborated in [DDD04] can be generalized to the minimization of (2.2).

2.2 Reformulation of the problem

We shall assume that C is a closed convex set in $\mathcal{H}$, C is symmetric, i.e. $h \in C \Rightarrow -h \in C$, and there exists finitely many vectors $a_1, \dots a_N \in \mathcal{H}$, and $r > 0$ so that $B_r(0) \cap \{a_1, \dots a_N\}^\perp \subset C$ (i.e. $\{h : \langle h, a_i \rangle = 0 \; ; i = 1, \dots, N \text{ and } \|h\| < r\} \subset C$). Note: we introduce the finite–dimensional subspace to which C is orthogonal for two reasons. First, there are cases of interest in which C consists of functions that have zero mean in $[0,1]^2$, e.g. if C contains only divergences of smooth periodic fields. Second, it will make it easier to restrict ourselves to only fine scale functions, below.

Defining the functionals

$$L(f,h) := \|g - Tf\|^2 + 2\alpha\langle f, h\rangle,$$

we can rewrite $\inf_{f\in\mathcal{H}} \mathcal{F}_C(f)$ as

$$\inf_{f\in\mathcal{H}} \sup_{h\in C} L(f,h) \; . \tag{2.3}$$

Lemma 2.1 *$L(f,h)$ is continuous in both arguments, it is also convex with respect to f, concave with respect to h.*

Proof. Continuity follows immediately. By

$$L(\lambda v + (1-\lambda)w, h) - \lambda L(v,h) - (1-\lambda)L(w,h) = -[\lambda(1-\lambda)]\|T(v-w)\|^2$$

the second assertion follows. ■

This means that (provided some technical conditions are fulfilled) we can apply the minimax theorem, which allows us to interchange inf and sup in (2.3). In this case the minimax theorem moreover asserts that inf and sup are achieved, i.e. the inf is a min, the sup is a max.

Remark In order to invoke the minimax theorem, we need the following Lemma:

Lemma 2.2 *For the minimization of (2.3), it suffices to consider f in some bounded set $B \subset \mathcal{H}$.*

Proof. We need to show that there is some $R > 0$ so that necessarily

$$\| \arg\min_f \sup_{h\in C} L(f,h)\| \le R \; .$$

First, we have $L(0,h) = \|g\|^2$, thus $\inf_f \sup_{h\in C} L(f,g) \le \|g\|^2$. If we define $V := \text{span}\{a_1, \dots, a_N\}$ (see conditions on C for the definition of the a_i), then we have $V^\perp \cap B_r(0) \subset C$ for all $f \in \mathcal{H}$, so that

$$\sup_{h\in C}\langle f, h\rangle \ge \langle f, \frac{rP_{V^\perp} f}{\|P_{V^\perp} f\|}\rangle = r\|P_{V^\perp} f\| \; ,$$

where $P_{V^\perp}$ denotes the orthogonal projection onto $V^\perp$. It follows that if $\sup_{h\in C}\langle f,h\rangle \leq 2\inf_\varphi \mathcal{F}_C(\varphi)$, then

$$\|P_{V^\perp} f\| \leq \frac{1}{r}\sup_{h\in C}\langle f,h\rangle \leq \frac{2}{r}\|g\|^2 \ .$$

Consequently,

$$\begin{aligned}
\|TP_V f\| &\leq \|Tf\| + \|TP_{V^\perp} f\| \\
&\leq \|g - Tf\| + \|g\| + \|TP_{V^\perp} f\| \\
&\leq \{\sup_{h\in C} L(f,h)\}^{1/2} + \|g\| + \|T\|\frac{2}{r}\|g\|^2 \\
&\leq (1+\sqrt{2})\|g\| + \|T\|\frac{2}{r}\|g\|^2 =: C_1 \ .
\end{aligned}$$

Since V is finite dimensional, $T|_V : V \to TV$ can be represented by an $N \times N$–matrix. If this matrix has a non–trivial kernel, then we can quotient it out. The quotient operator is then bounded below. It follows that there exists an $\tilde{f} \in V$ so that $T\tilde{f} = TP_V f$ and

$$\|\tilde{f}\| \leq \| \left(T|_{V/\ker T|_V}\right)^{-1} \| C_1 =: C_2 < \infty \ .$$

It follows that if $\sup_{h\in C}\langle F,h\rangle \leq 2\inf_f \sup_{h\in C}\langle f,h\rangle$, then there exists an $\bar{F} = \tilde{F} + P_{V^\perp}F$ so that for all $h \in C$ we have

$$L(\bar{F},h) = L(F,h) \ , \text{ and } \|\bar{F}\| \leq C_2 + \frac{2}{r}\|g\|^2 =: R < \infty \ .$$

Consequently, we have

$$\inf_{f\in\mathcal{H}}\sup_{h\in C} L(f,h) = \inf_{f\in\mathcal{H},\|f\|\leq R}\sup_{h\in C} L(f,h) \ .$$

■

Under a variety of different possible conditions on C (e.g. if the set C is bounded; see literature on convex optimization such as [BV05] or [HUL93] for this and other cases) it then follows that

$$\inf_{f\in\mathcal{H},\|f\|\leq R}\sup_{h\in C} L(f,h) = \max_{h\in C}\min_{\|f\|\leq R} L(f,h) \ .$$

In what follows we shall assume that we can indeed apply the minimax theorem and that we can first minimize $L(f,h)$ over f, and then maximize over h in C.

3 Solving the inverse problem for convex penalization

3.1 The well-posed case

Although the case where T^*T does not have a bounded inverse, i.e. where the inverse problem is ill–posed, is of most interest to us, we start by sketching the approach in the easier well–posed case.

