Maïtine Bergounioux

Basic variational tools for image processing

October 7, 2014

MAPMO - Université d'Orléans

Contents

1

Convex analysis

1.1 Functional analysis

Let V be a (real) Banach space and V' the dual space. We denote by $\| \ \|_V$ the V -norm and $\langle \cdot, \cdot \rangle$ the duality bracket between V and V':

$$\forall \varphi \in V', \forall x \in V \qquad \langle \varphi, x \rangle = \varphi(x) \ .$$

Any set which is closed for the weak (sequential) topology is closed for the strong one : indeed, if C is weakly closed, any strongly convergent sequence of C is also weakly convergent and its limit belongs to C. The converse is false. However it is true for the convex sets.

Theorem 1.1.1 *Let C be a convex subset of a Banach space V. Then C is (sequentially) weakly closed if and only if it is (sequentially) strongly closed.*

This is a consequence of the Hahn-Banach theorem 1.4.1.
Let us specify now the notion of continuity of a functional J from a Banach space V to $\mathbb{R} \cup \{+\infty\}$. We recall that a function is continuous at $x \in V$ for the sequential strong topology if

$$\forall x_n \to x \quad \text{(strongly)} \quad J(x_n) \to J(x) \ .$$

A functional J from V to $\mathbb{R} \cup \{+\infty\}$ will be continuous at $x \in V$ for the weak sequential topology if

$$\forall x_n \rightharpoonup x \quad \text{(weakly)} \quad J(x_n) \to J(x) \ .$$

A continuous function for the sequential weak convergence is also continuous for the strong sequential convergence because that strong convergence of a sequence implies its weak convergence. More exactly, if J is weakly continuous at x, for any sequence x_n which converges strongly towards x we have:

$$x_n \to x \Longrightarrow x_n \rightharpoonup x \Longrightarrow J(x_n) \to J(x) \ ,$$

and the function J is continuous for the strong sequential topology.

The converse is **false** in the general case. We will see that it is true, under certain conditions, in the case of linear operators. It is “ partly” true for convex functions. Specify what means to “partly”, defining the semi-continuity concept:

Definition 1.1.1 *A function J from V to $\mathbb{R} \cup \{+\infty\}$ is lower semi-continuous (lsc) on V if one of the two equivalent conditions is satisfied:*

- $\forall a \in \mathbb{R}, \qquad \{ \ u \in V \ \mid \ J(u) \leqslant a \ \}$ *is closed*
- $\forall \bar{u} \in V, \qquad \liminf_{u \to \bar{u}} J(u) \geqslant J(\bar{u}).$

Theorem 1.1.2 *Any convex function lsc for the strong topology of V is lsc for the weak topology as well.*

In practice this result is used in the form of the following corollary:

Corollary 1.1.1 *Let J be a convex functional from V to $\mathbb{R} \cup \{+\infty\}$ lsc (for example continuous) for the strong topology. If v_n is a weakly convergent sequence to $v \in V$ then*

$$J(v) \leqslant \liminf_{n \to +\infty} J(v_n).$$

We find that if $x_n \rightharpoonup x$ weakly in V then

$$\|x\|_V \leqslant \liminf_{n \to +\infty} \|x_n\|_V \ ;$$

indeed $x \to \|x\|_V$ is a continuous convex application. So it is strongly lsc and thus weakly lsc.

We now present a topology on V' which is called the weak-star (sequential) topology.

Definition 1.1.2 *Let $(\varphi_n)_{n \in \mathbb{N}}$ be a sequence of V'. We say that (φ_n) converges towards φ for the weak-star topology and we note $\varphi_n \overset{*}{\rightharpoonup} \varphi$ if*

$$\forall x \in V \qquad \langle \varphi_n, x \rangle \to \langle \varphi, x \rangle \ .$$

Be careful that the weak-star topology **is not** the weak topology of the dual space V'. It is true only if $V = V$" (up to an isomorphism) that is if V is a *reflexive* Banach space.

Theorem 1.1.3 *Let (φ_n) be sequence of V'. Then*

1. *if $\varphi_n \overset{*}{\rightharpoonup} \varphi$ for the weak-star topology, then $\|\varphi_n\|_{V'}$ is bounded and* $\|\varphi\| \leqslant \liminf_{n \to +\infty} \|\varphi_n\|_{V'}$.
2. *if $\varphi_n \overset{*}{\rightharpoonup} \varphi$ for the weak-star topology and if $x_n \to x$ strongly in V, then $\langle \varphi_n, x_n \rangle \to \langle \varphi, x \rangle$.*

1.1.1 Compactness theorem

We give now one of the most important compactness result of functional analysis, which motivates the introduction of the weak-star topology.

Theorem 1.1.4 (Banach-Alaoglu-Bourbaki) *Let V be a real normed space. The close unit ball of V'*

$$B_{V'} = \{ \varphi \in V' \mid \|\varphi\|_{V'} \leqslant 1 \} ,$$

is weakly star compact.
*In other words, from any bounded sequence in V', one can extract a sub-sequence converging for the weak * topology.*

When V is a reflexive space, V is identified to its bidual V" and the weak and weak * topologies coincide. Theorem 1.1.4 holds with V instead of V'. In fact, we even have a stronger result since it is reflexive spaces characterization.

Theorem 1.1.5 (Kakutani) *Let V be a Banach space. Then V is reflexive if and only if the unit closed ball of V*

$$B_V = \{ x \in V \mid \|x\|_V \leqslant 1 \} ,$$

is weakly compact, that is from any bounded sequence in V, one can extract a sub-sequence converging for the weak topology.

An immediate corollary is:

Corollary 1.1.2 *Let V be a Banach space. then V is reflexive if and only if V' is reflexive.*

1.2 Gâteaux-differentiability of convex functionals

Now let differentiability properties useful in the context of optimization in a Banach space.

Definition 1.2.1 *Let J a functional from V to $\mathbb{R} \cup \{+\infty\}$. We say that J is Gâteaux-differentiable at $u \in$ dom (J) if the directional derivative*

$$J'(u; v) = \lim_{t \to 0^+} \frac{J(u + tv) - J(u)}{t},$$

exits for any direction v of V and if

$$v \mapsto J'(u; v)$$

is linear continuous. Here dom (J) is the set of $u \in \mathcal{X}$ such that $J(u)$ is finite.

We shall denote $\nabla J(u)$ the Gâteaux-derivative of J at u. It belongs to the dual V'. If V is an Hilbert space, then with Riesz theorem (see [3]) one may identify V and its dual space; we have then

$$J'(u; v) = (\nabla J(u), v) ,$$

where $(\cdot,\cdot)$ is the inner product of V. The element $\nabla J(u) \in V$ is the *gradient* of J at u.

It is clear that if J is differentiable in the classical sense (say Fréchet - differentiable) at u, then J is Gâteaux-differentiable at u and the classical and Gâteaux derivatives coincide.

Theorem 1.2.1 *Let be $J : \mathcal{C} \subset V \to \mathbb{R}$, Gâteaux differentiable on $\mathcal{C}$, with $\mathcal{C}$ convex. J is convex if and only if*

$$\forall (u,v) \in \mathcal{C} \times \mathcal{C} \qquad J(v) \geqslant J(u) + \langle \nabla J(u), v-u \rangle \tag{1.1}$$

Theorem 1.2.2 *Let be $J : \mathcal{C} \subset V \to \mathbb{R}$, Gâteaux differentiable on $\mathcal{C}$, with $\mathcal{C}$ convex. J is convex if and only if ∇J is a monotone operator, that is*

$$\forall (u,v) \in \mathcal{C} \times \mathcal{C} \qquad \langle \nabla J(u) - \nabla J(v), u-v \rangle \geqslant 0. \tag{1.2}$$

Remark 1.2.1 *Assume that ∇ is a strictly monotone operator*

$$\forall (u,v) \in \mathcal{C} \times \mathcal{C},\ u \neq v, \qquad \langle \nabla J(u) - \nabla J(v), u-v \rangle > 0. \tag{1.3}$$

then J is strictly convex.

