ECOLE POLYTECHNIQUE
CENTRE DE MATHÉMATIQUES APPLIQUÉES
UMR CNRS 7641

91128 PALAISEAU CEDEX (FRANCE). Tél: 01 69 33 46 00. Fax: 01 69 33 46 46
http://www.cmap.polytechnique.fr/

Continuous Primal-Dual methods for Image Processing

Goldman Michael

R.I. 677 *Mars 2010*

Continuous Primal-Dual methods for Image Processing

M. Goldman*

March 17, 2010

Abstract

In this article we study a continuous Primal-Dual method proposed by Appleton and Talbot and generalize it to other problems in image processing. We interpret it as an Arrow-Hurwicz method which leads to a better description of the system of PDEs obtained. We show existence and uniqueness of solutions and get a convergence result for the denoising problem. Our analysis also yields new *a posteriori* estimates.

Acknowledgements. I would like to warmly thank my PhD advisor Antonin Chambolle for suggesting me this problem and for our fruitful discussions. This research was partially supported by ANR project MICA (2006-2009).

1 Introduction

In imaging, duality has been recognized as a fundamental ingredient for designing numerical schemes solving variational problems involving a total variation term. Primal-Dual methods were introduced in the field by Chan, Golub and Mulet in [10]. Afterwards, Chan and Zhu [17] proposed to rewrite the discrete minimization problem as a min-max and solve it using an Arrow-Hurwicz [5] algorithm which is a gradient ascent in one direction

*CMAP, CNRS UMR 7641, Ecole Polytechnique, 91128 Palaiseau Cedex, France. e-mail: goldman@cmap.polytechnique.fr

and a gradient descent in the other. Just as for the simple gradient descent, one can think of extending this method to the continuous framework. This is in fact what does the algorithm previously proposed by Appleton and Talbot in [4] derived by analogy with discrete graph cuts techniques. The first to notice the link between their method and Primal-Dual schemes were Chambolle and al. in [8]. This continuous framework leaves the way open to a wide variety of numerical schemes, ranging from finite differences to finite volumes. These numerical methods can also reduce the discretization bias and lead to more anisotropic results compared with classical discretizations of the total variation.
This paper proposes to study the continuous Primal-Dual algorithm following the philosophy of the work done for the gradient flow by Andreu and al. in [2]. We give a rigorous definition of the system of PDEs which is obtained and show existence and uniqueness of a solution to the Cauchy problem. We prove strong L^2 convergence to the minimizer for the Rudin-Osher-Fatemi model and derive some *a posteriori* estimates. As a byproduct of our analysis we also obtain *a posteriori* estimates for the numerical scheme proposed by Chan and Zhu.

1.1 Presentation of the problem

Many problems in image processing can be seen as minimizing in $BV \cap L^2$ an energy of the form

$$J(u) = \int_\Omega |Du| + G(u) + \int_{\partial\Omega_D} |u - \varphi|$$

We assume that Ω is a bounded Lipschitz open set and that $\partial\Omega_D$ is a subset of $\partial\Omega$. The function φ being given in $L^1(\partial\Omega_D)$, the term $\displaystyle\int_{\partial\Omega_D} |u-\varphi|$ is a Dirichlet condition on $\partial\Omega_D$. We call $\partial\Omega_N$ the complement of $\partial\Omega_D$ in $\partial\Omega$ and assume that G is convex and lower-semi-continuous (lsc) in L^2 with

$$G(u) \le C(1 + |u|_2^p) \qquad \text{with } 1 \le p \le +\infty$$

In this paper we note $|u|_2$ the L^2 norm of u. According to Giaquinta and al. [13], one has,

Proposition 1.1. *The functional J is convex and lsc in L^2.*

In the following, we also assume that J attains its minimum in $BV \cap L^2$. This is for example true if G satisfies some coercivity hypothesis or if G is non negative.

Two fundamental applications of our method are image denoising via total variation regularization and segmentation with geodesic active contours.

In the first problem, one starts with a corrupted image $f = \bar{u} + n$ and wants to find the clean image $\bar{u}$. Rudin, Osher and Fatemi proposed to look for an approximation of $\bar{u}$ by minimizing

$$\int_\Omega |Du| + \frac{\lambda}{2} \int_\Omega (u-f)^2$$

This corresponds to $G(u) = \frac{\lambda}{2} \int_\Omega (u-f)^2$ and $\partial\Omega_D = \emptyset$. For a comprehensive introduction to this subject, we refer to the lecture notes of Chambolle and al. [9]. Figure 1 shows the result of denoising using the algorithm of Chan and Zhu.

Figure 1: Denoising using the ROF model

The issue in the second problem is to extract automatically the boundaries of an object within an image. We suppose that we are given two subsets S and T of $\partial\Omega$ such that S lies inside the object that we want to segment and T lies outside. Caselles and al. proposed in [7] to associate a positive function g to the image in a way that g is high where the gradient of the image is low and vice versa. The object is then segmented by minimizing

$$\min_{E \supset S,\, E^c \supset T} \int_{\partial E} g(s)ds$$

In order to simplify the notations, we will only deal with $g = 1$ in the following. It is however straightforward to extend our discussion to general (continuous) g. The energy we want to minimize is thus $\int_\Omega |D\chi_E|$. This functional is non convex but by the coarea formula (see Ambrosio-Fusco-Pallara [1]), it can be relaxed to functions $u \in [0,1]$. Let $\varphi = 1$ on S and $\varphi = 0$ on T. Letting $\partial\Omega_D = S \cup T$, and f be a L^2 function, our problem can be seen as a special case of the prescribed mean curvature problem,

$$\inf_{\substack{0 \le u \le 1 \\ u=\varphi \text{ in } \partial\Omega_D}} \int_\Omega |Du| + \int_\Omega fu$$

The solution E is then any superlevel of u, namely $E = \{u > s\}$ for any $s \in]0,1[$.

It is however well known that the infimum is not attained in general because of the lack of compactness for the boundary conditions in BV. Following the ideas of Giaquinta and al. [13] we have to relax the boundary conditions by adding $\int_{\partial\Omega_D} |u - \varphi|$ to the functional. We also have to deal with the hard constraint on u. This will be discussed afterwards but it brings some mathematical difficulties that we were not able to solve. Fortunately, our problem is equivalent (see [8]) to the minimization of the unconstrained problem

$$J(u) = \inf_{u \in BV(\Omega)} \int_\Omega |Du| + \int_{\partial\Omega_D} |u - \varphi| + \int_\Omega f^+|u| + \int_\Omega f^-|1-u|$$

Here $f^+ = \max(f, 0)$ and $f^- = \max(-f, 0)$.

