Constant Along Primal Rays Conjugacies and the l0 Pseudonorm

Jean-Philippe Chancelier, Michel de Lara

Constant Along Primal Rays Conjugacies and the l_0 Pseudonorm

Jean-Philippe Chancelier and Michel De Lara*
Université Paris-Est, CERMICS (ENPC)

January 29, 2020

Abstract

The so-called ℓ_0 pseudonorm on $\mathbb{R}^d$ counts the number of nonzero components of a vector. It is used in sparse optimization, either as criterion or in the constraints, to obtain solutions with few nonzero entries. For such problems, the Fenchel conjugacy fails to provide relevant analysis: indeed, the Fenchel conjugate of the characteristic function of the level sets of the ℓ_0 pseudonorm is minus infinity, and the Fenchel biconjugate of the ℓ_0 pseudonorm is zero. In this paper, we display a class of conjugacies that are suitable for the ℓ_0 pseudonorm. For this purpose, we suppose given a (source) norm on $\mathbb{R}^d$. With this norm, we define, on the one hand, a sequence of so-called coordinate-k norms and, on the other hand, a coupling between $\mathbb{R}^d$ and $\mathbb{R}^d$, called Capra (constant along primal rays). Then, we provide formulas for the Capra-conjugate and biconjugate, and for the Capra subdifferentials, of functions of the ℓ_0 pseudonorm (hence, in particular, of the ℓ_0 pseudonorm itself and of the characteristic functions of its level sets), in terms of the coordinate-k norms. As an application, we provide a new family of lower bounds for the ℓ_0 pseudonorm, as a fraction between two norms, the denominator being any norm.

Key words: ℓ_0 pseudonorm, Fenchel-Moreau conjugacy, coordinate-k norm.
AMS classification: 46N10, 49N15, 46B99, 52A41, 90C46

1 Introduction

The *counting function*, also called *cardinality function* or ℓ_0 *pseudonorm*, counts the number of nonzero components of a vector in $\mathbb{R}^d$. It is used in sparse optimization, either as criterion or in the constraints, to obtain solutions with few nonzero entries. For such problems, the Fenchel conjugacy fails to provide relevant analysis: indeed, the Fenchel conjugate of the characteristic function of the level sets of the ℓ_0 pseudonorm is minus infinity, and the Fenchel

*delara@cermics.enpc.fr

biconjugate of the ℓ_0 pseudonorm is zero. In this paper, we display a class of conjugacies that are suitable for the ℓ_0 pseudonorm.

The paper is organized as follows. In Sect. 2, we recall the definition of the ℓ_0 pseudonorm, and we introduce the notion of sequence of norms on $\mathbb{R}^d$ that are (strictly or not) decreasingly graded with respect to the ℓ_0 pseudonorm. In Sect. 3, we introduce a sequence of *coordinate-k norms*, all generated from any (source) norm on $\mathbb{R}^d$, and their dual norms. In Sect. 4, we define a so-called *Capra coupling* between $\mathbb{R}^d$ and $\mathbb{R}^d$, that depends on any (source) norm on $\mathbb{R}^d$. Then, we provide formulas for the Capra-conjugate and biconjugate, and for the Capra subdifferentials, of functions of the ℓ_0 pseudonorm (hence, in particular, of the ℓ_0 pseudonorm itself and of the characteristic functions of its level sets), in terms of the coordinate-k norms. In Sect. 5, as an application, we provide a new family of lower bounds for the ℓ_0 pseudonorm, as a fraction between two norms, the denominator being any norm.

The Appendix A gathers background on Moreau upper and lower additions, and on Fenchel-Moreau conjugacies; it also provides results on what we call *one-sided linear couplings*.

2 The ℓ_0 pseudonorm and its level sets

First, we introduce basic notations regarding the ℓ_0 pseudonorm. Second, we recall the definition of a sequence of norms on $\mathbb{R}^d$ which is (strictly or not) decreasingly graded with respect to the ℓ_0 pseudonorm (as introduced in the companion paper [5]).

The ℓ_0 pseudonorm. For any vector $x \in \mathbb{R}^d$, we define its *support* by

$$\operatorname{supp}(x) = \big\{ j \in \{1, \ldots, d\} \,\big|\, x_j \neq 0 \big\} \subset \{1, \ldots, d\} \; . \tag{1}$$

The so-called ℓ_0 *pseudonorm* is the function $\ell_0 : \mathbb{R}^d \to \{0, 1, \ldots, d\}$ defined by

$$\ell_0(x) = |\operatorname{supp}(x)| = \text{number of nonzero components of } x \; , \;\; \forall x \in \mathbb{R}^d \; , \tag{2}$$

where $|K|$ denotes the cardinal of a subset $K \subset \{1, \ldots, d\}$. The ℓ_0 pseudonorm shares three out of the four axioms of a norm: nonnegativity, positivity except for $x = 0$, subadditivity. The axiom of 1-homogeneity does not hold true; in contrast, the ℓ_0 pseudonorm is 0-homogeneous:

$$\ell_0(\rho x) = \ell_0(x) \; , \;\; \forall \rho \in \mathbb{R}\backslash\{0\} \; , \;\; \forall x \in \mathbb{R}^d \; . \tag{3}$$

The level sets of the ℓ_0 pseudonorm. The ℓ_0 pseudonorm is used in exact sparse optimization problems of the form $\inf_{\ell_0(x) \leq k} f(x)$. Thus, we introduce the *level sets*

$$\ell_0^{\leq k} = \big\{ x \in \mathbb{R}^d \,\big|\, \ell_0(x) \leq k \big\} \; , \;\; \forall k \in \big\{0, 1, \ldots, d\big\} \; , \tag{4a}$$

and the *level curves*

$$\ell_0^{=k} = \big\{ x \in \mathbb{R}^d \,\big|\, \ell_0(x) = k \big\} \; , \;\; \forall k \in \big\{0, 1, \ldots, d\big\} \; . \tag{4b}$$

For any subset $K \subset \{1, \ldots, d\}$, we denote the subspace of $\mathbb{R}^d$ made of vectors whose components vanish outside of K by[1]

$$\mathcal{R}_K = \mathbb{R}^K \times \{0\}^{-K} = \{x \in \mathbb{R}^d \,\big|\, x_j = 0 \;,\; \forall j \notin K\} \subset \mathbb{R}^d \;, \tag{5}$$

where $\mathcal{R}_\emptyset = \{0\}$. We denote by $\pi_K : \mathbb{R}^d \to \mathcal{R}_K$ the *orthogonal projection mapping* and, for any vector $x \in \mathbb{R}^d$, by $x_K = \pi_K(x) \in \mathcal{R}_K$ the vector which coincides with x, except for the components outside of K that are zero. It is easily seen that the orthogonal projection mapping π_K is self-dual, giving

$$\langle x_K \,, y_K \rangle = \langle x_K \,, y \rangle = \langle \pi_K(x) \,, y \rangle = \langle x \,, \pi_K(y) \rangle = \langle x \,, y_K \rangle \;,\; \forall x \in \mathbb{R}^d \;,\; \forall y \in \mathbb{R}^d \;. \tag{6}$$

The level sets of the ℓ_0 pseudonorm in (4a) are easily related to the subspaces $\mathcal{R}_K$ of $\mathbb{R}^d$, as defined in (5), by

$$\ell_0^{\leq k} = \{x \in \mathbb{R}^d \,\big|\, \ell_0(x) \leq k\} = \bigcup_{|K| \leq k} \mathcal{R}_K \;,\; \forall k = 0, 1, \ldots, d \;, \tag{7}$$

where the notation $\bigcup_{|K| \leq k}$ is a shorthand for $\bigcup_{K \subset \{1,\ldots,d\}, |K| \leq k}$.

Decreasingly graded sequence of norms with respect to the ℓ_0 pseudonorm. Now, we introduce the notion of sequences of norms that are, strictly or not, decreasingly graded with respect to the ℓ_0 pseudonorm: in a sense, the monotone sequence detects the number of nonzero components of a vector in $\mathbb{R}^d$ when it becomes stationary.

Definition 1 *([5, Definition 20]) We say that a sequence $\{|||\cdot|||_k\}_{k=1,\ldots,d}$ of norms on $\mathbb{R}^d$ is* decreasingly graded *w.r.t. (with respect to) the ℓ_0 pseudonorm if, for any $x \in \mathbb{R}^d$, one of the three following equivalent statements holds true.*

1. *We have the implication, for any $l = 1, \ldots, d$,*

$$\ell_0(x) = l \Rightarrow |||x|||_1 \geq \cdots \geq |||x|||_{l-1} \geq |||x|||_l = \cdots = |||x|||_d \;. \tag{8a}$$

2. *The sequence $k \in \{1, \ldots, d\} \mapsto |||x|||_k$ is nonincreasing and we have the implication, for any $l = 1, \ldots, d$,*

$$\ell_0(x) \leq l \Rightarrow |||x|||_l = |||x|||_d \;. \tag{8b}$$

3. *The sequence $k \in \{1, \ldots, d\} \mapsto |||x|||_k$ is nonincreasing and we have the inequality*

$$\ell_0(x) \geq \min \{k \in \{1, \ldots, d\} \,\big|\, |||x|||_k = |||x|||_d\} \;. \tag{8c}$$

[1]Here, following notation from Game Theory, we have denoted by $-K$ the complementary subset of K in $\{1, \ldots, d\}$: $K \cup (-K) = \{1, \ldots, d\}$ and $K \cap (-K) = \emptyset$.

We say that a sequence $\{|||\cdot|||_k\}_{k=1,\ldots,d}$ *of norms on* $\mathbb{R}^d$ *is* strictly decreasingly graded *with respect to the* ℓ_0 *pseudonorm if, for any* $x \in \mathbb{R}^d$*, one of the three following equivalent statements holds true.*

1. *We have the equivalence, for any* $l = 1, \ldots, d$*,*

$$\ell_0(x) = l \iff |||x|||_1 \geq \cdots \geq |||x|||_{l-1} > |||x|||_l = \cdots = |||x|||_d \ . \tag{9a}$$

2. *The sequence* $k \in \{1, \ldots, d\} \mapsto |||x|||_k$ *is nonincreasing and we have the equivalence, for any* $l = 1, \ldots, d$*,*

$$\ell_0(x) \leq l \iff |||x|||_l = |||x|||_d \quad \big(\iff |||x|||_l \leq |||x|||_d \big) \ . \tag{9b}$$

3. *The sequence* $k \in \{1, \ldots, d\} \mapsto |||x|||_k$ *is nonincreasing and we have the equality*

$$\ell_0(x) = \min \big\{ k \in \{1, \ldots, d\} \, \big| \, |||x|||_k = |||x|||_d \big\} \ . \tag{9c}$$

3 Coordinate-k norms and dual coordinate-k norms

In § 3.1, we provide background on norms. Then, we define *coordinate-k norms* and *dual coordinate-k norms*, that are constructed from a *source norm*, in § 3.2. We provide some of their properties in § 3.3 and in § 3.4.

3.1 Background on norms

For any norm $|||\cdot|||$ on $\mathbb{R}^d$, we denote the unit sphere and the unit ball of the norm $|||\cdot|||$ by

$$\mathbb{S} = \big\{ x \in \mathbb{R}^d \, \big| \, |||x||| = 1 \big\} \ , \tag{10a}$$
$$\mathbb{B} = \big\{ x \in \mathbb{R}^d \, \big| \, |||x||| \leq 1 \big\} \ . \tag{10b}$$

Dual norms. We recall that the following expression

$$|||y|||_\star = \sup_{|||x||| \leq 1} \langle x \, , y \rangle \ , \ \ \forall y \in \mathbb{R}^d \tag{11}$$

defines a norm on $\mathbb{R}^d$, called the *dual norm* $|||\cdot|||_\star$. We denote the unit sphere and the unit ball of the dual norm $|||\cdot|||_\star$ by

$$\mathbb{S}_\star = \big\{ y \in \mathbb{R}^d \, \big| \, |||y|||_\star = 1 \big\} \ , \tag{12a}$$
$$\mathbb{B}_\star = \big\{ y \in \mathbb{R}^d \, \big| \, |||y|||_\star \leq 1 \big\} \ . \tag{12b}$$

We have

$$|||\cdot||| = \sigma_{\mathbb{B}_\star} = \sigma_{\mathbb{S}_\star} \ \text{ and } \ |||\cdot|||_\star = \sigma_{\mathbb{B}} = \sigma_{\mathbb{S}} \ , \tag{13a}$$

where σ_S denotes the support function of the set $S \subset \mathbb{R}^d$ ($\sigma_S(y) = \sup_{x \in S} \langle x \, , y \rangle$), and where $\mathbb{B}_\star$, the unit ball of the dual norm, is the polar set $\mathbb{B}^\odot$ of the unit ball $\mathbb{B}$:

$$\mathbb{B}_\star = \mathbb{B}^\odot = \big\{ y \in \mathbb{R}^d \, \big| \, \langle x \, , y \rangle \leq 1 \ , \ \ \forall x \in \mathbb{B} \big\} \ . \tag{13b}$$

Since the set $\mathbb{B}$ is closed, convex and contains 0, we have [2, Theorem 5.103]

$$\mathbb{B}^{\odot\odot} = \big(\mathbb{B}^\odot \big)^\odot = \mathbb{B} \ , \tag{13c}$$

hence the *bidual norm* $|||\cdot|||_{\star\star} = \big(|||\cdot|||_\star \big)_\star$ is the original norm:

$$|||\cdot|||_{\star\star} = \big(|||\cdot|||_\star \big)_\star = |||\cdot||| \ . \tag{13d}$$

$|||\cdot|||$-duality, normal cone. By definition of the dual norm in (11), we have the inequality

$$\langle x \, , y \rangle \leq |||x||| \times |||y|||_\star \ , \ \ \forall (x, y) \in \mathbb{R}^d \times \mathbb{R}^d \ . \tag{14a}$$

We are interested in the case where this inequality is an equality. One says that $y \in \mathbb{R}^d$ is $|||\cdot|||$-*dual* to $x \in \mathbb{R}^d$, denoted by $y \parallel_{|||\cdot|||} x$, if equality holds in Inequality (14a), that is,

$$y \parallel_{|||\cdot|||} x \iff \langle x \, , y \rangle = |||x||| \times |||y|||_\star \ . \tag{14b}$$

It will be convenient to express this notion of $|||\cdot|||$-duality in terms of geometric objects of convex analysis. For this purpose, we recall that the *normal cone* $N_C(x)$ to the (nonempty) closed convex subset $C \subset \mathbb{R}^d$ at $x \in C$ is the closed convex cone defined by [7, p.136]

$$N_C(x) = \big\{ y \in \mathbb{R}^d \, \big| \, \langle x' - x \, , y \rangle \leq 0 \ , \ \ \forall x' \in C \big\} \ . \tag{15}$$

Now, easy computations show that the notion of $|||\cdot|||$-duality can be rewritten in terms of normal cones $N_{\mathbb{B}}$ and $N_{\mathbb{B}_\star}$ as follows:

$$\Big(y \parallel_{|||\cdot|||} x \iff y \in N_{\mathbb{B}}\big(\frac{x}{|||x|||} \big) \iff x \in N_{\mathbb{B}_\star}\big(\frac{y}{|||y|||} \big) \Big) \ , \ \ \forall (x, y) \in \mathbb{R}^d \backslash \{0\} \times \mathbb{R}^d \backslash \{0\} \ . \tag{16}$$

Restriction norms.