Similarly, we define the (Gâteaux) second derivative of J at u, as the derivative of the (vectorial) function $u \mapsto \nabla J(u)$. We denote $D^2 J(u)$ and called it Hessian by analogy with the Hessian in the sense of Fréchet; this Hessian is identifiable to a $n \times n$ square matrix when $V = \mathbb{R}^n$.

1.3 Minimization in a reflexive Banach space

Unless otherwise stated, we now assume that V is a reflexive Banach space with (topological) dual V'.
Let us begin with a general minimization result of a semi-continuous functional on a closed set of V .

Definition 1.3.1 *We say that $J : V \to \mathbb{R}$ is coercive if*

$$\lim_{\|x\|_V \to +\infty} J(x) = +\infty.$$

Theorem 1.3.1 *Assume that V is a reflexive Banach space. Let J be a functional from V to $\mathbb{R} \cup \{+\infty\}$, lower semi-continuous for the weak topology of V. Let K be a non empty weakly closed subset of V. Assume that J is proper (there exists $v_o \in K$ such that $J(v_o) < +\infty$). Then the following minimization problem*

$$(\mathcal{P}) \qquad \begin{cases} \text{Find } u \text{ such that} \\ J(u) = \inf \ \{ \ J(v) \mid v \in K \ \}, \end{cases} \tag{1.4}$$

has at least a solution if one of the following conditions is verified :

- *either J is coercive i.e.* $\lim_{\|v\|_V \to +\infty} J(v) = +\infty$,
- *or K is bounded.*

An important corollary concerns the convex case.

Corollary 1.3.1 *Assume V is a reflexive Banach space. Let J be a functional from V to $\mathbb{R} \cup \{+\infty\}$, proper, convex and lower semi-continuous and K be a non empty closed convex subset of V. If J is coercive or K is bounded, then the minimization problem has at least a solution. Moreover, if J is strictly convex the solution is unique.*

We end with first order optimality conditions.

Theorem 1.3.2 *Let K be non empty convex subset of V and J a functional from K to $\mathbb{R}$ Gâteaux-differentiable on K. Let be $u \in K$ a solution to problem $(\mathcal{P})$. Then*

$$\forall v \in K, \qquad < \nabla J(u), v - u > \geqslant 0. \tag{1.5}$$

1.4 Convex and non smooth analysis

1.4.1 Hahn -Banach Theorem

In what follows $\mathcal{X}$ is a real Banach space with dual $\mathcal{X}'$ (not necessarily reflexive). We note $\langle \cdot, \cdot \rangle$ the duality product between $\mathcal{X}$ and $\mathcal{X}'$:

$$\forall \varphi \in \mathcal{X}', \forall x \in \mathcal{X} \qquad \langle \varphi, x \rangle = \varphi(x)\ .$$

The geometrical form of Hahn-Banach theorem separates convex sets. It is very important in convex analysis and is used in particular to exhibit lagrange multipliers in optimization. We recall here the geometrical forms of this theorem (which is the only useful in our case) and significant corollaries as well. For demonstrations and more details we refer to [3].

Definition 1.4.1 (Affine hyperplane) *An affine closed hyperplane is defined as*

$$H = \{\ x \in \mathcal{X}\ \mid\ \langle \alpha, x \rangle + \beta = 0\ \},$$

where $\alpha \in \mathcal{X}'$ is non zero and $\beta \in \mathbb{R}$.

When $\mathcal{X}$ is an Hilbert space V (in particular if $V = \mathbb{R}^n$), then $V \simeq V'$ and any affine closed hyperplane takes the form

$$H = \{\ x \in V\ \mid (\alpha, x) + \beta = 0\ \},$$

where $(\cdot, \cdot)$ is the inner product V, $\alpha \in V$, $\alpha \neq 0$ and $\beta \in \mathbb{R}$.

Definition 1.4.2 (Separation) *Let A and B be two non empty subsets of $\mathcal{X}$. The affine hyperplane H $\langle \alpha, x\rangle + \beta = 0$, separates A and B if*

$$\forall x \in A \quad \langle \alpha, x\rangle + \beta \leqslant 0 \quad \text{and} \quad \forall y \in B \quad \langle \alpha, y\rangle + \beta \geqslant 0.$$

H strictly separates A and B if there exists $\varepsilon > 0$ such that

$$\forall x \in A \quad \langle \alpha, x\rangle + \beta \leqslant -\varepsilon \quad \text{and} \quad \forall y \in B \quad \langle \alpha, y\rangle + \beta \geqslant \varepsilon.$$

We may now give the first geometrical form of Hahn-Banach theorem:

Theorem 1.4.1 *Let A and B be two non empty convex subsets of $\mathcal{X}$ such that $A \cap B = \emptyset$. Assume A is* ***open****. Then, there exists an affine closed hyperplane which separates A and B.*

Corollary 1.4.1 *Let C be a non empty closed convex subset of $\mathbb{R}^n$ and $x^* \in C$. Then $x^* \in$ Int (C) if and only if there is no linear form that separates x^* and C.*

The second geometrical form of Hahn-Banach theorem writes :

Theorem 1.4.2 *Let A and B be two non empty, convex, subsets of $\mathcal{X}$ such that $A \cap B = \emptyset$.Assume that A is* ***closed*** *and B is* ***compact****. Then, there exists an affine closed hyperplane which strictly separates A and B.*

1.4.2 Subdifferential

Definition 1.4.3 *Let be $f : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ and $u \in$ dom f (i.e. $f(u) < +\infty$). The subdifferential of f at u is the set $\partial f(u)$ (possibly empty) of elements $u^* \in \mathcal{X}'$ such that*

$$\forall v \in \mathcal{X} \qquad f(v) \geqslant f(u) + \langle u^*, v - u\rangle .$$

These elements u^* are the subgradients.

Remark 1.4.1 *1. $f : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ achieves its minimum at $u \in$ dom f if and only if*

$$0 \in \partial f(u).$$

2. if $f, g : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ and $u \in$ dom $f \cap$ dom g, we get

$$\partial f(u) + \partial g(u) \subset \partial (f + g)(u).$$

3. As

$$\partial f(u) = \bigcap_{v \in \mathcal{X}} \{u^* \in \mathcal{X}' \mid \langle u^*, v - u\rangle \leqslant f(v) - f(u) \},$$

*then $\partial f(u)$ is convex, weak * closed set.*
4. for every $\lambda > 0$ we get $\partial(\lambda f)(u) = \lambda \partial f(u)$.

Theorem 1.4.3 (Link with Gâteaux-differentiability) *Let be* $f : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ *convex.*
If f *is Gâteaux-differentiable at* $u \in$ *dom* f, *it is subdifferentiable and* $\partial f(u) = \{f'(u)\}$.
Conversely, if f *is finite, continuous at* u *and has only one subgradient, then* f *is Gâteaux-differentiable at* u *and* $\partial f(u) = \{f'(u)\}$.

Theorem 1.4.4 (Subdifferential of a sum) *Let* f *and* g *convex, lower semi-continuous with values in* $\mathbb{R} \cup \{+\infty\}$. *Assume there exists* $u_o \in$*dom* $f \cap$ *dom* g *such that* f *is continuous at* u_o. *Then*

$$\forall u \in \mathcal{X} \qquad \partial(f+g)(u) = \partial f(u) + \partial g(u).$$

We end with a chain rule result

Theorem 1.4.5 *Let* Λ *be a linear continuous operator from* V *to* W *(Banach spaces). Let* f *be convex, lower semi-continuous from* W *dans* $\mathbb{R} \cup \{+\infty\}$ *continuous at (at most) a point of its (non empty) domain. Then*

$$\forall u \in V \qquad \partial(f \circ \Lambda)(u) = \Lambda^* \partial f(\Lambda u),$$

where Λ^* *is the adjoint operator of* Λ.

Details on these notions can be found [2, 5].
We conclude with an important example.