We give in figure 2 the result of this segmentation on yeasts. The small square is the set S and the set T is taken to be the image boundary. The study of this problem was in fact our first motivation for this work.

1.2 Idea of the Primal-Dual method

Formally, the idea behind the Primal-Dual method is using the definition of $\int_\Omega |Du|$ in order to write J as

$$J(u) = \sup_{\substack{\xi \in \mathcal{C}^1_c(\Omega) \\ |\xi|_\infty \le 1}} K(u, \xi)$$

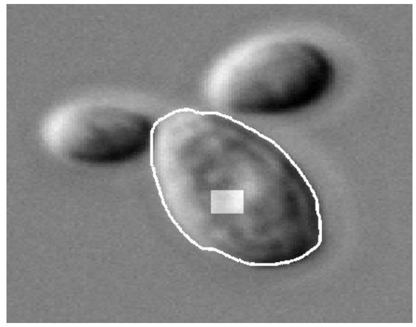

Figure 2: Yeast segmentation

Where $K(u,\xi) = -\int_\Omega u\,\mathrm{div}(\xi) + \int_{\partial\Omega_D} |u-\varphi| + G(u)$. Then, one seeks for a saddle point of K by a gradient descent in u and a gradient ascent in ξ, which amounts to solve the system of PDEs:

$$\begin{cases} \partial_t u = \mathrm{div}(\xi) - \partial G(u) \\ \partial_t \xi = Du - \partial I_{B(0,1)}(\xi) \\ + \text{ boundary conditions} \end{cases} \tag{1}$$

Here $I_{B(0,1)}(\xi)$ is the indicator function of the unit ball in L^∞ (it takes the value 0 if $|\xi|_\infty \leq 1$ and $+\infty$ otherwise) and ∂ stands for the subdifferential (see Ekeland-Temam for the definition [12]). This system is almost the one proposed by Appleton and Talbot in [4] for the segmentation problem.

Let us remark that, at least formally, the differential operator $A(u,\xi) = \begin{pmatrix} -\mathrm{div}\,\xi + \partial G(u) \\ -Du + \partial I_{B(0,1)}(\xi) \end{pmatrix}$ verifies $\langle A(u,\xi),(u,\xi)\rangle \geq 0$, which means that A is monotone.

We recall some facts about the theory of maximal monotone operators and its applications for finding saddle points in the next section. In the last section we use it to give a rigourous meaning to the hyperbolic system (1) together with existence and uniqueness of solutions of the Cauchy problem.

2 Maximal Monotone Operators

Following Brézis [6], we present briefly in the first part of this section the theory of maximal monotone operators. In the second part we show how this theory shade light on the general Arrow-Hurwicz method. We mainly give results found in Rockafellar's paper [16].

2.1 Definitions and first properties of maximal monotone operators

Definition 2.1. *Let X be an Hilbert space. An operator is a multivaluated mapping A from X into $\mathcal{P}(X)$. We call $D(A) = \{x \in X \,/\, A(x) \neq \varnothing\}$ the domain of A and $R(A) = \bigcup_{x \in X} A(x)$ its range. We identify A and its graph in $X \times X$.*

Definition 2.2. *An operator A is monotone if :*

$$\forall x_1, x_2 \in D(A), \qquad \langle A(x_1) - A(x_2), x_1 - x_2 \rangle \geq 0$$

or more precisely if for all $x_1^* \in A(x_1)$ et $x_2^* \in A(x_2)$,

$$\langle x_1^* - x_2^*, x_1 - x_2 \rangle \geq 0$$

It is maximal monotone if it is maximal in the set of monotone operators. The maximality is to be understood in the sense of graph inclusion.

One of the essential results for us is the maximal monotonicity of the subgradient for convex functions.

Proposition 2.3. [6] *Let φ be a proper lower-semi-continuous convex function on X then $\partial\varphi$ is a maximal monotone operator.*

Before stating the main theorem of this theory, namely the existence of solutions of the Cauchy problem $-u' \in A(u(t))$ we need one last definition.

Definition 2.4. *Let A be maximal monotone. For $x \in D(A)$ we call $A^\circ(x)$ the projection of 0 on $A(x)$ (it exists since $A(x)$ is closed and convex, see Brézis [6] p. 20).*

We now turn to the theorem.

Theorem 2.5. [6] *Let A be maximal monotone then for all $u_0 \in D(A)$, there exists a unique function $u(t)$ from $[0, +\infty[$ into X such that*

- $u(t) \in D(A)$ *for all* $t > 0$
- $u(t)$ *is Lipschitz continous on* $[0, +\infty[$*, i.e* $u' \in L^\infty(0, +\infty; X)$ *(in the sense of distributions) and*

$$|u'|_{L^\infty(0,+\infty;X)} \leq |A^\circ(u_0)|$$

- $-u' \in A(u(t))$ *for almost every* t
- $u(0) = u_0$

Moreover u verifies,

- u *has a right derivative for every* $t \in [0, +\infty[$ *and* $-\dfrac{d^+u}{dt} \in A^\circ(u(t))$
- *the function* $t \to A^\circ(u(t))$ *is right continuous and* $t \to |A^\circ(u(t))|$ *is non increasing*
- *if u and $\hat{u}$ are two solutions then* $|u(t) - \hat{u}(t)| \leq |u(0) - \hat{u}(0)|$

2.2 Application to Arrow-Hurwicz methods

Let us now see how this theory can be applied for tracking saddle points. As mentioned before, we follow here [16]. We start with some definitions.

Definition 2.6. *Let $X = Y \oplus Z$ where Y and Z are two Hilbert spaces. A proper saddle function on X is a function K such that :*

- *for all* $z \in Z$*, the function* $K(\cdot, z)$ *is convex*
- *for all* $y \in Y$*, the function* $K(y, \cdot)$ *is concave*
- *there exists* $x = (y, z)$ *such that* $K(y, z') < +\infty$ *for all* $z' \in Z$ *and* $K(y', z) > -\infty$ *for all* $y' \in Y$*. The set of x for which it holds, is called the effective domain of K and is noted* $\operatorname{dom} K$*.*

Definition 2.7. *A point $(y, z) \in X$ is called a saddle point of K if*

$$K(y, z') \leq K(y, z) \leq K(y', z) \qquad \forall y' \in Y, \forall z' \in Z$$

We then have,

Proposition 2.8. *A point (y,z) is a saddle point of a saddle function K, if and only if*

$$K(y,z) = \sup_{z' \in Z} \inf_{y' \in Y} K(y', z') = \inf_{y' \in Y} \sup_{z' \in Z} K(y', z')$$

The proof of this proposition is easy and can be found in Rockafellar's book [15] p. 380.