Theorem 3.1 *Suppose that all assumptions made above hold true, and T^*T has bounded inverse in its range. If we define $A := (T^*T)^{-1/2}$ and, for an arbitrary closed convex set $K \subset \mathcal{H}$, $S_K := Id - P_K$, where P_K is the (nonlinear) projection on K, i.e. $P_K\varphi = \arg\min_{h\in C} \|h - \varphi\|$, then the minimizing f is given by*

$$f = AS_{\alpha AC}\{AT^*g\} \ .$$

Proof. According to the last section we compute first

$$\min_{\|f\|\leq R} \|g - Tf\|^2 + 2\alpha\langle f, h\rangle.$$

Since T is well-posed, the minimizer is easily found

$$f = (T^*T)^{-1}(T^*g - \alpha h) \ .$$

In this case it remains to determine

$$\begin{aligned} \max_{h\in C} L((T^*T)^{-1}(T^*g - \alpha h), h) &= \max_{h\in C}\{\|g\|^2 - \|(T^*T)^{-1/2}(T^*g - \alpha h)\|^2\} \\ &= -\min_{h\in C} \|(T^*T)^{-1/2}(T^*g - \alpha h)\|^2 \ . \end{aligned}$$

This minimum is achieved when $\alpha(T^*T)^{-1/2}h$ is the convex projection onto the convex set $\alpha(T^*T)^{-1/2}C$ of $(T^*T)^{-1/2}T^*g$,

$$h = \frac{1}{\alpha}(T^*T)^{1/2}P_{\alpha(T^*T)^{-1/2}C}\{(T^*T)^{-1/2}T^*g\}.$$

Finally, the corresponding function f is given by

$$f = (T^*T)^{-1/2}\left[(T^*T)^{-1/2}T^*g - P_{\alpha(T^*T)^{-1/2}C}\{(T^*T)^{-1/2}T^*g\}\right] \ .$$

■

An obvious example is the case where we just need to denoise an image, without deblurring:

Example 3. Consider the denoising problem with an ℓ_1–constraint in the basis $\{\phi_\lambda\}_{\lambda\in\Lambda}$. In this case $T = Id$, so that $A = Id$ as well, and

$$C = \{f;\ \sup_\lambda |\langle f, \phi_\lambda\rangle| \leq 1\} \ .$$

Moreover, in the real case we have

$$\langle P_C f, \phi_\lambda\rangle = \begin{cases} \langle f, \phi_\lambda\rangle & \text{if } |\langle f, \phi_\lambda\rangle| \leq 1 \\ \text{sign}\langle f, \phi_\lambda\rangle & \text{if } |\langle f, \phi_\lambda\rangle| > 1 \ . \end{cases}$$

This implies that $S_{\alpha AC} \circ AT^*$ is exactly the soft thresholding operator

$$\langle S_{\alpha AC}(AT^*g), \phi_\lambda\rangle = \mathrm{S}_\alpha(\langle g, \phi_\lambda\rangle) \ .$$

In the complex case, we have

$$\langle P_C f, \phi_\lambda\rangle = \begin{cases} \langle f, \phi_\lambda\rangle & \text{if } |\langle f, \phi_\lambda\rangle| \leq 1 \\ \frac{\langle f, \phi_\lambda\rangle}{|\langle f, \phi_\lambda\rangle|} & \text{if } |\langle f, \phi_\lambda\rangle| > 1 \ , \end{cases}$$

and the $S_{\alpha AC} \circ AT^*$ reduces to the "complex soft thresholding operator", i.e.

$$\langle S_{\alpha AC}(AT^*g), \phi_\lambda\rangle = \mathrm{S}^c_\alpha(\langle g, \phi_\lambda\rangle) = \begin{cases} \langle g, \phi_\lambda\rangle - \alpha\frac{\langle g, \phi_\lambda\rangle}{|\langle g, \phi_\lambda\rangle|} & \text{if } |\langle g, \phi_\lambda\rangle| > \alpha \\ 0 \quad \text{if } |\langle g, \phi_\lambda\rangle| \leq \alpha \ , \end{cases}$$

3.2 The ill–posed case

In the most interesting problems, the operator T^*T does not have a bounded inverse. We can then use the surrogate functionals introduced in [DDD04]. We replace (1.1) by a family of surrogate functionals

$$\begin{aligned}\mathcal{G}_{n,C}(f) &= \mathcal{F}_C(f) + \|f_n - f\|^2 - \|T(f_n - f)\|^2 \\ &= -2\langle f, T^*g\rangle + 2\alpha \sup_{h\in C}\langle f, h\rangle - 2\langle f, f_n\rangle + 2\langle f, T^*Tf_n\rangle \\ &\qquad + \|f\|^2 + \|g\|^2 + \|f_n\|^2 - \|Tf_n\|^2 \ ,\end{aligned}$$

and we have

Proposition 3.1 *Let C be as assumed in Section 2.2. Then the minimizer of $\mathcal{G}_{n,C}$ is given by*

$$f_{n+1} := (Id - P_{\alpha C})(f_n + T^*g - T^*Tf_n) \ . \tag{3.1}$$

Proof. Under the assumptions made on C, we can apply the minimax theorem to the problem of finding the minimizer of $\mathcal{G}_{n,C}$, so that we have to determine

$$\max_{h\in C} \min_{f\in\mathcal{H}, \|f\|\le\tilde{R}} \left\{ \|f\|^2 - 2\langle f, f_n + T^*g - T^*Tf_n - \alpha h\rangle \right\} \ .$$

For fixed h, the argument is minimized for

$$f = f_n + T^*g - T^*Tf_n - \alpha h \ ,$$

and we need to determine

$$\max_{h\in C} \left\{ -\|f_n + T^*g - T^*Tf_n - \alpha h\|^2 \right\} \ ,$$

which is achieved for

$$h = \frac{1}{\alpha} P_{\alpha C}(f_n + T^*g - T^*Tf_n) \ .$$

■

Next, we discuss convergence properties of the iteration (3.1). We mentioned above that [CW05] contains an extensive discussion, including (not easily verifiable) conditions that ensure strong convergence. The full generality of [CW05] makes it less easy to read if one is mainly interested in the special case presented in this paper. For this reason, we sketch here a more direct (but more specialized) statement. Since the iteration is very similar to the one in [DDD04], a very similar strategy for the proof of convergence holds as well. Up to strong convergence the techniques apply in almost the same way (thus we just sketch the steps). In order to achieve norm convergence, we have to pay more attention to the structure of C, however.