1.4.3 Application to a set indicator

When f is the indicator function of a non empty subset K of $\mathcal{X}$:

$$f(u) \stackrel{def}{=} 1_K(u) = \begin{cases} 0 & \text{if } u \in K, \\ +\infty & \text{else} \end{cases} \tag{1.6}$$

the subdifferential of f at u is called the *normal cone* of K at u:

$$\partial 1_K(u) = N_K(u) = \{\ u^* \in \mathcal{X}' \ \mid\ \langle u^*, v-u\rangle \leqslant 0 \text{ for every } v \in K\ \}.$$

When $\mathcal{X}$ is an Hilbert space (identified to its dual,) and K is a non empty, convex closed subset of $\mathcal{X}$, we may describe the subdifferential of 1_K at u (that is the normal cone to K at u):

Proposition 1.4.1 *Let be* $u \in K$, *where* K *is a non empty, convex closed subset of* $\mathcal{X}$ *(Hilbert space). Then for any* $c > 0$,

$$\lambda \in \partial 1_K(u) \Longleftrightarrow \lambda = c[u + \frac{\lambda}{c} - P_K(u + \frac{\lambda}{c})]$$

where P_K *is the projection of* $\mathcal{X}$ *onto the convex set* K.

Proof. We first note that $\partial 1_K(u)$ is a subset of $\mathcal{X}$. We recall that if P_K is the projection of $\mathcal{X}$ onto the closed, convex set K, the image $P_K(w)$ of any $w \in \mathcal{X}$ is characterized by

$$\forall v \in K \qquad (w - P_K(w), v - P_K(w))_{\mathcal{X}} \leqslant 0,$$

where $(\cdot,\cdot)_{\mathcal{X}}$ is the inner product of $\mathcal{X}$. Let be $\lambda \in \partial 1_K(u)$: λ is characterized by

$$\forall v \in K \qquad (\lambda, v - u)_{\mathcal{X}} \leqslant 0$$

that is, for any $c > 0$

$$\forall v \in K \qquad \left(u + \frac{\lambda}{c} - u, v - u\right)_{\mathcal{X}} \leqslant 0.$$

Setting $w = u + \dfrac{\lambda}{c}$ we get

$$\lambda \in \partial 1_K(u) \iff u = P_K(u + \frac{\lambda}{c}) \iff \lambda = c[u + \frac{\lambda}{c} - P_K(u + \frac{\lambda}{c})].$$

1.4.4 Legendre-Fenchel transformation

Definition 1.4.4 *Let be $f : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$. The Legendre-Fenchel transformed or conjugate function of f is $f^* : \mathcal{X}' \to \bar{\mathbb{R}}$ defined as*

$$\forall \ell \in \mathcal{X}' \qquad f^*(\ell) = \sup_{u \in \mathcal{X}} \{\ \ell(u) - f(u)\ \}. \tag{1.7}$$

Remark 1.4.2 *(a) if f "takes " the value $-\infty$, then $f^* \equiv +\infty$. If f is proper (that is non identically equal to $+\infty$) then f^* takes its values in $\mathbb{R} \cup \{+\infty\}$.*
(a) We shall note $\ell(u) =< \ell, u >$, where $< \cdot,\cdot >$is the duality product between $\mathcal{X}$ and $\mathcal{X}'$. Equation (1.7) reads

$$\forall u^* \in \mathcal{X}' \qquad f^*(u^*) = \sup_{u \in \mathcal{X}} \{\ \langle u^*, u\rangle - f(u)\ \}.$$

Definition 1.4.5 *Let $A \subset \mathcal{X}$ be a non empty set. The support function of A is $\sigma_A : \mathcal{X}' \to \mathbb{R} \cup \{+\infty\}$ defined by $\sigma_A = (1_A)^*$*

Example 1.4.1 *Let A be a set and $f(x) = d(x, A)$. Then $f^* = \sigma_A + 1_{B*}$ where B is the unit ball of $\mathcal{X}'$.*
If $f : u \mapsto \|u\|_{\mathcal{X}}$ (where $\|\cdot\|_{\mathcal{X}}$ is the norm of $\mathcal{X}$), then $f^* = 1_{B*}$

Proposition 1.4.2 *For every function $f : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$, the function f^* is convex and lower semi-continuous for the weak * topology.*

Proof. The definition gives

$$f^* = \sup_{u \in dom f} \varphi_u,$$

where dom f is the domain of f (i.e. the set of $u \in \mathcal{X}$such that $f(u)$ is finite) and $\varphi_u : \mathcal{X}' \to \mathbb{R}$ is defined by

$$\varphi_u(u^*) =< u^*, u > -f(u).$$

Every function φ_u is affine and continuous, so convex and lower semi-continuous for the weak * topology of $\mathcal{X}'$. It is the same for the supremum.

□

More generally

Proposition 1.4.3 *Let f be a positively homogeneous (proper) function from $\mathcal{X}$ to $\mathbb{R} \cup \{+\infty\}$, that is such that*

$$\forall \lambda \in \mathbb{R}, \forall x \in \mathcal{X} \qquad f(\lambda x) = |\lambda| f(x)\ .$$

Then the conjugate f^* is the indicator function of a closed, convex subset K of $\mathcal{X}'$.

Proof. Let f be a positively homogeneous (proper) function from $\mathcal{X}$ to $\mathbb{R} \cup \{+\infty\}$. Let be $u^* \in \mathcal{X}'$. Two cases occur:
• $\exists u_o \in \mathcal{X}$ such that $\langle u^*, u_o\rangle - f(u_o) > 0$. Then, by homogeneity, for every $\lambda > 0$

$$\langle u^*, \lambda u_o\rangle - f(\lambda u_o) = \lambda[\langle u^*, u_o\rangle - f(u_o)] \leqslant f^*(u^*).$$

Passing to the limit as $\lambda \to +\infty$ we get $f^*(u^*) = +\infty$.
• On the contrary

$$\forall u \in \mathcal{X} \qquad \langle u^*, u\rangle - f(u) \leqslant 0,$$

and $f^*(u^*) \leqslant 0$. The definition of f^* yields

$$\langle u^*, 0\rangle - f(0) \leqslant f^*(u^*)\ ;$$

as f is a positively homogeneous $f(0) = f(n \cdot 0) = nf(0)$ for every $n \in \mathbb{N}$ and $f(0) = 0$. We eventually obtain $f^*(u^*) = 0$.
Set $K = \{u^* \in \mathcal{X}^* \mid f^*(u^*) = 0\ \}$. We just proved that $f^* = 1_K$. As f^* is convex and lower semi-continuous, K is convex and closed. □

We now give a result between $f + g$ and $f^* + g^*$ which is the basis for the theory of duality in convex analysis:

Theorem 1.4.6 *Let $f, g : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ be convex functions such that, there exists $u_o \in$ dom g with f continuous at u_o. Then*

$$\inf_{u \in \mathcal{X}} (f(u) + g(u)) = \max_{u^* \in \mathcal{X}'} (-f^*(u^*) - g^*(-u^*))\ .$$

Proof. Set

$$\alpha = \inf_{u \in \mathcal{X}} \left(f(u) + g(u) \right) \text{ and } \beta = \sup_{u^* \in \mathcal{X}'} \left(-f^*(u^*) - g^*(-u^*) \right).$$

Let be $u \in \mathcal{X}$ and $u^* \in \mathcal{X}'$: with the definition we get

$$-f^*(u^*) \leqslant - < u^*, u > + f(u) \text{ and } - g^*(-u^*) \leqslant < u^*, u > + g(u),$$

so

$$-f^*(u^*) - g^*(-u^*) \leqslant f(u) + g(u) \; ;$$

passing to the sup in the left hand side and to the inf in the right hand side, we get

$$\beta \leqslant \alpha.$$

Let us show the converse inequality. As $u_o \in$ dom $f \cap$ dom g, $\alpha \in \mathbb{R} \cup \{-\infty\}$.
If $\alpha = -\infty$, the theorem is proved and we may assume that $\alpha \in \mathbb{R}$. Let be

$$C = \text{int}(\{(u,t) \in \mathcal{X} \times \mathbb{R} \mid \; f(u) \leqslant t \; \}),$$

and

$$D = \{(u,t) \in \mathcal{X} \times \mathbb{R} \mid \; t \leqslant \alpha - g(u) \; \}) \neq \emptyset.$$