The next theorem shows that the Arrow-Hurwicz method always provides a monotone operator.

Theorem 2.9. [16] *Let K be a proper saddle function. For $x = (y,z)$ let*

$$T(x) = \left\{ (y^*, z^*) \in Y^* \oplus Z^* / \begin{array}{c} y^* \textit{ is a subgradient of } K(\cdot, z) \textit{ in } y \\ z^* \textit{ is a subgradient of } -K(y, \cdot) \textit{ in } z \end{array} \right\}$$

Then T is a monotone operator with $D(T) \subset \operatorname{dom} K$.

We can now characterize the saddle points of K using the operator T.

Proposition 2.10. [16] *Let K be a proper saddle function then a point x is a saddle point of K if and only if $0 \in T(x)$.*

Remark . *This property is to be compared with the minimality condition $0 \in \partial f(x)$ for convex functions f.*

The next theorem shows that for regular enough saddle functions, the corresponding operator T is maximal.

Theorem 2.11. [16] *Let K be a proper saddle function on X. Suppose that K is lsc in y and upper-semi-continuous in z then T is maximal monotone.*

Proof. We just sketch the proof because it will inspire us in the following. The idea is to use the equivalent theorem for convex functions. For this we "invert" the operator T in the second variable. Let

$$H(y, z^*) = \sup_{z \in X} \langle z^*, z \rangle + K(y,z)$$

The proof is then based on the following lemma :

Lemma 2.12. *H is a convex lsc function on X and*

$$(y^*, z^*) \in T(y,z) \Leftrightarrow (y^*, z) \in \partial H(y, z^*)$$

It is then not too hard to prove that T is maximal.

□

3 Study of the Primal-Dual Method

Before starting the study of the Primal-Dual method, let us remind some facts about functions with bounded variation and pairings between measures and bounded functions.

Definition 3.1. *Let $BV(\Omega)$ be the space of functions u in L^1 for which*

$$\int_\Omega |Du| := \sup_{\substack{\xi \in \mathcal{C}^1_c(\Omega) \\ |\xi|_\infty \le 1}} - \int_\Omega u \operatorname{div} \xi < +\infty$$

With the norm $|u|_{BV} = \int_\Omega |Du| + |u|_{L^1}$ it is a Banach space. We note the functional space $BV^2 = BV(\Omega) \cap L^2$.

More informations about functions with bounded variation, can be found in the books [1] or [14].
Following Anzellotti [3], we now define $\int_\Omega [\xi, Du]$ which has to be understood as $\int_\Omega \xi \cdot Du$, for functions u with bounded variation and bounded functions ξ with divergence in L^2.

Definition 3.2.
- *Let $X^2 = \{\xi \in (L^\infty(\Omega))^n \,/\, \operatorname{div} \xi \in L^2(\Omega)\}$.*
- *For $(u, \xi) \in BV^2 \times X^2$ we define the distribution $[\xi, Du]$ by*

$$\langle [\xi, Du], \varphi \rangle = - \int_\Omega u \varphi \operatorname{div}(\xi) - \int_\Omega u \, \xi \cdot \nabla \varphi \qquad \forall \varphi \in \mathcal{C}^\infty_c(\Omega)$$

Theorem 3.3. [3] *The distribution $[\xi, Du]$ is a bounded Radon measure on Ω and if ν is the outward unit normal to Ω, we have Green's formula,*

$$\int_\Omega [\xi, Du] = - \int_\Omega u \operatorname{div}(\xi) + \int_{\partial\Omega} (\xi \cdot \nu) u$$

Using the ideas of Andreu and al. [2] it can be shown that

Proposition 3.4. *Let $J(u) = \displaystyle\int_\Omega |Du| + G(u) + \int_{\partial\Omega_D} |u - \varphi|$ then u is a minimizer of J in BV^2 if and only if there exists $\xi \in X^2$ such that*

$$\begin{cases} \operatorname{div}(\xi) \in \partial G(u) \\ \displaystyle\int_\Omega |Du| = \int_\Omega [\xi, Du] \\ \xi \cdot \nu = 0 \ \textit{in } \partial\Omega_N \ \textit{ and } \ (\xi \cdot \nu) \in \textit{sign}(\varphi - u) \ \textit{in } \partial\Omega_D \end{cases}$$

Applying the approximations lemmas in [3], it is not hard to prove the following

Proposition 3.5. *Let* $u \in BV(\Omega)$ *then*

$$\int_\Omega |Du| = \sup_{\substack{\xi \in X^2 \\ |\xi|_\infty \leq 1}} \int_\Omega [Du, \xi]$$

We thus want to find a saddle point of

$$K(u, \xi) = \int_\Omega [Du, \xi] + G(u) + \int_{\partial\Omega_D} |u - \varphi|$$

The saddle function K does not fulfill the assumptions of Rockafellar's result but if we remember Lemma 2.12 and set

$$\begin{aligned} H(u, \xi^*) &= \sup_{\substack{\xi \in X^2 \\ |\xi|_\infty \leq 1}} \langle \xi, \xi^* \rangle + K(u, \xi) \\ &= \sup_{\substack{\xi \in X^2 \\ |\xi|_\infty \leq 1}} \langle \xi, \xi^* \rangle + \int_\Omega [Du, \xi] + G(u) + \int_{\partial\Omega_D} |u - \varphi| \\ &= \int_\Omega |Du + \xi^*| + G(u) + \int_{\partial\Omega_D} |u - \varphi| \end{aligned}$$

Then H is a convex lsc function on $L^2 \times (L^2)^n$ hence ∂H is maximal monotone. We are now able to define a maximal monotone operator T by

$$T(u, \xi) = \{(u^*, \xi^*) \,/\, (u^*, \xi) \in \partial H(u, \xi^*)\}$$

In order to compute ∂H which gives the expression of T, we use the characterization

$$(u^*, \xi) \in \partial H(u, \xi^*) \Leftrightarrow \langle u^*, u \rangle + \langle \xi^*, \xi \rangle = H(u, \xi^*) + H^*(u^*, \xi)$$

Hence we have to determine what H^* is.