We start by sketching how to achieve weak convergence for our scheme. To this end, we define the nonlinear operator $\mathbf{T}$ by

$$\mathbf{T}f := (Id - P_{\alpha C})(f + T^*g - T^*Tf) \ .$$

In order to establish weak convergence of the $f_n = \mathbf{T}^n f_0$ we apply Opial's Theorem (see [Opi67]) which reads as follows:

Theorem 3.2 *Let the mapping* $\mathbf{A}$ *from* $\mathcal{H}$ *to* $\mathcal{H}$ *satisfy the following conditions:*

(i) $\mathbf{A}$ *is non-expansive:* $\forall v, v' \in \mathcal{H},\ \|\mathbf{A}v - \mathbf{A}v'\| \leq \|v - v'\|$,

(ii) $\mathbf{A}$ *is asymptotically regular:* $\forall v \in H,\ \|\mathbf{A}^{n+1}v - \mathbf{A}^n v\| \overset{n\to\infty}{\longrightarrow} 0$,

(iii) the set $\mathcal{G}$ *of the fixed points of* $\mathbf{A}$ *in* $\mathcal{H}$ *is not empty.*

Then, $\forall v \in \mathcal{H}$*, the sequence* $(\mathbf{A}^n v)_{n\in\mathbb{N}}$ *converges weakly to a fixed point in* $\mathcal{G}$*.*

To apply this to $\mathbf{T}$, we thus need to verify that this operator satisfies the three conditions *(i), (ii)* and *(iii)* above.

First of all, condition *(iii)* is verified, by the following argument. We know that (1.1) has a minimizer $\tilde{f}$. Then $\tilde{f}$ minimizes not only $\mathcal{F}_C$, but also

$$\mathcal{F}_C(f) + \|f - \tilde{f}\|^2 - \|T(f - \tilde{f})\|^2 \ .$$

Consequently, the analysis above implies that $\mathbf{T}\tilde{f} = \tilde{f}$, so $\mathbf{T}$ does indeed have at least one fixed point.

Next, we ensure that $\mathbf{T}$ is asymptotically regular (condition *ii*): we observe that

$$\mathcal{F}_C(f_{n+1}) \leq \mathcal{G}_{n,C}(f_{n+1}) \leq \mathcal{G}_{n,C}(f_n) = \mathcal{F}_C(f_n) \leq \mathcal{G}_{n-1,C}(f_n) \ .$$

This implies

$$\begin{aligned} \left(1 - \|T\|^2\right) \|f_n - f_{n+1}\|^2 &\leq \langle f_n - f_{n+1}, (Id - T^*T)(f_n - f_{n+1})\rangle \\ &= \mathcal{G}_{n,C}(f_{n+1}) - \mathcal{F}_C(f_{n+1}) \leq \mathcal{F}_C(f_n) - \mathcal{F}_C(f_{n+1}) \ . \end{aligned}$$

Consequently,

$$\begin{aligned} \sum_{n=0}^{\infty} \|f_n - f_{n+1}\|^2 &= \lim_{N\to\infty} \sum_{n=0}^{N} \|f_n - f_{n+1}\|^2 \\ &\leq \lim_{N\to\infty} \left(1 - \|T\|^2\right)^{-1} \left[\mathcal{F}_C(f_0) - \mathcal{F}_C(f_{N+1})\right] \\ &\leq \left(1 - \|T\|^2\right)^{-1} \mathcal{F}_C(f_0) < \infty \ , \end{aligned}$$

and hence we deduce $\|f_n - f_{n+1}\| \overset{n\to\infty}{\longrightarrow} 0$.

Finally, we need to check that $\mathbf{T}$ is non–expansive (condition *i*): in order to show this property we need the following standard properties of convex sets.

Lemma 3.1 *Let* K *be a closed and convex set, then for all* $u \in H$ *and all* $k \in K$ *the inequality* $\langle u - P_K u, k - P_K u\rangle \leq 0$ *holds.*

Proof. For all $\lambda \in [0, 1]$ one has

$$\|u - [(1 - \lambda)P_K u + \lambda k]\|^2 \geq \|u - P_K u\|^2 \ .$$

Thus, for all $\lambda \in [0, 1]$

$$-2\lambda\langle u - P_K u, k - P_K u\rangle + \lambda^2 \|k - P_K u\|^2 \geq 0,$$

so that we have $\langle u - P_K u, k - P_K u\rangle \leq 0$. ■

Lemma 3.2 *Let K be a closed and convex set, then for all $u, v \in H$ the inequality $\|u - v - (P_K u - P_K v)\| \leq \|u - v\|$ holds.*

Proof. We need to prove

$$-2\langle u - v, P_K u - P_K v\rangle + \|P_K u - P_K v\|^2 \leq 0 \ .$$

By Lemma 3.1 we have $\langle u - P_K u, P_K v - P_K u\rangle \leq 0$, or equivalently

$$-\langle u, P_K u\rangle + \langle u, P_K v\rangle + \|P_K u\|^2 - \langle P_K u, P_K v\rangle \leq 0 \ .$$

