Journal of Convex Analysis
Volume 10 (2003), No. 1, 1–26

Minimizers of Energy Functional under not very Integrable Constraints

Christian Léonard
Equipe Modal-X. Université de Paris X, Bâtiment G. 200,
Avenue de la République, 92001 Nanterre Cedex, France;
Centre de Mathématiques Appliquées, Ecole Polytechnique,
91128 Palaiseau Cedex, France
Christian.Leonard@u-paris10.fr

Received July 12, 2001
Revised manuscript received March 1, 2002

We consider a general class of problems of minimization of convex integral functionals (such as entropy maximization) subject to linear constraints. Under general assumptions, the minimizers are characterized. Our results improve previous literature on the subject in the following directions:

- necessary and sufficient conditions for the shape of the minimizers are proved
- without constraint qualification
- under infinitely many linear constraints subject to natural integrability conditions (no topological restrictions).

This paper extends previous results of the author by relaxing some integrability conditions on the constraint. As a consequence, the minimizers may admit a singular component. Our proofs mainly rely on convex duality.

Keywords: Maximum entropy method, relative entropy, convex integral functionals, convex conjugacy, Orlicz spaces

1991 Mathematics Subject Classification: 49K22, 52A41, 46B10

1. Introduction

In this article, we characterize the solutions to a general problem of minimization of an energy functional under "bad" linear constraints, see (3). This badness is a consequence of a lack of integrability of the moment function defining the constraint. As a consequence, the minimizers may exhibit a singular component. This characterization is obtained without assuming any constraint qualification for a general linear constraint. In [18], the author had already obtained the characterization of the minimizers under more integrable constraints.

Let us begin describing the minimization problem.

The energy functional. Let $(\Omega, \mathcal{A}, R)$ be a measure space where R is a nonnegative σ-finite measure. Let U be a vector space of measurable functions $u : \Omega \to \mathrm{I\!R}$ and U^* be the space of all (algebraic) linear forms on U. We denote $\langle \ell, u\rangle \in \mathrm{I\!R}$ the action of $\ell \in U^*$ on $u \in U$. Let $\gamma : \mathrm{I\!R} \to [0, \infty]$ be a nonnegative convex function such that

$$\gamma(0) = 0 \text{ and } \lim_{s\to\pm\infty} \gamma(s) = +\infty.$$

ISSN 0944-6532 / $ 2.50

The energy functional to be minimized is

$$\Phi^*(\ell) = \sup_{u \in U}\{\langle \ell, u\rangle - \int_\Omega \gamma(u(\omega))\, R(d\omega)\} \in [0,\infty], \ell \in U^* \tag{1}$$

It is the convex conjugate for the duality (U, U^*) of

$$\Phi(u) = \int_\Omega \gamma(u)\, dR \in [0,\infty], u \in U.$$

Let us note that, as $\gamma \geq 0$, $\Phi(u)$ is defined for every measurable function u (if $\mathcal{A}$ is rich enough).

The constraints. Let a vector space $\mathcal{X}$ be given, together with a function $\varphi : \Omega \to \mathcal{X}$. The linear constraints subject to which $\Phi^*(\ell)$ will be minimized have the following form: $\ell \in U^*$ is subject to "$\langle \varphi, \ell\rangle = x$" where $x \in \mathcal{X}$. For instance, a finite number K of moment constraints: $\langle \ell, \varphi_k\rangle = x_k$ with $\varphi_k \in U$, $x_k \in \mathbb{R}$, $1 \leq k \leq K$, is expressed by $\langle \varphi, \ell\rangle = x$ with $\varphi = (\varphi_k)_{k\leq K}$ and $x = (x_k)_{k\leq K} \in \mathcal{X} = \mathbb{R}^K$.

More generally, let us consider $(\mathcal{X}, \mathcal{Y})$ a couple of vector spaces in separating duality such that for each $y \in \mathcal{Y}$, the application $\langle y, \varphi\rangle : \omega \in \Omega \mapsto \langle y, \varphi(\omega)\rangle_{\mathcal{Y},\mathcal{X}} \in \mathbb{R}$ stands in U. The constraint $\langle \varphi, \ell\rangle = x$ signifies:

$$\langle\langle y, \varphi\rangle, \ell\rangle_{U,U^*} = \langle y, x\rangle_{\mathcal{Y},\mathcal{X}}, \forall y \in \mathcal{Y}. \tag{2}$$

The minimization problem. The general form of our minimization problem is

$$\text{minimize } \ell \in U^* \mapsto \Phi^*(\ell) \text{ subject to } \langle \varphi, \ell\rangle = x \tag{3}$$

where Φ^* is given at (1) and $\langle \varphi, \ell\rangle = x$ signifies (2). Let us note that we assumed

$$\langle y, \varphi\rangle \in U, \forall y \in \mathcal{Y} \tag{4}$$

to insure the expression (2) to be meaningful.

The aim of this paper is to give a characterization of the minimizers of (3) without constraint qualification, in situations where the constraint is "bad": that is when φ is not very integrable. This characterization is the object of our *main results:* Theorems 3.2 and 3.4 which are stated in Section 3.

When the constraints are integrable enough, one expects the minimizers to be measures which are absolutely continuous with respect to the reference measure R, see [18]. It will be proved in this paper, that the lack of integrability of "bad" constraints is responsible for the coming out of a singular component of the minimizers (which is not even a measure!).

In order to make precise the notion of bad constraint at Section 2.3, we are going to introduce at Section 2.1 an Orlicz space which is naturally associated with γ and we shall present at Section 2.2 a simple example of entropy maximization. This example will allow us to illustrate the notion of "bad constraint" and to motivate the study of (3) under a bad constraint.

To characterize the minimizers, it is usual to identify them as subgradients of a convex conjugate $\overline{\Phi}$ of the objective function Φ^*. Since we do not require any constraint qualification, to build $\overline{\Phi}$, we have to consider the richest duality involving U^*. This duality is (U^*, U^{**}) where U^{**} is the algebraic bidual space of U. Indeed, the geometric version of Hahn-Banach theorem insures that at any internal point of its effective domain, the convex function Φ^* has a nonempty algebraic subdifferential (in U^{**}, not requiring any a priori topological regularity on the subgradients). Therefore, one has to compute $\overline{\Phi}$ with respect to this "saturated" duality. Fortunately, it appears that the effective domain of Φ^* and its subdifferentials are included in topological dual spaces of Orlicz spaces naturally associated with the integrand γ. This will allow us to obtain an explicit expression of $\overline{\Phi}$. Our approach is described in details at Section 4.

About the literature. The problem (3) leads to solutions of some moment problems under the additional constraint: $\Phi^*(\ell) < \infty$, which implies in particular that ℓ is absolutely continuous with respect to R, if φ is integrable enough (in a sense to be made precise). Among others, let us refer to: [1-5], [8-10], [11], [14], [23], [26] and more recently [7], [18] and [19].

The special case where Φ^* corresponds to the *relative entropy* is of special interest and is extensively studied in the literature. For more details on the relative entropy, see the examples treated in Sections 2 and 6. It arises naturally in the *theory of large deviations*, see [12], which is a part of probability theory that is mostly connected with statistical physics and information theory. More precisely, the minimization problem (3) in this context is a consequence of the contraction principle applied to an extended Sanov's theorem to obtain the usual statement of Cramér's theorem (see [22], [21]). This is closely related to the work of I. Csiszár about the generalized I-projections. Indeed, it appears that the absolutely continuous part of the minimizers are precisely the generalized I-projections which were discovered by I. Csiszár ([9], [10]). Another example is given by the large deviations of renormalized Lévy processes (with very frequent and very small jumps) in the situation where the Lévy kernel only integrates some (but not every) exponential moments of the jumps (see [20]). Applying the contraction principle to the renormalized underlying Poisson measures to obtain the large deviation principle for the Lévy process, leads us to a minimization problem (3).

For an interesting connection between large deviations and energy minimization problems under good constraints, the reader may have a look at [14] and [6], among others, and [22] where bad constraints are considered.

Outline of the paper. The main results are stated in Section 3, they are Theorems 3.2 and 3.4. The method of their proofs is described in details at Section 4, and their proofs are completed in Section 5. The notion of bad constraint is defined and illustrated in Section 2. Simple illustrations of Theorem 3.4 are presented in Section 6. The proofs of preliminary technical results are collected in the Appendix. In particular, the first section of the Appendix is deserved to the statements of basic results on the duality of Orlicz spaces.

2. Good and bad constraints. An example

Before giving a simple example, let us state some notations and introduce a natural Orlicz space.

An Orlicz space. The natural domain of Φ is the set of all measurable functions u such that $\int_\Omega \gamma(u)\, dR < \infty$. In this paper, we shall not work with this natural domain, but with a smaller one which is the largest vector space contained in the cone with vertex 0 generated by the natural domain. It is the Orlicz space

$$L_{\gamma_o} := \{u : \Omega \to \mathbb{R}; \exists\lambda > 0, \int_\Omega \gamma_o(\lambda u)\, dR < \infty\}$$

where

$$\gamma_o(s) := \max(\gamma(s), \gamma(-s)), s \in \mathbb{R}$$

is a Young function, under our assumptions on γ. As the evaluation of $\Phi(u)$ is insensitive to R-almost everywhere equality, we identify R-almost everywhere equal functions. We assume that

$$U \subset L_{\gamma_o}. \tag{5}$$

One shows with (5) that the effective domain *dom* $\Phi^* = \{\ell \in U^*; \Phi^*(\ell) < \infty\}$ is a subset of the topological dual L'_{γ_o} of the Banach space L_{γ_o} endowed with the gauge norm: $\|u\|_{\gamma_o} := \inf\{a > 0; \int_\Omega \gamma_o(u/a)\, dR \le 1\}$. Because of (5), we obtain that Φ^* is $\sigma(L'_{\gamma_o}, L_{\gamma_o})$-lower semicontinuous. We shall only consider $\sigma(L'_{\gamma_o}, L_{\gamma_o})$-closed constraints since it is assumed that φ satisfies

$$\langle y, \varphi\rangle \in L_{\gamma_o}, \forall y \in \mathcal{Y} \tag{6}$$

which is (4) with (5).

A simple example. (Entropy maximization). In what follows we do not go into the details, such as the existence of the encountered integrals, and stay at a formal level. Let R be a bounded nonnegative measure on Ω. We search a probability measure which attains the infimum of

$$\text{minimize } Q \mapsto \int_\Omega \{\frac{dQ}{dR}\log\frac{dQ}{dR} - \frac{dQ}{dR} + 1\}\, dR \tag{P_1}$$

$$\text{subject to } Q \text{ is a probability measure, } Q \ll R, \int_\Omega \varphi dQ = x + \int_\Omega \varphi\, dR$$

where $\varphi : \Omega \to \mathbb{R}$ is measurable and $x \in \mathbb{R}$. Applying formally Lagrange multipliers method, one gets a solution the form of which is (if it exists): $Q_y(d\omega) = \exp(y\varphi(\omega) - \Lambda(y))R(d\omega)$ where $y \in \mathbb{R}$ satisfies $x = \int_\Omega \varphi\, d(Q_y - R)$ and $\Lambda(y) := \log \int_\Omega e^{y\varphi}\, dR$. But $\int_\Omega \varphi\, dQ_y = \Lambda'(y)$, hence when *dom* $\Lambda = (-\infty, y^*]$ with $\Lambda'(y^*) := x^* < \infty$ and $x + \int_\Omega \varphi\, dR > x^*$, (P_1) has no solution.

Let us analyse this in our framework. Identifying Q with $f = \frac{d(Q-R)}{dR}$, (P_1) becomes

$$\text{minimize } f \mapsto \int_\Omega \gamma^*(f)\, dR \text{ subject to } \int_\Omega f\, dR = 0 \text{ and } \int_\Omega \varphi f\, dR = x \tag{P_2}$$

where $\gamma^*(t) = (t+1)\log(t+1) - t$ is the convex conjugate of $\gamma(s) = e^s - s - 1$.The Young function γ_o is $\gamma_o(s) = e^{|s|} - |s| - 1$, $s \in \mathbb{R}$. With $U = L_{\gamma_o}$, as R is bounded, (4) is: $\int_\Omega e^{\alpha|\varphi|}\, dR < \infty$, for some $\alpha > 0$. The energy functional is $\Phi^*(\ell) = \sup_{u \in L_{\gamma_o}}\{\langle \ell, u\rangle - \Phi(u)\}$ with $\Phi(u) = \int_\Omega (e^u - u - 1)\, dR$. As a consequence of the results of the present paper (see Theorem 3.4), the problem

$$\text{minimize } \ell \in L^*_{\gamma_o} \mapsto \Phi^*(\ell) \text{ subject to } \langle \ell, 1\rangle = 0 \text{ and } \langle \ell, \varphi\rangle = x \tag{P_3}$$

has infinitely many solutions when $x+\int_\Omega \varphi\, dR > x^*$. These solutions are in *dom* $\Phi^* \subset L'_{\gamma_o}$ and L'_{γ_o} is decomposed (see Theorem 7.2) into $L'_{\gamma_o} = L_{\gamma_o^*} \oplus L^s_{\gamma_o}$. This means that each $\ell \in L'_{\gamma_o}$ can be uniquely decomposed as $\ell = \ell^a + \ell^s$ where $\ell^a = \frac{d\ell^a}{dR}\cdot R$, $\frac{d\ell^a}{dR} \in L_{\gamma_o^*}$ is the absolutely continuous component of ℓ and ℓ^s, which is called the singular component of ℓ, is not a measure (it isn't σ-additive, unless it is null). Moreover, it is shown in (12), that for all $\ell \in$ *dom* $\Phi^* \subset L'_{\gamma_o}$, $\Phi^*(\ell) = \int_\Omega \gamma^*(\frac{d\ell^a}{dR})\, dR + \sup\{\langle \ell^s, u\rangle; u, \Phi(u) < \infty\}$. This means that (P_3) is an extension of (P_2). The main point is that, if

$$\int_\Omega e^{\alpha|\varphi|}\, dR \qquad \int_\Omega e^{\beta|\varphi|}\, dR = \infty, \text{ for some } 0 < \alpha < \beta < \infty, \tag{7}$$

then (P_2) may not be attained while (P_3) is. Let us note that if

$$\int_\Omega e^{\lambda|\varphi|}\, dR < \infty, \forall \lambda > 0, \tag{8}$$

then (P_2) is attained and the problems (P_2) and (P_3) are equivalent. Indeed, any singular form ℓ^s is such that $\langle \ell^s, u\rangle = 0$, as soon as u satisfies (8) (see Proposition 7.3). In particular, this implies that, although ℓ^s may not be null, its mass $\langle |\ell^s|, 1\rangle = 0$.

For additional details about this example, see Section 6 below.