Proposition 3.6. *We have*

$$D(H^*) = \big\{(u^*, \xi) \,/\, u^* \in L^2(\Omega) \text{ and } \xi \in X^2 \,,\, \xi \cdot \nu = 0 \text{ in } \partial\Omega_N \,,\ |\xi|_\infty \leq 1\big\}$$

and

$$H^*(u^*,\xi) = G^*(u^* + \operatorname{div}(\xi)) - \int_{\partial\Omega_D} (\xi \cdot \nu)\varphi.$$

Proof. We start by computing the domain of H^*.
If $(u^*,\xi) \in D(H^*)$ then there exists a constant C such that for every $(u,\xi^*) \in BV^2 \times (L^2)^n$,

$$\langle u^*, u\rangle + \langle \xi^*, \xi\rangle - H(u,\xi^*) \leq C$$

Restraining to $u \in H^1(\Omega)$ with $u_{|\partial\Omega_D} = 0$ and $\xi^* \in (L^2)^n$, we find that

$$\langle u^*, u) + \langle \xi^*, \xi\rangle - \int_\Omega |\nabla u + \xi^*| - G(u) \leq C$$

from which

$$\langle \nabla u + \xi^*, \xi\rangle - \langle \nabla u, \xi\rangle + \langle u^*, u\rangle - \int_\Omega |\nabla u + \xi^*| - G(u) \leq C$$

Setting $\xi' = \nabla u + \xi^*$ and taking the supremum over all $\xi' \in (L^2)^n$ we have that $|\xi|_\infty \leq 1$ and for all $u \in H^1(\Omega)$ with $u_{|\partial\Omega_D} = 0$,

$$-\langle \nabla u, \xi\rangle + \langle u^*, u\rangle \leq C + G(u)$$

Taking now $u = \lambda v$ with λ positive and reminding the form of G, it can be shown that

$$-\langle \nabla u, \xi\rangle + \langle u^*, u\rangle \leq C|u|_2$$

This implies that $u^* + \operatorname{div}\xi \in L^2$ hence $\operatorname{div}\xi \in L^2$. Then by Green's formula in $H^1(\operatorname{div})$ (see Dautray-Lions [11] p. 205) we have $\xi \cdot \nu = 0$ in $\partial\Omega_N$.

Let us now compute H^*.
Let $(u^*,\xi) \in D(H^*)$,

$$H^*(u^*,\xi) = \sup_{\xi^* \in L^2} \sup_{u \in BV^2} \left\{ \langle u^*, u\rangle + \langle \xi^*, \xi\rangle - \int_\Omega |Du + \xi^*| - G(u) - \int_{\partial\Omega_D} |u - \varphi| \right\}$$

Let $\xi^* \in L^2$ be fixed. Then by Lemma 5.2 p. 316 of Anzellotti's paper [3], for every $u \in BV^2$ there exists $u_n \in \mathcal{C}^\infty \cap BV^2$ such that

$$u_n \xrightarrow{L^2} u\,, \quad (u_n)_{|\partial\Omega_D} = u_{|\partial\Omega_D} \qquad \text{and}$$
$$\int_\Omega |Du_n + \xi^*| \to \int_\Omega |Du + \xi^*|$$

We can thus restrict the supremum to functions u of class $\mathcal{C}^\infty(\Omega)$. We then have

$$\begin{aligned}
H^*(u^*, \xi) &= \sup_{u \in BV^2 \cap \mathcal{C}^\infty} \sup_{\xi \in L^2} \left\{ \langle u^*, u \rangle + \langle \xi^*, \xi \rangle - \int_\Omega |Du + \xi^*| - G(u) - \int_{\partial\Omega_D} |u - \varphi| \right\} \\
&= \sup_{u \in BV^2 \cap \mathcal{C}^\infty} \left\{ \langle u^*, u \rangle - \langle \nabla u, \xi \rangle - G(u) - \int_{\partial\Omega_D} |u - \varphi| \right\} \\
&= \sup_{u \in BV^2} \left\{ \langle u^*, u \rangle - \int_\Omega [Du, \xi] - G(u) - \int_{\partial\Omega_D} |u - \varphi| \right\} \\
&= \sup_{u \in BV^2} \left\{ \langle u, u^* + \operatorname{div} \xi \rangle - G(u) - \int_{\partial\Omega_D} \{ |u - \varphi| + (\xi \cdot \nu) u \} \right\}
\end{aligned}$$

Beware that $u \in BV^2 \cap \mathcal{C}^\infty$ implies that $\nabla u \in L^1$ and not $\nabla u \in L^2$ but the density of L^2 in L^1 allows us to pass from the first equality to the second. The third equality follows from Lemma 1.8 of [3]. We now have to show that we can take separately the supremum in the interior of Ω and on the boundary $\partial\Omega_D$.

Let f be in $L^1(\partial\Omega)$ and v be in $L^2(\Omega)$. We want to find $u_\varepsilon \in BV^2$ converging to v in L^2 and such that $(u_\varepsilon)_{|\partial\Omega_D} = f$.

By Lemma 5.5 of [3] there is a $w_\varepsilon \in W^{1,1}$ with $(w_\varepsilon)_{|\partial\Omega_D} = f$ and $|w_\varepsilon|_2 \leq \varepsilon$. By density of $\mathcal{C}_c^\infty(\Omega)$ in L^2 we can find $v_\varepsilon \in \mathcal{C}_c^\infty(\Omega)$ with $|v_\varepsilon - v|_2 \leq \varepsilon$ We can then take $u_\varepsilon = v_\varepsilon + w_\varepsilon$.

This shows that

$$\begin{aligned}
H^*(u^*, \xi) &= \sup_{u \in L^2(\Omega)} \{ \langle u, u^* + \operatorname{div} \xi \rangle - G(u) \} - \inf_{u \in L^1} \int_{\partial\Omega_D} \{ |u - \varphi| + (\xi \cdot \nu) u \} \\
&= G^*(u^* + \operatorname{div}(\xi)) - \int_{\partial\Omega_D} (\xi \cdot \nu) \varphi
\end{aligned}$$

$\square$

We can now compute T

Proposition 3.7. *Let* $(u, \xi) \in BV^2 \times X^2$ *then,* $(u^*, \xi^*) \in T(u, \xi)$ *if and only if*

$$\begin{cases} u^* + \mathrm{div}(\xi) \in \partial G(u) \\ \displaystyle\int_\Omega |\xi^* + Du| = \langle \xi^*, \xi \rangle + \int_\Omega [\xi, Du] \\ \xi \cdot \nu = 0 \ \textit{in} \ \partial\Omega_N \ \ \textit{and} \ \ (\xi \cdot \nu) \in \textit{sign}(\varphi - u) \ \textit{in} \ \partial\Omega_D \end{cases}$$