By symmetry we have

$$-\langle v, P_K v\rangle + \langle v, P_K u\rangle + \|P_K v\|^2 - \langle P_K v, P_K u\rangle \leq 0 \ .$$

Summing the two inequalities leads to

$$-\langle u - v, P_K u - P_K v\rangle + \|P_K u - P_K v\|^2 \leq 0 \ ,$$

and thus

$$-2\langle u - v, P_K u - P_K v\rangle + \|P_K u - P_K v\|^2 \leq -\|P_K u - P_K v\|^2 \leq 0 \ .$$

■

Now with the help of Lemma 3.2 and since $0 \leq T^*T \leq Id$ we obtain

$$\begin{aligned} \|\mathbf{T}u - \mathbf{T}v\| &= \|(Id - P_{\alpha C})(u + T^*g - T^*Tu) - (Id - P_{\alpha C})(v + T^*g - T^*Tv)\| \\ &\leq \|(u + T^*g - T^*Tu) - (v + T^*g - T^*Tv)\| \\ &= \|(Id - T^*T)(u - v)\| \leq \|u - v\| \ . \end{aligned}$$

Thus, we have

Proposition 3.2 *Opial's Theorem* 3.2 *applies, i.e. $f_n = \mathbf{T}^n f_0$ converges weakly to a fixed point $\bar{f}$ of $\mathbf{T}$.*

One can argue that this weak convergence theorem suffices for practical purposes, because every numerical computation is always finite–dimensional so that weak and strong (i.e. norm) convergence of the f_n are equivalent. However, it is often useful to establish norm convergence for the infinite dimensional Hilbert space as well, since this then implies that the rate of convergence, and the other constants involved, do not "blow up" as the dimensionality of the discretization increases.

To obtain norm convergence, we need to do some more work. In summary, we have established the following facts:

- $f_n \xrightarrow{\text{weak}} \bar{f}, \quad$ for $n \to \infty$,
- $\bar{f} = \bar{f} + T^*g - T^*T\bar{f} - P_{\alpha C}(\bar{f} + T^*g - T^*T\bar{f})$,
- $f_{n+1} = f_n + T^*g - T^*Tf_n - P_{\alpha C}(f_n + T^*g - T^*Tf_n)$,

- $\|f_{n+1} - f_n\| \to 0, \quad \text{for } n \to \infty.$

Defining

$$u_n := f_n - \bar{f} \;\text{ and }\; v := \bar{f} + T^*g - T^*T\bar{f}, \tag{3.2}$$

we can recast the results as follows:

$$\begin{aligned} &u_n \xrightarrow{\text{weak}} 0, \quad \text{for } n \to \infty \\ &\|P_{\alpha C}(v) - P_{\alpha C}(v + u_n - T^*Tu_n) - T^*Tu_n\| \to 0, \quad \text{for } n \to \infty\ . \end{aligned}$$

We can then apply, without any change, Lemmas 3.15, 3.17 of [DDD04], leading to

$$\|T^*Tu_n\| \to 0, \quad \text{for } n \to \infty,$$

so that we obtain the equivalent formulation

$$\begin{aligned} &u_n \xrightarrow{\text{weak}} 0, \quad \text{for } n \to \infty \\ &\|P_{\alpha C}(v) - P_{\alpha C}(v + u_n)\| \to 0, \quad \text{for } n \to \infty. \end{aligned} \tag{3.3}$$

To obtain norm convergence of the f_n, we must establish $\|u_n\| \to 0$. For general convex sets C the conditions (3.3), where $\alpha > 0$ and $v \in \mathcal{H}$ are arbitrary (but fixed) actually do not imply norm convergence of the u_n to 0. Abstract sufficient and necessary conditions for norm convergence are given in [CW05]; the following theorem gives a more concrete restriction on C under which we can establish norm convergence.

Theorem 3.3 *Suppose $u_n \xrightarrow{\text{weak}} 0$ and $\|P_{\alpha C}(v) - P_{\alpha C}(v + u_n)\| \to 0$. Moreover, assume that u_n is orthogonal to v, $P_C(v)$. If for some sequence γ_n (with $\gamma_n \to \infty$) the convex set C satisfies $\gamma_n u_n \in C$ then $\|u_n\| \to 0$.*

Proof.

Since $\gamma_n u_n \in C$,

$$\frac{\gamma_n}{1+\gamma_n} P_C(v) + \frac{1}{1+\gamma_n}(\gamma_n u_n) \in C\ .$$

Thus,

$$\begin{aligned} \|\frac{\gamma_n}{1+\gamma_n}(P_C(v) + u_n) - (v + u_n)\| &\geq \|P_C(v + u_n) - (v + u_n)\| \\ &\geq \|P_C(v) - (v + u_n)\| - \|P_C(v + u_n) - P_C(v)\|. \end{aligned}$$

This yields

$$\begin{aligned} \|P_C(v) - (v + u_n)\| &\leq \|P_C(v + u_n) - P_C(v)\| \\ &\quad + \|P_C(v) - v - \frac{1}{1+\gamma_n} u_n - \frac{1}{1+\gamma_n} P_C(v))\|. \end{aligned}$$