As f and g are convex, C and D are convex. As f is continuous at u_o, C is non empty. Moreover $C \cap D = \emptyset$. We may apply Hahn-Banach theorem: there exists $(u_o^*, s_o) \in \mathcal{X}' \times \mathbb{R} \backslash \{0, 0\}$ and $c \in \mathbb{R}$ such that

$$\forall (v,s) \in D \qquad < u_o^*, v > + s s_o \geqslant c,$$

and

$$\forall (w,\sigma) \in C \qquad c \geqslant < u_o^*, w > + \sigma s_o. \tag{1.8}$$

As σ may go to $+\infty$ with the definition of C, we get $s_o \leqslant 0$.
Assume that $s_o \neq 0$. In that case, $s_o < 0$ and (up to a division by $|s_o|$) we may assume that $s_o = -1$. We obtain

$$\forall (v,s) \in D \qquad - < u_o^*, v > + s \leqslant -c.$$

Let $u \in \mathcal{X}$ and $s = \alpha - g(u)$: the couple (u,s) belongs to D. So

$$\forall u \in \mathcal{X} \qquad - < u_o^*, u > + \alpha - g(u) \leqslant -c.$$

On the other hand, relation (1.8) may be extended to $\bar{C}$ and, by convexity

$$\bar{C} = \{(u,t) \in \mathcal{X} \times \mathbb{R} \mid \; f(u) \leqslant t \; \} \; ;$$

we may apply it to $(u, f(u)$ for every $u \in \mathcal{X}$ which gives

$$c \geqslant < u_o^*, u > -f(u).$$

Finally,

$$g^*(-u_o^*) \leqslant -c - \alpha \text{ and } f^*(u_o^*) \leqslant c.$$

So

$$\alpha \leqslant -f^*(u_o^*) - g^*(-u_o^*) \leqslant \beta \leqslant \alpha$$

which ends the proof.
Case where $s_o = 0$: as f is continuous at u_o one may find a ball $B(u_o, R)$ with $R > 0$ included in dom f. Then $(u_o, \alpha - g(u_o)) \in D$ and for every $w \in B(u_o, R)$, $(u_o + w, f(u_o) + \varepsilon_o) \in C$: this gives

$$< u_o^*, u_o + w > \leqslant c \leqslant < u_o^*, u_o > .$$

This implies that $u_o^* = 0$ and a contradiction since $(u_o^*, s_o) \neq (0, 0)$. □

Remark 1.4.3 *Note that in the theorem the right hand term "sup" is always achieved (it is a maximum) which is not the case in the left hand term where the infimum is not necessarily achieved.*

Corollary 1.4.2 *Let* $f : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ *be convex continuous at* $u \in \mathcal{X}$. *Then*

$$f(u) = \max_{u^* \in \mathcal{X}'} (< u^*, u > -f^*(u^*)) \ .$$

Proof. Set $g = 1_{\{u\}}$. We get $g^*(u^*) = < u^*, u >$ for every $u^* \in \mathcal{X}'$. Functions f and g are convex and f is continuous at $u \in$dom g. The previous theorem gives

$$\begin{aligned} f(u) &= \inf_{u \in \mathcal{X}} (f + g)(u) \\ &= \max_{u^* \in \mathcal{X}'} (-f^*(u^*) - g^*(-u^*)) = \max_{u^* \in \mathcal{X}'} (< u^*, u > -f^*(u^*)) \,. \end{aligned}$$

□

This result can be generalized to convex lower semi-continuous functions

Theorem 1.4.7 *Let* $f : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ *be convex and lower semi-continuous. Then, for every* $u \in \mathcal{X}$

$$f(u) = \max_{u^* \in \mathcal{X}'} (< u^*, u > -f^*(u^*)) \,.$$

Proof. See [2] p. 89. □

We end with a biduality result which a corollary of the previous theorem if $\mathcal{X}$ is reflexive. This result still holds true even if $\mathcal{X}$ is not reflexive.

Theorem 1.4.8 *Let* f *be a proper, convex and lower semi-continuous function from* $\mathcal{X}$ *to* $\mathbb{R} \cup \{+\infty\}$. *Then* $f^{**} = f$.

1.4.5 Link with the subdifferential

Theorem 1.4.9 *Let be* $f : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ *and* f^* *the conjugate function. Then*

$$u^* \in \partial f(u) \Longleftrightarrow f(u) + f^*(u^*) = \langle u^*, u\rangle \, .$$

Proof. Let $u^* \in \partial f(u)$:

$$\forall v \in \mathcal{X} \qquad f(v) \geqslant f(u) + \langle u^*, v - u\rangle \, .$$

So

$$f^*(u^*) \geqslant \langle u^*, u\rangle - f(u) \geqslant \sup\{\langle u^*, v\rangle - f(v) \mid v \in \mathcal{X}\} = f^*(u^*).$$

We get :$f(u) + f^*(u^*) = \langle u^*, u\rangle$.
Conversely, if $f(u) + f^*(u^*) = \langle u^*, u\rangle$ we get, for every $v \in \mathcal{X}$

$$\langle u^*, u\rangle - f(u) = f^*(u^*) \geqslant \langle u^*, v\rangle - f(v),$$

$$\langle u^*, v - u\rangle \leqslant f(v) - f(u),$$

that is $u^* \in \partial f(u)$. □

Corollary 1.4.3 *If* $f : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ *is convex, proper and lower semi-continuous, then*

$$u^* \in \partial f(u) \Longleftrightarrow u \in \partial f^*(u^*).$$

Proof. It suffices to use the previous theorem to f^* and use that if f is convex, proper and lsc then $f = f^{**}$. □

2

The space of functions with bounded variation

2.1 Sobolev spaces

For more details, one may refer to [4]. Let Ω be a bounded open subset of $\mathbb{R}^n$, ($n \leqslant 3$ in practical cases) with regular boundary Γ. We call $\mathcal{D}(\Omega)$ the space of functions $\mathcal{C}^\infty$ with compact support in Ω. The dual space $\mathcal{D}'(\Omega)$ is the space of *distributions* on Ω.

For any distribution $u \in \mathcal{D}'(\Omega)$, the derivative $\dfrac{\partial u}{\partial x_i}$ is defined (by duality) as :

$$\forall \varphi \in \mathcal{D}(\Omega) \qquad \left\langle \frac{\partial u}{\partial x_i}, \varphi \right\rangle_{\mathcal{D}'(\Omega),\mathcal{D}(\Omega)} \overset{def}{\equiv} - \left\langle u, \frac{\partial \varphi}{\partial x_i} \right\rangle_{\mathcal{D}'(\Omega),\mathcal{D}(\Omega)} .$$

We denote the derivative of u in the sense of distributions $D_i u = \dfrac{\partial u}{\partial x_i} = \partial_i u$.

if $\alpha \in \mathbb{N}^n$, we note $D^\alpha u = \partial_1^{\alpha_1} u \cdots \partial_n^{\alpha_n} u$ and $|\alpha| = \alpha_1 + \cdots + \alpha_n$; we get

$$\forall \varphi \in \mathcal{D}(\Omega) \qquad \langle D^\alpha u, \varphi \rangle_{\mathcal{D}'(\Omega),\mathcal{D}(\Omega)} = (-1)^{|\alpha|} \langle u, D^\alpha \varphi \rangle_{\mathcal{D}'(\Omega),\mathcal{D}(\Omega)} .$$

Definition 2.1.1 *We define the Sobolev spaces $W^{m,p}(\Omega)$ as following:*

$$W^{m,p}(\Omega) = \{ \ u \in L^p(\Omega) \mid D^\alpha u \in L^p(\Omega), \ |\alpha| \leqslant m \ \}$$

$$H^m(\Omega) = W^{m,2}(\Omega) \{ \ u \in \mathcal{D}'(\Omega) \mid D^\alpha u \in L^2(\Omega), \ |\alpha| \leqslant m \ \} .$$

Remark 2.1.1 $H^o(\Omega) = L^2(\Omega)$.

We will enunciate a series of properties of Sobolev spaces without proofs. We can see [4] for example.