Proof. Let us first note that,

$$G(u) + G^*(u^* + \mathrm{div}(\xi)) \geq \langle u, u^* + \mathrm{div}(\xi) \rangle \tag{2}$$

$$\int_\Omega |Du + \xi^*| \geq \int_\Omega [\xi, Du] + \int_\Omega \xi^* \xi \tag{3}$$

$$|u - \varphi| \geq (\xi \cdot \nu)(\varphi - u) \tag{4}$$

where the second inequality is obtained arguing as in Proposition 3.5.
By definition, $(u^*, \xi^*) \in T(u, \xi)$ if and only if

$$\begin{aligned} \langle u, u^* \rangle + \langle \xi, \xi^* \rangle =& H(u, \xi^*) + H^*(u^*, \xi) \\ =& \int_\Omega |Du + \xi^*| + G(u) + \int_{\partial\Omega_D} |u - \varphi| \\ &+ G^*(u^* + \mathrm{div}(\xi)) - \int_{\partial\Omega_D} (\xi \cdot \nu)\varphi \end{aligned}$$

This shows that (2), (3) and (4) must be equalities which is exactly

$$\begin{cases} u^* + \mathrm{div}(\xi) \in \partial G(u) \\ \displaystyle\int_\Omega |\xi^* + Du| = \langle \xi^*, \xi \rangle + \int_\Omega [\xi, Du] \\ (\xi \cdot \nu) \in \textit{sign}(\varphi - u) \text{ in } \partial\Omega_D \end{cases}$$

$\square$

Remark .

- *Whenever it has a meaning, one can show that the condition*

$$\int_\Omega |\xi^* + Du| = \langle \xi^*, \xi \rangle + \int_\Omega [\xi, Du]$$

is equivalent to

$$\xi^* + Du \in I_{B(0,1)}(\xi)$$

so that we won't distinguish between these two.

- *This analysis shows why the constraint $u \in [0,1]$ is hard to deal with. In fact, it imposes that* $\operatorname{div}(\xi)$ *is a measure but not necessarily a* L^2 *function. It is not easy to give a meaning to* $\int_\Omega Du \cdot \xi$ *or to* $(\xi \cdot \nu)$ *on the boundary for such functions. However, when dealing with numerical implementations, it is better to keep the constraint on* u.

We can summarize those results in the following theorem which says that the Primal-Dual Method is well-posed.

Theorem 3.8. *For all* $(u_0, \xi_0) \in \operatorname{dom}(T)$*, there exists a unique* $(u(t), \xi(t))$ *such that*

$$\begin{cases} \partial_t u \in \operatorname{div}(\xi) - \partial G(u) \\ \partial_t \xi \in Du - \partial I_{B(0,1)}(\xi) \\ (\xi \cdot \nu) \in sign(\varphi - u) \ \text{in} \ \partial\Omega_D \qquad \xi \cdot \nu = 0 \ \text{in} \ \partial\Omega_N \\ (u(0), \xi(0)) = (u_0, \xi_0) \end{cases} \tag{5}$$

Moreover, the energy $|\frac{d^+u}{dt}|_2^2 + |\frac{d^+\xi}{dt}|_2^2$ *is non increasing and if* $(\bar{u}, \bar{\xi})$ *is a saddle point of* K, $|u - \bar{u}|_2^2 + |\xi - \bar{\xi}|_2^2$ *is also non increasing.*

Proof. Apply theorem 2.5. □

Remark . *This theorem also shows that whenever* J *has a minimizer,* K *has saddle points. This is because stationnary points of the system (5) are minimizers of* J *(verifying the Euler-Lagrange equation for* J*).*

For the Rudin-Osher-Fatemi model, one can show that there is convergence of u to the minimizer of the functional J and obtain *a posteriori* estimates.

Proposition 3.9. *Let* $G = \dfrac{\lambda}{2}\displaystyle\int_\Omega (u-f)^2$ *and* $\partial\Omega_D = \emptyset$. *Then if* $\bar{u}$ *is the minimizer of* J*, every solution of (5) converges in* L^2 *to* $\bar{u}$*. Furthermore one has*

$$|u-\bar{u}|_2 \leq \frac{1}{2}\left(\frac{1}{\lambda}|\partial_t u|_2 + \sqrt{\frac{|\partial_t u|_2^2}{\lambda^2} + \frac{8|\Omega|^{\frac{1}{2}}}{\lambda}|\partial_t \xi|_2}\right)$$

Proof. Let $(\bar{u},\bar{\xi})$ be such that $0 \in T(\bar{u},\bar{\xi})$. Let $e(t) = |u(t)-\bar{u}|_2^2$ and $g(t) = |\xi(t)-\bar{\xi}|_2^2$. We can show that

$$\frac{1}{2}(e+g)' \leq -\lambda e$$

Indeed, by definition of the flow,

$$\begin{aligned} &\int_\Omega [\xi, Du] - \langle \xi, \partial_t \xi\rangle \geq \int_\Omega [\bar{\xi}, Du] - \langle \bar{\xi}, \partial_t \xi\rangle \qquad \text{and} \\ &\int_\Omega [\bar{\xi}, D\bar{u}] - \langle \bar{\xi}, \partial_t \bar{\xi}\rangle \geq \int_\Omega [\xi, D\bar{u}] - \langle \xi, \partial_t \bar{\xi}\rangle \end{aligned}$$

Summing these two we find,

$$\int_\Omega [\xi-\bar{\xi}, D(u-\bar{u})] \geq \langle \xi-\bar{\xi}, \partial_t\xi - \partial_t\bar{\xi}\rangle$$

We thus have

$$\begin{aligned} \frac{1}{2}(e+g)' &= \langle u-\bar{u}, \partial_t u - \partial_t \bar{u}\rangle + \langle \xi-\bar{\xi}, \partial_t\xi - \partial_t\bar{\xi}\rangle \\ &\leq \langle u-\bar{u}, \operatorname{div}(\xi-\bar{\xi}) - \lambda(u-\bar{u})\rangle + \int_\Omega [\xi-\bar{\xi}, D(u-\bar{u})] \\ &= -\lambda e \end{aligned}$$

The functions e and g are Lipschitz continuous. Let L be the Lipschitz constant of e and let $h = e+g$.