Good and bad constraints. Considering the previous example, it appears that not enough integrable constraints may generate minimizers with a singular component. Let us be more precise and introduce the following vector subspace of L_{γ_o}

$$M_{\gamma_o} := \{u : \Omega \to \mathbb{R}; \forall \lambda > 0, \int_\Omega \gamma_o(\lambda u)\, dR < \infty\}$$

Taking (7) and (8) into account, one says that the constraint function φ is good if

$$\langle y, \varphi\rangle \in M_{\gamma_o}, \forall y \in \mathcal{Y}, \tag{9}$$

and that it is bad if

$$\varphi \text{ satisfies (6) and } \exists y_o \in \mathcal{Y}, \langle y_o, \varphi\rangle \notin M_{\gamma_o}. \tag{10}$$

If γ_o doesn't satisfy the Δ_2-condition (see (33)), M_{γ_o} is a proper subset of L_{γ_o}. A bad constraint is *not very integrable*.

Let us notice that if (9) holds, one can choose $U = M_{\gamma_o}$ which yields $\Phi^*(\ell) = \int_\Omega \gamma^*(\frac{d\ell}{dR})\, dR$ if $\ell \in L_{\gamma_o^*}$ and $+\infty$ otherwise. This gives the following problem, with $f = \frac{d\ell}{dR}$:

$$\text{minimize } f \in L_{\gamma_o^*} \mapsto \int_\Omega \gamma^*(f)\, dR = \sup_{u \in M_{\gamma_o}} \{\langle f, u\rangle - \Phi(u)\} \text{ subject to } \int_\Omega f\varphi\, dR = x \quad (P_x)$$

where $\int_\Omega f\varphi\, dR = x$ means: $\int_\Omega f\langle y, \varphi\rangle\, dR = \langle y, x\rangle$, for all $y \in \mathcal{Y}$.

Only assuming (6), these integrals are still well defined as $f \in L_{\gamma_o^*}$ and $\langle y, \varphi\rangle \in L_{\gamma_o}$: the problem (P_x) is meaningful. The following problem

$$\text{minimize } \ell \in L^*_{\gamma_o} \mapsto \Phi^*(\ell) = \sup_{u \in L_{\gamma_o}} \{\langle f, u\rangle - \Phi(u)\} \text{ subject to } \langle \ell, \varphi\rangle = x \qquad (\overline{P}_x)$$

is an extension of (P_x). When (10) holds, it may happen that for some x, (P_x) is not attained. While $(\overline{P}_x)$ is always attained in $L'_{\gamma_o} \subset L^*_{\gamma_o}$.

3. The main results

We shall use basic results on the duality of Orlicz spaces. The interested reader may have a look at the Appendix 7.1 where such results are collected. If L is an Orlicz space, let us denote L^* its algebraic dual space, L' its topological dual space and L^s the space of all continuous singular forms on L.

Recall that the integrand $\gamma : \mathbb{R} \to [0, \infty]$ is a function which satisfies

$$\begin{aligned}&\gamma \text{ is convex, lower semicontinuous, } \gamma(\mathbb{R}) \subset [0, +\infty],\ 0 < \gamma(-a) + \gamma(a) < \infty \\ &\text{for some } a > 0 \text{ and } \gamma(0) = 0.\end{aligned} \tag{11}$$

The dual space L'_{γ_o} is decomposed as $L'_{\gamma_o} = L_{\gamma_o^*} \oplus L^s_{\gamma_o}$. Writing this decomposition: $\ell = \ell^a + \ell^s$, $\ell \in L'_{\gamma_o}$, $\ell^a \in L_{\gamma_o^*}$, $\ell^s \in L^s_{\gamma_o}$, the objective function Φ^* is given for any $\ell \in L^*_{\gamma_o}$, by

$$\Phi^*(\ell) = \begin{cases} \int_\Omega \gamma^* \left(\frac{d\ell^a}{dR}\right) dR + \sup\{\langle \ell^s, u\rangle \,;\, u \in U, \Phi(u) < \infty\} & \text{if } \ell \in L'_{\gamma_o} \\ +\infty & \text{otherwise.} \end{cases} \tag{12}$$

where γ^* is the convex conjugate of γ. This result is a direct consequence of ([17], Lemma 2.1) and Proposition 7.2. Note that Φ^* is not strictly convex when the singular contribution is nontrivial. The primal problem

$$\text{minimize } \ell \in L^*_{\gamma_o} \mapsto \Phi^*(\ell) \text{ subject to } \langle \ell, \varphi\rangle = x \tag{$\overline{P}_x$}$$

admits the following formal dual problem

$$\text{maximize } y \in \mathcal{Y} \mapsto \langle x, y\rangle - \int_\Omega \gamma(\langle y, \varphi\rangle)\, dR. \tag{D_x}$$

Let us denote its value

$$\Lambda^*(x) = \sup D_x = \sup_{y\in\mathcal{Y}}\{\langle x, y\rangle - \int_\Omega \gamma(\langle y, \varphi\rangle)\, dR\}, x \in \mathcal{X}$$

which is the convex conjugate for the duality $(\mathcal{X}, \mathcal{Y})$ of $\Lambda(y) := \int_\Omega \gamma(\langle y, \varphi\rangle)\, dR$, $y \in \mathcal{Y}$. We have proved in [19] that $(\overline{P}_x)$ has at least one solution if and only if $\Lambda^*(x) < \infty$.

We shall say that $x \in \mathcal{X}$ is a *subgradient constraint* if Λ^* is subdifferentiable at x for the algebraic duality $(\mathcal{X}, \mathcal{X}^*)$, that is if $\partial\Lambda^*(x) := \{\eta \in \mathcal{X}^* \,;\, \forall x' \in \mathcal{X}, \Lambda^*(x') \geq \Lambda^*(x) + \langle \eta, x' - x\rangle\}$ is not empty. Because of the geometric version of Hahn-Banach theorem, this holds in particular when x belongs to the geometric relative interior of the effective domain of Λ^*.

The case of an even integrand. To make our statements simpler, let us first assume that γ is an even function. This restriction will be removed later on. The characterization of the minimizers will be expressed in terms of dual parameters that are called *admissible dual parameters.* Here is their definition.

Definition 3.1. (γ is even). Let z_1 be a linear form on $\mathcal{X}$. One says that (z_1, ζ_2) is an <u>admissible dual parameter</u>, if the following properties are satisfied.

1. $\langle z_1, \varphi\rangle$ is a measurable function
2. $\int_\Omega \gamma(\langle z_1, \varphi\rangle)\, dR < \infty$ and $\int_\Omega \gamma^* \circ \gamma'(\langle z_1, \varphi\rangle)\, dR < \infty$
3. ζ_2 stands in $L_\gamma^{s\prime}$
4. for all $\varepsilon > 0$, $K \geq 1$, $f_k \in L_{\gamma^*}$, $\ell_k \in L_\gamma^s$ $k \leq K$, there exists $y \in \mathcal{Y}$ such that $\int_\Omega \gamma(\langle y, \varphi\rangle)\, dR < \infty$ and for all $k \leq K$, $|\int_\Omega(\langle z_1, \varphi\rangle - \langle y, \varphi\rangle) f_k\, dR| \leq \varepsilon$, $|\langle \zeta_2 - \langle y, \varphi\rangle, \ell_k\rangle| \leq \varepsilon$.

Remarks.

- As γ is even, we have $\gamma = \gamma_o$.
- In 3, $L_\gamma^{s\prime}$ is the topological dual space of L_γ^s endowed with the relative topology of the uniform dual topology on L'_γ.
- Property 4 means that $\langle z_1, \varphi\rangle \in L_\gamma$ and $\zeta_2 \in L_\gamma^{s\prime}$ can be approximated simultaneously by some $\langle y, \varphi\rangle$'s in the effective domain of Φ with respect to the topologies $\sigma(L_\gamma, L_{\gamma^*})$ and $\sigma(L_\gamma^{s\prime}, L_\gamma^s)$, respectively.

To be able to characterize the solutions of (P_x) at Theorem 3.2 below, we also need to introduce a normal cone $K(u, \zeta)$. Let us describe it now.

Consider the equivalence relation on $L_\gamma : u_1 \sim u_2 \Leftrightarrow u_1 - u_2 \in M_\gamma$. The quotient space is $N_\gamma := L_\gamma / M_\gamma$. Because of the orthogonality property of Proposition 7.3, up to some isomorphisms, L_γ^s is equal to N'_γ and N_γ is a subset of $L_\gamma^{s\prime}$.

Denote $\mathcal{D}_\gamma = \mathit{dom}\ \Phi = \{u \in L_\gamma; \int_\Omega \gamma(u)\, dR < \infty\}$ and $D_\gamma = \mathcal{D}_\gamma / M_\gamma \subset N_\gamma$. The geometric closure of $\mathcal{D}_\gamma$ is $\overline{\mathcal{D}}_\gamma$. As $\mathcal{D}_\gamma$ has a nonempty $\|\cdot\|_\gamma$-interior, $\overline{\mathcal{D}}_\gamma$ is also the $\|\cdot\|_\gamma$-closure of $\mathcal{D}_\gamma$. The $\sigma(L_\gamma^{s\prime}, L_\gamma^s)$-closure of D_γ is denoted $\overline{D}_\gamma^\sigma$.

For all $u \in L_\gamma$, $\zeta \in L_\gamma^{s\prime}$, let us define

$$K(u, \zeta) = \{\ell \in L'_\gamma; \langle \ell, h\rangle \leq 0; \forall h \in L_\gamma, u + h \in \mathcal{D}_\gamma, \zeta + \dot{h} \in \overline{D}_\gamma^\sigma\}$$

where $\dot{h}$ is the equivalence class of h. It is a $\sigma(L'_\gamma, L_\gamma)$-closed convex cone with vertex 0.

We are now ready to state our first main result. Let us denote $\mathcal{S}(x)$ the set of solutions to $(\overline{P}_x)$.

Theorem 3.2. *(γ is even). Let us assume that γ satisfies* (11) *and is even.*

The set of minimizers $\mathcal{S}(x)$ of $(\overline{P}_x)$ is nonempty if and only if $x \in \mathrm{dom}\,\Lambda^*$.

Let x be a subgradient constraint (in particular, this holds for any x in the geometric relative interior of $\mathrm{dom}\,\Lambda^*$). One can associate with it an admissible dual parameter (z_1, ζ_2) in the sense of Definition 3.1, such that

$$x = \int_\Omega \varphi\gamma'(\langle z_1, \varphi\rangle)\, dR + \langle \varphi, \tilde{\ell}\rangle, \text{ for some } \tilde{\ell} \in K(\langle z_1, \varphi\rangle, \zeta_2). \tag{13}$$

Moreover, the set of all minimizers of $(\overline{P}_x)$ is

$$\mathcal{S}(x) = \{\gamma'(\langle z_1, \varphi\rangle) \cdot R + \tilde{\ell}\,;\ \tilde{\ell} \in K(\langle z_1, \varphi\rangle, \zeta_2) \text{ such that (13) holds}\}. \tag{14}$$

Conversely, for any admissible dual parameter (z_1, ζ_2), x being defined by (13) *is in* $\mathrm{dom}\,\Lambda^*$. *If in addition x is subgradient, then $\mathcal{S}(x)$ is given by* (14).

Of course, (13) means that for all $y \in \mathcal{Y}$, $\langle x, y\rangle = \int_\Omega \langle y,\varphi\rangle \gamma'(\langle z_1,\varphi\rangle)\, dR + \langle\langle y,\varphi\rangle, \tilde{\ell}\rangle$, for some ...

Finitely many constraints. Let us have a look at the special case of finitely many constraints. This means that $\mathcal{Y} = \mathcal{X} = \mathbb{R}^d$ with $1 \le d < \infty$ and $\varphi = (\varphi_1, \dots, \varphi_d) : \Omega \to \mathbb{R}^d$ is measurable. It is supposed that $\varphi_i \in L_{\gamma_o}$, for all $1 \le i \le d$. A constraint is given by: $\langle \varphi_i, \ell\rangle = x_i \in \mathbb{R}$, $1 \le i \le d$. It will be proved (see the remark below Proposition 5.3) that provided that γ is even (to simplify), the dual parameter (z_1, ζ_2) of any subgradient constraint satisfies $\langle z_1, \varphi\rangle = \zeta_2$. More precisely,

$$\langle z_1, \varphi\rangle \in L_\gamma \quad \langle \zeta_2, \ell\rangle = \langle\langle z_1,\varphi\rangle, \ell\rangle, \forall \ell \in L^s_\gamma \tag{15}$$

Saturated constraints. In the opposite direction, one may consider the saturated constraints $\ell = x \in \mathcal{X} = \mathcal{L}^*_{\gamma_o}$, where $\mathcal{L}_{\gamma_o}$ is the function space corresponding to L_γ without R-a.e. everywhere equality. This corresponds to $\varphi : \omega \in \Omega \mapsto \delta_\omega \in \mathcal{X}$ and $\mathcal{Y} = \mathcal{L}_{\gamma_o}$, since $\langle y, \varphi(\omega)\rangle = y(\omega)$ for any ω and $\langle \ell, \varphi\rangle = \ell$ for any $y \in \mathcal{L}_{\gamma_o}$, $\ell \in \mathcal{L}^*_{\gamma_o}$. To simplify, let us suppose that γ is even. The set of admissible constraints is the domain of Φ^* : ℓ is admissible if and only if $\ell \in L'_\gamma$, $I_{\gamma^*}(\frac{d\ell^a}{dR}) < \infty$ and $\delta^*(\ell^s \mid \mathit{dom}\ \Phi) < \infty$. In this situation, (z_1, ζ_2) is an admissible dual parameter in the sense of Definition 3.1 if and only if $z_1 \in L_\gamma$, $\int_\Omega \gamma(z_1)\, dR < \infty$, $\int_\Omega \gamma^* \circ \gamma'(z_1)\, dR < \infty$ and $\zeta_2 \in \overline{D}^\sigma_\gamma$. As $\ell^a = \gamma'(z_1) \cdot R$ and $\ell^s \in K(z_1, \zeta_2)$ may be chosen indepently to make $\ell = \ell^a + \ell^s$ admissible, the corresponding dual parameters z_1 and ζ_2 are not linked to each other. For a similar behavior, see the end of Section 6.

A more general integrand. When γ is not supposed to be even, the situation is technically more delicate to handle. One has to separate the positive and negative contributions of the linear forms (see Propositions 4.4 and 4.5 below for more details). We are going to state at Theorem 3.4 below a result, analogous to Theorem 3.2, in the case where γ satisfies the following hypothesis. Let us denote

$$\gamma_+(s) = \gamma(|s|) \qquad \gamma_-(s) = \gamma(-|s|), s \in \mathbb{R}$$

and in addition to (11), let us assume that

$$[\gamma^*_+ \text{ is } \Delta_2 \quad \gamma_+ \text{ is steep}] \quad [\gamma^*_-(\kappa_-) < \infty \quad \gamma_- \text{ is } \Delta_2] \tag{16}$$

where

$$\kappa_- = \lim_{s \to -\infty} \gamma(s)/|s| > 0.$$

As a special important example, we think of the relative entropy (see Section 6) which corresponds to $\gamma^*(t) = (t+1)\log(t+1) - t$ and $\gamma(s) = e^s - s - 1$, as already used at Section 2.