Proposition 2.1.1 *$H^m(\Omega)$ endowed with the scalar product*

$$(u, v)_m = \sum_{|\alpha| \leqslant m} \int_\Omega D^\alpha u(x) \ D^\alpha v(x) \ dx ,$$

is an Hilbert space.

Proposition 2.1.2

$$H^m(\Omega) \subset H^{m'}(\Omega)$$

and the injection is continuous, for $m \geqslant m'$.

Definition 2.1.2

$$H^1_o(\Omega) = \{ \ u \in H^1(\Omega) \mid u_{|\Gamma} = 0 \ \} \ .$$

It is also the closure of $\mathcal{D}(\Omega)$ *in* $H^1(\Omega)$.

$$H^m_o(\Omega) = \{ \ u \in H^1(\Omega) \mid \frac{\partial^j u}{\partial n^j}_{|\Gamma} = 0, \ j = 1, \cdots, m-1\} \ ,$$

where $\dfrac{\partial}{\partial n}$ *is the derivative of* u *along the exterior normal to the boundary* Γ:

$$\frac{\partial u}{\partial n} = \sum_{i=1}^{n} \frac{\partial u}{\partial x_i} cos(\mathbf{n}, \mathbf{e_i}) \ ,$$

where $\mathbf{n}$ *is the exterior normal vector to* Γ *and* Ω *is smooth enough (*Γ *is* $\mathcal{C}^\infty$ *for example).*

Definition 2.1.3 *For every* $m \in \mathbb{N}$, *we note* $H^{-m}(\Omega)$ *the dual space of* $H^m_o(\Omega)$.

Theorem 2.1.1 (Rellich) *If* Ω *is a bounded open subset of* $\mathbb{R}^n$, *then for every* $m \in \mathbb{N}$, *the embedding of* $H^{m+1}_o(\Omega)$ *in* $H^m_o(\Omega)$ *is compact.*

In particular the embedding of $H^1_o(\Omega)$ in $L^2(\Omega)$ is compact.
Practically, this means that any bounded sequence for the $H^1_o(\Omega)$ *norm, weakly converges in* $H^1_o(\Omega)$ *(up to a subsequence) and strongly in* $L^2(\Omega)$.

2.2 The space of functions of bounded variation $BV(\Omega)$

2.2.1 Definition

In the sequel Ω is a bounded open subset of $\mathbb{R}^2$ with Lipschitz boundary and $\mathcal{C}^1_c(\Omega, \mathbb{R}^2)$ is the space of $\mathcal{C}^1$ functions with compact support in Ω with values in $\mathbb{R}^2$.

Definition 2.2.1 *A function* $f \in L^1(\Omega)$ *(with values in* $\mathbb{R}$*) is with bounded variation in* Ω *if* $\Phi(f) < +\infty$ *where*

$$\Phi(f) = \sup \left\{ \int_\Omega f(x) \, div \, \varphi(x) \, dx \mid \varphi \in \mathcal{C}^1_c(\Omega, \mathbb{R}^2) \ , \ \|\varphi\|_\infty \leqslant 1 \right\}. \tag{2.1}$$

We denote

$$BV(\Omega) = \{ f \in L^1(\Omega) \mid \Phi(f) < +\infty \}$$

the space of such functions.

Remark 2.2.1 *Recall that if* $\varphi = (\varphi_1, \varphi_2) \in \mathcal{C}^1_c(\Omega, \mathbb{R}^2)$ *then*

$$\forall x = (x_1, x_2) \in \Omega \qquad div\ \varphi(x) = \frac{\partial \varphi_1}{\partial x_1}(x) + \frac{\partial \varphi_2}{\partial x_2}(x)\ .$$

So, integrating by parts

$$\begin{aligned}\int_\Omega f(x)\, div\ \varphi(x)\, dx &= \int_\Omega \left(f(x)\, \frac{\partial \varphi_1}{\partial x_1}(x) + \frac{\partial \varphi_2}{\partial x_2}(x) \right)\, dx \\ &= -\int_\Omega \left(\frac{\partial f}{\partial x_1}(x)\varphi_1(x) + \frac{\partial f}{\partial x_2}(x)\varphi_2(x) \right)\ dx \\ &= -\int_\Omega \nabla f(x) \cdot \varphi(x)\ dx\ ,\end{aligned}$$

where $\cdot$ *is the scalar product of* $\mathbb{R}^2$ *:* $x \cdot y = x_1 y_1 + x_2 y_2$ *.*

Definition 2.2.2 (Perimeter) *A set measurable* E *(for the Lebesgue measure) of* $\mathbb{R}^2$ *is of finite perimeter (or length) if its characteristic function* χ_E *(*$\chi_E = 1$ *on* E *and* 0 *elsewhere) belongs to* $BV(\Omega)$*.*

Recall that a Radon measure is a measure finite on any compact set and with the Riesz theorem, any linear continuous form $\mathcal{L}$ on $\mathcal{C}^o_c(\Omega)$ (continuous functions with compact support) may be written as

$$\mathcal{L}(f) = \int_\Omega f(x)\ d\mu\ ,$$

where μ is the (unique) Radon measure associated to $\mathcal{L}$. More precisely

Theorem 2.2.1 ([9] p. 126, [6] p. 49) *For every linear bounded form* $\mathcal{L}$ *on* $\mathcal{C}^o_c(\Omega, \mathbb{R}^2)$*, that is*

$$\forall K \text{ compact subset of } \Omega,\ \sup \left\{ \mathcal{L}(\varphi) \mid \varphi \in \mathcal{C}^o_c(\Omega, \mathbb{R}^2)\ ,\ \|\varphi\|_\infty \leqslant 1\ ,\ supp\, \varphi \subset K\ \right\} < +\infty\ ,$$

there corresponds a unique positive Radon measure μ *on* Ω *and a function* μ*-measurable* σ *(" sign" function) such that*
(i) $|\sigma(x)| = 1,\ \mu$ *a. e. , and*
(ii) $\mathcal{L}(\varphi) = \int_\Omega \varphi(x)\, \sigma(x)\, d\mu$ *for every function* $\varphi \in \mathcal{C}^o_c(\Omega, \mathbb{R}^2)$ *.*
(iii) Moreover μ *is the variation measure and verifies*

$$\mu(\Omega) = \sup \left\{\ \mathcal{L}(\varphi) \mid \varphi \in \mathcal{C}^o_c(\Omega, \mathbb{R}^2)\ ,\ \|\varphi\|_\infty \leqslant 1\ ,\ supp\, \varphi \subset \Omega\ \right\}\ . \qquad (2.2)$$

We may give a structural property of $BV(\Omega)$ functions.

Theorem 2.2.2 *Let be $f \in BV(\Omega)$. There exists a positive Radon measure μ on Ω and a function μ-measurable $\sigma : \Omega \to \mathbb{R}$ such that*
(i) $|\sigma(x)| = 1,\ \mu$ a. e. , and
(ii) $\displaystyle\int_\Omega f(x)\, div\, \varphi(x)\, dx = -\int_\Omega \varphi(x)\, \sigma(x)\, d\mu$ for every function $\varphi \in \mathcal{C}_c^1(\Omega, \mathbb{R}^2)$.