Suppose that there exists $\alpha > 0$ and $T > 0$ such that $e \geq \alpha$ for all $t > T$, then we would have $h' \leq -\lambda\alpha$ and h would tend to $-\infty$ which is impossible by positivity of h. Hence

$$\forall \alpha > 0 \; \forall T > 0 \; \exists t \geq T \qquad e(t) \leq \alpha$$

Suppose now the existence of $\varepsilon > 0$ such that for all $T \geq 0$ there exists $t \geq T$ with $e(t) \geq \varepsilon$.
By continuity of e, there exists a sequence (t_n) with $\lim_{n\to+\infty} t_n = +\infty$ such that

$$e(t_{2n}) = \frac{\varepsilon}{2} \qquad e(t_{2n+1}) = \varepsilon$$

Moreover, on $[t_{2n-1}, t_{2n}]$, we have $e(t) \geq \frac{\varepsilon}{2}$. We then find that

$$|e(t_{2n}) - e(t_{2n-1})| \leq L(t_{2n} - t_{2n-1}) \quad \text{so}$$
$$\frac{\varepsilon}{2L} \leq t_{2n} - t_{2n-1}$$

From which we see that,

$$\begin{aligned} h(t_{2n+2}) &= h(t_{2n+1}) + \int_{t_{2n+1}}^{t_{2n+2}} h'(t)\,dt \\ &\leq h(t_{2n+1}) - \varepsilon\lambda(t_{2n+2} - t_{2n+1}) \\ &\leq h(t_{2n}) - \frac{\lambda\varepsilon^2}{2L} \end{aligned}$$

This shows that $\lim_{t\to+\infty} e(t) = 0$.

We now prove the *a posteriori* error estimate.

We have that

$$\begin{aligned} u &= f + \frac{1}{\lambda}(\operatorname{div}\xi - \partial_t u) \\ \bar{u} &= f + \frac{1}{\lambda}\operatorname{div}\bar{\xi} \end{aligned}$$

Which leads to

$$\begin{aligned} |u - \bar{u}|_2^2 &= \frac{1}{\lambda}\langle \operatorname{div}(\xi - \bar{\xi}) - \partial_t u, u - \bar{u}\rangle \\ &= \frac{1}{\lambda}\left[\langle \operatorname{div}(\xi - \bar{\xi}), u - \bar{u}\rangle - \langle \partial_t u, u - \bar{u}\rangle\right] \\ &= \frac{1}{\lambda}\left[-\langle \xi - \bar{\xi}, Du - D\bar{u}\rangle - \langle \partial_t u, u - \bar{u}\rangle\right] \\ &\leq \frac{1}{\lambda}\left[\int_\Omega |Du| - \int_\Omega [\xi, Du] + |\partial_t u|_2 |u - \bar{u}|_2\right] \end{aligned}$$

From which we deduce that

$$|u-\bar{u}|_2 \leq \frac{1}{2}\left(\frac{1}{\lambda}|\partial_t u|_2 + \sqrt{\frac{|\partial_t u|_2^2}{\lambda^2} + \frac{4}{\lambda}(\int_\Omega |Du| - \int_\Omega [\xi, Du])}\right)$$

The estimate follows from the fact that

$$\begin{aligned}
\int_\Omega |-\partial_t\xi + Du| &= \int_\Omega [\xi, Du] - \int_\Omega \partial_t\xi\cdot\xi \quad \text{thus}\\
\int_\Omega |Du| - \int_\Omega |\partial_t\xi| &\leq \int_\Omega [\xi, Du] - \int_\Omega \partial_t\xi\cdot\xi \quad \text{hence}\\
\int_\Omega |Du| - \int_\Omega [\xi, Du] &\leq 2\int_\Omega |\partial_t\xi| \leq 2|\Omega|^{\frac{1}{2}}|\partial_t\xi|_2
\end{aligned}$$

□

Following the same lines, one can show *a posteriori* error estimates for general finite difference scheme. Indeed if ∇^h is any discretization of the gradient and if div^h is defined as $-(\nabla^h)^*$, the associated algorithm is

$$\begin{cases} \xi^n = P_{B(0,1)}(\xi^{n-1} + \delta\tau^n \nabla^h u^{n-1}) \\ u^n = u^{n-1} + \delta t^n(\mathrm{div}^h\, \xi^n - \lambda(u^{n-1} - f)) \end{cases}$$

Where $P_{B(0,1)}(\xi)_{i,j} = \dfrac{\xi_{i,j}}{\max(|\xi_{i,j}|, 1)}$ is the componentwise projection of ξ on the unit ball. We can associate to this system a discrete energy,

$$J_h(u) = \sum_{i,j} |\nabla^h u|_{i,j} + \frac{\lambda}{2}\sum_{i,j} |u_{i,j} - f_{i,j}|^2$$

The algorithm presented above could have been directly derived from the discrete energy using the method of Chan and Zhu [17] (which is just the discrete counterpart of our continuous method). Hence, the next proposition gives a stopping criterion for their algorithm.

Proposition 3.10. *If $\bar{u}$ is the minimizer of J_h then*

$$|u^n - \bar{u}|_2 \leq \frac{1}{2}\left(\frac{1}{\lambda}|\partial_t u^n|_2 + \sqrt{\frac{|\partial_t u^n|_2^2}{\lambda^2} + \frac{8}{\lambda}|\xi_t^n|_2}\right)$$

Where $\partial_t u^n = \dfrac{u^{n+1} - u^n}{\delta t^{n+1}}$ *and* $\partial_t \xi^n = \dfrac{\xi^{n+1} - \xi^n}{\delta\tau^{n+1}}$.

The proof of this dicrete estimate is almost the same as for the continuous one. We give it in the appendix.

For the general problem, there is no uniqueness for the minimizer (for example in the segmentation problem) and hence convergence may not occur or be hard to prove. Indeed, even when uniqueness holds, we can have non vanishing oscillations. For example in the simpler one dimensional problem

$$\min_{u \in BV([0,1])} \int_0^1 |u'|$$

the unique minimizer is $u = 0$ but $u(t,x) = \frac{1}{2}\cos(\pi x)\sin(\pi t)$ and $\xi(t,x) = \frac{1}{2}\sin(\pi x)\cos(\pi t)$ gives a solution to the associated system. In this example, the energy is constant hence not converging to 0. We can however show general *a posteriori* estimates for the energy.