Note that $\gamma^*_-(\kappa_-) = \gamma^*(-\kappa_-)$ and that if R is a bounded measure and $\gamma^*_-(\kappa_-) < \infty$, then γ_- is Δ_2.

We have proved in [18] that in the case where $\gamma^*_-(\kappa_-) < \infty$, *infinite force fields* may be part of the admissible dual parameters attached to the negative component of the minimizers of $(\overline{P}_x)$. On the contrary, as γ^*_+ is assumed to be Δ_2, the representation of the positive part of the minimizers doesn't require any infinite force field. But, if γ_+ is not Δ_2, the positive part of the minimizers may admit a singular component.

To be able to state the characterization of $\mathcal{S}(x)$ under the assumption (16), we have to recall the definition of a force field $z_1 + \infty(n)$.

Let $\mathcal{J}$ be a totally ordered countable index set which admits a smaller element: $\flat$. We consider a family $(n) = (n^j)_{j\in\mathcal{J}}$ of linear forms on $\mathcal{X}$ such that for each j, $\langle n^j, \varphi\rangle$ is measurable. For any $j \in \mathcal{J}$, let us denote $\mathcal{T}^j_+ = \{\langle n^j, \varphi\rangle > 0\} \bigcap \cap_{i<j}\{\langle n^i, \varphi\rangle = 0\}$ and $\mathcal{T}^j_- = \{\langle n^j, \varphi\rangle < 0\} \bigcap \cap_{i<j}\{\langle n^i, \varphi\rangle = 0\}$ with the convention: $\bigcap_{i<\flat}\{\langle n^i, \varphi\rangle = 0\} = \mathcal{X}$, so that $\mathcal{T}^\flat_+ = \{n^\flat > 0\}$ and $\mathcal{T}^\flat_- = \{n^\flat < 0\}$. We define

$$\mathcal{S} = \bigcap_{j\in\mathcal{J}} \{\langle n^j, \varphi\rangle = 0\}, \quad \mathcal{T}_+ = \bigcup_{j\in\mathcal{J}} \mathcal{T}^j_+ \qquad \mathcal{T}_- = \bigcup_{j\in\mathcal{J}} \mathcal{T}^j_- \tag{17}$$

Up to a R-negligible set, $\mathcal{S}, \mathcal{T}_+$ and $\mathcal{T}_-$ form a measurable partition of Ω.

Let us introduce a notation for the force fields. Let z_1 be a measurable linear form on $\mathcal{X}$ and $(n) = (\langle n^j, \varphi\rangle)_{j\in\mathcal{J}}$ as above. We define the application $\langle z_1 + \infty(n), \varphi\rangle : \Omega \to [-\infty, +\infty]$, for any $\omega \in \Omega$, by

$$\langle z_1 + \infty(n), \varphi(\omega)\rangle = \begin{cases} +\infty & \text{if } \omega \in \mathcal{T}_+ \\ -\infty & \text{if } \omega \in \mathcal{T}_- \\ \langle z_1, x\rangle & \text{if } \omega \in \mathcal{S}. \end{cases}$$

It is a measurable application. If $(n) = 0$, $z_1 + \infty(n) = z_1$ has no infinite value. We are now ready to state the definition of an admissible dual parameter.

Definition 3.3. (γ satisfies (16)). Let $z_1 + \infty(n)$ be a force field. Its infinite component: $\infty(n)$, determines the measurable subsets $\mathcal{S}$, $\mathcal{T}_+$ and $\mathcal{T}_-$, see (17). Its finite component z_1 is a linear form on $\mathcal{X}$.

Under assumption (16), $(z_1 + \infty(n), \zeta_2)$ is said to be an admissible dual parameter, if the following properties are satisfied.

1. $\langle z_1, \varphi\rangle$ is a measurable function whose support is included in $\mathcal{S}$
2. $\int_{\mathcal{S}} \gamma(\langle z_1, \varphi\rangle)\, dR < \infty$ and $\int_{\mathcal{S}} \gamma^* \circ \gamma'(\langle z_1, \varphi\rangle)\, dR < \infty$
3. for any $j \in \mathcal{J}$, $\int_{\cap_{i<j}\{\langle n^i,\varphi\rangle=0\}} \langle n^j, \varphi\rangle_-\, dR < \infty$
4. $R(\mathcal{T}_-) < \infty$ and $R(\mathcal{T}_+) = 0$
5. $\zeta_2 \in L^{s\prime}_{\gamma_+}$ and $\zeta_2 \geq 0$
6. for all $\varepsilon > 0$, $K \geq 1$, $f_k \in L_{\gamma^*_+}$, $g_k \in L_{\gamma^*_-}$, $\ell_k \in L^s_{\gamma_+}$, $k \leq K$, there exists $y \in \mathcal{Y}$ such that $\int_\Omega \gamma(\langle y, \varphi\rangle)\, dR < \infty$ and for all $k \leq K$:
   $|\int_{\mathcal{S}}(\langle z_1, \varphi\rangle_+ - \langle y, \varphi\rangle_+) f_k\, dR| + |\int_{\mathcal{S}}(\langle z_1, \varphi\rangle_- - \langle y, \varphi\rangle_-) g_k\, dR| \leq \varepsilon$
   $|\langle \zeta_2 - \langle y, \varphi\rangle_+, \ell_k\rangle| \leq \varepsilon$

In the above definition, we have used the usual convention: $f_+ = \max(f, 0)$ and $f_- = \max(-f, 0)$. Beware of the notational conflict with γ_+ and γ_-. We are now ready to state our main result.

Theorem 3.4. *Let us assume that γ satisfies* (11) *and* (16)*.*

The set $\mathcal{S}(x)$ is nonempty if and only if x is in dom Λ^*.

Let x be any such constraint. One can associate with it an admissible dual parameter

$(z_1+\infty(n),\zeta_2)$ *in the sense of Definition 3.3, such that*

$$x=\int_{\mathcal{S}}\varphi\gamma'(\langle z_1,\varphi\rangle)\,dR-\kappa_-\int_{\mathcal{T}_-}\varphi\,dR+\langle\varphi,\tilde\ell\rangle,\ \textit{for some } \tilde\ell\in K(\mathbb{1}_{\mathcal{S}}\langle z_1,\varphi\rangle_+,\zeta_2). \tag{18}$$

The set of minimizers of $(\overline{P}_x)$ *is*

$$\mathcal{S}(x)=\Big\{[\gamma'(\langle z_1,\varphi\rangle)\mathbb{1}_{\mathcal{S}}-\kappa_-\mathbb{1}_{\mathcal{T}_-}]\cdot R+\tilde\ell\,;\,\tilde\ell\in K(\mathbb{1}_{\mathcal{S}}\langle z_1,\varphi\rangle_+,\zeta_2)\ \textit{such that (18) holds}\Big\} \tag{19}$$

Moreover, $\mathcal{S}(x)$ *is a* $\sigma(L'_{\gamma_o},L_{\gamma_o})$*-compact convex subset of* L'_{γ_o}*. All the elements of* $\mathcal{S}(x)$ *share the same absolutely continuous part:* $[\gamma'(\langle z_1,\varphi\rangle)\mathbb{1}_{\mathcal{S}}-\kappa_-\mathbb{1}_{\mathcal{T}_-}]\cdot R$ *and have a nonnegative singular part:* $\tilde\ell$*, in* $L^s_{\gamma+}$.

Furthermore, if x *is a subgradient constraint (in particular if* x *is in the relative geometric interior of* dom Λ^**), no infinite force field enters its dual representation:* $\mathcal{S}=\Omega$ *and* $\mathcal{T}_-=\emptyset$ *up to R-negligible sets.*

Conversely, for any admissible dual parameter $(z_1+\infty(n),\zeta_2)$*,* x *being defined by* (18)*, we have* $x\in$ *dom* Λ^* *and* $\mathcal{S}(x)$ *is given by* (19).

The cones. We give below a description of the cone $K(u,\zeta)$ that arises in the dual description of the minimizers at Theorems 3.2 and 3.4.

Let us denote for any $u\in\overline{\mathcal{D}}_\gamma$ and $\zeta\in\overline{D}^\sigma_\gamma$:

$$K_1(u)=\{\ell\in L'_\gamma\,;\,\langle\ell,h\rangle\le 0,\forall h\in L_\gamma, u+h\in \textit{dom } I_\gamma\}$$
$$K_2(\zeta)=\{\ell\in L'_\gamma\,;\,\langle\ell,h\rangle\le 0,\forall h\in L_\gamma, \zeta+\dot h\in\overline{D}^\sigma_\gamma\}$$

These are $\sigma(L'_\gamma,L_\gamma)$-closed convex cones with vertex 0. $K_1(u)$ and $K_2(\zeta)$ are respectively the polar cones (for the duality (L'_γ,L_γ)) of the cones: $\{h\in L_\gamma\,;\,\exists t>0,u+th\in \textit{dom } I_\gamma\}$ and $\{h\in L_\gamma\,;\,\exists t>0,\zeta+t\dot h\in\overline{D}^\sigma_\gamma\}$. More,

$$K(u,\zeta)=\mathrm{cl}(K_1(u)+K_2(\zeta))$$

where the closure is in $\sigma(L'_\gamma,L_\gamma)$. For a proof of this equality, see ([Ro2], Corollary 16.4.2).

The following proposition specifies the "supports" of the elements of $K(u,\zeta)$.

Proposition 3.5. *Let* $u\in\overline{\mathcal{D}}_\gamma$ *be such that* $u\ge 0$*. Then,* $\ell\in K_1(u)$ *if and only if* $\ell\ge 0$ *and for all* $h\in L_\gamma$*,* $(h\ge 0$ *and* $\exists t>0,\ u+th\in\overline{\mathcal{D}}_\gamma)\Rightarrow\langle\ell,h\rangle=0$.

Let $\zeta\in\overline{D}^\sigma_\gamma$ *be such that* $\zeta\ge 0$*. Then,* $\ell\in K_2(\zeta)$ *if and only if* $\ell\ge 0$ *and for all* $h\in N_\gamma$*,* $(h\ge 0$ *and* $\exists t>0,\ \zeta+th\in\overline{D}^\sigma_\gamma)\Rightarrow\langle\ell,h\rangle=0$.

Let us assume that the function γ^* *is stricly convex at infinity, or equivalently that* γ *is steep. Then, for any* $u\in\overline{\mathcal{D}}_\gamma$ *and* $\zeta\in\overline{D}^\sigma_\gamma$*,* $K(u,\zeta)$*,* $K_1(u)$ *and* $K_2(\zeta)$ *are subsets of* L^s_γ *: their elements are singular.*

4. The method of proof

The approach is classical: If ℓ is such that $\partial\Phi^*(\ell)$ is nonempty, we pick some ζ in $\partial\Phi^*(\ell)$ to obtain $\ell\in\partial\overline{\Phi}(\zeta)$, where $\overline{\Phi}$ is a convex conjugate of Φ^*. Therefore, ζ is our dual parameter and $\ell \in \partial\overline{\Phi}(\zeta)$ is a representation of ℓ. Now, if ℓ is a minimizer of Φ^* under a linear constraint, ζ must have a special form connected with the constraint. With the formal Lagrange multipliers method in mind, one expects that ζ must be something like a linear combination of the components of the constraint function φ. As a prototype, let us think of $\zeta = \langle y, \varphi\rangle$ for some $y \in \mathcal{Y}$. Of course, if $\mathcal{Y}$ separates $\mathcal{X}$, but is a "small" space, one cannot expect that the $\langle y, \varphi\rangle$'s will dually describe all the minimizers (for all admissible constraints x.) Nevertheless, one can guess that the $\langle y, \varphi\rangle$'s will be dense (in some sense) in the set of all admissible dual parameters. This means that the dual problem is usually not attained, but that one can expect that the maximizing (generalized) sequences of y's are convergent in some sense.

But the implementation is unusual: Since we do not assume any constraint qualification and we do not seek approximations of the dual parameters but exact (nonasymptotic) dual representations of the minimizers, our strategy is to consider the algebraic duality (U^*, U^{**}), so that the ζ's are attained in U^{**} (without approximating them by maximizing sequences of the dual problem) for all the "internal" constraints, and even other ones (the so-called subgradient constraints). Of course, in such a program, the difficult task is to compute the convex conjugate $\overline{\Phi}$ of Φ^* for the duality (U^*, U^{**}).

An abstract solution. To describe the linear constraints, let us introduce the following vector space of "prototypes" of the dual parameters:

$$V = \{\langle y, \varphi\rangle; y \in \mathcal{Y}\}.$$

Because of assumption (4), we have $V \subset U$. The algebraic dual and bidual spaces of V are denoted by V^* and V^{**}.

Let us consider the relations between the vector spaces. We define the equivalence relation on $U^* : \ell \sim \ell'$ for any $\ell, \ell' \in U^*$ if and only if $\ell(u) = \ell'(u), \forall u \in V$. In other words: $\ell \sim \ell' \iff \ell_{|V} = \ell'_{|V}$. We identify V^* with the factor space:

$$V^* = U^*/\sim$$

and $\dot{\ell} \in V^*$ stands for the equivalence class of $\ell \in U^*$. Therefore, one can identify V^{**} with a vector subspace of $U^{**} : V^{**} \subset U^{**}$ as follows. For any $\zeta \in U^{**}$,

$$\zeta \in V^{**} \iff \Big(\forall \ell, \ell' \in U^*, \ell \sim \ell' \implies \langle \zeta, \ell - \ell'\rangle = 0\Big).$$

We are going to solve the minimization problem:

$$\text{minimize } \Phi^*(\ell) \text{ subject to } \ell \in \alpha \tag{20}$$

with $\alpha \in V^*$. Since α is an affine subspace of U^*, $\ell \in \alpha$ is a linear constraint. In order to solve it, let us introduce the restriction Ψ of Φ to $V \subset U$:

$$\Psi : u \in V \mapsto \Phi(u) \in [0, \infty].$$

Its conjugate is

$$\Psi^* : \alpha \in V^* \mapsto \sup_{u \in V}\{\langle \alpha, u\rangle - \Phi(u)\} \in [0, \infty].$$

The convex biconjugate of Φ is

$$\overline{\Phi} : \zeta \in U^{**} \mapsto \sup_{\ell \in U^*} \{\langle \zeta, \ell \rangle - \Phi^*(\ell)\} \in [0, \infty].$$

Let us introduce the following algebraic subdifferentials:

$$\begin{aligned} \partial\Psi^*(\alpha) &= \{\zeta \in V^{**}\, ;\, \Psi^*(\alpha+\eta) \geq \Psi^*(\alpha) + \langle \zeta, \eta \rangle, \forall \eta \in V^*\}, \alpha \in dom\ \Psi^*. \\ \partial\overline{\Phi}(\zeta) &= \{\ell \in U^*\, ;\, \overline{\Phi}(\zeta+u) \geq \overline{\Phi}(\zeta) + \langle \ell, u \rangle, \forall u \in U\}, \zeta \in dom\ \overline{\Phi} \end{aligned}$$

Notice that in the definition of $\partial\overline{\Phi}(\zeta)$, the increment u stands in U, rather than in U^{**}.