Relation (ii) is a *weak* integration by parts formula. This theorem shows that the weak derivative (in the sense of distributions) of a $BV(\Omega)$ function is a Radon measure.
Proof - Let be $f \in BV(\Omega)$ and consider the linear form $\mathcal{L}$ defined on $\mathcal{C}_c^1(\Omega, \mathbb{R}^2)$ by

$$\mathcal{L}(\varphi) = \int_\Omega f(x) \text{div } \varphi(x)\, dx \ .$$

As $f \in BV(\Omega)$,

$$C_\mathcal{L} := \sup \left\{ \mathcal{L}(\varphi) \mid \varphi \in \mathcal{C}_c^1(\Omega, \mathbb{R}^2) \ , \ \|\varphi\|_\infty \leqslant 1 \ \right\} < +\infty$$

for every function $\varphi \in \mathcal{C}_c^1(\Omega, \mathbb{R}^2)$, so that

$$\forall \varphi \in \mathcal{C}_c^1(\Omega, \mathbb{R}^2) \qquad \mathcal{L}(\varphi) \leqslant C_\mathcal{L} \|\varphi\|_\infty \ . \tag{2.3}$$

Let K be a compact subset of Ω. For every function $\varphi \in \mathcal{C}_c^o(\Omega, \mathbb{R}^2)$ with compact support in K, one may find (by density) a sequence of functions $\varphi_k \in \mathcal{C}_c^1(\Omega, \mathbb{R}^2)$ uniformly convergent to φ. Let us set

$$\bar{\mathcal{L}}(\varphi) = \lim_{k \to +\infty} \mathcal{L}(\varphi_k) \ .$$

With (2.3) the limit exists and does not depend on the sequence (φ_k). One may extend $\mathcal{L}$ by density to a linear form $\bar{\mathcal{L}}$ on $\mathcal{C}_c^o(\Omega, \mathbb{R}^2)$ such that

$$\sup \left\{ \bar{\mathcal{L}}(\varphi) \mid \varphi \in \mathcal{C}_c^o(\Omega, \mathbb{R}^2) \ , \ \|\varphi\|_\infty \leqslant 1 \ , \ \text{supp}\, \varphi \subset K \ \right\} < +\infty \ .$$

We conclude with the Riesz theorem. □
With (2.2), $\Phi(u) = \mu(\Omega) \geqslant 0$: it is the *total variation* of f. The application

$$\begin{aligned} BV(\Omega) &\to \mathbb{R}^+ \\ u &\mapsto \|u\|_{BV(\Omega)} = \|u\|_{L^1} + \Phi(u) \ . \end{aligned}$$

is a norm. We endow the space $BV(\Omega)$ with this norm.

Example 2.2.1 *Assume that*

$$f \in W^{1,1}(\Omega) = \{ \ f \in L^1(\Omega) \mid Df \in L^1(\Omega) \ \} \ ,$$

where Df is the derivative of f (in the sense of distributions). Let $\varphi \in \mathcal{C}_c^1(\Omega, \mathbb{R}^2)$ be such that $\|\varphi\|_\infty \leqslant 1$. Then

$$\int_\Omega f\, div\, \varphi\, dx = -\int_\Omega Df \cdot \varphi\, dx \leqslant \|\varphi\|_\infty \int_\Omega |Df|\, dx \leqslant \|Df\|_{L^1} < +\infty \ .$$

Therefore $f \in BV(\Omega)$. *Moreover*

$$\Phi(f) = \sup\{-\int_\Omega Df \cdot \varphi\, dx \mid \|\varphi\|_\infty \leqslant 1 \} = \|Df\|_{L^1} \ ,$$

and

$$\sigma = \begin{cases} \dfrac{Df}{|Df|} & \text{if } Df \neq 0 \ , \\ 0 & \text{else.} \end{cases}$$

So $W^{1,1}(\Omega) \subset BV(\Omega)$. *In particular, as* Ω *is bounded*

$$\forall 1 \leqslant p \leqslant +\infty \quad W^{1,p}(\Omega) \subset BV(\Omega) \ .$$

Remark 2.2.2 *With Radon-Nikodym theorem that describes measures decomposition, then for every function* $u \in BV(\Omega)$, *the following decomposition of* Du *holds:*

$$Du = \nabla u\, dx + D^s u \ ,$$

where $\nabla u\, dx$ *is the absolutely continuous part of* Du *with respect to the Lebesgue measure and* $D^s u$ *is the singular part.*

2.2.2 Approximation and compactness

Theorem 2.2.3 (Lower semi-continuity of the total variation) *The application* $u \mapsto \Phi(u)$ *from* $BV(\Omega)$ *to* $\mathbb{R}^+$ *is lower semi-continuous for the sequential topology of* $L^1(\Omega)$.
More precisely, if (u_k) *is a sequence of functions in* $BV(\Omega)$ *that converges to* u *strongly in* $L^1(\Omega)$ *then*

$$\Phi(u) \leqslant \liminf_{k \to +\infty} \Phi(u_k) \ .$$

Proof - Let $\varphi \in \mathcal{C}^1_c(\Omega, \mathbb{R}^2)$ be such that $\|\varphi\|_\infty \leqslant 1$. Then

$$\int_\Omega u(x)\, \mathrm{div}\ \varphi(x)\, dx = \lim_{k \to +\infty} \int_\Omega u_k(x)\, \mathrm{div}\ \varphi(x)\, dx \ .$$

So, for every $\varepsilon > 0$, there exists $k[\varphi, \varepsilon]$ such that for every $k \geqslant k[\varphi, \varepsilon]$:

$$\int_\Omega u(x)\, \mathrm{div}\ \varphi(x)\, dx - \varepsilon \leqslant \int_\Omega u_k(x)\, \mathrm{div}\ \varphi(x)\, dx \leqslant \int_\Omega u(x)\, \mathrm{div}\ \varphi(x)\, dx + \varepsilon \ .$$

As

$$\int_\Omega u_k(x)\, \mathrm{div}\ \varphi(x)\, dx \leqslant \Phi(u_k)$$

it comes

$$\forall k \geqslant k[\varphi,\varepsilon] \quad \int_\Omega u(x)\,\mathrm{div}\;\varphi(x)\,dx - \varepsilon \leqslant \Phi(u_k)\ ,$$

and therefore

$$\int_\Omega u(x)\,\mathrm{div}\;\varphi(x)\,dx \leqslant \liminf_{k\to+\infty} \Phi(u_k)\ .$$

As it true for every φ, we obtain

$$\Phi(u) \leqslant \liminf_{k\to+\infty} \Phi(u_k)\ .$$

□

We admit the following result

Theorem 2.2.4 (Approximation) *For every function $u \in BV(\Omega)$, there exists a sequence of functions $(u_k)_{k\in\mathbb{N}}$ of $BV(\Omega)\cap\mathcal{C}^\infty(\Omega)$ such that*
(i) $u_k \to u$ in $L^1(\Omega)$ and
(ii) $\Phi(u_k) \to \Phi(u)$ (in $\mathbb{R}$).

The proof use a classical regularization process (by convolution). On can refer to [6] p.172.
The previous result **is not** a density result of $BV(\Omega)\cap\mathcal{C}^\infty(\Omega)$ in $BV(\Omega)$ since we do not have $\Phi(u_k - u) \to 0$ but only $\Phi(u_k) \to \Phi(u)$.

Theorem 2.2.5 *The space $BV(\Omega)$ endowed with the norm*

$$u \mapsto \|u\|_{BV(\Omega)} = \|u\|_{L^1} + \Phi(u)$$

is a Banach space.

Proof - Let $(u_n)_{n\in\mathbb{N}}$ be a Cauchy sequence of $BV(\Omega)$. It is also a Cauchy sequence in $L^1(\Omega)$: so it converges to some $u \in L^1(\Omega)$. On the other hand, it is bounded in $BV(\Omega)$ (as a Cauchy sequence), so

$$\exists M > 0, \forall n \qquad \Phi(u_n) \leqslant M\ .$$

With theorem 2.2.3,

$$\Phi(u) \leqslant \liminf_{n\to\infty} \Phi(u_n) \leqslant M < +\infty\ .$$

This implies that $u \in BV(\Omega)$. Let be $\varepsilon > 0$ and N such that

$$\forall n, k \geqslant N \qquad \|u_n - u_k\|_{BV(\Omega)} \leqslant \varepsilon\ .$$

So

$$\forall n, k \geqslant N \qquad \Phi(u_n - u_k) \leqslant \varepsilon$$

and with the lower semi-continuity of Φ (fixing n) we get

$$\forall n \geqslant N \qquad \Phi(u_n - u) \leqslant \liminf_{k\to\infty} \Phi(u_n - u_k) \leqslant \varepsilon\ .$$

This proves $\Phi(u_n - u) \to 0$. □
We end with a compactness result that we will admit

Theorem 2.2.6 (compactness) *The space $BV(\Omega)$ is compactly embedded in $L^1(\Omega)$. More precisely, if $(u_n)_{n\in\mathbb{N}}$ is a bounded sequence in $BV(\Omega)$*

$$\sup_{n\in\mathbb{N}} \|u_n\|_{BV(\Omega)} < +\infty \ ,$$

then , there exists a subsequence $(u_{n_k})_{k\in\mathbb{N}}$ and a function $u \in BV(\Omega)$ such that u_{n_k} strongly converges to u in $L^1(\Omega)$.