Definition 2 *For any norm $|||\cdot|||$ on $\mathbb{R}^d$ and any subset $K \subset \{1, \ldots, d\}$, we define*

- *the K-*restriction norm *$|||\cdot|||_K$ on the subspace $\mathcal{R}_K$ of $\mathbb{R}^d$, as defined in (5), by*

$$|||x|||_K = |||x||| \ , \ \ \forall x \in \mathcal{R}_K \ . \tag{17}$$

- *the $(K, \star)$-*norm *$|||\cdot|||_{K,\star}$, on the subspace $\mathcal{R}_K$ of $\mathbb{R}^d$, which is the norm $\big(|||\cdot|||_K \big)_\star$, given by the dual norm (on the subspace $\mathcal{R}_K$) of the restriction norm $|||\cdot|||_K$ to the subspace $\mathcal{R}_K$ (first restriction, then dual).*

We have that [5, Equation (14b)]

$$|||y|||_{K,\star} = \sigma_{\mathcal{R}_K \cap \mathbb{B}}(y) = \sigma_{\mathcal{R}_K \cap \mathbb{S}}(y) \ , \ \ \forall y \in \mathcal{R}_K \ . \tag{18}$$

source norm $\|\!|\cdot|\!\|$	$\|\!|\cdot|\!\|^{\mathcal{R}}_{(k)}$	$\|\!|\cdot|\!\|^{\mathcal{R}}_{(k),\star}$
$\|\cdot\|_p$	(p,k)-support norm $\|\|x\|\|^{\text{sn}}_{p,k}$	top (k,q)-norm $\|\|y\|\|^{\text{tn}}_{k,q}$ $= \big(\sum_{j=1}^{k} \|y_{\nu(j)}\|^q\big)^{1/q}$, $1/p + 1/q = 1$
$\|\cdot\|_1$	$(1,k)$-support norm ℓ_1-norm $\|\|x\|\|^{\text{sn}}_{1,k} = \|x\|_1$	top (k,∞)-norm ℓ_∞-norm $\|\|y\|\|^{\text{tn}}_{k,\infty} = \|y_{\nu(1)}\| = \|y\|_\infty$
$\|\cdot\|_2$	$(2,k)$-support norm	top $(k,2)$-norm $\|\|y\|\|^{\text{tn}}_{k,2} = \sqrt{\sum_{j=1}^{k} \|y_{\nu(j)}\|^2}$
$\|\cdot\|_\infty$	(∞,k)-support norm	top $(k,1)$-norm $\|\|y\|\|^{\text{tn}}_{k,1} = \sum_{j=1}^{k} \|y_{\nu(j)}\|$

Table 1: Examples of coordinate-k and dual coordinate-k norms generated by the ℓ_p source norms $|\!|\!|\cdot|\!|\!| = \|\cdot\|_p$ for $p \in [1,\infty]$

3.2 Definition of coordinate-k and dual coordinate-k norms

Source norm. Let $|\!|\!|\cdot|\!|\!|$ be a norm on $\mathbb{R}^d$, that we will call the *source norm.*

Definition of coordinate-k and dual coordinate-k norms.

Definition 3 *For $k \in \{1,\ldots,d\}$, we call* coordinate-k norm *the norm $|\!|\!|\cdot|\!|\!|^{\mathcal{R}}_{(k)}$ whose dual norm is the* dual coordinate-k norm*, denoted by $|\!|\!|\cdot|\!|\!|^{\mathcal{R}}_{(k),\star}$, with expression*

$$|\!|\!|y|\!|\!|^{\mathcal{R}}_{(k),\star} = \sup_{|K|\leq k} |\!|\!|y_K|\!|\!|_{K,\star} \ , \ \ \forall y \in \mathbb{R}^d \ , \tag{19}$$

where the $(K,\star)$-norm $|\!|\!|\cdot|\!|\!|_{K,\star}$ is given in Definition 2, and where the notation $\sup_{|K|\leq k}$ *is a shorthand for* $\sup_{K\subset\{1,\ldots,d\},|K|\leq k}$.

It is easily verified that $|\!|\!|\cdot|\!|\!|^{\mathcal{R}}_{(k),\star}$ indeed is a norm. We will adopt the convention $|\!|\!|\cdot|\!|\!|^{\mathcal{R}}_{(0),\star} = 0$ (although this is not a norm on $\mathbb{R}^d$, but a seminorm).

Examples. Table 1 provides examples [5, 6]. For $y \in \mathbb{R}^d$, ν denotes a permutation of $\{1,\ldots,d\}$ such that $|y_{\nu(1)}| \geq |y_{\nu(2)}| \geq \cdots \geq |y_{\nu(d)}|$. With this, we define the top (k,q)-norms in the last right column of Table 1. The (p,k)-support norm, in the middle column of Table 1, is defined as the dual norm of the top (k,q)-norm, with $1/p + 1/q = 1$.

To prepare our results in Sect. 4, we provide properties of coordinate-k and dual coordinate-k norms.

3.3 Properties of dual coordinate-k norms

We denote the unit sphere and the unit ball of the dual coordinate-k norm $|||\cdot|||^{\mathcal{R}}_{(k),\star}$ in Definition 3 by

$$\mathbb{S}^{\mathcal{R}}_{(k),\star} = \left\{ y \in \mathbb{R}^d \,\middle|\, |||y|||^{\mathcal{R}}_{(k),\star} = 1 \right\} \;, \quad k = 1, \ldots, d \;, \tag{20a}$$

$$\mathbb{B}^{\mathcal{R}}_{(k),\star} = \left\{ y \in \mathbb{R}^d \,\middle|\, |||y|||^{\mathcal{R}}_{(k),\star} \leq 1 \right\} \;, \quad k = 1, \ldots, d \;. \tag{20b}$$

Proposition 4

- *For $k \in \{1, \ldots, d\}$, the dual coordinate-k norm satisfies*

$$|||y|||^{\mathcal{R}}_{(k),\star} = \sup_{|K| \leq k} \sigma_{(\mathcal{R}_K \cap \mathbb{S})}(y) = \sigma_{\ell_0^{\leq k} \cap \mathbb{S}}(y) = \sigma_{\ell_0^{= k} \cap \mathbb{S}}(y) \;, \quad \forall y \in \mathbb{R}^d \;. \tag{21}$$

- *We have the equality*

$$|||\cdot|||_{\star} = |||\cdot|||^{\mathcal{R}}_{(d),\star} \;. \tag{22}$$

- *The sequence $\left\{ |||\cdot|||^{\mathcal{R}}_{(j),\star} \right\}_{j=1,\ldots,d}$ of dual coordinate-k norms in Definition 3 is nondecreasing, in the sense that the following inequalities and equality hold true:*

$$|||y|||^{\mathcal{R}}_{(1),\star} \leq \cdots \leq |||y|||^{\mathcal{R}}_{(j),\star} \leq |||y|||^{\mathcal{R}}_{(j+1),\star} \leq \cdots \leq |||y|||^{\mathcal{R}}_{(d),\star} = |||y|||_{\star} \;, \quad \forall y \in \mathbb{R}^d \;. \tag{23}$$

- *The sequence $\left\{ \mathbb{B}^{\mathcal{R}}_{(j),\star} \right\}_{j=1,\ldots,d}$ of units balls of the dual coordinate-k norms in Definition 3 is nonincreasing, in the sense that the following equality and inclusions hold true:*

$$\mathbb{B}_{\star} = \mathbb{B}^{\mathcal{R}}_{(d),\star} \subset \cdots \subset \mathbb{B}^{\mathcal{R}}_{(j+1),\star} \subset \mathbb{B}^{\mathcal{R}}_{(j),\star} \subset \cdots \subset \mathbb{B}^{\mathcal{R}}_{(1),\star} \;. \tag{24}$$

Proof.

- For any $y \in \mathbb{R}^d$, we have

$$\begin{aligned}
|||y|||^{\mathcal{R}}_{(k),\star} &= \sup_{|K| \leq k} |||y_K|||_{K,\star} && \text{(by definition (19) of } |||y|||^{\mathcal{R}}_{(k),\star}\text{)} \\
&= \sup_{|K| \leq k} \sigma_{(\mathcal{R}_K \cap \mathbb{S})}(y_K) && \text{(as } |||y_K|||_{K,\star} = \sigma_{(\mathcal{R}_K \cap \mathbb{S})}(y_K) \text{ by (18))} \\
&= \sup_{|K| \leq k} \sup_{x \in \mathcal{R}_K \cap \mathbb{S}} \langle x \,, y_K \rangle && \text{(by definition of the support function } \sigma_{(\mathcal{R}_K \cap \mathbb{S})}\text{)} \\
&= \sup_{|K| \leq k} \sup_{x \in \mathcal{R}_K \cap \mathbb{S}} \langle x \,, y \rangle && \text{(by (6) as } x \in \mathcal{R}_K\text{)} \\
&= \sup_{|K| \leq k} \sigma_{(\mathcal{R}_K \cap \mathbb{S})}(y) && \text{(by definition of the support function } \sigma_{(\mathcal{R}_K \cap \mathbb{S})}\text{)} \\
&= \sigma_{\bigcup_{|K| \leq k}(\mathcal{R}_K \cap \mathbb{S})}(y) && \text{(as the support function turns a union of sets into a supremum)} \\
&= \sigma_{\ell_0^{\leq k} \cap \mathbb{S}}(y) \;. && \text{(as } \ell_0^{\leq k} \cap \mathbb{S} = \textstyle\bigcup_{|K| \leq k}(\mathcal{R}_K \cap \mathbb{S}) \text{ by (7))}
\end{aligned}$$

To finish, we will now prove that $\sigma_{\ell_0^{\leq k}\cap\mathbb{S}} = \sigma_{\ell_0^{=k}\cap\mathbb{S}}$. For this purpose, we show in two steps that $\ell_0^{\leq k}\cap\mathbb{S} = \overline{\ell_0^{=k}\cap\mathbb{S}}$.

First, we establish the (known) fact that $\overline{\ell_0^{=k}} = \ell_0^{\leq k}$. The inclusion $\overline{\ell_0^{=k}} \subset \ell_0^{\leq k}$ is easy because, on the one hand, $\ell_0^{=k} \subset \ell_0^{\leq k}$ and, on the other hand, the level set $\ell_0^{\leq k}$ in (4a) is closed, as follows from the well-known property that the pseudonorm ℓ_0 is lower semicontinuous. There remains to prove the reverse inclusion $\ell_0^{\leq k} \subset \overline{\ell_0^{=k}}$. For this purpose, we consider $x \in \ell_0^{\leq k}$. If $x \in \ell_0^{=k}$, obviously $x \in \overline{\ell_0^{=k}}$. Therefore, we suppose that $\ell_0(x) = l < k$. By definition of $\ell_0(x)$ in (2), there exists $L \subset \{1,\ldots,d\}$ such that $|L| = l < k$ and $x = x_L$. For $\epsilon > 0$, define x^ϵ as coinciding with x except for $k-l$ indices outside L for which the components are $\epsilon > 0$. By construction $\ell_0(x^\epsilon) = k$ and $x^\epsilon \to x$ when $\epsilon \to 0$. This proves that $\ell_0^{\leq k} \subset \overline{\ell_0^{=k}}$.

Second, we prove that $\ell_0^{\leq k}\cap\mathbb{S} = \overline{\ell_0^{=k}\cap\mathbb{S}}$. The inclusion $\overline{\ell_0^{=k}\cap\mathbb{S}} \subset \ell_0^{\leq k}\cap\mathbb{S}$, is easy. Indeed, $\overline{\ell_0^{=k}} = \ell_0^{\leq k} \Rightarrow \overline{\ell_0^{=k}\cap\mathbb{S}} \subset \overline{\ell_0^{=k}}\cap\overline{\mathbb{S}} = \ell_0^{\leq k}\cap\mathbb{S}$. To prove the reverse inclusion $\ell_0^{\leq k}\cap\mathbb{S} \subset \overline{\ell_0^{=k}\cap\mathbb{S}}$, we consider $x \in \ell_0^{\leq k}\cap\mathbb{S}$. As we have just seen that $\ell_0^{\leq k} = \overline{\ell_0^{=k}}$, we deduce that $x \in \overline{\ell_0^{=k}}$. Therefore, there exists a sequence $\{z_n\}_{n\in\mathbb{N}}$ in $\ell_0^{=k}$ such that $z_n \to x$ when $n \to +\infty$. Since $x \in \mathbb{S}$, we can always suppose that $z_n \neq 0$, for all $n \in \mathbb{N}$. Therefore $z_n/|||z_n|||$ is well defined and, when $n \to +\infty$, we have $z_n/|||z_n||| \to x/|||x||| = x$ since $x \in \mathbb{S} = \{x \in \mathbb{X} \,|\, |||x||| = 1\}$. Now, on the one hand, $z_n/|||z_n||| \in \ell_0^{=k}$, for all $n \in \mathbb{N}$, and, on the other hand, $z_n/|||z_n||| \in \mathbb{S}$. As a consequence $z_n/|||z_n||| \in \ell_0^{=k}\cap\mathbb{S}$, and we conclude that $x \in \overline{\ell_0^{=k}\cap\mathbb{S}}$. Thus, we have proved that $\ell_0^{\leq k}\cap\mathbb{S} \subset \overline{\ell_0^{=k}\cap\mathbb{S}}$.

From $\ell_0^{\leq k}\cap\mathbb{S} = \overline{\ell_0^{=k}\cap\mathbb{S}}$, we get that $\sigma_{\ell_0^{\leq k}\cap\mathbb{S}} = \sigma_{\overline{\ell_0^{=k}\cap\mathbb{S}}} = \sigma_{\ell_0^{=k}\cap\mathbb{S}}$, by [3, Proposition 7.13]. Thus, we have proved all equalities in (21).

- By the equality $|||y|||^{\mathcal{R}}_{(k),\star} = \sigma_{\ell_0^{\leq k}\cap\mathbb{S}}(y)$ in (21), we get that, for all $y \in \mathbb{R}^d$, $|||y|||^{\mathcal{R}}_{(d),\star} = \sigma_{\ell_0^{\leq d}\cap\mathbb{S}}(y) = \sigma_{\mathbb{S}}(y) = |||y|||_\star$ since $\ell_0^{\leq d} = \mathbb{R}^d$ and by (13a).

- The inequalities in (23) easily derive from the very definition (19) of the dual coordinate-k norms $|||\cdot|||^{\mathcal{R}}_{(k),\star}$. The last equality is just the equality (22).

- The equality and the inclusions in (24) directly follow from the inequalities and the equality between norms in (23).

This ends the proof. □

3.4 Properties of coordinate-k norms

We denote the unit sphere and the unit ball of the coordinate-k norm $|||\cdot|||^{\mathcal{R}}_{(k)}$ by

$$\mathbb{S}^{\mathcal{R}}_{(k)} = \left\{x \in \mathbb{R}^d \,\middle|\, |||x|||^{\mathcal{R}}_{(k)} = 1\right\}, \tag{25a}$$

$$\mathbb{B}^{\mathcal{R}}_{(k)} = \left\{x \in \mathbb{R}^d \,\middle|\, |||x|||^{\mathcal{R}}_{(k)} \leq 1\right\}. \tag{25b}$$

We will adopt the convention $\mathbb{B}^{\mathcal{R}}_{(0)} = \{0\}$ (although this is not the unit ball of a norm on $\mathbb{R}^d$).