Proposition 3.11. *For every saddle point $(\bar{u}, \bar{\xi})$ and every (u_0, ξ_0), the solution $(u(t), \xi(t))$ of (5) satisfies*

$$|J(u) - J(\bar{u})| \leq \left(\sqrt{|u_0 - \bar{u}|_2^2 + |\xi_0 - \bar{\xi}|_2^2}\right) |\partial_t u|_2 + 2|\Omega|^{\frac{1}{2}} |\partial_t \xi|_2$$

Proof. Let $(\bar{u}, \bar{\xi})$ be a saddle point and $(u(t), \xi(t))$ be a solution of (5).

$$J(u) - J(\bar{u}) = \int_\Omega |Du| + \int_{\partial\Omega_D} |u - \varphi| - \int_\Omega |D\bar{u}| - \int_{\partial\Omega_D} |\bar{u} - \varphi| + G(u) - G(\bar{u})$$

By definition of the operator T we have

$$\begin{aligned} \int_\Omega [\xi, Du] - \int_\Omega \partial_t \xi \cdot \xi &= \int_\Omega |Du - \partial_t \xi| \\ &\geq \int_\Omega |Du| - \int_\Omega |\partial_t \xi| \end{aligned}$$

This shows that

$$\int_\Omega |Du| \leq \int_\Omega [\xi, Du] + 2\int_\Omega |\partial_t \xi|$$

On the other hand,

$$\int_\Omega [\xi, Du] + \int_{\partial\Omega_D} |u - \varphi| = -\int_\Omega u \operatorname{div} \xi + \int_{\partial\Omega_D} \{(\xi \cdot \nu)u + |u - \varphi|\}$$

Applying $\int_{\partial\Omega_D} \{(\xi\cdot\nu)u + |u-\varphi|\} = \inf_v \{(\xi\cdot\nu)v + |v-\varphi|\}$ to $v = \bar{u}$ we have

$$\begin{aligned}
\int_\Omega [\xi, Du] + \int_{\partial\Omega_D} |u-\varphi| - \int_{\partial\Omega_D} |\bar{u}-\varphi| &\leq -\int_\Omega u \operatorname{div}\xi + \int_{\partial\Omega_D} (\xi\cdot\nu)\bar{u} \\
&= -\int_\Omega u \operatorname{div}\xi + \int_\Omega \bar{u} \operatorname{div}\xi + \int_\Omega [\xi, D\bar{u}] \\
&= \int_\Omega (\bar{u}-u) \operatorname{div}\xi + \int_\Omega [\xi, D\bar{u}]
\end{aligned}$$

If we now use

$$\langle \operatorname{div}(\xi) - \partial_t u, u - \bar{u}\rangle \geq G(u) - G(\bar{u})$$

and combine it with all these inequalities, we find

$$\begin{aligned}
J(u) - J(\bar{u}) &\leq \int_\Omega (\bar{u}-u)\partial_t u + 2|\partial_t \xi|_2 + \int_\Omega [\xi, D\bar{u}] - \int_\Omega |D\bar{u}| \\
&\leq |\bar{u}-u|_2 |\partial_t u|_2 + 2|\partial_t \xi|_2
\end{aligned}$$

Which gives the estimate reminding that $\sqrt{|u-\bar{u}|_2^2 + |\xi-\bar{\xi}|_2^2}$ is non increasing. □

Remark .
Supported by numerical evidence, we can conjecture that whenever the constraint on ξ is saturated somewhere, convergence of u occurs. It might however be also necessary to add the constraint $u \in [0,1]$ in order to have this convergence.

Considering a finite difference scheme, just as for the Rudin-Osher-Fatemi model, we can define a discrete energy J_h and show the corresponding *a posteriori* estimate.

Proposition 3.12. *If $\bar{u}$ is a minimizer of J_h and (u^n, ξ^n) is defined by*

$$\begin{cases}
\xi^n = P_{B(0,1)}(\xi^{n-1} + \delta\tau^n \nabla^h u^{n-1}) \\
u^n = u^{n-1} + \delta t^n (\operatorname{div}^h \xi^n - p^n)
\end{cases}$$

with $p^n \in \partial G^h(u^{n-1})$ then

$$|J_h(u^n) - J_h(\bar{u})| \leq 2|\partial_t \xi^n| + |\partial_t u^n||u^{n-1} - \bar{u}|$$

We omit the proof because it is exactly the same as for Proposition 3.11.

Remark .

- *The boundary conditions are hidden here in the operator* ∇^h.
- *In the discrete framework, the estimate involves* $|u^n - \bar{u}|$ *which can not be easily bounded by the initial error.*

4 Numerical Experiments

To illustrate the relevance of our *a posteriori* estimates, we first consider the simple example of denoising a rectangle (see figure 3). We then compare the *a posteriori* error bound with the "true" error. We use the relative L^2 error defined as $\dfrac{|u^n - \bar{u}|}{|\bar{u}|}$ and ran the algorithm of Chan and Zhu with fixed time steps $\delta t = 0.1$ and $\delta\tau = 4$. The minimizer $\bar{u}$ is computed by the algorithm after 15000 iterations. Figure 4 shows that the *a posteriori* bound is quite sharp even if there are some unexpected pikes on the curve.

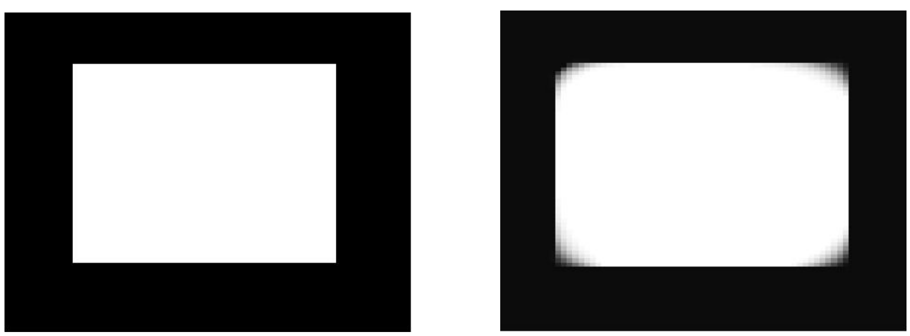

Figure 3: Denoising of a rectangle using the ROF model

The second experiment is performed on the yeast segmentation of figure 2. The solution was computed with the algorithm of Chan and Zhu using as weight function g the one proposed by Appleton and Talbot [4]. We used this time the error $|J_h(u^n) - J_h(\bar{u})|$ and ran the algorithm with $\delta t = 0.2$ and $\delta\tau = 0.2$. The minimizer $\bar{u}$ is computed by the algorithm after 15000 iterations. We can see on figure 5 that again the *a posteriori* estimate is sharp. We must notice that in general we do not know $\bar{u}$. However, this

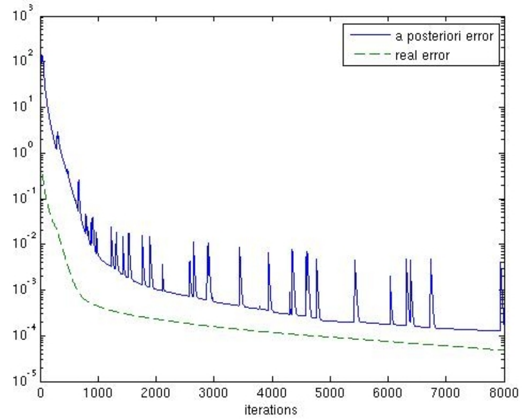

Figure 4: Comparison of the relative L^2 error with the predicted *a posteriori* bound.

experiment shows that using $|\partial_t \xi^n|$ and $|\partial_t u^n|$ as a stopping criterion makes sense.