Theorem 4.1. *Suppose that the following dual equality holds:*

$$\Psi^*(\alpha) = \inf\{\Phi^*(\ell); \ell \in \alpha\} \text{ for all } \alpha \in V^*. \tag{21}$$

Then, for any $\alpha \in V^*$ such that $\Psi^*(\alpha) < \infty$ and $\partial\Psi^*(\alpha) \neq \emptyset$, we have:

$$\mathcal{S}(\alpha) = \alpha \cap \partial\overline{\Phi}(\zeta), \forall \zeta \in \partial\Psi^*(\alpha),$$

where $\mathcal{S}(\alpha)$ is the set of solutions to (20).

Remarks.

- As $V^{**} \subset U^{**}$, $\overline{\Phi}(\zeta)$ is meaningful for every $\zeta \in V^{**}$ and in particular for $\zeta \in \partial\Psi^*(\alpha) \subset V^{**}$.
- The main interest of this result is that any α in *ridom* Ψ^* satisfies $\partial\Psi^*(\alpha) \neq \emptyset$. This is a consequence of the geometric version of Hahn-Banach's theorem. It is the reason why algebraic duality is considered here.
- Theorem 4.1 is general: U may be any vector space, Φ may be any $[0, \infty]$-valued convex function on U and V may be any subspace of U.

The main statement of our abstract result stated at Theorem 4.2 below, is a consequence of this result.

Theorem 4.2. *For any $x \in \mathcal{X}$, $\mathcal{S}(x)$ is nonempty if and only if x stands in dom Λ^*.*

The set of solutions $\mathcal{S}(x)$ is a convex $\sigma(L'_{\gamma_o}, L_{\gamma_o})$-compact subset of L'_{γ_o}.

For any subgradient constraint $x \in \mathcal{X}$, we have

$$\mathcal{S}(x) = \alpha_x \cap \partial\overline{\Phi}(\zeta), \forall \zeta \in \partial\Psi^*(\alpha_x),$$

where for any $x \in \mathcal{X}$, α_x in V^* is defined by: $\langle \alpha_x, \langle y, \varphi \rangle \rangle_{V^*,V} = \langle x, y \rangle_{\mathcal{X},\mathcal{Y}}$ for all $y \in \mathcal{Y}$.

The biconjugate of Ψ. As $\zeta \in \partial\Psi^*(\alpha)$, and $\partial\Psi^*(\alpha) \subset dom\ \overline{\Psi}$ where $\overline{\Psi}$ is the convex conjugare of Ψ^* for the duality (V^*, V^{**}), the dual parameters of interest stand in the effective domain of $\overline{\Psi}$. In ([17], Proposition 3.3), it is proved that, provided that the dual equality (21) holds, for any $\zeta \in V^{**}$, we have

$$\overline{\Psi}(\zeta) = \begin{cases} \overline{\Phi}(\zeta) & \text{if } \zeta \in \overline{V} \\ +\infty & \text{otherwise} \end{cases} \tag{22}$$

where $\overline{V}$ is the $\sigma(U^{**},U^*)$-closure of V.

The expression of $\overline{\Phi}$ when γ is even. In the situation where γ is even, we have shown in ([17], Lemma 2.1) that the effective domain of $\overline{\Phi}$ is a subset of the topological bidual L''_γ of the Orlicz space L_γ. This will allow us to invoke convex duality results for integral functionals on Orlicz spaces (see Proposition 7.2) to obtain the expression of $\overline{\Phi}$ stated at Proposition 4.3 below.

The appearence of Orlicz space structures is not an artefact. The topological structures of L_γ, L'_γ, and L''_γ do not carry more information than the geometrical structures of U, U^* and U^{**}, that are needed to perform the convex duality giving rise to Φ^* and $\overline{\Phi}$. In fact, it is proved in [17] that, provided that γ is even, *dom* $\Phi^* \subset L'_\gamma$, *dom* $\overline{\Phi} \subset L''_\gamma$ and the geometric interiors of these domains are their respective Orlicz-topological interiors. This means that the relevent algebraic linear forms share some regularity property: they are continuous with respect to some Orlicz topologies.

For any ζ in the algebraic dual space L'^*_γ of L'_γ, let us denote the restrictions

$$\zeta_1 = \zeta_{|L_{\gamma^*}} \in L^*_{\gamma^*} \qquad \zeta_2 = \zeta_{|L^s_\gamma} \in L^{s\,*}_\gamma. \tag{23}$$

If ζ_1 is continuous on L_{γ^*}, its decomposition into absolutely continuous and singular components is $\zeta_1 = \zeta_1^a + \zeta_1^s$, with $\zeta_1^a \in L_\gamma$ and $\zeta_1^s \in L^s_{\gamma^*}$.

Proposition 4.3. *Let us assume that γ is even. Then for any $\zeta \in U^{**}$,*

$$\overline{\Phi}(\zeta) = \begin{cases} \Phi(\frac{d\zeta_1^a}{dR}) + \sup\{\langle \zeta_1^s, f\rangle; f, \int_\Omega \gamma^*(f)\,dR < \infty\} + \delta(\zeta_2 \mid \overline{D}^\sigma_\gamma) & \textit{if } \zeta \in L''_\gamma \\ \infty & \textit{otherwise} \end{cases}$$

Moreover, $\overline{D}^\sigma_\gamma$ is norm bounded in $L^s_\gamma{}'$ and $\delta^*(\ell \mid \overline{D}_\gamma) < \infty$ for all ℓ in L^s_γ, where $\overline{D}_\gamma$ is the $\|\cdot\|_\gamma$-closure of D_γ in N_γ.

How to compute Φ^* when γ is not even. In this situation, let us consider the Young functions

$$\gamma_+(s) = \gamma(|s|) \qquad \gamma_-(s) = \gamma(-|s|), s \in \mathbb{R},$$

and the associated integral functionals Φ_+ and Φ_- defined on $U = L_{\gamma_o} = L_{\gamma_+} \cap L_{\gamma_-}$ where $\gamma_o = \max(\gamma_+, \gamma_-)$ as in Section 2.

Let us recall that there is a natural order on the algebraic dual space E^* of a Riesz vector space E which is defined by: $e^* \le f^*$ if and only if $\langle e^*, e\rangle \le \langle f^*, e\rangle$ for any $e \in E$ with $e \ge 0$. A linear form $e^* \in E^*$ is said to be *relatively bounded* if for any $f \in E$, $f \ge 0$, we have $\sup_{e\,;\,|e|\le f} |\langle e^*, e\rangle| < +\infty$. Although E^* may not be a Riesz space in general, the vector space: E^b, of all the relatively bounded linear forms on E is always a Riesz space. In particular, the elements of E^b admit a decomposition in nonnegative and nonpositive parts $e^* = e^*_+ - e^*_-$.

We have proved in ([17], Proposition 4.4) the following result.

Proposition 4.4. *For any $\zeta \in U^{**}$, $\overline{\Phi}(\zeta) = \overline{\Phi_+}(\zeta_+) + \overline{\Phi_-}(\zeta_-)$ if $\zeta \in L'^b_{\gamma_o}$ and $\overline{\Phi}(\zeta) = \infty$ otherwise, where $L'^b_{\gamma_o} \subset U^{**}$ stands for the subspace of relatively bounded linear forms on L'_{γ_o} : $\zeta \in U^{**}$ belongs to $L'^b_{\gamma_o}$ if its restriction to $L'_{\gamma_o} \subset U^*$ is relatively bounded.*

Remarks.

- Since the domains of Φ^* ($dom\ \Phi^* \subset L'_{\gamma_o}$) and $\overline{\Phi}$ consist of relatively bounded forms, one can consider their decompositions into nonnegative and nonpositive parts.
- As $\overline{\Phi}(\zeta)$ only depends on the restriction of $\zeta \in U^{**}$ to $L'_{\gamma_o} \subset U^*$ (see [17], Lemma 4.6), the above expression of $\overline{\Phi}(\zeta)$ is unambiguous.

Because of Theorem 4.2, one needs to compute the subdifferential $\partial\overline{\Phi}(\zeta)$.

Proposition 4.5. *For any $\ell \in \partial\overline{\Phi}(\zeta)$ with $\zeta \in dom\ \overline{\Phi}$, we have $\ell \in L^b_{\gamma_o}$ and $\zeta \in L'^b_{\gamma_o}$. In particular, ℓ and ζ admit positive and negative components.*

Moreover, $\ell_+ \in \partial\overline{\Phi_+}(\zeta_+)$ $\ell_- \in \partial\overline{\Phi_-}(\zeta_-)$.

As γ_+ and γ_- are even functions, Propositions 4.3, 4.4 and 4.5 lead us to the complete expression of $\overline{\Phi}$ and of its subgradients.

Remarks. For the decomposition $\overline{\Phi}(\zeta) = \overline{\Phi_+}(\zeta_+) + \overline{\Phi_-}(\zeta_-)$ at Proposition 4.4 to hold, it is necessary that $\gamma(0) = 0$. And for the decomposition $\ell_+ \in \partial\overline{\Phi_+}(\zeta_+)$ $\ell_- \in \partial\overline{\Phi_-}(\zeta_-)$ at Proposition 4.5 to hold, it is necessary that in addition γ is nondecreasing on $[0, \infty)$ and nonincreasing on $(-\infty, 0]$. This is the reason why we have assumed that γ is normalized as in (11).

5. The proofs

In this section, we give the proofs of Theorems 3.2, 3.4 and Proposition 3.5.

We shall use the notations $\zeta_1, \zeta_1^a, \zeta_1^s$ and ζ_2 introduced at (23). Let us introduce the notations

$$I_\gamma(u) = \int_\Omega \gamma(u)\, dR (= \Phi(u)) \qquad I_{\gamma^*} = \int_\Omega \gamma^*(f)\, dR$$

for the integral functionals associated with γ and γ^*.

Proof of Theorem 3.2. Theorem 3.2 is a direct consequence of Theorem 4.2 and Propositions 5.2 and 5.3 below.

In this section, it is assumed that γ is even.

We shall need Lemma 5.1 to prove Proposition 5.2.

For any $\ell \in L^*_\gamma$, ℓ belongs to $\partial\overline{\Phi}(\zeta)$ if and only if for all $u \in L_\gamma$, $\overline{\Phi}(\zeta + u) - \overline{\Phi}(\zeta) \geq \langle \ell, u\rangle$. By Proposition 4.3, for all $u \in L_\gamma$, $\zeta \in dom\ \overline{\Phi}$, $\overline{\Phi}(\zeta + u) - \overline{\Phi}(\zeta) = I_\gamma(\zeta_1^a + u) - I_\gamma(\zeta_1^a) + \delta^*((\zeta_1 + u)^s \mid dom\ I_{\gamma^*}) - \delta^*(\zeta_1^s \mid dom\ I_{\gamma^*}) + \delta(\zeta_2 + u \mid \overline{D}^\sigma_\gamma) - \delta(\zeta_2 \mid \overline{D}^\sigma_\gamma)$. But $(\zeta_1 + u)^s = \zeta_1^s$ for all $u \in L_\gamma$. Hence, the difference of δ^* vanishes. As $\overline{\Phi}(\zeta) < \infty$, we have: $\delta(\zeta_2 \mid \overline{D}^\sigma_\gamma) = 0$. It comes out that for all $u \in L_\gamma$, $\zeta \in dom\ \overline{\Phi}$,

$$\overline{\Phi}(\zeta + u) - \overline{\Phi}(\zeta) = I_\gamma(\zeta_1^a + u) - I_\gamma(\zeta_1^a) + \delta(\zeta_2 + u \mid \overline{D}^\sigma_\gamma). \tag{24}$$

Lemma 5.1. *Let $\zeta \in dom\ \overline{\Phi}$ be such that $\partial\overline{\Phi}(\zeta)$ is nonempty.*

(a) *We have $0 \leq \int_\Omega \zeta_1^a \gamma'(\zeta_1^a)\, dR < \infty$ and $0 \leq \int_\Omega \gamma^* \circ \gamma'(\zeta_1^a)\, dR < \infty$. In particular, $\int_\Omega u\gamma'(\zeta_1^a)\, dR$ is meaningful for any $u \in L_\gamma$.*

(b) *For any $u \in L_\gamma$ such that $\zeta_1^a + t_o u \in dom\ I_\gamma$ for some $t_o > 0$, we have:* $\lim_{t \downarrow 0} \frac{1}{t}[I_\gamma(\zeta_1^a + tu) - I_\gamma(\zeta_1^a)] = \int_\Omega u\gamma'(\zeta_1^a)\, dR$.

Proof. (a) is ([18], Lemma 3.2), noting that its proof still holds when γ is not finite and the restrictive assumption (3.1) in [18] is useless for this result.

(b) Let us introduce $g_t = [\gamma(\zeta_1^a + tu) - \gamma(\zeta_1^a) - tu\gamma'(\zeta_1^a)]/t$, $t > 0$. As γ is a convex function, $t \mapsto g_t$ is nondecreasing and $\lim_{t \downarrow 0} g_t = 0$. The result follows from (a) and the dominated convergence theorem. $\square$

With this lemma in hand, we are going to prove the following result.

Proposition 5.2. *Let $\zeta \in dom\ \overline{\Phi}$ be such that $\partial\overline{\Phi}(\zeta)$ is nonempty. Then, $\partial\overline{\Phi}(\zeta) = \gamma'(\zeta_1^a) \cdot R + K(\zeta_1^a, \zeta_2)$.*

Proof.
1. Let us show that $\partial\overline{\Phi}(\zeta) \subset \gamma'(\zeta_1^a) \cdot R + K(\zeta)$. Let $\ell \in \partial\overline{\Phi}(\zeta)$. We denote $\ell' = \ell - \gamma'(\zeta_1^a) \cdot R \in L'_\gamma$. Then, by (24), for all $t > 0$ and $u \in L_\gamma$: $\frac{1}{t}[\overline{\Phi}(\zeta + tu) - \overline{\Phi}(\zeta)] = \frac{1}{t}[I_\gamma(\zeta_1^a + tu) - I_\gamma(\zeta_1^a)] + \frac{1}{t}\delta(\zeta_2 + tu \mid \overline{D}_\gamma^\sigma) \geq \frac{1}{t}\langle \ell, tu \rangle = \langle \ell, u \rangle = \int_\Omega u\gamma'(\zeta_1^a)\, dR + \langle \ell', u \rangle$. But for any $u \in L_\gamma$ such that $\zeta_1^a + t_o u \in dom\ I_\gamma$ for some $t_o > 0$, we have by Lemma 5.1: $\lim_{t \to 0} \frac{1}{t}[I_\gamma(\zeta_1^a + tu) - I_\gamma(\zeta_1^a)] = \int_\Omega u\gamma'(\zeta_1^a)\, dR$. Therefore, $\liminf_{t \to 0} \frac{1}{t}\delta(\zeta_2 + tu \mid \overline{D}_\gamma^\sigma) \geq \langle \ell', u \rangle$. In particular, ℓ' satisfies

$$\langle \ell', u \rangle \leq 0 \text{ for all } u \in L_\gamma \text{ such that } \zeta_1^a + u \in dom\ I_\gamma \text{ and } \zeta_2 + \dot{u} \in \overline{D}_\gamma^\sigma. \tag{25}$$

2. Let us prove the converse inclusion: $\partial\overline{\Phi}(\zeta) \supset \gamma'(\zeta_1^a) \cdot R + K(\zeta)$. Let $\ell = \gamma'(\zeta_1^a) \cdot R + \ell'$ where ℓ' satisfies (25). By (24), for any $u \in L_\gamma$ such that $\zeta_1^a + u \in dom\ I_\gamma$ and $\zeta_2 + \dot{u} \in \overline{D}_\gamma^\sigma$, we have: $\overline{\Phi}(\zeta + u) - \overline{\Phi}(\zeta) \geq \int_\Omega u\gamma'(\zeta_1^a)\, dR \geq \int_\Omega u\gamma'(\zeta_1^a)\, dR + \langle \ell', u \rangle = \langle \ell, u \rangle$. The first inequality is a consequence of Lemma 5.1, the second one follows from $\langle \ell', u \rangle \leq 0$.