Proposition 5

- *For $k \in \{1, \ldots, d\}$, the coordinate-k norm $|||\cdot|||^{\mathcal{R}}_{(k)}$ has unit ball*

$$\mathbb{B}^{\mathcal{R}}_{(k)} = \overline{\mathrm{co}}\Big(\bigcup_{|K| \leq k} (\mathcal{R}_K \cap \mathbb{S}) \Big) , \tag{26}$$

  where $\overline{\mathrm{co}}(S)$ denotes the closed convex hull of a subset $S \subset \mathbb{R}^d$.

- *We have the equality*

$$|||\cdot|||^{\mathcal{R}}_{(d)} = |||\cdot||| . \tag{27}$$

- *The sequence $\Big\{ |||\cdot|||^{\mathcal{R}}_{(j)} \Big\}_{j=1,\ldots,d}$ of coordinate-k norms in Definition 3 is nonincreasing, in the sense that the following equality and inequalities hold true:*

$$|||x||| = |||x|||^{\mathcal{R}}_{(d)} \leq \cdots \leq |||x|||^{\mathcal{R}}_{(j+1)} \leq |||x|||^{\mathcal{R}}_{(j)} \leq \cdots \leq |||x|||^{\mathcal{R}}_{(1)} , \ \ \forall x \in \mathbb{R}^d . \tag{28}$$

- *The sequence $\Big\{ \mathbb{B}^{\mathcal{R}}_{(j)} \Big\}_{j=1,\ldots,d}$ of units balls of the coordinate-k norms in (26) is nondecreasing, in the sense that the following inclusions and equality hold true:*

$$\mathbb{B}^{\mathcal{R}}_{(1)} \subset \cdots \subset \mathbb{B}^{\mathcal{R}}_{(j)} \subset \mathbb{B}^{\mathcal{R}}_{(j+1)} \subset \cdots \subset \mathbb{B}^{\mathcal{R}}_{(d)} = \mathbb{B} . \tag{29}$$

Proof.

- For any $y \in \mathbb{R}^d$, we have

$$\begin{aligned} |||y|||^{\mathcal{R}}_{(k),\star} &= \sup_{|K| \leq k} \sigma_{(\mathcal{R}_K \cap \mathbb{S})}(y) && \text{(by (21))} \\ &= \sigma_{\bigcup_{|K| \leq k} (\mathcal{R}_K \cap \mathbb{S})}(y) && \text{(as the support function turns a union of sets into a supremum)} \\ &= \sigma_{\overline{\mathrm{co}}\left(\bigcup_{|K| \leq k} (\mathcal{R}_K \cap \mathbb{S})\right)}(y) && \text{(by [3, Proposition 7.13])} \end{aligned}$$

and we conclude that $\mathbb{B}^{\mathcal{R}}_{(k)} = \overline{\mathrm{co}}\big(\bigcup_{|K| \leq k} (\mathcal{R}_K \cap \mathbb{S}) \big)$ by (13a). Thus, we have proved (26).

- From the equality (22), we deduce the equality (27) between the dual norms by (11).
- The equality and inequalities between norms in (28) easily derive from the inclusions and equality between unit balls in (29).
- The inclusions and equality between unit balls in (29) directly follow from the inclusions and equality between unit balls in (24) and from (13b), as $\mathbb{B}^{\mathcal{R}}_{(j)} = \big(\mathbb{B}^{\mathcal{R}}_{(j),\star} \big)^{\odot}$, the polar set of $\mathbb{B}^{\mathcal{R}}_{(j),\star}$.

This ends the proof. □

We recall that the normed space $\big(\mathbb{R}^d, |||\cdot|||\big)$ is said to be *strictly convex* if the unit ball $\mathbb{B}$ (of the norm $|||\cdot|||$) is *rotund*, that is, if all points of the unit sphere $\mathbb{S}$ are extreme points of the unit ball $\mathbb{B}$. The normed space $\big(\mathbb{R}^d, \|\cdot\|_p\big)$, equipped with the ℓ_p-norm $\|\cdot\|_p$ (for $p \in [1, \infty]$), is strictly convex if and only if $p \in]1, \infty[$.

We now show that the sequences $\Big\{ |||\cdot|||^{\mathcal{R}}_{(j)} \Big\}_{j=1,\ldots,d}$ of coordinate-k norms (in Definition 3) are naturally decreasingly graded with respect to the ℓ_0 pseudonorm (as in Definition 1). Part of the proof relies upon the forthcoming Lemma 7.

Proposition 6

1. *The nonincreasing sequence* $\left\{ \|\!|\cdot|\!\|^{\mathcal{R}}_{(j)} \right\}_{j=1,\ldots,d}$ *of coordinate-k norms is decreasingly graded with respect to the* ℓ_0 *pseudonorm, that is, for any* $l = 1, \ldots, d$,

$$\ell_0(x) \leq l \Rightarrow |\!|\!| x |\!|\!| = |\!|\!| x |\!|\!|^{\mathcal{R}}_{(l)} \; , \; \forall x \in \mathbb{R}^d \; . \tag{30a}$$

2. *If the normed space* $\left(\mathbb{R}^d, |\!|\!|\cdot|\!|\!|\right)$ *is strictly convex, then the nonincreasing sequence* $\left\{ |\!|\!|\cdot|\!|\!|^{\mathcal{R}}_{(j)} \right\}_{j=1,\ldots,d}$ *of coordinate-k norms is strictly decreasingly graded with respect to the* ℓ_0 *pseudonorm, that is, for any* $l = 1, \ldots, d$,

$$\ell_0(x) \leq l \iff |\!|\!| x |\!|\!| = |\!|\!| x |\!|\!|^{\mathcal{R}}_{(l)} \; , \; \forall x \in \mathbb{R}^d \; . \tag{30b}$$

Proof.
• We prove Item 1. As the sequence $\left\{ |\!|\!|\cdot|\!|\!|^{\mathcal{R}}_{(j)} \right\}_{j=1,\ldots,d}$ of coordinate-k norms is nonincreasing by (23), it suffices to show that (8b) holds true — that is, that (30a) holds true — to prove that the sequence is decreasingly graded with respect to the ℓ_0 pseudonorm (see Definition 1).

Now, for any $x \in \mathbb{R}^d$ and for any $k \in \{1, \ldots, d\}$, we have[2]

$$
\begin{aligned}
x \in \ell_0^{\leq k} \iff\; & x = 0 \text{ or } \frac{x}{|\!|\!| x |\!|\!|} \in \ell_0^{\leq k} \\
& \text{(by 0-homogeneity (3) of the } \ell_0 \text{ pseudonorm, and by definition (4a) of } \ell_0^{\leq k}) \\
\iff\; & x = 0 \text{ or } \frac{x}{|\!|\!| x |\!|\!|} \in \ell_0^{\leq k} \cap \mathbb{S} \quad \text{(as } \tfrac{x}{|\!|\!| x |\!|\!|} \in \mathbb{S} \text{ by definition (10a) of the unit sphere } \mathbb{S}) \\
\iff\; & x = 0 \text{ or } \frac{x}{|\!|\!| x |\!|\!|} \in \bigcup_{|K| \leq k} (\mathcal{R}_K \cap \mathbb{S}) \quad \text{(as } \ell_0^{\leq k} = \textstyle\bigcup_{|K|\leq k} \mathcal{R}_K \text{ by (7))} \\
\Rightarrow\; & x = 0 \text{ or } \frac{x}{|\!|\!| x |\!|\!|} \in \mathbb{B}^{\mathcal{R}}_{(k)} \quad \text{(as } \mathbb{B}^{\mathcal{R}}_{(k)} = \overline{\mathrm{co}}\big(\textstyle\bigcup_{|K|\leq k} (\mathcal{R}_K \cap \mathbb{S})\big) \text{ by (26))} \\
\Rightarrow\; & x = 0 \text{ or } \left|\!\left|\!\left| \frac{x}{|\!|\!| x |\!|\!|} \right|\!\right|\!\right|^{\mathcal{R}}_{(k)} \leq 1 \quad \text{(since } \mathbb{B}^{\mathcal{R}}_{(k)} \text{ is the unit ball of the norm } |\!|\!|\cdot|\!|\!|^{\mathcal{R}}_{(k)} \text{ by (25b))} \\
\Rightarrow\; & |\!|\!| x |\!|\!|^{\mathcal{R}}_{(k)} \leq |\!|\!| x |\!|\!| \\
\Rightarrow\; & |\!|\!| x |\!|\!|^{\mathcal{R}}_{(k)} \leq |\!|\!| x |\!|\!| = |\!|\!| x |\!|\!|^{\mathcal{R}}_{(d)} \quad \text{(where the last equality comes from (28))} \\
\Rightarrow\; & |\!|\!| x |\!|\!|^{\mathcal{R}}_{(k)} = |\!|\!| x |\!|\!|^{\mathcal{R}}_{(d)} \; . \quad \text{(as } |\!|\!| x |\!|\!|^{\mathcal{R}}_{(k)} \geq |\!|\!| x |\!|\!|^{\mathcal{R}}_{(d)} \text{ by (28))}
\end{aligned}
$$

Therefore, we have obtained (30a).

• We prove Item 2. As the sequence $\left\{ |\!|\!|\cdot|\!|\!|^{\mathcal{R}}_{(j)} \right\}_{j=1,\ldots,d}$ of coordinate-k norms is nonincreasing by (23), it suffices to show that (9b) holds true — that is, that (30b) holds true — to prove that the sequence is strictly decreasingly graded with respect to the ℓ_0 pseudonorm (see Definition 1).

[2] In what follows, by "or", we mean the so-called *exclusive or* (exclusive disjunction). Thus, every "or" should be understood as "or $x \neq 0$ and".

We suppose that the normed space $\left(\mathbb{R}^d, |||\cdot|||\right)$ is strictly convex. Then, for any $x \in \mathbb{R}^d$ and for any $k \in \{1, \dots, d\}$, we have[3]

$$\begin{aligned}
x \in \ell_0^{\leq k} &\iff x = 0 \text{ or } \frac{x}{|||x|||} \in \ell_0^{\leq k} \\
&\qquad \text{(by 0-homogeneity (3) of the } \ell_0 \text{ pseudonorm, and by definition (4a) of } \ell_0^{\leq k}\text{)} \\
&\iff x = 0 \text{ or } \frac{x}{|||x|||} \in \ell_0^{\leq k} \cap \mathbb{S} \qquad \text{(as } \tfrac{x}{|||x|||} \in \mathbb{S} \text{ by definition (10a) of the unit sphere } \mathbb{S}\text{)} \\
&\iff x = 0 \text{ or } \frac{x}{|||x|||} \in \mathbb{B}^{\mathcal{R}}_{(k)} \cap \mathbb{S}
\end{aligned}$$

as $\ell_0^{\leq k} \cap \mathbb{S} = \mathbb{B}^{\mathcal{R}}_{(k)} \cap \mathbb{S}$ by (32) since the assumption of Lemma 7 is satisfied, that is, the normed space $\left(\mathbb{R}^d, |||\cdot|||\right)$ is strictly convex

$$\begin{aligned}
&\iff x = 0 \text{ or } \left|\left|\left|\frac{x}{|||x|||}\right|\right|\right|^{\mathcal{R}}_{(k)} \leq 1 \quad \text{(since } \mathbb{B}^{\mathcal{R}}_{(k)} \text{ is the unit ball of the norm } |||\cdot|||^{\mathcal{R}}_{(k)} \text{ by (25b))} \\
&\iff |||x|||^{\mathcal{R}}_{(k)} \leq |||x||| \\
&\iff |||x|||^{\mathcal{R}}_{(k)} \leq |||x||| = |||x|||^{\mathcal{R}}_{(d)} \qquad \text{(where the last equality comes from (28))} \\
&\iff |||x|||^{\mathcal{R}}_{(k)} = |||x|||^{\mathcal{R}}_{(d)} \; . \qquad \text{(as } |||x|||^{\mathcal{R}}_{(k)} \geq |||x|||^{\mathcal{R}}_{(d)} \text{ by (28))}
\end{aligned}$$

Therefore, we have obtained (30b).

This ends the proof. □

	$\left\{|||\cdot|||^{\mathcal{R}}_{(j)}\right\}_{j=1,\dots,d}$ graded	strictly graded
$|||\cdot|||$ is any norm	✓	
$(\mathbb{R}^d, |||\cdot|||)$ is strictly convex		✓

Table 2: Table of results. It reads as follows: to obtain that the sequence $\left\{|||\cdot|||^{\mathcal{R}}_{(j)}\right\}_{j=1,\dots,d}$ be graded (second column), it suffices that $|||\cdot|||$ be any norm; to obtain that the sequence $\left\{|||\cdot|||^{\mathcal{R}}_{(j)}\right\}_{j=1,\dots,d}$ be strictly graded (third column), it suffices that $(\mathbb{R}^d, |||\cdot|||)$ be strictly convex.

Table 2 summarizes the results of Proposition 6. As an application with any ℓ_p-norm $\|\cdot\|_p$ for source norm (for $p \in [1, \infty]$), we obtain that the nonincreasing sequence $\left\{||\cdot||^{\mathrm{sn}}_{p,j}\right\}_{j=1,\dots,d}$ of (p, k)-support norms (see Table 1) is strictly decreasingly graded w.r.t. the ℓ_0 pseudonorm for $p \in]1, \infty[$. This gives, by (9c):

$$\ell_0(x) = \min\left\{k \in \{1, \dots, d\} \,\middle|\, ||x||^{\mathrm{sn}}_{p,k} = ||x||_p\right\} \; , \; \forall x \in \mathbb{R}^d \; , \; \forall p \in]1, \infty[\; . \tag{31a}$$

[3]See Footnote 2.

We also have that the sequence $\big\{ ||\cdot||^{\mathrm{sn}}_{p,j} \big\}_{j=1,\ldots,d}$ is decreasingly graded with respect to the ℓ_0 pseudonorm for $p \in [1,\infty]$. Looking at Table 1, the only interesting case is for $p=\infty$, giving, by (8c):

$$\ell_0(x) \geq \min \Big\{ k \in \{1,\ldots,d\} \,\Big|\, ||x||^{\mathrm{sn}}_{\infty,k} = ||x||_\infty \Big\} \; , \;\; \forall x \in \mathbb{R}^d \; . \tag{31b}$$

Lemma 7 *Let* $|||\cdot|||$ *be a norm on* $\mathbb{R}^d$*. If the normed space* $\big(\mathbb{R}^d, |||\cdot|||\big)$ *is strictly convex, we have the equality*

$$\ell_0^{\leq k} \cap \mathbb{S} = \mathbb{B}^{\mathcal{R}}_{(k)} \cap \mathbb{S} \; , \;\; \forall k \in \{0,1,\ldots,d\} \; , \tag{32}$$

where $\ell_0^{\leq k}$ *is the level set in* (4a) *of the* ℓ_0 *pseudonorm in* (2)*, where* $\mathbb{S}$ *is the unit sphere in* (10a)*, and where* $\mathbb{B}^{\mathcal{R}}_{(k)}$ *in* (25b) *is the unit ball of the norm* $|||\cdot|||^{\mathcal{R}}_{(k)}$*.*

Proof. It is proved in [5, Proposition 16] that, if the unit ball $\mathbb{B}$ is rotund — that is, if the normed space $\big(\mathbb{R}^d, |||\cdot|||\big)$ is strictly convex — and if A is a closed subset of $\mathbb{S}$, then $A = \overline{\mathrm{co}}(A) \cap \mathbb{S}$.