Since the u_n are uniformly bounded there is some C_1 such that

$$\|P_C(v + u_n) - P_C(v)\| + 2\|P_C(v) - v - \frac{1}{1+\gamma_n} u_n - \frac{1}{1+\gamma_n} P_C(v))\| \leq C_1$$

and because $u_n \perp P_C(v), v$, we obtain from the latter inequalities

$$\begin{aligned}
&\|P_C(v) - v\|^2 + \|u_n\|^2 \\
&\leq \; C_1\|P_C(v+u_n) - P_C(v)\| + \|\frac{\gamma_n}{1+\gamma_n}P_C(v) - v\|^2 + \frac{1}{(1+\gamma_n)^2}\|u_n\|^2,
\end{aligned}$$

which, in turn, gives

$$\begin{aligned}
\left(1 - \frac{1}{(1+\gamma_n)^2}\right)\|u_n\|^2 \;&\leq\; C_1\|P_C(v+u_n) - P_C(v)\| \\
&\quad + \|\frac{\gamma_n}{1+\gamma_n}P_C(v) - v\|^2 - \|P_C(v) - v\|^2 \; .
\end{aligned}$$

With

$$\frac{1}{1+\gamma_n}\|P_C(v)\| + 2\|P_C(v) - v\| \leq C_2$$

we have

$$\|\frac{\gamma_n}{1+\gamma_n}P_C(v) - v\|^2 \leq C_1\frac{1}{1+\gamma_n}\|P_C(v)\| + \|P_C(v) - v\|^2$$

and consequently,

$$\left(1 - \frac{1}{(1+\gamma_n)^2}\right)\|u_n\|^2 \leq C_1\|P_C(v+u_n) - P_C(v)\| + C_2\frac{1}{1+\gamma_n}\|P_C(v)\|$$

Which shows that if $n \to \infty$ then $\|u_n\| \to 0$. ■

Unfortunately, this theorem is not sufficiently strong to apply to the BV–functional of *Example 2*, above. Without going in full detail, we sketch here how it (just) falls short.

The set C in *Example 2* is (loosely speaking) the set of all divergences of 2-dim. fields that are uniformly bounded by 1. It contains, in particular, the functions

$$\begin{aligned}
h_{\mathbf{n}}(x,y) \;&:=\; \sqrt{2}\pi(|n_1| + |n_2|)e^{2\pi i(n_1x + n_2y)} \\
&=\; -i\mathrm{div}\Big(\frac{1}{\sqrt{2}}\mathrm{sign}(n_1)e^{2\pi i(n_1x+n_2y)}, \frac{1}{\sqrt{2}}\mathrm{sign}(n_2)e^{2\pi i(n_1x+n_2y)}\Big) \; ,
\end{aligned}$$

where $|n_1| + |n_2| \neq 0$. Because C is closed and convex, it also contains all the

$$\sum_{\mathbf{n}\in\mathbb{Z}^2} \alpha_{\mathbf{n}} h_{\mathbf{n}} \; ,$$

with $\sum_{\mathbf{n}\in\mathbb{Z}^2} \alpha_{\mathbf{n}} = 1$. Suppose now (just for the sake of simplifying the argument, which can also be made, a bit more lengthily, without this assumption) that

$$\|P_C(u_n)\| \to 0 \text{ as } n \to \infty \; ,$$

i.e. that the condition

$$\|P_{\alpha C}(v) - P_{\alpha C}(v+u_n)\| \to 0 \text{ as } n \to \infty$$

holds true for $v = 0$. That would mean that, for all $w \in C$

$$\begin{aligned}
&\lim_{n\to\infty} \langle u_n - P_C(u_n), w - P_C(u_n)\rangle \\
&\\
&= \lim_{n\to\infty} \langle u_n, w\rangle + \|P_C(u_n)\|^2 - \langle u_n, P_C(u_n)\rangle - \langle P_C(u_n), w\rangle \\
&= \lim_{n\to\infty} \langle u_n, w\rangle \ ,
\end{aligned}$$

which implies that $\lim_{n\to\infty}\langle u_n, w\rangle$ is nonpositive. Since the same is true for $-w \in C$, if follows that $\lim_{n\to\infty}\langle u_n, w\rangle = 0$ for all $w \in C$. Consequently, $\langle u_n, h_{\mathbf{k}}\rangle \to 0$ as $n \to \infty$, or even, for all sequences $(\alpha_{\mathbf{k}})_{\mathbf{k}\in\mathbb{Z}^2}$ with $\sum_{\mathbf{k}\in\mathbb{Z}^2} \alpha_{\mathbf{k}} = 1$, $\alpha_{\mathbf{k}} \in [0,1]$ $\forall \mathbf{k}$,

$$\sum_{\mathbf{k}} \alpha_{\mathbf{k}}(|k_1| + |k_2|)\langle u_n, e_{\mathbf{k}}\rangle \to 0 \text{ as } n \to \infty \ ,$$

where $e_{\mathbf{k}}(x,y) = e^{2\pi i(k_1 x + k_2 y)}$. This just misses ensuring that

$$\sum_{\mathbf{k}} |\langle u_n, e_{\mathbf{k}}\rangle|^2 \to 0 \text{ as } n \to \infty \ .$$

This concludes our theoretical analysis of our first case described in the introduction, i.e. the case where $J(f)$ in (1.1) is convex. We now turn to the second approach.