On the other hand, if $u \in L_\gamma$ is such that $\zeta_1^a + u \notin dom\ I_\gamma$ or $\zeta_2 + \dot{u} \notin \overline{D}_\gamma^\sigma$, then $\overline{\Phi}(\zeta + u) = \infty$ and the above inequality is trivial. Therefore, for all $u \in L_\gamma$, $\overline{\Phi}(\zeta + u) - \overline{\Phi}(\zeta) \geq \langle \ell, u \rangle$. $\square$

By (22), we have $dom\ \overline{\Psi} = \overline{V} \cap dom\ \overline{\Phi}$ with $dom\ \overline{\Phi} \subset U^{**}$ and $\overline{V}$ the $\sigma(U^{**}, U^*)$-closure of V in U^{**}. For any $\zeta \in U^{**}$, one writes $\zeta \in L''_\gamma$ to signify that the restriction to L'_γ of ζ stands in L''_γ. More, we have: for all $\zeta, \xi \in U^{**}$, if ζ and ξ match on L'_γ, then $\overline{\Phi}(\zeta) = \overline{\Phi}(\xi)$. Therefore, we can identify any element of $dom\ \overline{\Phi}$ with its restriction to L'_γ : $dom\ \overline{\Phi} \subset L''_\gamma$.

As $\overline{\Psi}$ is the "restriction" of $\overline{\Phi}$ to $\overline{V}$, we also have: for all $\zeta, \xi \in U^{**}$, if ζ and ξ match on L'_γ, then $\overline{\Psi}(\zeta) = \overline{\Psi}(\xi)$. Hence, one can identify $dom\ \overline{\Psi}$ with $dom\ \overline{\Phi} \cap \widetilde{V} \subset L''_\gamma$, where $\widetilde{V}$ is the $\sigma(L''_\gamma, L'_\gamma)$-closure of V in L''_γ. It is clear that for any $\zeta \in dom\ \overline{\Phi}$, $\zeta_1^a \in dom\ I_\gamma \subset L_\gamma$ and $\zeta_2 \in \overline{D}_\gamma^\sigma \subset L_\gamma^s$. Taking these identifications into account, we have shown the following

Proposition 5.3. *For all $\zeta \in U^{**}$, we have: $\zeta \in dom\ \overline{\Psi}$ and $\zeta_1^s = 0$ if and only if $\zeta_1 \in dom\ I_\gamma$, $\zeta_2 \in \overline{D}_\gamma^\sigma$ and for all $\varepsilon > 0$, $K \geq 1$, $f_1, \dots, f_K \in L_{\gamma^*}$, $\ell_1, \dots, \ell_K \in L_\gamma^s$, there exists $v \in V$ such that $\sum_{k \leq K} |\int_\Omega (v - \zeta_1) f_k\, dR| + |\langle \ell_k, (v - \zeta_2) \rangle| \leq \varepsilon$.*

In ([18], Lemma 3.1), it is proved that ζ_1 can be written as $\zeta_1 = \langle z_1, \varphi \rangle$ for some $z_1 \in \mathcal{X}^*$.

Remark. In the case of finitely many constraints, we have $V = \overline{V}$. Therefore, $\zeta_1 = \zeta_2 = v$ for some v in V.

Proof of Theorem 3.4. Theorem 3.4 is a direct consequence of Theorem 4.2, Propositions 4.4 and 4.5, for the abstract approach; and of Theorems 3.2 and Theorem 5.5 (stated below), for the application in function spaces. The statement that the $\tilde{\ell}$'s are singular, is proved at Lemma 5.6 below.

Theorem 5.5 is proved in [18]. We recall it for the convenience of the reader. We first need a definition.

Definition 5.4. Let $z_1 + \infty(n)$ be a force field. Its infinite component: $\infty(n)$, determines the measurable subsets $\mathcal{S}$, $\mathcal{T}_+$ and $\mathcal{T}_-$ (see (17)). Its finite component z_1 is a linear form on $\mathcal{X}$.

Under the assumption that γ is even and $\gamma^*(\gamma'(\infty)) < \infty$, $z_1 + \infty(n)$ is said to be an <u>admissible dual parameter</u>, if the following properties are satisfied.

1. $\langle z_1, \varphi \rangle$ is a measurable function whose support is included in $\mathcal{S}$.
2. $\int_{\mathcal{S}} \gamma(\langle z_1, \varphi\rangle)\, dR < \infty$ and $\int_{\mathcal{S}} \gamma^* \circ \gamma'(\langle z_1, \varphi\rangle)\, dR < \infty$
3. for any $j \in \mathcal{J}$, there exists $\alpha > 0$ such that $\int_{\cap_{i<j}\{n^i=0\}} \gamma(\alpha\langle n^j, \varphi\rangle)\, dR < \infty$
4. $R(\mathcal{T}_+) < \infty$ and $R(\mathcal{T}_-) < \infty$
5. for all $\varepsilon > 0$, $K \geq 1$, $f_k \in L_{\gamma^*}$, $k \leq K$, there exists $y \in \mathcal{Y}$ such that $\int_\Omega \gamma(\langle y, \varphi\rangle)\, dR < \infty$ and for all $k \leq K$, $|\int_{\mathcal{S}} (\langle z_1, \varphi\rangle - \langle y, \varphi\rangle) f_k\, dR| \leq \varepsilon$

With this definition in hand, we can state the following result.

Theorem 5.5. *(γ is an even function). Let us assume that $\gamma^*(\gamma'(\infty)) < \infty$.*

Let x be any admissible constraint. One can associate with it an admissible dual parameter $\bar{z}_1 = z_1 + \infty(n)$, in the sense of Definition 5.4, such that

$$x = \int_\Omega \varphi \gamma'(\langle \bar{z}_1, \varphi\rangle)\, dR \tag{26}$$

Moreover, the set of minimizers is

$$\mathcal{S}(x) = \{\gamma'(\langle \bar{z}_1, \varphi\rangle) \cdot R\} \tag{27}$$

Conversely, for any admissible force field $\bar{z}_1$, x being defined by (26), we have $\Lambda^*(x) < \infty$ and $\mathcal{S}(x)$ is given by (27).

Furthermore, any admissible constraint is interior if and only if the force field is finite R-a.e..

Proof. See ([18], Theorems 4.4 and 4.5). □

If x stands in the relative geometric interior of the effective domain of Λ^*, $\bar{z}_1 = z_1$ has no infinite component. If it stands on the geometric boundary of the effective domain of Λ^*, the field of ordered collections of outward normal vector (n) characterizes the minimal face of the boundary on which x stands. Note that x is in the relative geometric interior of this face and z_1 characterizes x in this face. For more details, see ([18], Section 4).

Proof of Proposition 3.5. The last statement of Proposition 3.5 is Lemma 5.6. Its first and second statements are respectively Proposition 5.10 and 5.12, that we are going to prove. Lemmas 5.7, 5.9 and Corollary 5.8 are preliminary results for the proof of Proposition 5.10; Lemma 5.11 is a preliminary result for the proof of Proposition 5.12.

Lemma 5.6. *The function γ^* is stricly convex at infinity if and only if γ is steep.*

If γ is even and steep, for any $\zeta \in$ dom $\overline{\Phi}$ such that $\partial\overline{\Phi}(\zeta)$ is nonempty, we have: $K(\zeta) \subset L^s_\gamma$.

Proof. The first statement is easy.

Suppose that there exists $k \in K(\zeta)$ such that k is absolutely continuous and different from zero. Because of Proposition 5.2, for all $t \geq 0$, $\ell_t = \gamma'(\zeta_1^a) \cdot R + tk$ is absolutely continuous and belongs to $\partial\overline{\Phi}(\zeta)$. With (12), we obtain: $\Phi^*(\ell_t) = \int_\Omega \gamma^*(\gamma'(\zeta_1^a) + t\frac{dk}{dR})\, dR < \infty, \forall t \geq 0$. On the other hand, $\ell_0, \ell_t \in \partial\overline{\Phi}(\zeta)$ implies that $\Phi^*(\ell_t) = \Phi^*(\ell_0) + \langle \zeta, \ell_t - \ell_0 \rangle = \Phi^*(\ell_0) + t\langle \zeta, k \rangle$, $\forall t \geq 0$. These two expressions of $\Phi^*(\ell_t)$ imply that γ^* is affine at infinity. □

Let us go on with some considerations about the gauge function of $\overline{\mathcal{D}}_\gamma$. As the geometric and topological closures of $\mathcal{D}_\gamma$ are equal, the gauge functions of $\mathcal{D}_\gamma$ and $\overline{\mathcal{D}}_\gamma$ are equal. We have: $p_{\overline{\mathcal{D}}_\gamma}(u) = \inf\{\lambda > 0\,;\, u/\lambda \in \overline{\mathcal{D}}_\gamma\}$, for any $u \in L_\gamma$. As the $\|\cdot\|_\gamma$-interior of $\mathcal{D}_\gamma$ is nonempty, $p_{\overline{\mathcal{D}}_\gamma}(u)$ is finite for all $u \in L_\gamma$. Note that $p_{\overline{\mathcal{D}}_\gamma}(u) \leq 1 \Leftrightarrow \int_\Omega \gamma(u/\lambda)\, dR < \infty, \forall \lambda > 1 \Leftrightarrow u \in \overline{\mathcal{D}}_\gamma$.

We derive a description of $K_1(u)$, which is stated at Proposition 5.10 below.

Lemma 5.7. *Let $u \in \overline{\mathcal{D}}_\gamma$, $u \geq 0$. For all $h \in L_\gamma$, $h \geq 0$, there exists $t_o > 0$ which only depends on h such that $u - t_o h$ belongs to $\overline{\mathcal{D}}_\gamma$.*

Proof. For any $t > 0, \lambda > 1$, $\int_\Omega \gamma(\frac{u-th}{\lambda})\, dR \leq \int_\Omega \gamma(\frac{(u-th)_+}{\lambda})\, dR + \int_\Omega \gamma((u-th)_-)\, dR \leq \int_\Omega \gamma(u/\lambda)\, dR + \int_\Omega \gamma(th)\, dR$. But $\int_\Omega \gamma(u/\lambda)\, dR < \infty$ and $\int_\Omega \gamma(t_o h)\, dR < \infty$ for $t_o > 0$ small enough. Therefore, for all $\lambda > 1$, $\int_\Omega \gamma(\frac{u-t_o h}{\lambda})\, dR < \infty$. That is $p_{\overline{\mathcal{D}}_\gamma}(u - t_o h) \leq 1$, which is the desired result. □

Corollary 5.8. *For any $u \in \overline{\mathcal{D}}_\gamma$ and $\ell \in L'_\gamma$, we have: $(u \geq 0$ and $\ell \in K_1(u)) \Rightarrow \ell \geq 0$.*

Proof. For all $h \in L_\gamma, h \geq 0$, with Lemma 5.7, we have: $u - t_o h \in \overline{\mathcal{D}}_\gamma$. As $\ell \in K_1(u)$, we get $\langle \ell, -t_o h \rangle \leq 0$. Hence, $\langle \ell, h \rangle \geq 0$. □

Lemma 5.9. *Let $u \in \overline{\mathcal{D}}_\gamma$ and $w \in L_\gamma$ be such that $u + w \in \overline{\mathcal{D}}_\gamma$, then: $u + w_+ \in \overline{\mathcal{D}}_\gamma$.*

Proof. For all $\lambda > 1$, we have: $\int_\Omega \gamma(u/\lambda)\, dR < \infty$, $\int_\Omega \gamma(\frac{u+w}{\lambda})\, dR < \infty$ and $\int_{\{w\geq 0\}} \gamma(\frac{u+w_+}{\lambda})\, dR = \int_{\{w\geq 0\}} \gamma(\frac{u+w}{\lambda})\, dR \leq \int_\Omega \gamma(\frac{u+w}{\lambda})\, dR$. Hence, $\int_\Omega \gamma(\frac{u+w_+}{\lambda})\, dR = \int_{\{w\geq 0\}} \gamma(\frac{u+w_+}{\lambda})\, dR + \int_{\{w<0\}} \gamma(\frac{u+w_+}{\lambda})\, dR = \int_{\{w\geq 0\}} \gamma(\frac{u+w_+}{\lambda})\, dR + \int_{\{w<0\}} \gamma(u/\lambda)\, dR \leq \int_\Omega \gamma(\frac{u+w}{\lambda})\, dR + \int_\Omega \gamma(u/\lambda)\, dR < \infty$. That is $p_{\overline{\mathcal{D}}_\gamma}(u + w_+) \leq 1$ which is the desired result. □

Proposition 5.10. *Let $u \in \overline{\mathcal{D}}_\gamma$, $u \geq 0$. Then, $\ell \in K_1(u)$ if and only if $\ell \geq 0$ and for all $h \in L_\gamma$, $(h \geq 0$ and $\exists t > 0$, $u + th \in \overline{\mathcal{D}}_\gamma) \Rightarrow \langle \ell, h \rangle = 0$.*

Proof. - Necessary condition. Let $\ell \in K_1(u)$. We have: $u + th \in \overline{\mathcal{D}}_\gamma \Rightarrow \langle \ell, th \rangle \leq 0 \Rightarrow \langle \ell, h \rangle \leq 0$. As $\ell \geq 0$ (Corollary 5.8) and $h \geq 0$, we also obtain $\langle \ell, h \rangle \geq 0$, which proves $\langle \ell, h \rangle = 0$.