Now, we turn to the proof. First, we observe that the level set $\ell_0^{\leq k}$ is closed because the pseudonorm ℓ_0 is lower semi continuous. Second, we have

$$\ell_0^{\leq k} \cap \mathbb{S} = \overline{\mathrm{co}}\big(\ell_0^{\leq k} \cap \mathbb{S}\big) \cap \mathbb{S}$$

(because $\ell_0^{\leq k} \cap \mathbb{S} \subset \mathbb{S}$ and is closed, and because the unit ball $\mathbb{B}$ is rotund)

$$= \overline{\mathrm{co}}\big(\bigcup_{|K| \leq k} (\mathcal{R}_K \cap \mathbb{S}) \big) \cap \mathbb{S} \qquad \text{(by (7))}$$

$$= \mathbb{B}^{\mathcal{R}}_{(k)} \cap \mathbb{S} \; . \qquad \text{(by (26))}$$

This ends the proof. □

4 The Capra-conjugacy and the ℓ_0 pseudonorm

We introduce the coupling Capra in §4.1. Then, we provide formulas for Capra-conjugates of functions of the ℓ_0 pseudonorm in §4.2, for Capra-subdifferentials of functions of the ℓ_0 pseudonorm in §4.4, and for Capra-biconjugates of functions of the ℓ_0 pseudonorm in §4.3.

We work on the Euclidian space $\mathbb{R}^d$ (with $d \in \mathbb{N}^*$), equipped with the scalar product $\langle \cdot \,, \cdot \rangle$ (but not necessarily with the Euclidian norm).

4.1 Constant along primal rays coupling (Capra)

Following [4], we introduce the coupling Capra, which is a special case of one-sided linear coupling, as defined in §A.3. Fenchel-Moreau conjugacies are recalled in §A.2.

Definition 8 *Let* $|||\cdot|||$ *be a norm on* $\mathbb{R}^d$*. We define the* constant along primal rays coupling ¢*, or Capra, between* $\mathbb{R}^d$ *and* $\mathbb{R}^d$ *by*

$$\forall y \in \mathbb{R}^d \; , \; \begin{cases} ¢(x,y) & = \dfrac{\langle x \, , y \rangle}{|||x|||} \; , \; \forall x \in \mathbb{R}^d \backslash \{0\} \; , \\ ¢(0,y) & = 0. \end{cases} \tag{33}$$

We stress the point that, in (33), the Euclidian scalar product $\langle x \, , y \rangle$ and the norm term $|||x|||$ need not be related, that is, the norm $|||\cdot|||$ is not necessarily Euclidian.

The coupling Capra has the property of being constant along primal rays, hence the acronym Capra (Constant Along Primal RAys). We introduce the primal *normalization mapping* n, from $\mathbb{R}^d$ towards the unit sphere $\mathbb{S}$ united with $\{0\}$, as follows:

$$n : \mathbb{R}^d \to \mathbb{S} \cup \{0\} \; , \; n(x) = \begin{cases} \frac{x}{|||x|||} & \text{if } x \neq 0 \; , \\ 0 & \text{if } x = 0 \; . \end{cases} \tag{34}$$

With these notations, the coupling Capra in (33) is a special case of one-sided linear coupling c_n, as in (57b) with $\theta = n$, the *Fenchel coupling after primal normalization*:

$$¢(x,y) = c_n(x,y) = \langle n(x) \, , y \rangle \; , \; \forall x \in \mathbb{R}^d \; , \; \forall y \in \mathbb{R}^d \; .$$

We will see below that the Capra-conjugacy, induced by the coupling Capra, shares some relations with the Fenchel conjugacy (see §A.2.2).

Capra-conjugates and biconjugates. Here are expressions for the Capra-conjugates and biconjugates of a function. The following Proposition simply is Proposition 19 (in the Appendix) in the case where the mapping θ is the normalization mapping n in (34).

In the whole paper, we use $\overline{\mathbb{R}} = [-\infty, +\infty]$.

Proposition 9 *For any function* $g : \mathbb{R}^d \to \overline{\mathbb{R}}$*, the* $¢'$*-Fenchel-Moreau conjugate is given by*

$$g^{¢'} = g^{\star'} \circ n \; . \tag{35a}$$

For any function $f : \mathbb{R}^d \to \overline{\mathbb{R}}$*, the* ¢*-Fenchel-Moreau conjugate is given by*

$$f^{¢} = \big(\inf \big[f \mid n \big] \big)^{\star} \; , \tag{35b}$$

where the epi-composition $\inf \big[f \mid n \big]$*, defined in* (55a)*, has here the expression*

$$\inf \big[f \mid n \big](x) = \inf \big\{ f(x') \, \big| \, n(x') = x \big\} = \begin{cases} \inf_{\lambda > 0} f(\lambda x) & \text{if } x \in \mathbb{S} \cup \{0\} \; , \\ +\infty & \text{if } x \notin \mathbb{S} \cup \{0\} \; , \end{cases} \tag{35c}$$

and the ¢*-Fenchel-Moreau biconjugate is given by*

$$f^{¢¢'} = \big(f^{¢} \big)^{\star'} \circ n = \big(\inf \big[f \mid n \big] \big)^{\star\star'} \circ n \; . \tag{35d}$$

We observe that the ¢-Fenchel-Moreau conjugate $f^{¢}$ is a closed convex function on $\mathbb{R}^d$ (see §A.2.2).

Capra-convex functions. We recall that so-called ¢-*convex functions* are all functions of the form $g^{¢'}$, for any $g : \mathbb{R}^d \to \overline{\mathbb{R}}$, or, equivalently, all functions of the form $f^{¢¢'}$, for any $f : \mathbb{R}^d \to \overline{\mathbb{R}}$, or, equivalently, all functions that are equal to their ¢-biconjugate ($f^{¢¢'} = f$) [13, 12, 8]. From §A.3 in the Appendix, and especially Proposition 20, we easily deduce the following result.

We recall that a function is closed convex on $\mathbb{R}^d$ if and only if it is either a proper convex lower semi continuous (lsc) function or one of the two constant functions $-\infty$ or $+\infty$ (see §A.2.2).

Proposition 10 *A function is ¢-convex if and only if it is the composition of a closed convex function on* $\mathbb{R}^d$ *with the normalization mapping* (34)*. More precisely, for any function* $h : \mathbb{R}^d \to \overline{\mathbb{R}}$*, we have the equivalences*

$$
\begin{aligned}
& h \textit{ is ¢-convex} \\
& \quad \Leftrightarrow h = h^{¢¢'} \\
& \quad \Leftrightarrow h = \big(h^{¢}\big)^{\star'} \circ n \textit{ (where } \big(h^{¢}\big)^{\star'} \textit{ is a closed convex function)} \\
& \quad \Leftrightarrow \textit{there exists a closed convex function } f : \mathbb{R}^d \to \overline{\mathbb{R}} \textit{ such that } h = f \circ n \; .
\end{aligned}
$$

Capra-subdifferential. Following the definition of the subdifferential of a function with respect to a duality in [1], and the expressions in (61) for a one-sided linear coupling, the *Capra-subdifferential* of the function $f : \mathbb{R}^d \to \overline{\mathbb{R}}$ at $x \in \mathbb{R}^d$ has the following expressions

$$
\partial_{¢} f(x) = \big\{ y \in \mathbb{R}^d \,\big|\, ¢(x', y) \dotplus \big(-f(x')\big) \leq ¢(x, y) \dotplus \big(-f(x)\big) \; , \; \forall x' \in \mathbb{R}^d \big\} \tag{37a}
$$

$$
= \big\{ y \in \mathbb{R}^d \,\big|\, f^{¢}(y) = ¢(x, y) \dotplus \big(-f(x)\big) \big\} \tag{37b}
$$

$$
= \big\{ y \in \mathbb{R}^d \,\big|\, \big(\inf \big[f \mid n\big]\big)^{\star}(y) = \langle n(x) \, , y \rangle \dotplus \big(-f(x)\big) \big\} \; , \tag{37c}
$$

so that, thanks to the definition (34) of the normalization mapping n, we deduce that

$$
\partial_{¢} f(0) = \big\{ y \in \mathbb{R}^d \,\big|\, \big(\inf \big[f \mid n\big]\big)^{\star}(y) = -f(0) \big\} \tag{37d}
$$

$$
\partial_{¢} f(x) = \Big\{ y \in \mathbb{R}^d \,\Big|\, \big(\inf \big[f \mid n\big]\big)^{\star}(y) = \frac{\langle x \, , y \rangle}{|||x|||} \dotplus \big(-f(x)\big) \Big\} \; , \; \forall x \in \mathbb{R}^d \backslash \{0\} \; . \tag{37e}
$$

Now, we turn to analyze the ℓ_0 pseudonorm by means of the Capra conjugacy.

4.2 Capra-conjugates related to the ℓ_0 pseudonorm

With the Fenchel conjugacy, we easily get that $\delta^{\star}_{\ell_0^{\leq k}} = +\infty$, for all $k = 0, 1, \ldots, d$ — where $\delta_{\ell_0^{\leq k}}$ is the characteristic function of the level sets (4a) as defined in (56) — and that $\ell_0^{\star} = \delta_{\{0\}}$. Hence, the Fenchel conjugacy does not seem to be suitable to handle the ℓ_0 pseudonorm. We will see that we obtain more interesting formulas with the Capra-conjugacy.

More precisely, we will now show that functions of the ℓ_0 pseudonorm in (2) — including the ℓ_0 pseudonorm itself and the characteristic functions $\delta_{\ell_0^{\leq k}}$ of its level sets (4a) — are related to the sequence of dual coordinate-k norms in Definition 3 by the following Capra-conjugacy formulas.

Proposition 11 *Let $|||\cdot|||$ be a norm on $\mathbb{R}^d$, with associated sequence $\left\{|||\cdot|||_{(j),\star}^{\mathcal{R}}\right\}_{j=1,\ldots,d}$ of dual coordinate-k norms in Definition 3, and associated coupling ¢ in (33).*

For any function $\varphi : \{0,1,\ldots,d\} \to \overline{\mathbb{R}}$, we have

$$(\varphi \circ \ell_0)^{¢} = \sup_{j=0,1,\ldots,d} \left[|||\cdot|||_{(j),\star}^{\mathcal{R}} - \varphi(j) \right] , \tag{38}$$

where we adopt the convention $|||\cdot|||_{(0),\star}^{\mathcal{R}} = 0$.

Proof. We prove (38):

$$(\varphi \circ \ell_0)^{¢} = \Big(\inf_{j=0,1,\ldots,d} \big[\delta_{\ell_0^{=j}} \dotplus \varphi(j) \big] \Big)^{¢}$$

because $\varphi \circ \ell_0 = \inf_{j=0,1,\ldots,d} \big[\delta_{\ell_0^{=j}} \dotplus \varphi(j) \big]$ since $\varphi \circ \ell_0$ takes the values $\varphi(j)$ on the level curves $\ell_0^{=j}$ of ℓ_0 in (4b)

$$\begin{aligned}
&= \sup_{j=0,1,\ldots,d} \big[\delta_{\ell_0^{=j}} \dotplus \varphi(j) \big]^{¢} && \text{(as conjugacies, being dualities, turn infima into suprema)} \\
&= \sup_{j=0,1,\ldots,d} \big[\delta_{\ell_0^{=j}}^{¢} \dotplus (-\varphi(j)) \big] && \text{(by property of conjugacies)} \\
&= \sup_{j=0,1,\ldots,d} \big[\sigma_{n(\ell_0^{=j})} \dotplus (-\varphi(j)) \big] && \text{(as } \delta_{\ell_0^{=j}}^{¢} = \sigma_{n(\ell_0^{=j})} \text{ by (59d))} \\
&= \sup_{j=0,1,\ldots,d} \Big\{ \sup \big\{ 0, \sigma_{\ell_0^{=j} \cap \mathbb{S}} \big\} \dotplus (-\varphi(j)) \Big\}
\end{aligned}$$

as $n(\ell_0^{=j}) = \{0\} \cup \big(\ell_0^{=j} \cap \mathbb{S}\big)$ by (34), and as the support function turns a union of sets into a supremum

$$\begin{aligned}
&= \sup_{j=0,1,\ldots,d} \big\{ \sigma_{\ell_0^{=j} \cap \mathbb{S}} \dotplus (-\varphi(j)) \big\} && \text{(as } \sigma_{\ell_0^{=j} \cap \mathbb{S}} \geq 0 \text{ since } \ell_0^{=j} \cap \mathbb{S} = -\big(\ell_0^{=j} \cap \mathbb{S}\big)) \\
&= \sup \Big\{ -\varphi(0), \sup_{j=1,\ldots,d} \Big[|||y|||_{(j),\star}^{\mathcal{R}} - \varphi(j) \Big] \Big\} && \text{(as } \sigma_{\ell_0^{=j} \cap \mathbb{S}} = |||\cdot|||_{(j),\star}^{\mathcal{R}} \text{ by (21))} \\
&= \sup_{j=0,1,\ldots,d} \Big[|||y|||_{(j),\star}^{\mathcal{R}} - \varphi(j) \Big] . && \text{(using the convention that } |||\cdot|||_{(0),\star}^{\mathcal{R}} = 0)
\end{aligned}$$

This ends the proof. □

With φ the identity function on $\{0,1,\ldots,d\}$, we find the Capra-conjugate of the ℓ_0 pseudonorm. With the functions $\varphi = \delta_{\{0,1,\ldots,k\}}$ (for any $k \in \{0,1,\ldots,d\}$), we find the Capra-conjugates of the characteristic functions $\delta_{\ell_0^{\leq k}}$ of its level sets (4a). The corresponding expressions are given in Table 3.

4.3 Capra-biconjugates related to the ℓ_0 pseudonorm

With the Fenchel conjugacy, we easily get that $\delta_{\ell_0^{\leq k}}^{\star\star'} = -\infty$, for all $k = 0, 1, \ldots, d$, and that $\ell_0^{\star\star'} = 0$. Hence, the Fenchel conjugacy does not seem to be suitable to handle the ℓ_0 pseudonorm. We will see that we obtain more interesting formulas with the Capra-conjugacy.

More precisely, we will now show that functions of the ℓ_0 pseudonorm in (2) — including the ℓ_0 pseudonorm itself and the characteristic functions $\delta_{\ell_0^{\leq k}}$ of its level sets (4a) — are related to the sequences of coordinate-k norms and dual coordinate-k norms in Definition 3 by the following Capra-biconjugacy formulas.