4 Iterative algorithm for PDE–based deblurring and denoising.

We start by recalling briefly the framework of [RO94], in which the inverse problem $g = Tf + e$, which edge-preserving regularization, is cast as the minimization of an energy functional of the form (assuming Gaussian noise for simplicity)

$$\inf_f \mathcal{F}_\phi(f) = \|Tf - g\|^2_{L^2(\Omega)} + 2\alpha \int_\Omega \phi(\nabla f) dx \ ; \tag{4.1}$$

here the potential $\phi : \mathbb{R}^2 \to \mathbb{R}$ is typically a positive continuous increasing function, with at most linear growth at infinity, therefore satisfying $|\phi(\xi)| \le c(1 + |\xi|)$, for some positive constant c (note that we have replaced the subscript C by ϕ describing now the penalty and not the convex set of the previous approach). Convex examples include (note that, only for illustration reasons, we give also examples beyond the one–homogeneous case)

- $\phi(\xi) = |\xi|$ (the total variation minimization of Rudin-Osher-Fatemi [ROF92], [RO94]),
- $\phi(\xi) = |\xi_1| + |\xi_2|$,
- $\phi(\xi) = \sqrt{1 + |\xi|^2}$ (the function of minimal surfaces),
- $\phi(\xi) = \log\cosh(1 + |\xi|^2)$, or

- $\phi(\xi) = \begin{cases} \frac{1}{2}|\xi|^2 \text{ if } |\xi| \leq 1 \\ |\xi| - \frac{1}{2} \text{ if } |\xi| \geq 1 \end{cases}$ (used in [DT84], [CL97]).

In the non-convex case, examples of the potential ϕ are

- $\phi(\xi) = \frac{|\xi|^p}{1+|\xi|^p}$ or
- $\phi(\xi) = \log(1+|\xi|^p)$, with $p = 1$ or $p = 2$ for instance, see Perona-Malik [PM90], Geman-Geman, [GG84], Geman-Reynolds [GR92], and more recently [CL97], [AV97], [Ves01].

Let us now restrict again to the one–homogeneous case and assume in addition that ϕ is differentiable. Then the Euler-Lagrange equation associated with the minimization problem (4.1), that must be satisfied by a minimizer f, if such a minimizer exists, is given by

$$T^*Tf - T^*g = \alpha \text{div}\Big(\nabla_\xi \phi(\nabla f)\Big), \text{ in } \Omega, \tag{4.2}$$

where $\nabla \phi_\xi = (\phi_{\xi_1}, \phi_{\xi_2})$, and with the boundary conditions $\nabla_\xi \phi(\nabla f) \cdot \vec{n} = 0$ on $\partial\Omega$, where $\vec{n}$ is the unit exterior normal to the boundary. In case $\alpha > 0$, the partial differential equation (4.2) is non-linear for the examples of potential ϕ given above. Moreover, the presence of the term T^*Tf makes it computationally expensive and numerically nontrivial.

In order to overcome these problems, we propose here to not directly solve (4.2) numerically, but to apply the surrogate functional algorithm (see [DDD04], or the previous sections), i.e. we construct a sequence of iterates f_n that approximate f, without having to invert T^*T at every iteration. On the other hand, the direct implementation of the projection $P_{\alpha C}$ associated to our minimization is rather complicated; in this case we prefer to avoid it by switching to an expression based on the Euler-Lagrange equation. The total iteration goes thus as follows: start with an initial f_0; find f_n, $n > 0$ as a minimizer of the surrogate functionals

$$\mathcal{G}_{n-1}(f_n) = \|Tf_n - g\|^2_{L^2(\Omega)} - \|Tf_n - Tf_{n-1}\|^2_{L^2(\Omega)} + \|f_n - f_{n-1}\|^2_{L^2(\Omega)} + 2\alpha \int_\Omega \phi(\nabla f_n), \tag{4.3}$$

where we have assumed that $\|T^*T\| < 1$. The associated Euler-Lagrange equation in f_n, now easily solved in practice, is:

$$f_n = f_{n-1} + T^*g - T^*Tf_{n-1} + \alpha \text{div}\Big(\nabla_\xi \phi(\nabla f_n)\Big), \tag{4.4}$$

together with the same boundary conditions. One then simply carries out this iterative algorithm to find (an approximation to) desired minimizer.

Remark The same idea can be also applied to other cases, when the $L^2(\Omega)$ norm in (4.1) is replaced by another norm in a Hilbert space.

For instance, let us consider the minimization problem from Osher-Solé-Vese [OSV02] for denoising, for simplicity,

$$\inf_{f \in BV(\Omega)} \mathcal{F}(f) \text{ with } \mathcal{F}(f) = \int_\Omega |\nabla \triangle^{-1}(f - g)|^2 dx + 2\alpha \int_\Omega |\nabla f| dx. \tag{4.5}$$

This model is a refinement over the total variation minimization [ROF92], and decomposes the data g into a $BV(\Omega)$ component f and an oscillatory component $g-f$ that belongs to $(H_0^1(\Omega))'$, the dual of $H_0^1(\Omega)$ when endowed with the standard seminorm. Even in the case without blur, the minimization of the convex functional (4.5) leads to a computationally expensive partial differential equation in f,

$$\triangle^{-1}(g-f) = \alpha \mathrm{div}\Big(\frac{\nabla f}{|\nabla f|}\Big), \tag{4.6}$$

or to a fourth order non-linear partial differential equation

$$g - f = \alpha \triangle \Big[\mathrm{div}\Big(\frac{\nabla f}{|\nabla f|}\Big)\Big], \tag{4.7}$$

with associated boundary conditions. By applying the surrogate functional approach to the minimization of (4.5), we can avoid solving this PDE directly; instead compute f_n as minimizer of the convex functional

$$\begin{aligned}\mathcal{G}_{n-1}(f_n) &= \int_\Omega |\nabla \triangle^{-1}(f_n - g)|^2 dx + 2\alpha \int_\Omega |\nabla f| dx \\ &\quad - \int_\Omega |\nabla \triangle^{-1}(f_n - f_{n-1})|^2 dx + \mu \int_\Omega |f_n - f_{n-1}|^2 dx.\end{aligned}$$

The minimization of the energy $\mathcal{G}_{n-1}$ with respect to f_n leads to a simpler partial differential equation

$$\mu f_n = \mu f_{n-1} - \triangle^{-1} g + \triangle^{-1} f_{n-1} + \alpha \mathrm{div}\Big(\frac{\nabla f_n}{|\nabla f_n|}\Big). \tag{4.8}$$

5 Numerical Illustrations

In this section, we present numerical results of the two approaches. We assume the linear degradation model $g = Tf + e$, where g is the given data, as a square integrable function in $L^2(\Omega)$, f is the unknown true image, e is additive noise of zero mean. The operator $T : L^2(\Omega) \to L^2(\Omega)$ models a linear and continuous degradation operator, by a convolution with a Gaussian kernel.