- Sufficient condition. Let w be such that $u + w \in \overline{\mathcal{D}}_\gamma$. Because of Lemma 5.9, we have $u + w_+ \in \overline{\mathcal{D}}_\gamma$. Together with our assumption, this yields $\langle \ell, w_+ \rangle = 0$. Therefore, as $\ell \geq 0$: $\langle \ell, w \rangle = \langle \ell, w_+ \rangle - \langle \ell, w_- \rangle = -\langle \ell, w_- \rangle \leq 0$. Hence, $\ell \in K_1(u)$. □

We derive a description of $K_2(\xi)$, which is stated at Proposition 5.12 below.

Recall that $\overline{D}_\gamma$ is the $\|\cdot\|_\gamma$-closure of $E_\gamma := \{\dot{u} \in N_\gamma \,;\, u \in \mathcal{D}_\gamma\}$ in N_γ. It is a convex set. Indeed, E_γ is the image of the convex set $\mathcal{D}_\gamma$ by the canonical projection $u \mapsto \dot{u}$. As the $\|\cdot\|_\gamma$-interior of $\mathcal{D}_\gamma$ is nonempty, the gauge functions of E_γ and $\overline{D}_\gamma$ are equal. This gauge function is $p_{\overline{D}_\gamma}(v) = \inf\{\lambda > 0 \,;\, v/\lambda \in \overline{D}_\gamma\} = \inf_{u \in v, u \in L_\gamma} \inf\{\lambda > 0 \,;\, \int_\Omega \gamma(u/\lambda)\, dR < \infty\}, v \in N_\gamma$. It is finite for all $v \in N_\gamma$.

From now on, any equivalence class $v \in N_\gamma$ is identified with one of its element $v \in L_\gamma$.

Lemma 5.11. *Let us denote* $\overline{D}_\gamma^+ = \{v \in \overline{D}_\gamma \,;\, v \geq 0\} \subset N_\gamma$ *and* $\overline{D}_\gamma^{\sigma +} = \{\xi \in \overline{D}_\gamma^\sigma \,;\, \xi \geq 0\} \subset L_\gamma^{s\prime}$. *The set* $\overline{D}_\gamma^+$ *is* $\sigma(L_\gamma^{s\prime}, L_\gamma^s)$*-dense in* $\overline{D}_\gamma^{\sigma +}$.

Proof. (Ad absurdum). Let $\xi \in \overline{D}_\gamma^{\sigma +}$ be isolated from the closed convex set $\overline{D}_\gamma^+$. By Hahn-Banach's theorem, there exists $\ell_o \in L_\gamma^s$ which separates ξ from $\overline{D}_\gamma^+$: $\forall v \in \overline{D}_\gamma^+$, $\langle \ell_o, v \rangle \leq \alpha < \langle \ell_o, \xi \rangle$. Since for all $v \geq 0$, $\langle \ell_o^+, v \rangle = \sup\{\langle \ell_o, w \rangle \,;\, 0 \leq w \leq v\}$, we have: $\sup_{v \in \overline{D}_\gamma} \langle \ell_o^+, v \rangle = \sup_{v \in \overline{D}_\gamma^+} \langle \ell_o, v \rangle \leq \alpha < \langle \ell_o, \xi \rangle = \langle \ell_o^+, \xi \rangle - \langle \ell_o^-, \xi \rangle \leq \langle \ell_o^+, \xi \rangle$. This is a contradiction, since $\xi \in \overline{D}_\gamma^\sigma$ and $\ell_o^+ \in L_\gamma^s$. □

Proposition 5.12. *Let* $\xi \in \overline{D}_\gamma^\sigma$, $\xi \geq 0$. *Then,* $\ell \in K_2(\xi)$ *if and only if* $\ell \geq 0$ *and for all* $h \in N_\gamma$, ($h \geq 0$ *and* $\exists t > 0$, $\xi + th \in \overline{D}_\gamma^\sigma$) $\Rightarrow \langle \ell, h \rangle = 0$.

Proof. The analogues of Lemma 5.7, Corollary 5.8, Lemma 5.9 and Proposition 5.10 hold mutatis mutandi for $(K_2(\xi), \overline{D}_\gamma, N_\gamma)$ instead of $(K_1(u), \mathcal{D}_\gamma, L_\gamma)$. Let us call them Lemma 5.7',... Thanks to Lemmas 5.7' and 5.11', we obtain the following extension of Lemma 5.7': Let $\xi \in \overline{D}_\gamma^\sigma$, $\xi \geq 0$, for all $h \in N_\gamma$, $h \geq 0$, there exists $t_o > 0$ such that $\xi - t_o h \in \overline{D}_\gamma^\sigma$. As in Corollary 5.8', it follows that $\ell \in K_2(\xi)$, whenever $\ell \geq 0$.

As $\overline{D}_\gamma$ is dense in $\overline{D}_\gamma^\sigma$, by Lemma 5.9', we obtain the following extension of Lemma 5.9': Let $\xi \in \overline{D}_\gamma^\sigma$ and $w \in N_\gamma$ be such that $\xi + w \in \overline{D}_\gamma^\sigma$, then, $\xi + w^+/2 \in \overline{D}_\gamma^\sigma$. One considers $w/2$ rather than w to insure the existence of a generalized sequence (v_α) which converges towards ξ with $v_\alpha \in \overline{D}_\gamma$ and $v_\alpha + w/2 \in \overline{D}_\gamma$. One completes the proof of the proposition as in Proposition 5.10'. □

6. Examples

For the sake of clarity, we present easy examples of entropy maximisation. Let R be a probability measure on Ω. As already noticed at Section 2, entropy corresponds to $\gamma(s) = e^s - s - 1$. Let $\gamma_o(s) = e^{|s|} - |s| - 1$ and $\rho^*(t) = t \log t - t + 1$ with *dom* $\rho^* = [0, \infty)$.

We are interested in the minimisation of the extended relative entropy, defined for all $k \in L^*_{\gamma_o}$, by

$$I(k) = \begin{cases} \int_\Omega \rho^*(\frac{dk^a}{dR})\, dR + \sup\{\langle k^s, u\rangle; u, \int_\Omega e^u\, dR < \infty\} & \text{if } \ell \in L'_{\gamma_o} \\ +\infty & \text{otherwise} \end{cases}$$

In the case where $\langle k, 1\rangle = 1$, we obtain $\int_\Omega \rho^*(\frac{dk^a}{dR})\, dR = \int_\Omega \log(\frac{dk^a}{dR})\, dk^a$ that is the relative entropy of k^a with respect to R. Our problem is

$$\text{minimize } I(k) \text{ subject to } \langle k, 1\rangle = 1 \text{ and } \langle k, \theta\rangle = c \tag{P_1}$$

where θ belongs to L_{γ_o}. This means that there exists $\alpha > 0$ such that $\int_\Omega e^{\alpha|\theta|}\, dR < \infty$. Considering $\ell = k - R$, (P_1) becomes

$$\text{minimize } \Phi^*(\ell) \text{ subject to } \langle \ell, 1\rangle = 0 \text{ and } \langle \ell, \theta\rangle = a := c - \int_\Omega \theta\, dR \tag{P_2}$$

where Φ^* is associated with $\gamma^*(t) = (t+1)\log(t+1) - t$ with *dom* $\gamma^* = [-1, \infty)$, the convex conjugate of $\gamma(s) = e^s - s - 1$. Let us take the constraint function $\varphi = (1, \theta) : \Omega \to \mathcal{X} = \mathbb{R}^2$. With $x = (0, a)$, an easy computation leads us to

$$\Lambda^*(0, a) = \sup_{y\in\mathbb{R}}\{yc - \log \int_\Omega e^{y\theta}\, dR\} := \Gamma^*(c)$$

which is the convex conjugate of the convex function $\Gamma(y) = \log \int_\Omega e^{y\theta}\, dR$.

To specify our example, we take $\Omega = \mathbb{R}$ and $R(d\omega) = \frac{1}{\sigma\sqrt{2\pi}} \exp(-\frac{\omega^2}{2\sigma^2})\, d\omega$: the centered normal distribution with variance σ^2, $(\sigma > 0)$. We are going to consider the constraint functions $\theta(\omega) = \omega$, $\theta(\omega) = \omega^2$ and $\theta(\omega) = \omega^2 - |\omega|$.

In these three cases, the domain of Γ^* is open. This implies that there is no infinite component entering into the representation of the minimizers of (P_1). By Theorem 3.4 with $k = \ell + R$, the minimizers k_* of (P_1) can be written as

$$k_* = P_y + \tilde{\ell} \tag{28}$$

where $P_y := \exp(y\theta - \Gamma(y)) \cdot R$ (which is a probability measure) and $\tilde{\ell}$ belongs to the cone $K((y\theta)_+, (y\theta)_+)$. The admissibility condition for y is (see Definition 3.3.2)

$$\int_\Omega (y\theta)_+ e^{(y\theta)_+}\, dR < \infty. \tag{29}$$

As the constraint is finite dimensional, the cone of the $\tilde{\ell}$'s has the form $K(u, u)$, see (15).

When c is an internal value of *dom* Γ^*, we denote by $y(c)$ the unique solution to the equation

$$\Gamma'(y) = \int_\Omega \theta\, dP_y = c. \tag{30}$$

This solution is unique since Γ is strictly convex, because $\Gamma''(y) = \mathrm{Var}_{P_y}(\theta) > 0$.

The case $\theta(\omega) = \omega$. In this case *dom* $\Gamma^* = \mathbb{R}$, so that (P_1) has a solution for each $c \in \mathbb{R}$. Since θ belongs to M_{γ_o}, no singular component enters into (28), i.e.: $\tilde{\ell} = 0$; this has been proved in [18]. One recovers this result, noticing that θ is internal to $\mathcal{D}_\gamma$. Therefore, for all $c \in \mathbb{R}$,

$$k_* = P_{y(c)}.$$

Since $y(c) = \Gamma^{*\prime}(c) = c/\sigma^2$, k_* is the normal distribution with variance σ^2 and mean c.

The case $\theta(\omega) = \omega^2$. In this case $dom\,\Gamma^* = (0, \infty)$, so that (P_1) has a solution for each $c > 0$. This time, θ belongs to L_{γ_o}, but not to M_{γ_o}. The admissible y's (see (29)) are

$$y \in (-\infty, y_o) \text{ with } y_o = 1/\sigma^2.$$

Since for each $y < y_o$, $y\theta$ is internal to $\mathcal{D}_\gamma$, k_* has no singular component. Therefore

$$k_* = P_{y(c)}.$$

The case $\theta(\omega) = \omega^2 - |\omega|$. In this case $dom\,\Gamma^* = (0, \infty)$, so that (P_1) has a solution for each $c > 0$. Again, θ belongs to L_{γ_o}, but not to M_{γ_o}. The admissible y's (see (29)) are

$$y \in (-\infty, y_o] \text{ with } y_o = 1/\sigma^2.$$

Let us set $c_o = \int_\Omega \theta\, dP_{y_o}$ which is finite and solution to $y(c_o) = y_o$. Then, for all $0 < c < c_o$, we have $k_* = P_{y(c)}$ as before, since $y(c) < y_o$ and $y\theta$ is internal to $\mathcal{D}_\gamma$ for all $y < y_o$.

For each $c \geq c_o$, we have $k_* = P_{y_o} + \tilde{\ell}$ where $P_{y_o}(d\omega) = \frac{e^{-y_o|\omega|}}{Z(y_o)}\, d\omega$ ($Z(y_o) = 2/y_o$ is the normalizing constant) and $\tilde{\ell}$ belongs to $L^s_{\gamma_o}$ is such that: $\tilde{\ell} \geq 0$, for each $h \in L_{\gamma_o}(R) \cap L_{\gamma_o}(P_{y_o})$, $\langle \tilde{\ell}, h\rangle = 0$ (as a consequence of Proposition 3.5) and $\langle \tilde{\ell}, \theta\rangle = c - c_o$.

An infinite-dimensional constraint. We keep the previous setting where R is the normal distribution on $\mathbb{R}$. Let us take the moment function θ in M_{γ_o}. We denote

$$\beta_o(\omega) = y_o(\omega^2 - |\omega|)$$

the "critical" moment function of the previous example. Let us consider the following $k_* \in L'_{\gamma_o}$:

$$k_* = P_{y(c)} + k^s_*, \text{ where } k^s_* \in K_2(\beta_o). \tag{31}$$

We have already seen that $K_2(\beta_o)$ is not reduced to $\{0\}$. By Theorem 3.4, k_* is a minimizer of $I(k)$ subject to the constraints:

$$\langle k, 1\rangle = 1, \quad \langle k, \theta\rangle = c, \quad \langle k, \beta_n\rangle = \int \beta_n\, dP_{y(c)} + \langle k^s_*, \beta_o\rangle, \forall n \geq 1$$

with $\beta_n := \mathbb{1}_{\{|\beta_o| \geq n\}}\beta_o$, $n \geq 1$. To prove this, it is sufficient to check that $k_* - R$ has the form (19) for an admissible dual parameter (z_1, ζ_2) in the sense of Definition 3.3. Notice that $\gamma_o = \gamma_+$. Let us take $\varphi = (1, \theta, \beta_1, \beta_2, \dots)$, with $\mathcal{X}$ and $\mathcal{Y}$ both equal to the space of numerical sequences with finitely many nonzero terms so that (6) holds. (31) corresponds to some $\langle z_1, \varphi\rangle = y_1 + y_2\theta$ and to $\zeta_2 = \beta_o$. But, for all $n \geq 1$, we have $\beta_o = \beta_n$ in $L^{s\prime}_{\gamma_o}$, since, by the orthogonality property of Proposition 7.3, we have for all $\ell^s \in L^s_{\gamma_o}$, $\langle \beta_o - \beta_n, \ell^s\rangle = \langle \mathbb{1}_{\{|\beta_o| < n\}}\beta_o, \ell^s\rangle = 0$ as $\mathbb{1}_{\{|\beta_o|<n\}}\beta_o$ is bounded. Therefore,

$$\langle \beta_o, \ell^s\rangle = \langle \beta_n, \ell^s\rangle, \forall n \geq 1 \tag{32}$$

Consequently, the sequence $v_n = y_1 + y_2\theta + \beta_n \in V$ tends to $(\langle z_1, \varphi\rangle, \zeta_2)$ in the sense of Definition 3.3.6, since for all $f \in L_{\gamma_o^*}$, $\lim_{n\to\infty}\langle f, \beta_n\rangle = 0$ by dominated convergence and for all $\ell^s \in L^s_{\gamma_o}$, $\langle v_n, \ell^s\rangle = \langle \beta, \ell^s\rangle$, for all $n \geq 1$, because of (32) and $y_1 + y_2\theta$ is assumed to be in M_{γ_o} so that $\langle y_1 + y_2\theta, \ell^s\rangle = 0$.

7. Appendix

7.1. Duality in Orlicz spaces.

The function $\gamma : \mathbb{R} \to [0, +\infty]$ is called a *Young function* if it is convex, even and satisfies $\gamma(0) = 0$, $\lim_{s\to\infty} \gamma(s) = +\infty$ and there exists $s_o > 0$ such that $0 \le \gamma(s_o) < \infty$.