Proposition 12 *Let $|||\cdot|||$ be a norm on $\mathbb{R}^d$, with associated sequence $\left\{|||\cdot|||_{(j)}^{\mathcal{R}}\right\}_{j=1,\ldots,d}$ of coordinate-k norms and sequence $\left\{|||\cdot|||_{\star(j)}^{\mathcal{R}}\right\}_{j=1,\ldots,d}$ of dual coordinate-k norms, as in Definition 3, and with associated Capra coupling ¢ in (33).*

1. *For any function $\varphi : \{0, 1, \ldots, d\} \to \overline{\mathbb{R}}$, we have*

$$(\varphi \circ \ell_0)^{¢¢'}(x) = \big((\varphi \circ \ell_0)^{¢}\big)^{\star'}\big(\frac{x}{|||x|||}\big) \; , \quad \forall x \in \mathbb{R}^d \backslash \{0\} \; , \tag{39a}$$

where the closed convex function $\big((\varphi \circ \ell_0)^{¢}\big)^{\star'}$ has the following expression as a Fenchel conjugate

$$\big((\varphi \circ \ell_0)^{¢}\big)^{\star'} = \Big(\sup_{j=0,1,\ldots,d} \big[|||\cdot|||_{(j),\star}^{\mathcal{R}} - \varphi(j) \big] \Big)^{\star'} \; , \tag{39b}$$

and also has the following four expressions as a Fenchel biconjugate

$$= \Big(\inf_{j=0,1,\ldots,d} \big[\delta_{\mathbb{B}_{(j)}^{\mathcal{R}}} \dot{+} \varphi(j) \big] \Big)^{\star\star'} \; , \tag{39c}$$

hence the function $\big((\varphi \circ \ell_0)^{¢}\big)^{\star'}$ is the largest closed convex function below the integer valued function $\inf_{j=0,1,\ldots,d} \big[\delta_{\mathbb{B}_{(j)}^{\mathcal{R}}} \dot{+} \varphi(j) \big]$, which is such that $x \in \mathbb{B}_{(j)}^{\mathcal{R}} \backslash \mathbb{B}_{(j-1)}^{\mathcal{R}} \mapsto \varphi(j)$ for $l = 1, \ldots, d$, and $x \in \mathbb{B}_{(0)}^{\mathcal{R}} = \{0\} \mapsto \varphi(0)$, the function being infinite outside $\mathbb{B}_{(d)}^{\mathcal{R}} = \mathbb{B}$ (the above construction makes sense as $\mathbb{B}_{(1)}^{\mathcal{R}} \subset \cdots \subset \mathbb{B}_{(j-1)}^{\mathcal{R}} \subset \mathbb{B}_{(j)}^{\mathcal{R}} \subset \cdots \subset \mathbb{B}_{(d)}^{\mathcal{R}} = \mathbb{B}$ by (24)), that is,

$$= \Big(x \mapsto \inf \big\{ \varphi(j) \, \big| \, x \in \mathbb{B}_{(j)}^{\mathcal{R}} \; , \; j \in \{0, 1, \ldots, d\} \big\} \Big)^{\star\star'} \; , \tag{39d}$$

with the convention that $\mathbb{B}_{(0)}^{\mathcal{R}} = \{0\}$ and that $\inf \emptyset = +\infty$

$$= \Big(\inf_{j=0,1,\ldots,d} \big[\delta_{\mathbb{S}_{(j)}^{\mathcal{R}}} \dot{+} \varphi(j) \big] \Big)^{\star\star'} \; , \tag{39e}$$

hence the function $\big((\varphi\circ\ell_0)^{¢}\big)^{\star'}$ *is the largest closed convex function below the integer valued function* $\inf_{j=0,1,\ldots,d}\big[\delta_{\mathbb{S}^{\mathcal{R}}_{(j)}} \dotplus \varphi(j)\big]$, *that is,*

$$= \Big(x \mapsto \inf\big\{\varphi(j)\,\big|\, x\in\mathbb{S}^{\mathcal{R}}_{(j)}\ ,\ \ j\in\{0,1,\ldots,d\}\big\}\Big)^{\star\star'} , \tag{39f}$$

with the convention that $\mathbb{S}^{\mathcal{R}}_{(0)} = \{0\}$ *and that* $\inf\emptyset = +\infty$.

(39g)

2. *For any function* $\varphi : \{0,1,\ldots,d\} \to \mathbb{R}$, *that is, with finite values, the function* $\big((\varphi\circ\ell_0)^{¢}\big)^{\star'}$ *is proper convex lsc and has the following variational expression*

$$\big((\varphi\circ\ell_0)^{¢}\big)^{\star'}(x) = \min_{\substack{(\lambda_0,\lambda_1,\ldots,\lambda_d)\in\Delta_{d+1}\\ x\in\sum_{j=1}^{d}\lambda_j\mathbb{B}^{\mathcal{R}}_{(j)}}} \sum_{j=0}^{d}\lambda_j\varphi(j)\ ,\ \ \forall x\in\mathbb{R}^d\ , \tag{39h}$$

where Δ_{d+1} *denotes the simplex of* $\mathbb{R}^{d+1}$.

3. *For any function* $\varphi : \{0,1,\ldots,d\} \to \mathbb{R}_+$, *that is, with nonnegative finite values, and such that* $\varphi(0) = 0$, *the function* $\big((\varphi\circ\ell_0)^{¢}\big)^{\star'}$ *is proper convex lsc and has the following two variational expressions*[4]

$$\big((\varphi\circ\ell_0)^{¢}\big)^{\star'}(x) = \min_{\substack{(\lambda_0,\lambda_1,\ldots,\lambda_d)\in\Delta_{d+1}\\ x\in\sum_{j=1}^{d}\lambda_j\mathbb{S}^{\mathcal{R}}_{(j)}}} \sum_{j=1}^{d}\lambda_j\varphi(j)\ ,\ \ \forall x\in\mathbb{R}^d\ , \tag{39i}$$

$$= \min_{\substack{z^{(1)}\in\mathbb{R}^d,\ldots,z^{(d)}\in\mathbb{R}^d\\ \sum_{j=1}^{d}|||z^{(j)}|||^{\mathcal{R}}_{(j)}\leq 1\\ \sum_{j=1}^{d}z^{(j)}=x}} \sum_{j=1}^{d}\varphi(j)|||z^{(j)}|||^{\mathcal{R}}_{(j)}\ ,\ \ \forall x\in\mathbb{R}^d\ , \tag{39j}$$

and the function $(\varphi\circ\ell_0)^{¢¢'}$ *has the following variational expression*

$$(\varphi\circ\ell_0)^{¢¢'}(x) = \frac{1}{|||x|||}\min_{\substack{z^{(1)}\in\mathbb{R}^d,\ldots,z^{(d)}\in\mathbb{R}^d\\ \sum_{j=1}^{d}|||z^{(j)}|||^{\mathcal{R}}_{(j)}\leq |||x|||\\ \sum_{j=1}^{d}z^{(j)}=x}} \sum_{j=1}^{d}|||z^{(j)}|||^{\mathcal{R}}_{(j)}\varphi(j)\ ,\ \ \forall x\in\mathbb{R}^d\backslash\{0\}\ . \tag{40}$$

Proof. We first note that $(\varphi\circ\ell_0)^{¢¢'} = \big((\varphi\circ\ell_0)^{¢}\big)^{\star'}\circ n$, by (35d), and we study $\big((\varphi\circ\ell_0)^{¢}\big)^{\star'}$.

[4] In (39h), the sum starts from $j = 0$, whereas in (39i) and in (39j), the sum starts from $j = 1$.

1. Let the function $\varphi : \{0,1,\ldots,d\} \to \overline{\mathbb{R}}$ be given. The equality (39a) is a straightforward consequence of the expression (35d) for a Capra-biconjugate, and of the fact that $n(x) = \frac{x}{|||x|||}$ when $x \neq 0$ by (34).

   We have

$$\big((\varphi \circ \ell_0)^{¢}\big)^{\star'} = \Big(\sup_{j=0,1,\ldots,d} \big[|||\cdot|||^{\mathcal{R}}_{(j),\star} - \varphi(j)\big]\Big)^{\star'} \qquad \text{(by (38))}$$
$$= \Big(\sup_{j=0,1,\ldots,d} \big[\sigma_{\mathbb{B}^{\mathcal{R}}_{(j)}} - \varphi(j)\big]\Big)^{\star'}$$

(by (13a) as $\mathbb{B}^{\mathcal{R}}_{(j)}$ is the unit ball of the norm $|||\cdot|||^{\mathcal{R}}_{(j)}$ by (25b), and with the convention $\mathbb{B}^{\mathcal{R}}_{(0)} = \{0\}$)

$$= \Big(\sup_{j=0,1,\ldots,d} \big[\delta^{\star}_{\mathbb{B}^{\mathcal{R}}_{(j)}} - \varphi(j)\big]\Big)^{\star'} \qquad \text{(because } \delta^{\star}_{\mathbb{B}^{\mathcal{R}}_{(j)}} = \sigma_{\mathbb{B}^{\mathcal{R}}_{(j)}}\text{)}$$
$$= \Big(\sup_{j=0,1,\ldots,d} \big(\delta_{\mathbb{B}^{\mathcal{R}}_{(j)}} + \varphi(j)\big)^{\star}\Big)^{\star'} \qquad \text{(by property of conjugacies)}$$
$$= \bigg(\Big(\inf_{j=0,1,\ldots,d} \big[\delta_{\mathbb{B}^{\mathcal{R}}_{(j)}} + \varphi(j)\big]\Big)^{\star}\bigg)^{\star'}$$

(as conjugacies, being dualities, turn infima into suprema)

$$= \Big(\inf_{j=0,1,\ldots,d} \big[\delta_{\mathbb{B}^{\mathcal{R}}_{(j)}} + \varphi(j)\big]\Big)^{\star\star'} . \qquad \text{(by (54c))}$$

   Thus, we have obtained (39c) and (39d). Now, if we follow again the above sequence of equalities, we see that, everywhere, we can replace the balls $\mathbb{B}^{\mathcal{R}}_{(j)}$ by the spheres $\mathbb{S}^{\mathcal{R}}_{(j)}$, since $|||\cdot|||^{\mathcal{R}}_{(j),\star} = \sigma_{\mathbb{S}^{\mathcal{R}}_{(j)}} = \delta^{\star}_{\mathbb{S}^{\mathcal{R}}_{(j)}}$. Thus, we obtain (39e) and (39f).

2. Let the function $\varphi : \{0,1,\ldots,d\} \to \mathbb{R}$ be given. Then the closed convex function $\big((\varphi \circ \ell_0)^{¢}\big)^{\star'}$ is proper. Indeed, on the one hand, it is easily seen that the function $(\varphi \circ \ell_0)^{¢}$ takes finite values, from which we deduce that the function $\big((\varphi \circ \ell_0)^{¢}\big)^{\star'}$ never takes the value $-\infty$. On the other hand, by (39a) and by the inequality $(\varphi \circ \ell_0)^{¢¢'} \leq \varphi \circ \ell_0$ obtained from (53e), we deduce that the function $\big((\varphi \circ \ell_0)^{¢}\big)^{\star'}$ never takes the value $+\infty$ on the unit sphere. Therefore, the $\big((\varphi \circ \ell_0)^{¢}\big)^{\star'}$ is proper.

   For the remaining expressions for $\big((\varphi \circ \ell_0)^{¢}\big)^{\star'}$, we use a general formula [14, Corollary 2.8.11] for the Fenchel conjugate of the supremum of proper convex functions $f_j : \mathbb{R}^d \to \overline{\mathbb{R}}$, $j = 0,1,\ldots,n$:

$$\bigcap_{j=0,1,\ldots,n} \mathrm{dom} f_j \neq \emptyset \Rightarrow \big(\sup_{j=0,1,\ldots,n} f_j\big)^{\star} = \min_{(\lambda_0,\lambda_1,\ldots,\lambda_n)\in\Delta_{n+1}} \Big(\sum_{j=0}^{n} \lambda_j f_j\Big)^{\star} , \qquad (41)$$

   where $\mathrm{dom} f = \big\{x \in \mathbb{R}^d \,\big|\, f(x) < +\infty\big\}$ is the effective domain (see §A.2.2), and where Δ_{n+1} is the simplex of $\mathbb{R}^{n+1}$.

We obtain

$$\big((\varphi \circ \ell_0)^{¢}\big)^{\star'} = \Big(\sup_{j=0,1,\dots,d} \Big[|||\cdot|||^{\mathcal{R}}_{(j),\star} - \varphi(j) \Big] \Big)^{\star'} \qquad \text{(by (38))}$$

$$= \Big(\sup_{j=0,1,\dots,d} \Big[\sigma_{\mathbb{B}^{\mathcal{R}}_{(j)}} - \varphi(j) \Big] \Big)^{\star'}$$

(by (13a) as $\mathbb{B}^{\mathcal{R}}_{(j)}$ is the unit ball of the norm $|||\cdot|||^{\mathcal{R}}_{(j)}$ by (25b), and with the convention $\mathbb{B}^{\mathcal{R}}_{(0)} = \{0\}$)

$$= \min_{(\lambda_0,\lambda_1,\dots,\lambda_d)\in\Delta_{d+1}} \Big(\sum_{j=0}^{d} \lambda_j \Big[\sigma_{\mathbb{B}^{\mathcal{R}}_{(j)}} - \varphi(j) \Big] \Big)^{\star'} \qquad \text{(by (41))}$$

by [14, Corollary 2.8.11], as the functions $f_j = \sigma_{\mathbb{B}^{\mathcal{R}}_{(j)}} - \varphi(j)$ are proper convex (they even take finite values), for $j = 0, 1, \dots, d$

$$= \min_{(\lambda_0,\lambda_1,\dots,\lambda_d)\in\Delta_{d+1}} \Big(\sigma_{\sum_{j=0}^{d} \lambda_j \mathbb{B}^{\mathcal{R}}_{(j)}} - \sum_{j=0}^{d} \lambda_j \varphi(j) \Big)^{\star'}$$

as, for all $j = 1, \dots, d$, $\lambda_j \sigma_{\mathbb{B}^{\mathcal{R}}_{(j)}} = \sigma_{\lambda_j \mathbb{B}^{\mathcal{R}}_{(j)}}$ since $\lambda_j \geq 0$, and then using the well-known property that the support function of a Minkowski sum of subsets is the sum of the support functions of the individual subsets [7, p. 226]

$$= \min_{(\lambda_0,\lambda_1,\dots,\lambda_d)\in\Delta_{d+1}} \Big(\sigma_{\sum_{j=1}^{d} \lambda_j \mathbb{B}^{\mathcal{R}}_{(j)}} - \sum_{j=0}^{d} \lambda_j \varphi(j) \Big)^{\star'}$$

(thanks to the convention $\mathbb{B}^{\mathcal{R}}_{(0)} = \{0\}$)

$$= \min_{(\lambda_0,\lambda_1,\dots,\lambda_d)\in\Delta_{d+1}} \Big(\big(\sigma_{\sum_{j=1}^{d} \lambda_j \mathbb{B}^{\mathcal{R}}_{(j)}} \big)^{\star'} + \sum_{j=0}^{d} \lambda_j \varphi(j) \Big)$$

(by property of conjugacies)

$$= \min_{(\lambda_0,\lambda_1,\dots,\lambda_d)\in\Delta_{d+1}} \Big(\delta_{\sum_{j=1}^{d} \lambda_j \mathbb{B}^{\mathcal{R}}_{(j)}} + \sum_{j=0}^{d} \lambda_j \varphi(j) \Big)$$

(because $\sum_{j=1}^{d} \lambda_j \mathbb{B}^{\mathcal{R}}_{(j)}$ is a closed convex set.)

Therefore, we deduce that, for all $x \in \mathbb{R}^d$,

$$\big((\varphi \circ \ell_0)^{¢}\big)^{\star'}(x) = \min_{\substack{(\lambda_0,\lambda_1,\dots,\lambda_d)\in\Delta_{d+1} \\ x \in \sum_{j=1}^{d} \lambda_j \mathbb{B}^{\mathcal{R}}_{(j)}}} \sum_{j=0}^{d} \lambda_j \varphi(j) \; ,$$

which is (39h).