Proof - [6] p.176. □

More generally we have Sobolev type injections :

Theorem 2.2.7 (Embedding in L^p spaces) *Assume Ω is a smooth open subset of $\mathbb{R}^N$. then*

- *The space $BV(\Omega)$ is continuously embedded in $L^p(\Omega)$ for $1 \leqslant p \leqslant \dfrac{N}{N-1}$*
- *The space $BV(\Omega)$ is compactly embedded in $L^p(\Omega)$ for $1 \leqslant p < \dfrac{N}{N-1}$*

In particular, for $N = 2$ the space $BV(\Omega)$ is continuously embedded in $L^2(\Omega)$. For details on the functions with bounded variation, reference may be made to [6].

2.3 Variational method

Variational methods proposed to minimize noise while adding a priori on the desired image. We will clarify this idea: we work now in a continuous (infinite-dimensional) framework, and then we will make a discretization. Given an original image u_0 we assume it has been degraded by additive noise w, and possibly by a blur operator R. From the observed image $u_d = Ru_0 + v$ (which is a degraded version of the original image u_0), we attempt to reconstruct u_0. If we assume that the additive noise w is Gaussian, the method of maximum likehood leads to seek u_0 as the solution of minimization problem

$$\inf_u \|u_d - Ru\|_2^2,$$

where $\|\cdot\|_2$ denotes the norm in $L^2(\Omega)$. This is an ill-posed inverse problem : the operator is not necessarily invertible (and even when invertible, its inverse is often difficult numerically to compute). In other words, the existence and /or uniqueness of solutions is not ensured and even if this is the case, the solution may not be stable (that is continuous with respect to the data). To solve it numerically, we introduce a regularization term (prior on the image), and consider the problem

$$\inf_u \underbrace{\|u_d - Ru\|_2^2}_{\text{data fitting}} + \underbrace{L(u)}_{\text{Regularization}} \ .$$

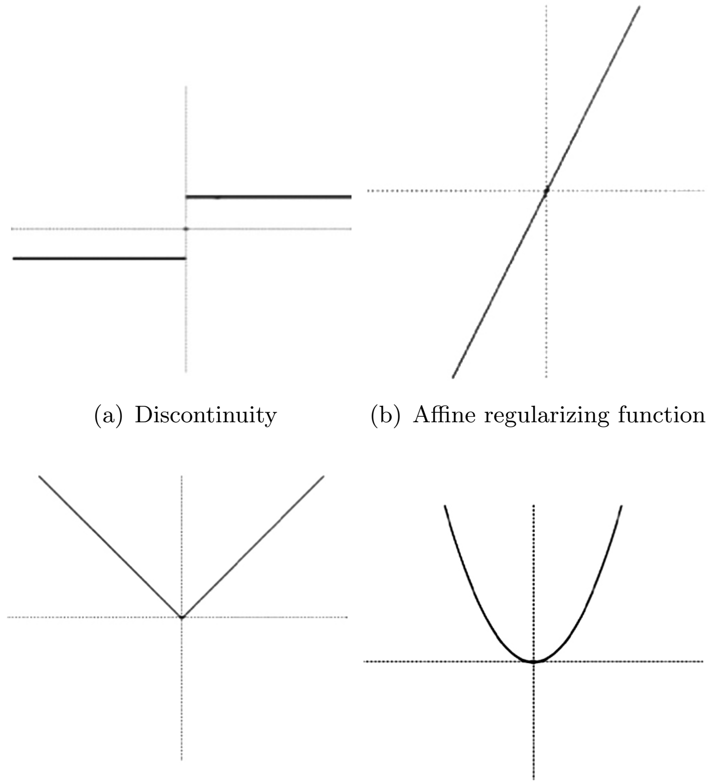
2.3.1 The continuous model of Rudin-Osher-Fatemi

We will first motivate the model that follows. In the discrete part, it consists of replacing the square of the norm of the gradient of the standard image by (1 to the power). The continuous model presented in next subsection is the rigorous mathematical formalization of the change in the standard regularization term.

Recall that we wish to suppress noise while preserving the edges of the image that is to say the discontinuity set of the function describing the image. Let us consider the one-dimensional case, for example by taking a section of the image below.

Fig. 2.1. Original image and smoothed image

White/black outline corresponds to the discontinuity in the original image. In the smoothed image the discontinuity is regulated by an affine function. The gradient of the image is the quantity describing the contours, the choice of the regularization term is done by taking a primitive. Then we see that the standard (1D absolute value) is preferable to the choice of the norm squared (in the 1D function $x \mapsto x^2$).



(a) Discontinuity (b) Affine regularizing function

(c) Primitive: $|x|$ (d) Primitive: x^2

Fig. 2.2. Regularization $\|x\|$ versus$\|x\|^2$

An alternative to the H^1 -Tychonov regularization (which is too violent) is to replace the regularization term $\|\nabla u\|_2^2$ by a less restrictive regularizing term. Rudin, Osher and Fatemi [7] proposed a model where the image is decomposed into two parts: $u_d = u + w$ where w is the noise and the u part is *regular*. We will therefore seek the solution of the problem in the form $u + w$ with $u \in BV(\Omega)$ and $w \in L^2(\Omega)$. This leads to :

$$(\mathcal{P}_{ROF}) \qquad \min \left\{ \frac{1}{2}\|w\|_2^2 + \varepsilon \Phi(u) \mid u \in BV(\Omega),\ w \in L^2(\Omega),\ u + w = u_d \right\}.$$

Here the regularization term is the total variation of u : $\Phi(u)$ and $\varepsilon > 0$. If the function u belongs to

$$W^{1,1}(\Omega) = \{\ u \in L^1(\Omega) \mid \nabla u \in L^1(\Omega)\ \}(\subset BV(\Omega)),$$

the total variation is $\Phi(u) = \|\nabla u\|_1$. Here $\|\cdot\|_p$ stands for the usual norm in $L^p(\Omega)$.

Theorem 2.3.1 *Problem* $(\mathcal{P}_{ROF})$*has a unique solution.*

Proof. Let $(u_n, w_n) \in BV(\Omega) \times L^2(\Omega)$ be a minimizing sequence. As w_n is bounded in $L^2(\Omega)$ one may extract a subsequence weakly convergent to w^* in $L^2(\Omega)$. As the $L^2(\Omega)$ norm is convex and lower semi-continuous, we get

$$\|w^*\|_2^2 \leqslant \liminf_{n\to+\infty} \|w_n\|_2^2.$$

Similarly $u_n = u_d - w_n$ is bounded in $L^2(\Omega)$ and thus in $L^1(\Omega)$ since Ω is bounded. As $\Phi(u_n)$ is bounded, this yields that u_n is bounded in $BV(\Omega)$. With the compact embedding of $BV(\Omega)$ in $L^1(\Omega)$ (Theorem 2.2.6), we obtain that u_n converges (up to a subsequence) strongly in $L^1(\Omega)$ to $u^* \in BV(\Omega)$.
Moreover Φ is lower semi-continuous (Theorem 2.2.3), so that

$$\Phi(u^*) \leqslant \liminf_{n\to+\infty} \Phi(u_n),$$

and

$$\Phi(u^*) + \frac{1}{2\varepsilon}\|w^*\|_2^2 \leqslant \liminf_{n\to+\infty} \Phi(u_n) + \frac{1}{2\varepsilon}\|w_n\|_2^2 = \inf(\mathcal{P}_{ROF}).$$

As $u_n + w_n = u_d$ for every n, we have $u^* + w^* = u_d$. Therefore u^* is a solution to problem $(\mathcal{P}_{ROF})$.
As the functional is strictly convex with respect to (u, w) and the constraint is affine we get uniqueness.

We will need to establish optimality conditions for optimal solutions of the proposed models. However, Φ is not Gâteaux-differentiable and we have to use the notions of non-smooth analysis.