Let Ω be an arbitrary set, $\mathcal{A}$ be a σ-field of subsets of Ω and let R be a nonnegative σ-finite measure on $\mathcal{A}$. In this section, all the numerical functions on Ω are $\mathcal{A}$-measurable and R-almost everywhere equal functions are identified.

The *Orlicz space* associated with γ is defined by: $L_\gamma := \{u : \Omega \to \mathbb{R} \,; \|u\|_\gamma < +\infty\}$ with $\|u\|_\gamma = \inf\left\{\beta > 0 \,;\, \int_\Omega \gamma\left(\frac{|u(\omega)|}{\beta}\right) R(d\omega) \le 1\right\}$. The function $\|\cdot\|_\gamma$ is a norm (the Luxemburg norm) and

$$L_\gamma = \{u : \Omega \to \mathbb{R} \,; \exists \lambda_o > 0, \int_\Omega \gamma(\lambda_o u)\, dR < \infty\}.$$

A subspace of interest is

$$M_\gamma := \{u : \Omega \to \mathbb{R} \,; \forall \lambda > 0, \int_\Omega \gamma(\lambda u)\, dR < \infty\}.$$

Of course: $M_\gamma \subset L_\gamma$. The function γ is said to satisfy the Δ_2-condition if

$$\text{there exist } C > 0, s_o \ge 0 \text{ such that } \forall s \ge s_o, \gamma(2s) \le C\gamma(s) \tag{33}$$

If $s_o = 0$, the Δ_2-condition is said to be global. When R is bounded, in order that $M_\gamma = L_\gamma$, it is enough that γ satisfies the Δ_2-condition. When R is unbounded, this equality still holds if the Δ_2-condition is global.

Note that if $\gamma(s) = \infty$ for some $s > 0$, M_γ reduces to the null space. If in addition R is bounded, L_γ is L_∞.

On the other hand, when γ is a finite function, M_γ contains all the bounded functions.

Duality in Orlicz spaces is intimately linked with the convex conjugacy. The convex conjugate γ^* of γ is also a Young function so that one may consider the Orlicz space L_{γ^*}.

A continuous linear form $\ell \in L'_\gamma$ is said to be *singular* if for all $u \in L_\gamma$, there exists a nonincreasing sequence of measurable sets (A_n) such that $R(\cap_n A_n) = 0$ and for all $n \ge 1$, $\langle \ell, u\mathbb{1}_{\Omega\setminus A_n}\rangle = 0$. Let us denote L^s_γ the subspace of L'_γ of all singular forms.

Theorem 7.1. *(Representation of L'_γ). Let γ be any Young function. Any $\ell \in L'_\gamma$ is uniquely decomposed as*

$$\ell = \ell^a + \ell^s \tag{34}$$

with ℓ^a in $L_{\gamma^*} \cdot R$ and ℓ^s in L^s_γ (the space of all continuous L_{γ^*}-singular forms on L_γ). This means that L'_γ is the direct sum $L'_\gamma = (L_{\gamma^*} \cdot R) \oplus L^s_\gamma$.

If γ satisfies the Δ_2-condition $L'_\gamma = L_{\gamma^*} \cdot R$, so that L^s_γ reduces to the null vector space.

In the decomposition (34), ℓ^a is called the *absolutely continuous part* of ℓ while ℓ^s is its *singular part*.

Proof. For a proof of this result, see ([15], Theorems 6.4 and 7.2bis), or for an almost complete result in this direction, see ([16], Theorem 2.2). □

We denote $I_\gamma(f) = \int_\Omega \gamma(f)\, dR \in [0,\infty]$ and $I_{\gamma^*}(f) = \int_\Omega \gamma^*(f)\, dR \in [0,\infty]$. Let A be a subset of a vector space X in duality with Y. The indicator function of A is $\delta(x \mid A) = \begin{cases} 0 & \text{if } x \in A \\ \infty & \text{if } x \notin A \end{cases}, x \in X$ and its support function is $\delta^*(y \mid A) = \sup_{x\in X}\{\langle x, y\rangle - \delta(x \mid A)\} = \sup_{x\in A}\langle x, y\rangle, y \in Y$.

Proposition 7.2. *Let I_γ^* be the convex conjugate of I_γ for the duality (L_γ, L'_γ). For any $\ell \in L'_\gamma$, $I_\gamma^*(\ell) = I_{\gamma^*}(\frac{d\ell^a}{dR}) + \delta^*(\ell^s \mid \text{dom}\, I_\gamma)$ where $\ell = \ell^a + \ell^s$ is the decomposition* (34)*.*

Proof. This result is ([16], Thm 2.6) when γ is a finite Young function, it is ([24], Thm 1) when $L_\gamma = L_\infty$. For the general case, see ([13], Thm 3.2). □

Proposition 7.3. *Let us assume that γ is finite. Then, $\ell \in L'_\gamma$ is singular if and only if $\langle \ell, u\rangle = 0$, for all u in M_γ.*

Proof. This result is ([13], Cor 4.5). □

7.2. The proofs of Theorems 4.1 and 4.2.

Theorem 4.2 is a direct consequence of Theorem 4.1. The only thing to be proved is the dual equality (21) that is assumed in Theorem 4.1. But, this dual equality has been derived at ([19], Theorem 3.3).

Now, let us prove Theorem 4.1. We begin with a preliminary result.

Lemma 7.4.

(a) *For any $\alpha \in V^*, \Psi^*(\alpha) \le \inf\{\Phi^*(\ell)\,;\, \ell \in \alpha\}$*

(b) *For any $\alpha \in V^*, \Psi^*(\alpha) = \infty \Rightarrow \mathcal{S}(\alpha) = \emptyset$*

(c) *For any $\ell \in U^*, \Psi^*(\dot{\ell}) = \Phi^*(\ell) < \infty \Rightarrow \ell \in \mathcal{S}(\dot{\ell})$*

(d) *For any $\alpha \in V^*$ such that $\Psi^*(\alpha) = \inf\{\Phi^*(\ell)\,;\, \ell \in \alpha\} < \infty$ and $\partial\Psi^*(\alpha) \neq \emptyset$, we have*
$$\mathcal{S}(\alpha) \subset \partial\overline{\Phi}(\zeta),\ \forall \zeta \in \partial\Psi^*(\alpha)$$

(e) *For all $\zeta \in \overline{V}, \ell \in U^*, \ell \in \partial\overline{\Phi}(\zeta) \Rightarrow \Psi^*(\dot{\ell}) = \Phi^*(\ell) < \infty$*

Proof of Theorem 4.1. Let us admit the lemma for a while. As $\mathcal{S}(\alpha) \subset \alpha$, (d) leads us to: $\mathcal{S}(\alpha) \subset \alpha \cap \partial\overline{\Phi}(\zeta)$ for all $\zeta \in \partial\Psi^*(\alpha)$.

Let us prove the converse inclusion. As $\zeta \in \partial\Psi^*(\alpha) \Rightarrow \overline{\Psi}(\zeta) < \infty \Rightarrow \zeta \in \overline{V}$ (see (22)) and $\ell \in \alpha \Leftrightarrow \dot{\ell} = \alpha$, with (e) and (c) one obtains that for any $\ell \in U^*, \zeta \in \partial\Psi^*(\alpha)$: $\ell \in \alpha \cap \partial\overline{\Phi}(\zeta) \Rightarrow \Psi^*(\alpha) = \Phi^*(\ell) < \infty \Rightarrow \ell \in \mathcal{S}(\alpha)$. □

Proof of Lemma 7.4. Proof of (a). For any $\ell \in \alpha$, we have: $\Psi^*(\alpha) = \sup_{v\in V}\{\langle \alpha, v\rangle - \Psi(v)\} = \sup_{v\in V}\{\langle \ell, v\rangle - \Phi(v)\} \le \sup_{u\in U}\{\langle \ell, u\rangle - \Phi(u)\} = \Phi^*(\ell)$, from which it follows that $\Psi^*(\alpha) \le \inf_{\ell\in\alpha} \Phi^*(\ell)$.

(b) and (c) are direct consequences of (a).

Proof of (d). Let us begin noting that any $\zeta \in \partial\Psi^*(\alpha)$ belongs to *dom* $\overline{\Psi}$ and therefore to $\overline{V}$. As a consequence, $\overline{\Psi}(\zeta) = \overline{\Phi}(\zeta)$ (see (22)) and $\langle \zeta, \alpha \rangle = \langle \zeta, \ell \rangle$ for all $\ell \in \alpha$. Indeed, let (v_α) be a generalized sequence in V converging to $\zeta \in \overline{V}$ for $\sigma(V^{**}, V^*)$. Then, $\langle \zeta, \alpha \rangle = \lim_\alpha \langle v_\alpha, \alpha \rangle = \lim_\alpha \langle v_\alpha, \ell \rangle$ for any $\ell \in \alpha$.

Let ℓ stand in $\mathcal{S}(\alpha)$. As $\Psi^*(\alpha) = \inf\{\Phi^*(\ell)\,;\, \ell \in \alpha\} < \infty$, we have $\Phi^*(\ell) = \Psi^*(\alpha)$. Then, for any $\zeta \in \partial\Psi^*(\alpha)$, we get $\Phi^*(\ell) = \Psi^*(\alpha) = \langle \zeta, \alpha \rangle - \overline{\Psi}(\zeta) = \langle \zeta, \ell \rangle - \overline{\Phi}(\zeta)$. This implies that $\ell \in \partial\overline{\Phi}(\zeta)$, which is the desired result.

Proof of (e). For any $\zeta \in \overline{V}, \ell \in U^*$, $\ell \in \partial\overline{\Phi}(\zeta) \Rightarrow \forall u \in U, \overline{\Phi}(\zeta + u) \geq \overline{\Phi}(\zeta) + \langle \ell, u \rangle \Rightarrow \forall v \in V, \overline{\Phi}(\zeta + v) \geq \overline{\Phi}(\zeta) + \langle \ell, v \rangle \Rightarrow \forall v \in V, \overline{\Psi}(\zeta + v) \geq \overline{\Psi}(\zeta) + \langle \dot{\ell}, v \rangle$ since $\zeta \in \overline{V}$ and $\zeta + v \in \overline{V}$. This means that $\dot{\ell} \in \partial\overline{\Psi}(\zeta)$. Therefore, $\infty > \Psi^*(\dot{\ell}) = \langle \zeta, \dot{\ell} \rangle - \overline{\Psi}(\zeta) = \langle \zeta, \ell \rangle - \overline{\Phi}(\zeta)$ (since $\zeta \in \overline{V}$) and because $\ell \in \partial\overline{\Phi}(\zeta)$, one gets $\Psi^*(\dot{\ell}) = \langle \zeta, \ell \rangle - \overline{\Phi}(\zeta) = \Phi^*(\ell) < \infty$ which is the desired result. □

7.3. The proof of Proposition 4.3.

It is supposed that γ is even. The main reason for considering as a first step the case where γ is even, is provided by the proof of Lemma 7.6 below. The key result of this subsection is the following lemma.

Lemma 7.5.

(a) *Suppose that there exists a norm $\|\cdot\|$ on U such that* $\sup\{\Phi(u)\,;\, \|u\| \leq r\} \leq 1$, *for some $r > 0$. Then, Φ is $\|\cdot\|$-continuous on ridom Φ and dom $\Phi^* \subset U'$, where U' is the topological dual space of $(U, \|\cdot\|)$.*

(b) *If in addition,* $0 < \inf\{\Phi(u)\,;\, \|u\| = t_o\} < \infty$ *for some $t_o > 0$, Φ^* is $\|\cdot\|^*$-continuous on ridom Φ^* and dom $\overline{\Phi} \subset U''$, where $\|\cdot\|^*$ is the dual norm on U' and U'' is the topological bidual space of $(U, \|\cdot\|)$.*

Proof. See ([17], Lemma 2.1). □

We are going to consider the two cases where the property

$$\|u_o\|_\gamma > 1 \quad \Phi(u_o) < \infty \text{ for some } u_o \in L_\gamma \tag{35}$$

holds or fails.

Lemma 7.6. *Let us assume that γ is even. If* (35) *holds, then there exists $t_o > 0$ such that* $0 < \inf\{\Phi(u)\,;\, u \in L_\gamma, \|u\|_\gamma = t\} < \infty$.

Proof. Take $t_o = \|u_o\|_\gamma > 1$. For any $u \in L_\gamma$, $\|u\|_\gamma = t_o \Rightarrow \forall 0 < a < t_o, \int_\Omega \gamma(u/a)\, dR \geq 1 \Rightarrow \Phi(u) \geq 1$ (with $a = 1$). Therefore, $1 \leq \inf\{\Phi(u)\,;\, \|u\|_\gamma = t_o\} \leq \Phi(u_o) < \infty$. □

Lemma 7.7. *Let assume that γ is even.*

(a) *Φ is $\|\cdot\|_\gamma$-continuous on ridom Φ.*

(b) *dom $\Phi^* \subset L'_\gamma$*

(c) *If* (35) *holds, Φ^* is $\|\cdot\|^*_\gamma$-continuous on ridom Φ^**

(d) *If* (35) *holds, dom $\overline{\Phi} \subset L''_\gamma$, where for any $\zeta \in U^{**}$, $\zeta \in L''_\gamma$ means that $\zeta_{|L'_\gamma} \in L''_\gamma$.*

Proof. To prove (a) and (b) it is enough to check that $\|\cdot\|_\gamma$ satisfies the assumption of Lemma 7.5.(a). With $r = \frac{1}{2}$, for any $u \in L_\gamma$, $\|u\|_\gamma \leq \frac{1}{2} \Rightarrow \forall \beta > \frac{1}{2}$, $\int_\Omega \gamma(u/\beta)\, dR \leq 1 \Rightarrow \Phi(u) \leq 1$ (with $\beta = 1$).

Proof of (c) and (d). By Lemma 7.6, the assumption of Lemma 7.5.(b) is satisfied for $\|\cdot\|_\gamma$. This completes the proof of the lemma. □

Lemma 7.8. *Let us assume that γ is even. If (35) fails, then $L_\gamma = L_\infty$ and dom $\Phi \subset \{u \in L_\infty\,;\, \|u\|_\gamma \leq 1\}$.*

Proof. Since (35) fails, for all $u \in L_\gamma$, we have: $\Phi(u) < \infty \Rightarrow \|u\|_\gamma \leq 1 \Rightarrow \forall a < 1, \int_\Omega \gamma(au)\, dR \leq 1 \Rightarrow \Phi(u) = \lim_{a \uparrow 1} \Phi(au) \leq 1$ (monotone convergence). Hence,

$$\Phi(L_\gamma) \subset [0,1] \cup \{\infty\}. \tag{36}$$

Suppose that: $\sup \mathit{dom}\, \gamma := a_+ = +\infty$. Let $A \in \mathcal{A}$ such that $0 < R(A) < \infty$ (R is σ-finite and non-zero) and $n \geq 1$, let us take $u_n = n\mathbb{1}_A$. Then, $\Phi(u_n) = \gamma(n)R(A) < \infty$, $\forall n \geq 1$ and $\lim_{n\to\infty} \Phi(u_n) = \infty$. This contradicts (36) and proves that $a_+ < \infty$.