3. Let the function $\varphi : \{0, 1, \dots, d\} \to \mathbb{R}_+$ be given, such that $\varphi(0) = 0$. Then the closed convex

function $\big((\varphi \circ \ell_0)^{¢}\big)^{\star'}$ is proper, as seen above. We go on with

$$\big((\varphi \circ \ell_0)^{¢}\big)^{\star'}(x) = \min_{\substack{(\lambda_0,\lambda_1,\ldots,\lambda_d)\in\Delta_{d+1} \\ x\in\sum_{j=1}^d \lambda_j \mathbb{B}^{\mathcal{R}}_{(j)}}} \sum_{j=1}^{d} \lambda_j \varphi(j) \qquad \text{(because } \varphi(0)=0\text{)}$$

$$= \min_{\substack{z^{(1)}\in\mathbb{B}^{\mathcal{R}}_{(1)},\ldots,z^{(d)}\in\mathbb{B}^{\mathcal{R}}_{(d)} \\ \lambda_1\geq 0,\ldots,\lambda_d\geq 0 \\ \sum_{j=1}^d \lambda_j \leq 1 \\ \sum_{j=1}^d \lambda_j z^{(j)}=x}} \sum_{j=1}^{d} \lambda_j \varphi(j)$$

because $(\lambda_0, \lambda_1, \ldots, \lambda_d) \in \Delta_{d+1}$ if and only if $\lambda_1 \geq 0, \ldots, \lambda_d \geq 0$ and $\sum_{j=1}^d \lambda_j \leq 1$ and $\lambda_0 = 1 - \sum_{j=1}^d \lambda_j$

$$= \min_{\substack{s^{(1)}\in\mathbb{S}^{\mathcal{R}}_{(1)},\ldots,s^{(d)}\in\mathbb{S}^{\mathcal{R}}_{(d)} \\ \mu_1\geq 0,\ldots,\mu_d\geq 0 \\ \sum_{j=1}^d \mu_j \leq 1 \\ \sum_{j=1}^d \mu_j s^{(j)}=x}} \sum_{j=1}^{d} \mu_j \varphi(j)$$

because, on the one hand, the inequality $\leq$ is obvious as the unit sphere $\mathbb{S}^{\mathcal{R}}_{(j)}$ in (20a) is included in the unit ball $\mathbb{B}^{\mathcal{R}}_{(j)}$ in (20b) for all $j = 1, \ldots, d$; and, on the other hand, the inequality $\geq$ comes from putting, for $j = 1, \ldots, d$, $\mu_j = \lambda_j \vert\!\vert\!\vert z^{(j)} \vert\!\vert\!\vert^{\mathcal{R}}_{(j)}$ and observing that i) $\sum_{i=1}^d \mu_j = \sum_{i=1}^d \lambda_j \vert\!\vert\!\vert z^{(j)} \vert\!\vert\!\vert^{\mathcal{R}}_{(j)} \leq \sum_{i=1}^d \lambda_j \leq 1$ because $\vert\!\vert\!\vert z^{(j)} \vert\!\vert\!\vert^{\mathcal{R}}_{(j)} \leq 1$ as $z^{(j)} \in \mathbb{B}^{\mathcal{R}}_{(j)}$ ii) for all $j = 1, \ldots, d$, there exists $s^{(j)} \in \mathbb{S}^{\mathcal{R}}_{(j)}$ such that $\lambda_j z^{(j)} = \mu_j s^{(j)}$ (take any $s^{(j)}$ when $z^{(j)} = 0$ because $\mu_j = 0$, and take $s^{(j)} = \frac{z^{(j)}}{\vert\!\vert\!\vert z^{(j)} \vert\!\vert\!\vert^{\mathcal{R}}_{(j)}}$ when $z^{(j)} \neq 0$) iii) $\sum_{j=1}^d \lambda_j \varphi(j) \geq \sum_{j=1}^d \lambda_j \vert\!\vert\!\vert z^{(j)} \vert\!\vert\!\vert^{\mathcal{R}}_{(j)} \varphi(j) = \sum_{j=1}^d \mu_j \varphi(j)$ because $1 \geq \vert\!\vert\!\vert z^{(j)} \vert\!\vert\!\vert^{\mathcal{R}}_{(j)}$ and $\varphi(j) \geq 0$

$$= \min_{\substack{z^{(1)}\in\mathbb{R}^d,\ldots,z^{(d)}\in\mathbb{R}^d \\ \sum_{j=1}^d \vert\!\vert\!\vert z^{(j)} \vert\!\vert\!\vert^{\mathcal{R}}_{(j)} \leq 1 \\ \sum_{j=1}^d z^{(j)}=x}} \sum_{j=1}^{d} \varphi(j) \vert\!\vert\!\vert z^{(j)} \vert\!\vert\!\vert^{\mathcal{R}}_{(j)} ,$$

by putting $z^{(j)} = \mu_j s^{(j)}$, for all $j = 1, \ldots, d$. Thus, we have obtained (39i).

Finally, from $(\varphi \circ \ell_0)^{¢¢'} = \big((\varphi \circ \ell_0)^{¢}\big)^{\star'} \circ n$, by (35d), we get that

$$(\varphi \circ \ell_0)^{¢¢'}(x) = \frac{1}{\vert\!\vert\!\vert x \vert\!\vert\!\vert} \min_{\substack{z^{(1)}\in\mathbb{R}^d,\ldots,z^{(d)}\in\mathbb{R}^d \\ \sum_{j=1}^d \vert\!\vert\!\vert z^{(j)} \vert\!\vert\!\vert^{\mathcal{R}}_{(j)} \leq \vert\!\vert\!\vert x \vert\!\vert\!\vert \\ \sum_{j=1}^d z^{(j)}=x}} \sum_{j=1}^{d} \varphi(j) \vert\!\vert\!\vert z^{(j)} \vert\!\vert\!\vert^{\mathcal{R}}_{(j)} , \ \ \forall x \in \mathbb{R}^d \backslash \{0\} ,$$

where we have used that $n(x) = \frac{x}{\vert\!\vert\!\vert x \vert\!\vert\!\vert}$ when $x \neq 0$ by (34). Therefore, we have proved (40).

This ends the proof. □

Before finishing that part on Capra-biconjugates, we provide the following characterization of when the characteristic functions $\delta_{\ell_0^{\leq k}}$ are ¢-convex.

Corollary 13 *Let $|||\cdot|||$ be a norm on $\mathbb{R}^d$, with associated sequence $\left\{|||\cdot|||^{\mathcal{R}}_{(j)}\right\}_{j=1,\ldots,d}$ of coordinate-k norms in Definition 3 and associated Capra coupling ¢ in (33).*

The following statements are equivalent.

1. *The sequence $\left\{|||\cdot|||^{\mathcal{R}}_{(j)}\right\}_{j=1,\ldots,d}$ of coordinate-k norms is strictly decreasingly graded with respect to the ℓ_0 pseudonorm, as in Definition 1.*
2. *For all $k \in \{0,1,\ldots,d\}$, the characteristic functions $\delta_{\ell_0^{\leq k}}$ are ¢-convex, that is,*

$$\delta_{\ell_0^{\leq k}}^{¢¢'} = \delta_{\ell_0^{\leq k}} \ , \ \ k = 0,1,\ldots,d \ . \tag{42}$$

Proof.

For any $k = 0,1,\ldots,d$, we have

$$\begin{aligned}
\delta_{\ell_0^{\leq k}}^{¢¢'} &= \Big(\inf_{j=0,1,\ldots,d} \big[\delta_{\mathbb{B}^{\mathcal{R}}_{(j)}} \dotplus \delta_{\{0,1,\ldots,k\}}(j)\big]\Big)^{\star\star'} \circ n \qquad \text{(by (39c) with the functions } \varphi = \delta_{\{0,1,\ldots,k\}}) \\
&= \Big(\inf_{j=0,1,\ldots,k} \delta_{\mathbb{B}^{\mathcal{R}}_{(j)}}\Big)^{\star\star'} \circ n \\
&= \big(\delta_{\mathbb{B}^{\mathcal{R}}_{(k)}}\big)^{\star\star'} \circ n \\
&\qquad \text{(by the inclusions } \mathbb{B}^{\mathcal{R}}_{(1)} \subset \cdots \subset \mathbb{B}^{\mathcal{R}}_{(k)} \text{ in (29) and by the convention } \mathbb{B}^{\mathcal{R}}_{(0)} = \{0\}) \\
&= \delta_{\mathbb{B}^{\mathcal{R}}_{(k)}} \circ n \qquad \text{(because the unit ball } \mathbb{B}^{\mathcal{R}}_{(k)} \text{ is closed and convex)} \\
&= \delta_{n^{-1}(\mathbb{B}^{\mathcal{R}}_{(k)})}
\end{aligned}$$

where, by (34), $n^{-1}(\mathbb{B}^{\mathcal{R}}_{(k)}) = \{0\} \cup \{x \in \mathbb{R}^d \backslash \{0\} \mid |||\frac{x}{|||x|||}|||^{\mathcal{R}}_{(k)} \leq 1\}$, so that we go on with

$$\begin{aligned}
&= \delta_{\{x \in \mathbb{R}^d \mid |||x|||^{\mathcal{R}}_{(k)} \leq |||x|||\}} \\
&= \delta_{\{x \in \mathbb{R}^d \mid |||x|||^{\mathcal{R}}_{(k)} = |||x|||\}} \qquad \text{(using the equality and inequalities between norms in (28))}
\end{aligned}$$

Therefore, we have

$$\begin{aligned}
&\forall k \in \{0,1,\ldots,d\} \ , \ \delta_{\ell_0^{\leq k}}^{¢¢'} = \delta_{\ell_0^{\leq k}} \\
&\qquad \Leftrightarrow \forall k \in \{0,1,\ldots,d\} \ , \ \Big(x \in \ell_0^{\leq k} \iff |||x|||^{\mathcal{R}}_{(k)} = |||x||| \ , \ \forall x \in \mathbb{R}^d\Big) \\
&\qquad \Leftrightarrow \text{(9b) holds true for the sequence } \left\{|||\cdot|||^{\mathcal{R}}_{(j)}\right\}_{j=1,\ldots,d} \\
&\qquad\qquad \text{(because } x \in \ell_0^{\leq k} \iff \ell_0(x) \leq k \text{ by definition of the level sets in (4a))} \\
&\qquad \Leftrightarrow \left\{|||\cdot|||^{\mathcal{R}}_{(j)}\right\}_{j=1,\ldots,d} \text{ is strictly decreasingly graded w.r.t. the } \ell_0 \text{ pseudonorm}
\end{aligned}$$

because this sequence is nonincreasing by (23) (see Definition 1).

This ends the proof. □

Notice that, by Item 2 in Proposition 6, it suffices that the normed space $\big(\mathbb{R}^d, |||\cdot|||\big)$ be strictly convex to obtain that the characteristic functions $\delta_{\ell_0^{\leq k}}$ are ¢-convex, for all $k = 0, 1, \ldots, d$. This is the case when the source norm is the ℓ_p-norm $\|\cdot\|_p$ for $p \in]1, \infty[$.

Determinining sufficient conditions under which the ℓ_0 pseudonorm is ¢-convex requires additional concepts. This question is treated in the companion paper [6].

4.4 Capra-subdifferentials related to the ℓ_0 pseudonorm

With the Fenchel conjugacy, we easily get that $\partial\delta_{\ell_0^{\leq k}}(x) = \emptyset$, for all $x \in \mathbb{R}^d$ and for all $k = 0, 1, \ldots, d$ (indeed, this is a consequence of $\delta_{\ell_0^{\leq k}}^{\star\star'} = -\infty \neq \delta_{\ell_0^{\leq k}}$). We also easily get that $\partial\ell_0(0) = \{0\}$ and $\partial\ell_0(x) = \emptyset$, for all $x \in \mathbb{R}^d\backslash\{0\}$ (indeed, this is a consequence of $\ell_0^{\star\star'}(x) = 0 \neq \ell_0(x)$ when $x \in \mathbb{R}^d\backslash\{0\}$). Hence, the Fenchel conjugacy does not seem to be suitable to handle the ℓ_0 pseudonorm. We will see that we obtain more interesting formulas with the Capra-conjugacy.

More precisely, we will now show that functions of the ℓ_0 pseudonorm in (2) — including the ℓ_0 pseudonorm itself and the characteristic functions $\delta_{\ell_0^{\leq k}}$ of its level sets (4a) — display Capra-subdifferentials, as in (37b), that are related to the sequence of dual coordinate-k norms in Definition 3 as follows.

Proposition 14 *Let $|||\cdot|||$ be a norm on $\mathbb{R}^d$, with associated sequence $\Big\{|||\cdot|||_{\star(j)}^{\mathcal{R}}\Big\}_{j=1,\ldots,d}$ of dual coordinate-k norms, as in Definition 3, and associated coupling ¢ in* (33).

Let a function $\varphi : \{0, 1, \ldots, d\} \to \overline{\mathbb{R}}$ and a vector $x \in \mathbb{R}^d$ be given.

- *The Capra-subdifferential, as in* (37d), *of the function $\varphi \circ \ell_0$ at $x = 0$ is given by*

$$\partial_{¢}(\varphi \circ \ell_0)(0) = \bigcap_{j=1,\ldots,d} \big[\varphi(j) \dot{+} \big(-\varphi(0)\big)\big]\mathbb{B}_{(j),\star}^{\mathcal{R}} , \tag{43}$$

  where, by convention $\lambda\mathbb{B}_{(j),\star}^{\mathcal{R}} = \emptyset$, for any $\lambda \in [-\infty, 0[$, and $+\infty\mathbb{B}_{(j),\star}^{\mathcal{R}} = \mathbb{R}^d$.

- *The Capra-subdifferential, as in* (37e), *of the function $\varphi \circ \ell_0$ at $x \neq 0$ is given by the following cases*

  – *if $l = \ell_0(x) \geq 1$ and either $\varphi(l) = -\infty$ or $\varphi \equiv +\infty$, then $\partial_{¢}(\varphi \circ \ell_0)(x) = \mathbb{R}^d$,*

  – *if $l = \ell_0(x) \geq 1$ and $\varphi(l) = +\infty$ and there exists $j \in \{0, 1, \ldots, d\}$ such that $\varphi(j) \neq +\infty$, then $\partial_{¢}(\varphi \circ \ell_0)(x) = \emptyset$,*

– *if* $l = \ell_0(x) \geq 1$ *and* $-\infty < \varphi(l) < +\infty$*, then*

$$y \in \partial_{\text{Ç}}(\varphi \circ \ell_0)(x) \iff \begin{cases} y \in N_{\mathbb{B}^{\mathcal{R}}_{(l)}}\big(\frac{x}{\|x\|^{\mathcal{R}}_{(l)}}\big) \\ \text{and} \\ l \in \operatorname{argmax}_{j=0,1,\ldots,d}\left[|||y|||^{\mathcal{R}}_{(j),\star} - \varphi(j)\right] \end{cases} \tag{44}$$

where the normal cone $N_{\mathbb{B}^{\mathcal{R}}_{(l)}}$ *has been introduced in* (15).