In the first approach let now $\{\varphi_\lambda\}_{\lambda\in\Lambda}$ be a frame, i.e. there exists positive and bounded constants $0 < A \le B < \infty$, such that for $f \in L^2(\Omega)$,

$$A\|f\|_{L^2(\Omega)} \le \sum_{\lambda\in\Lambda} |\langle f, \varphi_\lambda\rangle|^2 \le B\|f\|_{L^2(\Omega)} \ ;$$

we define the corresponding frame operator F,

$$F : f \mapsto \{\langle f, \varphi_\lambda\rangle\}_{\lambda\in\Lambda} \ .$$

By means of the adjoint F^* the variational problem (2.1) reads in the ℓ_1 setting as follows

$$\mathcal{F}(\mathbf{f}) = \|g - TF^*\mathbf{f}\|^2_{L^2(\Omega)} + 2\alpha|\mathbf{f}|_{\ell_1} \ , \tag{5.1}$$

Figure 1: Top left: original image. Top right: blurred version.

Figure 2: Top left to bottom right: blurred image, several iterates: 1st, 100th, 500th, 1000th, 2000th, 3000th, 4000th.

with the shorthand notation $\mathbf{f} = \{f_\lambda\}_{\lambda\in\Lambda}$. The convex set C is now

$$C = \{\{h_\lambda\}_{\lambda\in\Lambda};\ \sup_\lambda |h_\lambda| \le 1\}.$$

In accordance with Proposition 3.1, the resulting iteration scheme reads as

$$\begin{aligned} \mathbf{f}_{n+1} &= (Id - P_{\alpha C})(\mathbf{f}_n + FT^*g - FT^*TF^*\mathbf{f}_n) \\ &= S_\alpha(\mathbf{f}_n + FT^*g - FT^*TF^*\mathbf{f}_n), \end{aligned} \tag{5.2}$$

where S_α denotes the soft shrinkage operator with shrinkage parameter α. Wavelet–based methods of this kind (in the context of image decomposition) are considered in [DT04, DT05]. In particular, we have chosen here a wavelet frame that is simply given by a translation invariant wavelet system. As the example image we consider a finger print and its blurred version, see Figure 1. The results obtained with iteration (5.2) are visualized in Figure 2 and the convergence rates are given in Table 1. The blur operator T used in the experiments has the discrete spatial representation given in Table 2. The blur convolution is easily implemented as a multiplication

ITER	CPU	RMSE $\|f_n - f_{orig}\|/\|f_{orig}\|$
1	0.28600841	0.17322202591876
100	11.0832407	0.16619377802673
200	21.3800888	0.16053661204032
350	36.9959693	0.15341521024097
500	52.2572860	0.14754897184399
1000	104.031806	0.13476786059559
2000	207.406313	0.12458884053496
3000	312.880765	0.11962522471006
4000	419.051499	0.11633033283685
5000	524.362921	0.11388107705039

Table 1: Convergence rates for the wavelet frame–based shrinkage algorithm. We give the number of iterations, the CPU time and the corresponding relative RMSE, applied to the blurry fingerprint image.

$\frac{1}{60}$

0	0	0	0	0	0.25	0	0	0	0	0
0	0	0.25	0.5	0.5	0.5	0.5	0.5	0.25	0	0
0	0.25	0.5	1	1	1	1	1	0.5	0.25	0
0	0.5	1	1	1	1	1	1	1	0.5	0
0	0.5	1	1	1	1	1	1	1	0.5	0
0.25	0.5	1	1	1	1	1	1	1	0.5	0.25
0	0.5	1	1	1	1	1	1	1	0.5	0
0	0.5	1	1	1	1	1	1	1	0.5	0
0	0.25	0.5	1	1	1	1	1	0.5	0.25	0
0	0	0.25	0.5	0.5	0.5	0.5	0.5	0.25	0	0
0	0	0	0	0	0.25	0	0	0	0	0

Table 2: Spatial discretization of the blur operator T.

in Fourier domain, which means that we switch between the wavelet and Fourier representation at every step of the iteration process.

Next, we present numerical results for the second (PDE) approach. In figure 3 we show the results of the iterative algorithm (4.4) on the same blurred and noisy image. For comparison with the purely PDE–based method (without the iterative approach corresponding to surrogate functionals) we show in Figure 4 the end results of methods (4.4) and (4.2); they look very similar. Table 3 lists the CPU time and the relative RMSE for the first 5000 iterations of both methods, illustrating that the surrogate functional method produces a better error decay for the same amount of CPU time. (These two computations were carried out on the same machine; note that the numerical results in Table 2 were obtained on a different computer and should thus not be compared with this Table.) Finally, we show in Figure 5 the result of a Cartoon+Texture decomposition of the same image (without blur) obtained by the surrogate–based iteration (4.8).

Figure 3: Top left to bottom right: blurred image, several iterates using (4.4): 1st, 100th, 500th, 1000th, 2000th, 3000th, 4000th.

With surrogate functional Without surrogate functional

Figure 4: Deblurring results obtained using the models (4.4) left and (4.2) right, with $\phi(\xi) = \sqrt{\epsilon + |\xi|^2}$ (total variation minimization with regularization).