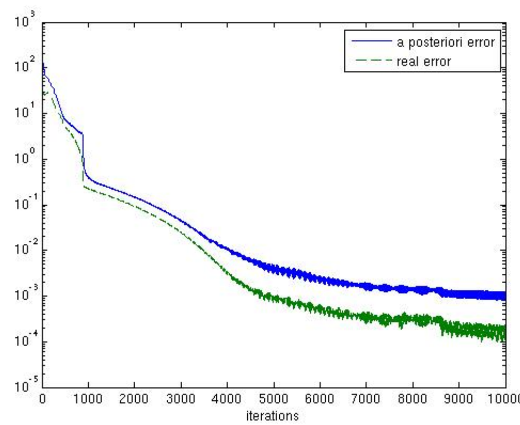

Figure 5: Comparison for the segmentation problem.

A Proof of Proposition 3.10

For notational convenience, we present the proof for $\lambda = 1$. Let $\bar{u}$ be the minimizer of J_h and $\bar{\xi}$ be such that $|\bar{\xi}|_\infty \leq 1$ then

$$\begin{cases} \sum_{i,j} |\nabla^h \bar{u}|_{i,j} = \langle \nabla^h \bar{u}, \bar{\xi} \rangle \\ \bar{u} = \operatorname{div}^h \bar{\xi} + f \end{cases}$$

Reminding that $u^n = f + \operatorname{div}^h \xi^n - \partial_t u^n$ we get

$$\begin{aligned} |u^n - \bar{u}|^2 &= \langle \operatorname{div}^h(\xi^{n+1} - \bar{\xi}) - \partial_t u^n, u^n - \bar{u} \rangle \\ &= -\langle \xi^{n+1} - \bar{\xi}, \nabla^h u^n - \nabla^h \bar{u} \rangle - \langle \partial_t u^n, u^n - \bar{u} \rangle \\ &\leq \langle \bar{\xi} - \xi^{n+1}, \nabla^h u^n \rangle + |\partial_t u^n||u^n - \bar{u}| \end{aligned}$$

We have that $\xi^{n+1} = P_{B(0,1)}(\xi^n + \delta\tau^{n+1}\nabla^h u^n)$ hence by definition of the projection,

$$\forall \bar{\xi} \in B(0,1) \qquad \langle \xi^{n+1} - (\xi^n + \delta\tau^{n+1}\nabla^h u^n), \bar{\xi} - \xi^{n+1} \rangle \geq 0$$

This gives us

$$\langle \nabla^h u^n, \bar{\xi} - \xi^{n+1} \rangle \leq \langle \partial_t \xi^n, \bar{\xi} - \xi^n \rangle$$

Combining this with $\langle \partial_t \xi^n, \bar{\xi} \rangle - \langle \partial_t \xi^n, \bar{\xi} \rangle \leq 2|\partial_t \xi^n|$ (which holds because $|\bar{\xi}|_\infty \leq 1$ and $|\xi^n|_\infty \leq 1$) we find that

$$|u^n - \bar{u}|^2 \leq 2|\partial_t \xi^n| + |\partial_t u^n||u^n - \bar{u}|$$

The announced inequality easily follows.

References

[1] L. Ambrosio, N. Fusco and D. Pallara, *Functions of Bounded Variation and Free Discontinuity Problems*, Oxford Science Publications, 2000.

[2] F. Andreu-Vaillo, V. Caselles and J.M. Mazòn, *Parabolic Quasilinear Equations Minimizing Linear Growth Functionals*, Birkhäuser, collection "Progress in Mathematics", no. 223, 2004.

[3] G. Anzellotti, *Pairings between measures and bounded functions and compensated compactness*, Annali di Matematica Pura ed Applicata, Vol. 135, No. 1, p. 293-318, 1983.

[4] B. Appleton and H. Talbot, *Globally Minimal Surfaces by Continuous Maximal Flows*, IEEE Trans. on Pattern Anal. and Mach. Intelligence, vol. 28, no. 1, p. 106-118, 2006.

[5] K. J. Arrow, L. Hurwicz and H. Uzawa (editors), *Studies in Linear and Nonlinear Programming*, Stanford Univ. Press, Stanford, 1958.

[6] H. Brézis, *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*, Amsterdam-London : North Holland. 1973.

[7] V. Caselles, R. Kimmel and G. Shapiro, *Geodesic Active Contours*, Int. J. Computer Vision, vol. 22, no. 1, p. 61-79, 1997.

[8] A. Chambolle, D. Cremers and T. Pock, *A convex approach for computing minimal partitions*, preprint CMAP, November 2008.

[9] A. Chambolle, V. Caselles,M. Novaga, D. Cremers and T. Pock, *An introduction to Total Variation for Image Analysis*, hal-00437581.

[10] T.F. Chan, G.H. Golub and P. Mulet, *A nonlinear primal dual method for total variation based image restoration*, SIAM J. Sci. Comput., 20, p. 1964-1977, 1999.

[11] R. Dautray and J.L. Lions, *Mathematical Analysis and Numerical Methods for Science and Technology*, vol. 3, Spectral Theory and Applications, Springer, 1990.

[12] I. Ekeland and R. Temam, *Analyse convexe et problèmes variationnels*, Dunod Gauthiers-Villars, collection "Études mathématiques", 1974.

[13] M. Giaquinta, G. Modica and J. Souček, *Functionals with linear growth in the calculus of variations*, Com. Math. Uni. Carolinae, 20, p. 143-171, 1979.

[14] E. Giusti, *Minimal Surfaces and Functions of Bounded Variation*, Birkhäuser, collection "Monographs in Mathematics", vol. 80, 1984.

[15] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, 1970.

[16] R. T. Rockafellar, *Monotone Operators Associated with Saddle-Functions and Minimax Problems*, Nonlinear Functional Analysis, F.E. Browder ed., Proc. of Symp. in Pure Math. of the Am. Math. Society, vol. XVIII, part. 1, p.241-250, 1968.

[17] M. Zhu and T. Chan, *An Efficient Primal-Dual Hybrid Gradient Algorithm for Total Variation Image Restoration*, cam report (08-34), May 2008.