Proof. We have

$$y \in \partial_{\text{Ç}}(\varphi \circ \ell_0)(x) \iff (\varphi \circ \ell_0)^{\text{¢}}(y) = \text{¢}(x,y) \dotplus \big(-(\varphi \circ \ell_0)(x)\big)$$

(by definition (37b) of the Capra-subdifferential)

$$\iff \sup_{j=0,1,\ldots,d}\left[|||y|||^{\mathcal{R}}_{(j),\star} - \varphi(j)\right] = \text{¢}(x,y) \dotplus \big(-(\varphi \circ \ell_0)(x)\big)$$

(as $(\varphi \circ \ell_0)^{\text{¢}}(y) = \sup_{j=0,1,\ldots,d}\left[|||y|||^{\mathcal{R}}_{(j),\star} - \varphi(j)\right]$ by (38))

$$\iff \Big(x = 0 \text{ and } \sup_{j=0,1,\ldots,d}\left[|||y|||^{\mathcal{R}}_{(j),\star} - \varphi(j)\right] = -\varphi(0)\Big)$$

$$\text{or } \Big(x \neq 0 \text{ and } \sup_{j=0,1,\ldots,d}\left[|||y|||^{\mathcal{R}}_{(j),\star} - \varphi(j)\right] = \frac{\langle x\,,y\rangle}{|||x|||} - \varphi\big(\ell_0(x)\big)\Big)$$

(by definition (33) of $\text{¢}(x,y)$)

Therefore, on the one hand, we obtain that

$$y \in \partial_{\text{Ç}}(\varphi \circ \ell_0)(0) \iff |||y|||^{\mathcal{R}}_{(j),\star} - \varphi(j) \leq -\varphi(0)\ ,\ \ \forall j = 1,\ldots,d \quad (\text{as } |||y|||^{\mathcal{R}}_{(0),\star} = 0 \text{ by convention})$$

$$\iff |||y|||^{\mathcal{R}}_{(j),\star} \leq \varphi(j) \dotplus \big(-\varphi(0)\big)\ ,\ \ \forall j = 1,\ldots,d \qquad (\text{using (51f)})$$

$$\iff y \in \bigcap_{j=1,\ldots,d}\big[\varphi(j) \dotplus \big(-\varphi(0)\big)\big]\mathbb{B}^{\mathcal{R}}_{(j),\star}\ ,$$

where, by convention $\lambda\mathbb{B}^{\mathcal{R}}_{(j),\star} = \emptyset$, for any $\lambda \in [-\infty, 0[$, and $+\infty\mathbb{B}^{\mathcal{R}}_{(j),\star} = \mathbb{R}^d$.

On the other hand, when $x \neq 0$, we get

$$y \in \partial_{\text{Ç}}(\varphi \circ \ell_0)(x) \iff \sup_{j=0,1,\ldots,d}\left[|||y|||^{\mathcal{R}}_{(j),\star} - \varphi(j)\right] = \frac{\langle x\,,y\rangle}{|||x|||} - \varphi\big(\ell_0(x)\big)\ . \tag{45a}$$

We now establish necessary and sufficient conditions for y to belong to $\partial_{\text{Ç}}(\varphi \circ \ell_0)(x)$ when $x \neq 0$. For this purpose, we consider $x \in \mathbb{R}^d\backslash\{0\}$, and we denote $L = \operatorname{supp}(x)$ and $l = |L| = \ell_0(x)$. We have

$$y \in \partial_{\text{Ç}}(\varphi \circ \ell_0)(x)$$

$$\iff \sup_{j=0,1,\ldots,d}\left[|||y|||^{\mathcal{R}}_{(j),\star} - \varphi(j)\right] = \frac{\langle x\,,y\rangle}{|||x|||} - \varphi(l) \qquad (\text{by (45a) with } \ell_0(x) = l)$$

$$\iff |||y|||^{\mathcal{R}}_{(l),\star} - \varphi(l) \leq \sup_{j=0,1,\ldots,d}\left[|||y|||^{\mathcal{R}}_{(j),\star} - \varphi(j)\right] = \frac{\langle x\,,y\rangle}{|||x|||} - \varphi(l)$$

$$\iff |||y_L|||_{L,\star} - \varphi(l) \leq |||y|||^{\mathcal{R}}_{(l),\star} - \varphi(l) \leq \sup_{j=0,1,\ldots,d}\left[|||y|||^{\mathcal{R}}_{(j),\star} - \varphi(j)\right] = \frac{\langle x\,,y\rangle}{|||x|||} - \varphi(l)$$

as $|\!|\!| y_L |\!|\!|_{L,\star} \leq |\!|\!| y |\!|\!|^{\mathcal{R}}_{(l),\star}$ by expression (21) of the dual coordinate-k norm $|\!|\!| y |\!|\!|^{\mathcal{R}}_{(l),\star}$, and because $l = |L|$

$$\iff |\!|\!| y_L |\!|\!|_{L,\star} - \varphi(l) \leq |\!|\!| y |\!|\!|^{\mathcal{R}}_{(l),\star} - \varphi(l) \leq \sup_{j=0,1,\ldots,d} \big[|\!|\!| y |\!|\!|^{\mathcal{R}}_{(j),\star} - \varphi(j) \big] = \frac{\langle x \,, y \rangle}{|\!|\!| x |\!|\!|} - \varphi(l) \leq |\!|\!| y_L |\!|\!|_{L,\star} - \varphi(l)$$

(as we have $\frac{\langle x \,, y \rangle}{|\!|\!| x |\!|\!|} = \frac{\langle x_L \,, y_L \rangle}{|\!|\!| x_L |\!|\!|} \leq |\!|\!| y_L |\!|\!|_{L,\star}$ since $x = x_L$ and by (14a))

$$\iff |\!|\!| y_L |\!|\!|_{L,\star} - \varphi(l) = |\!|\!| y |\!|\!|^{\mathcal{R}}_{(l),\star} - \varphi(l) = \sup_{j=0,1,\ldots,d} \big[|\!|\!| y |\!|\!|^{\mathcal{R}}_{(j),\star} - \varphi(j) \big] = \frac{\langle x \,, y \rangle}{|\!|\!| x |\!|\!|} - \varphi(l)$$

(as all terms in the inequalities are necessarily equal)

$$\iff \begin{cases} \text{either } \varphi(l) = -\infty \\ \text{or } \big(\varphi(l) = +\infty \text{ and } \varphi(j) = +\infty \;,\;\; \forall j = 0,1,\ldots,d \big) \\ \text{or } \Big(-\infty < \varphi(l) < +\infty \;\text{ and } \\ \qquad |\!|\!| y_L |\!|\!|_{L,\star} = |\!|\!| y |\!|\!|^{\mathcal{R}}_{(l),\star} = \frac{\langle x \,, y \rangle}{|\!|\!| x |\!|\!|} \;\text{ and }\; |\!|\!| y |\!|\!|^{\mathcal{R}}_{(l),\star} - \varphi(l) = \sup_{j=0,1,\ldots,d} \big[|\!|\!| y |\!|\!|^{\mathcal{R}}_{(j),\star} - \varphi(j) \big] \Big) \;. \end{cases}$$

Let us make a brief insert and notice that

$$\begin{aligned} & x = x_L \;,\;\; \ell_0(x) = l = |L| > 1 \;,\;\; \langle x \,, y \rangle = |\!|\!| x |\!|\!| \times |\!|\!| y |\!|\!|^{\mathcal{R}}_{(l),\star} \\ & \quad \Rightarrow \quad \ell_0(x) = l = |L| > 1 \;,\;\; \langle x_L \,, y_L \rangle = |\!|\!| x_L |\!|\!| \times |\!|\!| y |\!|\!|^{\mathcal{R}}_{(l),\star} \\ & \quad \Rightarrow \quad \ell_0(x) = l = |L| > 1 \;,\;\; |\!|\!| x_L |\!|\!| \times |\!|\!| y |\!|\!|^{\mathcal{R}}_{(l),\star} \leq |\!|\!| x_L |\!|\!| \times |\!|\!| y_L |\!|\!|_{L,\star} \qquad \text{(by (14a))} \\ & \quad \Rightarrow \quad l = |L| \;,\;\; |\!|\!| y |\!|\!|^{\mathcal{R}}_{(l),\star} \leq |\!|\!| y_L |\!|\!|_{L,\star} \\ & \quad \Rightarrow \quad |\!|\!| y |\!|\!|^{\mathcal{R}}_{(l),\star} = |\!|\!| y_L |\!|\!|_{L,\star} \end{aligned}$$

as $|\!|\!| y_L |\!|\!|_{L,\star} \leq |\!|\!| y |\!|\!|^{\mathcal{R}}_{(l),\star}$ by expression (21) of the dual coordinate-k norm $|\!|\!| y |\!|\!|^{\mathcal{R}}_{(l),\star}$, and because $l = |L|$.

Now, let us go back to the equivalences regarding $y \in \partial_{\text{Ç}}(\varphi \circ \ell_0)(x)$. Focusing on the case where $-\infty < \varphi(l) < +\infty$, we have

$$\begin{aligned} y \in \partial_{\text{Ç}}(\varphi \circ \ell_0)(x) & \Leftrightarrow |\!|\!| y_L |\!|\!|_{L,\star} = |\!|\!| y |\!|\!|^{\mathcal{R}}_{(l),\star} = \frac{\langle x \,, y \rangle}{|\!|\!| x |\!|\!|} \;\text{ and }\; l \in \operatorname*{argmax}_{j=0,1,\ldots,d} \big[|\!|\!| y |\!|\!|^{\mathcal{R}}_{(j),\star} - \varphi(j) \big] \\ & \Leftrightarrow |\!|\!| y_L |\!|\!|_{L,\star} = |\!|\!| y |\!|\!|^{\mathcal{R}}_{(l),\star} \;\text{ and }\; \langle x \,, y \rangle = |\!|\!| x |\!|\!| \times |\!|\!| y |\!|\!|^{\mathcal{R}}_{(l),\star} \;\text{ and }\; l \in \operatorname*{argmax}_{j=0,1,\ldots,d} \big[|\!|\!| y |\!|\!|^{\mathcal{R}}_{(j),\star} - \varphi(j) \big] \\ & \Leftrightarrow \langle x \,, y \rangle = |\!|\!| x |\!|\!| \times |\!|\!| y |\!|\!|^{\mathcal{R}}_{(l),\star} \;\text{ and }\; l \in \operatorname*{argmax}_{j=0,1,\ldots,d} \big[|\!|\!| y |\!|\!|^{\mathcal{R}}_{(j),\star} - \varphi(j) \big] \end{aligned}$$

(as just established in the insert)

$$\Leftrightarrow \langle x \,, y \rangle = |\!|\!| x |\!|\!|^{\mathcal{R}}_{(l)} \times |\!|\!| y |\!|\!|^{\mathcal{R}}_{(l),\star} \;\text{ and }\; l \in \operatorname*{argmax}_{j=0,1,\ldots,d} \big[|\!|\!| y |\!|\!|^{\mathcal{R}}_{(j),\star} - \varphi(j) \big]$$

(as $\ell_0(x) = l \Rightarrow |\!|\!| x |\!|\!| = |\!|\!| x |\!|\!|^{\mathcal{R}}_{(l)}$ by (30a))

$$\Leftrightarrow y \in N_{\mathbb{B}^{\mathcal{R}}_{(l)}}\Big(\frac{x}{|\!|\!| x |\!|\!|^{\mathcal{R}}_{(l)}}\Big) \;\text{ and }\; l \in \operatorname*{argmax}_{j=0,1,\ldots,d} \big[|\!|\!| y |\!|\!|^{\mathcal{R}}_{(j),\star} - \varphi(j) \big] \qquad \text{(by the equivalence in (16))}$$

This ends the proof. □

With φ the identity function on $\{0,1,\ldots,d\}$, we find the Capra-subdifferential of the ℓ_0 pseudonorm. With the functions $\varphi = \delta_{\{0,1,\ldots,k\}}$ (for any $k \in \{0,1,\ldots,d\}$), we find the Capra-subdifferentials of the characteristic functions $\delta_{\ell_0^{\leq k}}$ of its level sets (4a). The corresponding expressions are given in Table 3.

5 Norm ratio lower bounds for the l_0 pseudonorm

As an application, we provide a new family of lower bounds for the ℓ_0 pseudonorm, as a fraction between two norms, the denominator being any norm.

Proposition 15 *Let $|||\cdot|||$ be a norm on $\mathbb{R}^d$, with associated sequence of dual coordinate-k norms, as in Definition 3.*

For any function $\varphi : \{0,1,\ldots,d\} \to [0,+\infty[$, such that $\varphi(j) > \varphi(0) = 0$ for all $j = 1,\ldots,d$, there exists a norm $|||\cdot|||^{\mathcal{R}}_{(\varphi)}$ characterized

- *either by its dual norm $|||\cdot|||^{\mathcal{R}}_{(\varphi),\star}$, which has unit ball $\bigcap_{j=1,\ldots,d} \varphi(j)\mathbb{B}^{\mathcal{R}}_{(j),\star}$, that is,*

$$\mathbb{B}^{\mathcal{R}}_{(\varphi),\star} = \bigcap_{j=1,\ldots,d} \varphi(j)\mathbb{B}^{\mathcal{R}}_{(j),\star} \text{ and } |||\cdot|||^{\mathcal{R}}_{(\varphi)} = \sigma_{\mathbb{B}^{\mathcal{R}}_{(\varphi),\star}} , \tag{46a}$$

  or, equivalently,

$$|||y|||^{\mathcal{R}}_{(\varphi),\star} = \sup_{j=1,\ldots,d} \frac{|||y|||^{\mathcal{R}}_{(j),\star}}{\varphi(j)} , \ \forall y \in \mathbb{R}^d , \tag{46b}$$

- *or by the inf-convolution*

$$|||\cdot|||^{\mathcal{R}}_{(\varphi)} = \mathop{\square}_{j=1,\ldots,d} \Big(\varphi(j) |||\cdot|||^{\mathcal{R}}_{(j)} \Big) , \tag{46c}$$

  that is,

$$|||x|||^{\mathcal{R}}_{(\varphi)} = \inf_{\substack{z^{(1)}\in\mathbb{R}^d,\ldots,z^{(d)}\in\mathbb{R}^d \\ \sum_{j=1}^d z^{(j)}=x}} \sum_{j=1}^d \varphi(j) |||z^{(j)}|||^{\mathcal{R}}_{(j)} , \ \forall x \in \mathbb{R}^d . \tag{46d}$$

Proof.