2.3.2 First order optimality conditions

Problem $(\mathcal{P}_{ROF})$ equivalently writes

$$\min_{u\in BV(\Omega)} \mathcal{F}(u) := \Phi(u) + \frac{1}{2\varepsilon}\|u - u_d\|_2^2. \tag{2.4}$$

The functional $\mathcal{F}$ is convex and $\bar{u}$ is solution to $(\mathcal{P}_{ROF})$ if and only if $0 \in \partial\mathcal{F}(\bar{u})$. Relation: $0 \in \partial\mathcal{F}(\bar{u}) \iff \bar{u} \in \partial\mathcal{F}^*(0)$, is true even if the space is not reflexive (see Corollary1.4.3 et/and/or [1] Theorem 9.5.1 p. 333). Here $\mathcal{F}^*$ is the Fenchel conjugate of $\mathcal{F}$.

We use Theorem 1.4.4 to compute $\partial\mathcal{F}(u)$. The application $u \mapsto \|u - u_d\|_2^2$ is continuous on $L^2(\Omega)$ and Φ is (convex) with values in $\mathbb{R} \cup \{+\infty\}$ and finite on $BV(\Omega)$. Furthermore, $u \mapsto \|u - u_d\|_2^2$ is Gâteaux-differentiable on $L^2(\Omega)$. As we are going to use the convex duality we first compute the Legendre-Fenchel conjugate of Φ.

Theorem 2.3.2 *The Legendre-Fenchel conjugate Φ^* of the total variation Φ is the indicator function of the closure (in $L^2(\Omega)$) of the set*

$$\mathcal{K} := \left\{\xi = div\ \varphi \mid \varphi \in \mathcal{C}_c^1(\Omega, \mathbb{R}^2),\ \|\varphi\|_\infty \leqslant 1\ \right\}.$$

Proof. Ad Φ is positively homogeneous la conjugate Φ^* of Φ is the indicator function of a closed, convex set $\tilde{\mathcal{K}}$ in the dual of $BV(\Omega)$ (Proposition 1.4.3). As $L^2(\Omega)$ is continuously embedded in $BV(\Omega)$ (Theorem 2.2.7), the dual of $BV(\Omega)$ is continuously embedded in $L^2(\Omega)$. Therefore $\tilde{\mathcal{K}}$ is closed in $L^2(\Omega)$.
We first prove that $\mathcal{K} \subset \tilde{\mathcal{K}}$: Let $u \in \mathcal{K}$. By definition

$$\Phi(u) = \sup_{\xi \in \mathcal{K}} (\xi, u) , \tag{2.5}$$

where $(\cdot, \cdot)$ is the inner product of $L^2(\Omega)$. Therefore, $(\xi, u)) - \Phi(u) \leqslant 0$ for every $\xi \in \mathcal{K}$ and $u \in L^2(\Omega)$ (Note that if $u \in L^2(\Omega) \backslash BV(\Omega)$ then $\Phi(u) = +\infty$ by definition of $BV(\Omega)$). We deduce that, for every $u^* \in \mathcal{K}$

$$\Phi^*(u^*) = \sup_{u \in L^2(\Omega)} (u^*, u) - \Phi(u) = \sup_{u \in BV(\Omega)} (u^*, u) - \Phi(u) \leqslant 0.$$

As Φ^* has only one finite value we get $\Phi^*(u^*) = 0$, and $u^* \in \tilde{\mathcal{K}}$. Consequently $\mathcal{K} \subset \tilde{\mathcal{K}}$ and as $\tilde{\mathcal{K}}$ is closed:

$$\bar{\mathcal{K}} \subset \tilde{\mathcal{K}}.$$

In particular

$$\Phi(u) = \sup_{\xi \in \mathcal{K}} (u, \xi) \leqslant \sup_{\xi \in \bar{\mathcal{K}}} (u, \xi) \leqslant \sup_{\xi \in \tilde{\mathcal{K}}} (u, \xi) = \sup_{\xi \in \tilde{\mathcal{K}}} (u, \xi) - \Phi^*(\xi) = \Phi^{**}(u).$$

As $\Phi^{**} = \Phi$, we obtain

$$\sup_{\xi \in \mathcal{K}} (u, \xi) \leqslant \sup_{\xi \in \tilde{\mathcal{K}}} (u, \xi) \leqslant \sup_{\xi \in \mathcal{K}} (u, \xi) ,$$

and

$$\sup_{\xi \in \mathcal{K}} (u, \xi) = \sup_{\xi \in \bar{\mathcal{K}}} (u, \xi) = \sup_{\xi \in \tilde{\mathcal{K}}} (u, \xi) . \tag{2.6}$$

Assume there exists $u^* \in \tilde{\mathcal{K}}$ such that $u^* \notin \bar{\mathcal{K}}$. Then we may strictly separate u^* and the closed convex set $\bar{\mathcal{K}}$. There exists $\alpha \in \mathbb{R}$ and u_o such that

$$(u_o, u^*) > \alpha \geqslant \sup_{v \in \bar{\mathcal{K}}} (u_o, v) .$$

With (2.6) we get

$$\sup_{\xi \in \tilde{\mathcal{K}}} (u_o, \xi) \geqslant (u_o, u^*) > \alpha \geqslant \sup_{v \in \bar{\mathcal{K}}} (u_o, v) = \sup_{v \in \tilde{\mathcal{K}}} (u_o, v) .$$

We have a contradiction: $\tilde{\mathcal{K}} = \bar{\mathcal{K}}$. □

Then u is a solution de $(\mathcal{P}_{ROF})$ if and only if

$$0 \in \partial \left(\Phi(u) + \frac{1}{2\varepsilon}\|u - u_d\|_2^2 \right) = \frac{u - u_d}{\varepsilon} + \partial\Phi(u).$$

As Φ is convex, lower semi-continuous and proper we may apply corollary1.4.3. So

$$\frac{u_d - u}{\varepsilon} \in \partial\Phi(u) \Longleftrightarrow u \in \partial\Phi^*(\frac{u_d - u}{\varepsilon}) \Longleftrightarrow 0 \in -u + \partial\Phi^*(\frac{u_d - u}{\varepsilon}).$$

Finally, u is a solution de $(\mathcal{P}_{ROF})$ if and only

$$0 \in -u + \partial 1_{\overline{\mathcal{K}}}(\frac{u_d - u}{\varepsilon}),$$

that is

$$u \in \mathcal{N}_{\overline{\mathcal{K}}}(\frac{u_d - u}{\varepsilon}) \ ,$$

where $\mathcal{N}_{\overline{\mathcal{K}}}$ is the normal cone to $\overline{\mathcal{K}}$.

References

1. Attouch, H., Buttazzo, G., Michaille, G.: Variational analysis in Sobolev and BV spaces, *MPS/SIAM Series on Optimization*, vol. 6. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA (2006). Applications to PDEs and optimization
2. Azé, D.: Éléments d'analyse convexe et variationnelle. Ellipses (1997)
3. Brezis, H.: Analyse fonctionnelle. Collection Mathématiques Appliquées pour la Maîtrise. [Collection of Applied Mathematics for the Master's Degree]. Masson, Paris (1983). Théorie et applications. [Theory and applications]
4. Dautray, R., Lions, J.L.: Analyse mathématique et calcul numérique pour les sciences et techniques, vol. 9 volumes. Masson, Paris (1987)
5. Ekeland, I., Temam, R.: Analyse convexe et problèmes variationnels. Dunod, Paris (1973)
6. Evans, L.C., Gariepy, R.F.: Measure theory and fine properties of functions. Studies in Advanced Mathematics. CRC Press, Boca Raton, FL (1992)
7. Osher, S., Fatemi, E., Rudin, L.: Nonlinear total variation based noise removal algorithms. Physica D **60**, 259–268 (1992)
8. Pallara, D., Ambrosio, L., Fusco, N.: Functions of bounded variations and free discontinuity problems. Oxford Mathematical Monograph (2000)
9. Rudin, W.: Analyse réelle et complexe. Masson, Paris (1978)