Let us suppose that $R(\Omega) = \infty$ and that there exists $a_o > 0$ such that $0 < \gamma(a_o) < \infty$. Taking $u_p = a_o \mathbb{1}_{\Omega_p}$ where (Ω_p) is R-localizing, we obtain: $\Phi(u_p) = \gamma(a_o)R(\Omega_p) < \infty$, $\forall p \geq 1$ and

$\lim_{p\to\infty} \Phi(u_p) = +\infty$, which contradicts (36). Therefore, if (35) fails, we have $0 < a_+ < \infty$ and $[R(\Omega) < \infty$ or $\gamma = \delta(\cdot \mid [-a_+, a_+])]$. But, $[0 < a_+ < \infty$ and $R(\Omega) < \infty] \Rightarrow L_\gamma = L_\infty$ and $[0 < a_+ < \infty$ and $\gamma = \delta(\cdot \mid [-a_+, a_+])] \Rightarrow L_\gamma = L_\infty$, which is the desired result. □

We have $\Phi(u) = \Phi_1(u) + \Phi_2(\dot{u}), u \in L_\gamma$ where $\Phi_1(u) = I_\gamma(u), u \in L_\gamma$ and $\Phi_2(\dot{u}) = \delta(\dot{u} \mid D_\gamma), \dot{u} \in N_\gamma$ being the equivalence class of $u \in L_\gamma$. Their convex conjugates with respect to the dualities (L_γ, L_{γ^*}) and (N_γ, L_γ^s) are $\Phi_1^*(\ell) = I_{\gamma^*}(\ell), \ell \in L_{\gamma^*}$ and $\Phi_2^*(\ell) = \delta^*(\ell \mid D_\gamma), \ell \in L_\gamma^s$. Proposition 7.2 states that: $\Phi^*(\ell) = \Phi_1^*(\ell^a) + \Phi_2^*(\ell^s), \ell \in L'_\gamma$.

Proof of Proposition 4.3. As $\mathit{dom}\, \Phi^* \subset L'_\gamma$ (by Lemma 7.7.(b)), for any $\zeta \in U^{**}$, $\overline{\Phi}(\zeta) = \sup_{\ell \in U^*}\{\langle \zeta, \ell\rangle - \Phi^*(\ell)\} = \sup_{\ell \in L'_\gamma}\{\langle \zeta, \ell\rangle - \Phi^*(\ell)\}$. Thanks to the decomposition $L'_\gamma = L_{\gamma^*} \oplus L_\gamma^s$ (see Theorem 7.1) and Proposition 7.2, for any $\zeta = (\zeta_1, \zeta_2) \in L'^*_\gamma = L^*_{\gamma^*} \oplus L_\gamma^{s*}$ we have: $\overline{\Phi}(\zeta) = \overline{\Phi}_1(\zeta_1) + \overline{\Phi}_2(\zeta_2)$ with $\overline{\Phi}_1(\zeta_1) = \sup_{f_1 \in L_{\gamma^*}}\{\langle \zeta_1, f_1\rangle - I_{\gamma^*}(f_1)\}$ and $\overline{\Phi}_2(\zeta_2) = \sup_{\ell_2 \in L_\gamma^s}\{\langle \zeta_2, \ell_2\rangle - \delta^*(\ell_2 \mid \mathit{dom}\, I_\gamma)\}$.

(a) Let us first suppose that (35) holds. By Lemma 7.7.(d), $\mathit{dom}\, \overline{\Phi} \subset L''_\gamma$. Hence, one can apply Proposition 7.2 to $I_{\gamma^*} : \overline{\Phi}_1(\zeta_1) = I_\gamma(\frac{d\zeta_1^a}{dR}) + \delta^*(\zeta_1^s \mid \mathit{dom}\, I_{\gamma^*})$, and by Lemma 7.9 below (applied with $X = N_\gamma, Y = L_\gamma^s$ and $Z = L_\gamma^{s\prime}$) : $\overline{\Phi}_2(\zeta_2) = \delta(\zeta_2 \mid \overline{D}_\gamma^\sigma)$.

As Lemma 7.7.(c) holds, Φ_2^* is bounded above on a non empty $\|\cdot\|_\gamma^*$-open ball. As Φ_2^* is positively homogeneous, we obtain that its domain is the whole space L_γ^s, which in turn is equivalent to: $\overline{D}_\gamma^\sigma$ is $\sigma(L_\gamma^{s\prime}, L_\gamma^s)$-bounded, hence norm bounded in $L_\gamma^{s\prime}$ (by the uniform bound theorem).

(b) If (35) fails, by Lemma 7.8: $L_\gamma = L_\infty$ and $\{u \in L_\infty\,;\, \|u\|_\infty < a_+\} \subset \mathit{dom}\, \Phi \subset \{u \in L_\infty\,;\, \|u\|_\infty \leq a_+\}$ with $0 < a_+ = \sup \mathit{dom}\, \gamma < \infty$. More, $\Phi_1^*(\ell^a) = I_{\gamma^*}(\frac{d\ell^a}{dR}), \forall \ell^a \in L_1$ and $\Phi_2^*(\ell^s) = \delta^*(\ell^s \mid D_\gamma), \forall \ell^s \in L_\infty^s$. Applying Lemma 7.5.(a) with $\Phi_1^* = I_{\gamma^*}$ and $\|\cdot\| = \|\cdot\|_1$, we obtain: $\mathit{dom}\, \overline{\Phi}_1 \subset L'_1 = L_\infty$ and by Proposition 7.2: $\overline{\Phi}_1(\zeta_1) = I_\gamma(\zeta_1), \forall \zeta_1 \in L_\infty$.

We have $L_\gamma = N_\gamma = L_\infty$ with D_γ a $\|\cdot\|_\infty$-bounded convex set with a non empty interior. Therefore, Φ_2^* is a norm which is equivalent to the dual norm $\|\cdot\|_\infty^*$ of L'_∞, and $dom\ \Phi_2^* = L_\gamma^s$.

It follows that the condition of Lemma 7.5.(a) is fulfilled with Φ_2^* and $\|\cdot\|_\infty^*$. Therefore, $\zeta_2 \in dom\ \overline{\Phi}_2$ implies that $\ell \in L'_\infty \mapsto \langle \zeta_2, \ell^s\rangle \in \mathbb{R}$ belongs to L''_∞. One concludes using Lemma 7.9 with $X = L_\infty$, $Y = L_\infty^s$, $Z = L_\infty^{s\prime}$, which states that $\overline{\Phi}_2(\zeta_2) = \delta(\zeta_2 \mid \overline{D}_\gamma^\sigma)$ for all $\zeta_2 \in L_\infty^{s\prime}$. □

Lemma 7.9. *Let (X, Y) and (Y, Z) two dual pairs such that $X \subset Z$ and Y separates Z. Let A be a subset of X and $\overline{co}A$ be its $\sigma(Z, Y)$-closed convex hull in Z.*

Then, for any $z \in Z$, $\overline{\delta}(z \mid A) := \sup_{y\in Y}\{\langle y, z\rangle - \delta^*(y \mid A)\} = \delta(z \mid \overline{co}A)$.

Proof. Let $z \in Z$, $\overline{\delta}(z \mid A) = \sup_{y\in Y}\{\langle y, z\rangle - \sup_{x\in A}\langle x, y\rangle\} = \sup_{y\in Y}\inf_{x\in A}\langle z - x, y\rangle$. With $y = 0$, it comes out that $\overline{\delta}(z \mid A) \geq 0$.

If there exists $y_o \in Y$ such that $\inf_{x\in A}\langle z - x, y_o\rangle > 0$, with $y = \lambda y_o$ and $\lambda \to +\infty$, it comes out that $\overline{\delta}(z \mid A) = +\infty$. Otherwise, we have

$$\forall y \in Y, \sup_{x\in A}\langle x - z, y\rangle \geq 0 \tag{37}$$

which implies: $\overline{\delta}(z \mid A) \leq 0$, hence: $\overline{\delta}(z \mid A) = 0$.

Now, let us show that (37) is equivalent to $z \in \overline{co}A$. By linearity of $x \mapsto \langle x, y\rangle$, and then by $\sigma(Z, Y)$-continuity of $z' \mapsto \langle z', y\rangle \in \mathbb{R}$, we obtain: $\sup_{x\in A}\langle x - z, y\rangle = \sup_{x\in\overline{co}A}\langle x - z, y\rangle = \sup_{z'\in\overline{co}A}\langle z' - z, y\rangle$. Therefore, non (37) $\Leftrightarrow \exists y_o \in Y, \sup_{z'\in\overline{co}A}\langle z' - z, y_o\rangle < 0 \Leftrightarrow \exists y_o \in Y, \sup_{z'\in\overline{co}A}\langle z', y_o\rangle < \langle z, y_o\rangle \Leftrightarrow z \notin \overline{co}A$ by Hahn-Banach theorem. □

7.4. The proof of Proposition 4.5.

The first statement comes from $dom\ \Phi^* \subset L'_{\gamma_o} \subset L^b_{\gamma_o}$ and $dom\ \overline{\Phi} \subset L'^b_{\gamma_o}$.

Let us prove the second statement. Since $\langle \ell_+, \zeta_+\rangle \leq \Phi_+^*(\ell_+) + \overline{\Phi_+}(\zeta_+)$ and $\langle \ell_-, \zeta_-\rangle \leq \Phi_-^*(\ell_-) + \overline{\Phi_-}(\zeta_-)$, we have $\langle \ell, \zeta\rangle = \langle \ell_+ - \ell_-, \zeta_+ - \zeta_-\rangle = \langle \ell_+, \zeta_+\rangle + \langle \ell_-, \zeta_-\rangle - [\langle \ell_+, \zeta_-\rangle + \langle \ell_-, \zeta_+\rangle] \leq \langle \ell_+, \zeta_+\rangle + \langle \ell_-, \zeta_-\rangle \leq \Phi_+^*(\ell_+) + \Phi_-^*(\ell_-) + \overline{\Phi_+}(\zeta_+) + \overline{\Phi_-}(\zeta_-) = \Phi^*(\ell) + \overline{\Phi}(\zeta) = \langle \ell, \zeta\rangle$. Hence, $\langle \ell_+, \zeta_-\rangle = \langle \ell_-, \zeta_+\rangle = 0$.

Suppose that $\langle \ell_+, \zeta_+\rangle < \Phi_+^*(\ell_+) + \overline{\Phi_+}(\zeta_+)$. Then, $\langle \ell, \zeta\rangle = \langle \ell_+, \zeta_+\rangle + \langle \ell_-, \zeta_-\rangle < \Phi_+^*(\ell_+) + \overline{\Phi_+}(\zeta_+) + \Phi_-^*(\ell_-) + \overline{\Phi_-}(\zeta_-) = \Phi^*(\ell) + \overline{\Phi}(\zeta) = \langle \ell, \zeta\rangle$, which is impossible. Therefore, $\langle \ell_+, \zeta_+\rangle = \Phi_+^*(\ell_+) + \overline{\Phi_+}(\zeta_+)$, that is $\ell_+ \in \partial\overline{\Phi_+}(\zeta_+)$ (and similarly for the negative index). □

References

[1] J. M. Borwein, A. S. Lewis: Duality relationships for entropy-like minimization problems, SIAM J. Control and Optim. 29(2) (1991) 325–338.

[2] J. M. Borwein, A. S. Lewis: On the convergence of moment problems, Trans. Amer. Math. Soc. 325(1) (1991) 249–271.

[3] J. M. Borwein, A. S. Lewis: Convergence of best entropy estimates, SIAM J. Optim. 1(2) (1991) 191–205.

[4] J. M. Borwein, A. S. Lewis: Partially-finite programming in L_1 and the existence of maximum entropy estimates, SIAM J. Optim. 3(2) (1993) 248–267.

[5] J. M. Borwein, A. S. Lewis: Strong rotundity and optimization, SIAM J. Optim. 4(1) (1994) 146–158.

[6] P. Cattiaux, F. Gamboa: Large deviations and variational theorems for marginal problems, Bernoulli 5 (1999) 81–108.

[7] I. Csiszár, F. Gamboa, E. Gassiat: MEM pixel correlated solutions for generalized moment and interpolation problems, IEEE Trans. Inform. Theory. 45(7) (1999) 2253–2270.

[8] I. Csiszár: I-divergence geometry of probability distributions and minimization problems, Ann. Probab. 3 (1975) 146–158.

[9] I. Csiszár: Sanov property, generalized I-projection and a conditional limit theorem, Ann. Probab. 12 (1984) 768–793.

[10] I. Csiszár: Generalized projections for non-negative functions, Acta Math. Hungar. 68(1-2) (1995) 161–185.

[11] D. Dacunha-Castelle, F. Gamboa: Maximum d'entropie et problèmes de moments, Ann. Inst. Henri Poincaré 26 (1990) 567–596.

[12] A. Dembo, O. Zeitouni: Large Deviations Techniques and Applications, Springer-Verlag (1998).

[13] A. Fougères, E. Giner: Applications de la décomposition d'un espace d'Orlicz ..., Séminaire d'Analyse Convexe, Montpellier 76, Exposé 18 (1976).

[14] F. Gamboa, E. Gassiat: Bayesian methods and maximum entropy for ill posed problems, Ann. Statist. 25(1) (1997) 328–350.

[15] E. Giner: Décomposabilité et dualité des espaces d'Orlicz, Séminaire d'Analyse Convexe, Montpellier 76, Exposé 16 (1976).

[16] A. Kozek: Convex integral functionals on Orlicz spaces, Commentat. Math. 21 (1979) 109–135.

[17] C. Léonard: Convex conjugates of integral functionals, Acta Math. Hungar. 93(4) (2001) 253–280.

[18] C. Léonard: Minimizers of energy functionals, Acta Math. Hungar. 93(4) (2001) 281–325.

[19] C. Léonard: Minimization of energy functionals applied to some inverse problems, J. Appl. Math. Optim. 44 (2001) 273–297.

[20] C. Léonard: Large deviations for Poisson random measure and processes with independent increments, Stoch. Proc. and Appl. 85 (2000) 93–121.

[21] C. Léonard: Some results on entropic projections, Proceedings of the Conference on Stochastic Analysis and Mathematical Physics, Lisbon 1999, A. B. Cruzeiro, J. C. Zambrini (eds.), Progress in Probability, Birkhäuser (2001).

[22] C. Léonard, J. Najim: An extension of Sanov's theorem. Application to the Gibbs conditioning principle, Bernoulli (2002), to appear.

[23] F. Liese, I. Vajda: Convex Statistical Distances, B. G. Teubner, Leipzig (1987).

[24] R. T. Rockafellar: Integrals which are convex functionals II, Pacific J. Math. 39(2) (1971) 439–469.

[25] R. T. Rockafellar: Convex Analysis, Princeton University Press (1970).

[26] M. Teboulle, I. Vajda: Convergence of best ϕ-entropy estimates, IEEE Trans. Inform. Theory 39 (1993) 297–301.