OSV model with surrogate functional

Figure 5: Decomposition into cartoon (left) and texture (right) of the fingerprint image without blur, obtained using the model (4.8).

With surrogate functionals

ITER	CPU	RMSE$\|f_n - f_{orig}\|/\|f_{orig}\|$
1	0	0.256365806
100	53.7299995	0.22643654
200	107.440002	0.213100523
350	188.470001	0.19897379
500	268.699982	0.188747868
1000	542.880005	0.168787047
2000	1085.82996	0.153525516
3000	1625.08997	0.148305818
4000	2151.29004	0.146324202
5000	2684.62988	0.145639017

Without surrogate functionals

ITER	CPU	RMSE$\|f_n - f_{orig}\|$
1	0.400000036	0.257728785
100	55.1000023	0.236117765
200	109.139999	0.226667866
350	191.959991	0.216331303
500	276.119995	0.208168939
1000	551.720032	0.189164072
2000	1105.32996	0.169159457
3000	1666.43994	0.159271181
4000	2234.70996	0.15375042
5000	2790.92993	0.150457606

Table 3: Comparison of the convergence rates: for both algorithms, we give the number of iterations, the CPU time and the corresponding relative RMSE, applied to the blurry fingerprint image, using the total variation minimization. We notice that the new method using the surrogate functionals converges faster to the restored image: the relative RMSE $\|f_n - f_{orig}\|/\|f_{orig}\|$ hits the value 0.15 at 2.500 iterations instead of 5.000, and uses a CPU time of $\sim$ 1300 instead of 2790; there seems thus to be a speed–up factor 2.

6 Conclusion

In this paper, we have extended the approach of [DDD04], or, alternatively, illustrated with concrete examples the abstract analysis in [CW05]. In particular, we have written an iteration algorithm for solving linear inverse problems of the following general variational form

$$\mathcal{F}_C(f) = \|g - Tf\|^2 + 2\alpha \sup_{h \in C} \langle f, h \rangle \ ,$$

and we have applied it to sparse representation w.r.t. a wavelet frame. We have also shown that the iterative approach that naturally follows from introducing surrogate functionals, leads to a simplified solution algorithm for PDE–linked variational problems, resulting in a more efficient algorithm.

We have discussed convergence and a sequence of numerical illustrations that verify the usefulness of the iteration.

References

[AA03] G. Aubert and J.-F. Aujol. Modeling very oscillating signals. Applications to image processing. *INRIA report No 4878, ISSN 0249-6399*, 2003.

[AV97] G. Aubert and L. Vese. A variational method in image recovery. *SINUM*, 34(5):1948–1979, October 1997.

[BV05] S. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge Univ. Press, Cambridge, 2005.

[Cha] A. Chambolle. An Algorithm for total variation minimization and applications. *CEREMADE - CNRS UMR 7534*.

[CL97] A. Chambolle and P.-L. Lions. Image recovery via total variation minimization and related problems. *Numerische Mathematik*, 76(2):167–188, April 1997.

[CW05] P. L. Combettes and V. R. Wajs. Signal recovery by forward–backward splitting. March 2005.

[DDD04] I. Daubechies, M. Defrise, and C. DeMol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Comm. Pure Appl. Math*, 57:1413–1541, 2004.

[DT84] F. Demengel and R. Temam. Convex functions of a measure and applications. *Indiana University Mathematics Journal*, 33(5):673–709, 1984.

[DT04] I. Daubechies and G. Teschke. Wavelet–based image decomposition by variational functionals. *Proc. SPIE Vol. 5266, p. 94-105, Wavelet Applications in Industrial Processing; Frederic Truchetet; Ed.*, Feb. 2004.

[DT05] I. Daubechies and G. Teschke. Variational image restoration by means of wavelets: simultaneous decomposition, deblurring and denoising. *Applied and Computational Harmonic Analysis*, (19(1)):1–16, 2005.

[GG84] S. Geman and D. Geman. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Transactions on PAMI*, 6(6):721–741, 1984.

[GR92] D. Geman and G. Reynolds. Constrained restoration and the recovery of discontinuities. *IEEE Transactions on PAMI*, 14(3):367–383, March 1992.

[HUL93] J-B. Hiriart-Urruty and C. Lemaréchal. *Convex analysis and minimization algorithms I.* Springer, Berlin, 1993.

[Mey02] Y. Meyer. Oscillating Patterns in Image Processing and Nonlinear Evolution Equations. *University Lecture Series Volume 22, AMS*, 2002.

[Opi67] Z. Opial. Weak convergence of the sequence of successive approximations for nonexpansive mappings. *Bull. Amer. Math. Soc.*, (73):591–597, 1967.

[OSV02] S. Osher, A. Sole, and L. Vese. Image decomposition and restoration using total variation minimization and the H^{-1} norm. Technical Report 02-57, University of California Los Angeles C.A.M., 2002.

[PM90] P. Perona and J. Malik. Scale space and edge detection using anisotropic diffusion. *IEEE Trans. Pattern Anal. Mach. Intell.*, 12:629–639, 1990.

[RO94] L.I. Rudin and S. Osher. Total variation based image restoration with free local constraints. *Proceedings. ICIP-94.*, 1:31–35, November 1994.

[ROF92] L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60(1-4):259–268, November 1992.

[RW98] R. T. Rockafellar and R. J-B. Wets. *Variational Analysis*. Springer, Berlin, 1998.

[Ves01] L. Vese. A study in the bv space of a denoising-deblurring variational problem. *Applied Mathematics and Optimization*, 44(2):131–161, September-October 2001.

[VO02] L. Vese and S. Osher. Modeling textures with total variation minimization and oscillating patterns in image processing. Technical Report 02-19, University of California Los Angeles C.A.M., 2002.