• It is easily seen that $\sigma_{\mathbb{B}^{\mathcal{R}}_{(\varphi),\star}}$ in (46a) defines a norm, and that, for all $y \in \mathbb{R}^d$,

$$|||y|||^{\mathcal{R}}_{(\varphi),\star} = \inf\big\{ \lambda \geq 0 \,\big|\, y \in \lambda \bigcap_{j=1}^d \varphi(j)\mathbb{B}^{\mathcal{R}}_{(j),\star} \big\} = \inf\big\{ \lambda \geq 0 \,\big|\, \frac{|||y|||^{\mathcal{R}}_{(j),\star}}{\varphi(j)} \leq \lambda \big\} = \sup_{j=1,\ldots,d} \frac{|||y|||^{\mathcal{R}}_{(j),\star}}{\varphi(j)} .$$

- We have

$$
\begin{aligned}
|||\cdot|||^{\mathcal{R}}_{(\varphi)} &= \sigma_{\mathbb{B}^{\mathcal{R}}_{(\varphi),\star}} && \text{(by (46a))} \\
&= \delta^{\star}_{\mathbb{B}^{\mathcal{R}}_{(\varphi),\star}} && \text{(because } \mathbb{B}^{\mathcal{R}}_{(\varphi),\star} \text{ is closed and convex)} \\
&= \Big(\sum_{j=1,\ldots,d} \delta_{\varphi(j)\mathbb{B}^{\mathcal{R}}_{(j),\star}} \Big)^{\star}
\end{aligned}
$$

by (46a) and by expressing the characteristic function of an intersection of sets as a sum

$$
= \mathop{\Box}_{j=1,\ldots,d} \delta^{\star}_{\varphi(j)\mathbb{B}^{\mathcal{R}}_{(j),\star}}
$$

using [3, Proposition 15.3 and (v) in Proposition-15.5] because the intersection $\mathbb{B}^{\mathcal{R}}_{(\varphi),\star} = \bigcap_{j=1}^{d} \varphi(j)\mathbb{B}^{\mathcal{R}}_{(j),\star}$ of all the domains of the functions $\delta_{\varphi(j)\mathbb{B}^{\mathcal{R}}_{(j),\star}}$ contain a neighborhood of 0 since $\varphi(j) > 0$ for all $j = 1, \ldots, d$

$$
\begin{aligned}
&= \mathop{\Box}_{j=1,\ldots,d} \sigma_{\varphi(j)\mathbb{B}^{\mathcal{R}}_{(j),\star}} && \text{(as } \delta^{\star}_{\varphi(j)\mathbb{B}^{\mathcal{R}}_{(j),\star}} = \sigma_{\varphi(j)\mathbb{B}^{\mathcal{R}}_{(j),\star}} \text{, for all } j = 1, \ldots, d) \\
&= \mathop{\Box}_{j=1,\ldots,d} \varphi(j)|||\cdot|||^{\mathcal{R}}_{(j)} && \text{(by (13a))}
\end{aligned}
$$

This ends the proof. □

Proposition 16 *Let $|||\cdot|||$ be a norm on $\mathbb{R}^d$, with associated sequence of dual coordinate-k norms, as in Definition 3.*

For any function $\varphi : \{0, 1, \ldots, d\} \to [0, +\infty[$, such that $\varphi(j) > \varphi(0) = 0$ for all $j = 1, \ldots, d$, we have the inequalities

$$
\frac{|||x|||^{\mathcal{R}}_{(\varphi)}}{|||x|||} \leq \frac{1}{|||x|||} \min_{\substack{z^{(1)} \in \mathbb{R}^d, \ldots, z^{(d)} \in \mathbb{R}^d \\ \sum_{j=1}^{d} |||z^{(j)}|||^{\mathcal{R}}_{(j)} \leq |||x||| \\ \sum_{j=1}^{d} z^{(j)} = x}} \sum_{j=1}^{d} \varphi(j)|||z^{(j)}|||^{\mathcal{R}}_{(j)} \leq \varphi\big(\ell_0(x)\big) , \ \ \forall x \in \mathbb{R}^d \backslash \{0\} , \tag{47}
$$

where the norm $|||\cdot|||^{\mathcal{R}}_{(\varphi)}$ has been defined in Proposition 15.

Proof. We consider the coupling ¢ in (33).

By (40) — because the function $\varphi : \{0, 1, \ldots, d\} \to [0, +\infty[$ satisfies the assumption in Item 3 of Proposition 12 — and by the inequality $(\varphi \circ \ell_0)^{¢¢'} \leq \varphi \circ \ell_0$ obtained from (53e), we get that

$$
\frac{1}{|||x|||} \min_{\substack{z^{(1)} \in \mathbb{R}^d, \ldots, z^{(d)} \in \mathbb{R}^d \\ \sum_{j=1}^{d} |||z^{(j)}|||^{\mathcal{R}}_{(j)} \leq |||x||| \\ \sum_{j=1}^{d} z^{(j)} = x}} \sum_{j=1}^{d} j|||z^{(j)}|||^{\mathcal{R}}_{(j)} \leq \varphi\big(\ell_0(x)\big) , \ \ \forall x \in \mathbb{R}^d \backslash \{0\} . \tag{48}
$$

Thus, we have obtained the right hand side inequality in (47).

By relaxing one constraint in (48), we immediately get that

$$\inf_{\substack{z^{(1)}\in\mathbb{R}^d,\ldots,z^{(d)}\in\mathbb{R}^d \\ \sum_{j=1}^d z^{(j)}=x}} \sum_{j=1}^{d} \varphi(j) |||z^{(j)}|||_{(j)}^{\mathcal{R}} \leq \min_{\substack{z^{(1)}\in\mathbb{R}^d,\ldots,z^{(d)}\in\mathbb{R}^d \\ \sum_{j=1}^d |||z^{(j)}|||_{(j)}^{\mathcal{R}} \leq |||x||| \\ \sum_{j=1}^d z^{(j)}=x}} \sum_{j=1}^{d} \varphi(j) |||z^{(j)}|||_{(j)}^{\mathcal{R}} \leq \varphi\big(\ell_0(x)\big) \ , \ \ \forall x \in \mathbb{R}^d \ .$$

Thus, we have obtained the left hand side inequality in (47), thanks to (46d). □

For any function $\varphi : \{0, 1, \ldots, d\} \to [0, +\infty[$, such that $\varphi(j) > \varphi(0) = 0$ for all $j = 1, \ldots, d$, using Table 1 when the source norm $|||\cdot|||$ is the ℓ_p-norm $||\cdot||_p$, for $p \in [1, \infty]$ and $1/p + 1/q = 1$, we denote $|||\cdot|||_{(\varphi)}^{\mathcal{R}}$ by $||\cdot||_{p,\varphi}^{\text{sn}}$. The calculations show that $||\cdot||_{1,\varphi}^{\text{sn}} = ||\cdot||_1$, and that, when $p \in]1, \infty]$, we also have $||\cdot||_{p,\varphi}^{\text{sn}} = ||\cdot||_1$, whatever $p \in [1, \infty]$, if we suppose that $\big(\varphi(j)\big)^q \geq j$, for all $j = 1, \ldots, d$. As a consequence, when $p = 1$, the inequality (47) is trivial. When $p \in]1, \infty]$, if we take the function $\varphi(j) = j^{1/q}$ for all $j = 1, \ldots, d$, the inequality (47) yields that $\frac{||x||_1}{||x||_p} \leq \big(\ell_0(x)\big)^{1/q}$, which is easily obtained directly from the Hölder inequality.

6 Conclusion

As recalled in the introduction, the Fenchel conjugacy fails to provide relevant insight into the ℓ_0 pseudonorm. In this paper, we have presented a new family of conjugacies, which depend on a given general source norm, and we have shown that they are suitable for the ℓ_0 pseudonorm.

Indeed, given a (source) norm on $\mathbb{R}^d$, we have defined, on the one hand, a sequence of so-called coordinate-k norms and, on the other hand, a coupling between $\mathbb{R}^d$ and $\mathbb{R}^d$, called Capra (constant along primal rays). With this, we have provided formulas for the Capra-conjugate and biconjugate, and for the Capra subdifferentials, of functions of the ℓ_0 pseudonorm (hence, in particular, of the ℓ_0 pseudonorm itself and of the characteristic functions of its level sets), in terms of the coordinate-k norms. Table 3 provides the results of Proposition 11, Proposition 12, and Proposition 14, in the case of the ℓ_0 pseudonorm and of the characteristic functions $\delta_{\ell_0^{\leq k}}$ of its level sets (4a). It compares them with the Fenchel conjugates and biconjugates. As an application, we have provided a new family of lower bounds for the ℓ_0 pseudonorm, as a fraction between two norms, the denominator being any norm.

A Appendix

A.1 Background on J. J. Moreau lower and upper additions

When we manipulate functions with values in $\overline{\mathbb{R}} = [-\infty, +\infty]$, we adopt the following Moreau *lower addition* or *upper addition*, depending on whether we deal with sup or inf operations. We follow [9]. In the sequel, u, v and w are any elements of $\overline{\mathbb{R}}$.

Fenchel conjugacy	Capra conjugacy
$\delta^{(-\star)}_{\ell_0^{\leq k}} = +\infty$	$\delta^{-¢}_{\ell_0^{\leq k}} = \vert\vert\vert\cdot\vert\vert\vert^{\mathcal{R}}_{(k),\star}$
$\delta^{\star\star'}_{\ell_0^{\leq k}} = -\infty$	$\delta^{¢¢'}_{\ell_0^{\leq k}} = \delta_{\{x\in\mathbb{R}^d \mid \vert\vert\vert x\vert\vert\vert^{\mathcal{R}}_{(k)} = \vert\vert\vert x\vert\vert\vert\}}$
$\partial\delta_{\ell_0^{\leq k}}(x) = \emptyset$ $\forall x\in\mathbb{R}^d$	$\partial_{¢}\delta_{\ell_0^{\leq k}}(x) = \begin{cases} \emptyset & \text{if } \ell_0(x) = k+1,\ldots,d \; , \\ N_{\mathbb{B}^{\mathcal{R}}_{(k)}}(\frac{x}{\vert\vert\vert x\vert\vert\vert^{\mathcal{R}}_{(k)}}) & \text{if } \ell_0(x) = 1,\ldots,k \; , \\ \{0\} & \text{if } \ell_0(x) = 0 \end{cases}$ $\forall x\in\mathbb{R}^d$
$\ell_0^{\star} = \delta_{\{0\}}$	$\ell_0^{¢} = \sup_{j=0,1,\ldots,d}\big[\vert\vert\vert\cdot\vert\vert\vert^{\mathcal{R}}_{(j),\star} - j\big]$
$\ell_0^{\star\star'} = 0$	$\ell_0^{¢¢'}(x) = \frac{1}{\vert\vert\vert x\vert\vert\vert} \min_{\substack{z^{(1)}\in\mathbb{R}^d,\ldots,z^{(d)}\in\mathbb{R}^d \\ \sum_{j=1}^d \vert\vert\vert z^{(j)}\vert\vert\vert^{\mathcal{R}}_{(j)} \leq \vert\vert\vert x\vert\vert\vert \\ \sum_{j=1}^d z^{(j)} = x}} \sum_{j=1}^d j\vert\vert\vert z^{(j)}\vert\vert\vert^{\mathcal{R}}_{(j)} \; , \; \forall x\in\mathbb{R}^d\backslash\{0\}$ $\ell_0^{¢¢'}(0) = 0$
$\partial\ell_0(0) = \{0\}$ $\partial\ell_0(x) = \emptyset$ $\forall x\in\mathbb{R}^d\backslash\{0\}$	$\partial_{¢}\ell_0(0) = \bigcap_{j=1,\ldots,d} j\mathbb{B}^{\mathcal{R}}_{(j),\star} = \mathbb{B}^{\mathcal{R}}_{(\mathrm{Id}),\star}$ $y\in\partial_{¢}\ell_0(x) \iff \begin{cases} y\in N_{\mathbb{B}^{\mathcal{R}}_{(l)}}(\frac{x}{\vert\vert\vert x\vert\vert\vert^{\mathcal{R}}_{(l)}}) \\ \text{and } l\in \operatorname{argmax}_{j=0,1,\ldots,d}\big[\vert\vert\vert y\vert\vert\vert^{\mathcal{R}}_{(j),\star} - j\big] \end{cases}$ $\forall x\in\mathbb{R}^d\backslash\{0\}$, where $l = \ell_0(x) \geq 1$

Table 3: Comparison of Fenchel and Capra-conjugates, biconjugates and subdifferentials of the ℓ_0 pseudonorm in (2), and of the characteristic functions $\delta_{\ell_0^{\leq k}}$ of its level sets (4a), for $k = 0,1,\ldots,d$

Moreau lower addition

The Moreau *lower addition* extends the usual addition with

$$(+\infty) \dot{+} (-\infty) = (-\infty) \dot{+} (+\infty) = -\infty \ . \tag{49a}$$

With the lower addition, $(\overline{\mathbb{R}}, \dot{+})$ is a convex cone, and $\dot{+}$ is a commutative and associative operation. The lower addition displays the following properties

$$u \leq u' \ , \ \ v \leq v' \Rightarrow u \dot{+} v \leq u' \dot{+} v' \ , \tag{49b}$$

$$(-u) \dot{+} (-v) \leq -(u \dot{+} v) \ , \tag{49c}$$

$$(-u) \dot{+} u \leq 0 \ , \tag{49d}$$

and, for any functions $f : \mathbb{A} \to \overline{\mathbb{R}}$ and $g : \mathbb{B} \to \overline{\mathbb{R}}$,

$$\sup_{a\in\mathbb{A}} f(a) \dot{+} \sup_{b\in\mathbb{B}} g(b) = \sup_{a\in\mathbb{A}, b\in\mathbb{B}} \big(f(a) \dot{+} g(b) \big) \ , \tag{49e}$$

$$\inf_{a\in\mathbb{A}} f(a) \dot{+} \inf_{b\in\mathbb{B}} g(b) \leq \inf_{a\in\mathbb{A}, b\in\mathbb{B}} \big(f(a) \dot{+} g(b) \big) \ , \tag{49f}$$

$$t < +\infty \Rightarrow \inf_{a\in\mathbb{A}} f(a) \dot{+} t = \inf_{a\in\mathbb{A}} \big(f(a) \dot{+} t \big) \ . \tag{49g}$$

Moreau upper addition

The Moreau *upper addition* extends the usual addition with

$$(+\infty) \dot{+} (-\infty) = (-\infty) \dot{+} (+\infty) = +\infty \ . \tag{50a}$$

With the upper addition, $(\overline{\mathbb{R}}, \dot{+})$ is a convex cone, and $\dot{+}$ is a commutative and associative operation. The upper addition displays the following properties

$$u \leq u' \ , \ \ v \leq v' \Rightarrow u \dot{+} v \leq u' \dot{+} v' \ , \tag{50b}$$

$$(-u) \dot{+} (-v) \geq -(u \dot{+} v) \ , \tag{50c}$$

$$(-u) \dot{+} u \geq 0 \ , \tag{50d}$$

and, for any functions $f : \mathbb{A} \to \overline{\mathbb{R}}$ and $g : \mathbb{B} \to \overline{\mathbb{R}}$,

$$\inf_{a\in\mathbb{A}} f(a) \dot{+} \inf_{b\in\mathbb{B}} g(b) = \inf_{a\in\mathbb{A}, b\in\mathbb{B}} \big(f(a) \dot{+} g(b) \big) \ , \tag{50e}$$

$$\sup_{a\in\mathbb{A}} f(a) \dot{+} \sup_{b\in\mathbb{B}} g(b) \geq \sup_{a\in\mathbb{A}, b\in\mathbb{B}} \big(f(a) \dot{+} g(b) \big) \ , \tag{50f}$$

$$-\infty < t \Rightarrow \sup_{a\in\mathbb{A}} f(a) \dot{+} t = \sup_{a\in\mathbb{A}} \big(f(a) \dot{+} t \big) \ . \tag{50g}$$

Joint properties of the Moreau lower and upper addition

We obviously have that

$$u \underset{\cdot}{+} v \leq u \dot{+} v \; . \tag{51a}$$

The Moreau lower and upper additions are related by

$$-(u \dot{+} v) = (-u) \underset{\cdot}{+} (-v) \; , \;\; -(u \underset{\cdot}{+} v) = (-u) \dot{+} (-v) \; . \tag{51b}$$

They satisfy the inequality

$$(u \dot{+} v) \underset{\cdot}{+} w \leq u \dot{+} (v \underset{\cdot}{+} w) \; . \tag{51c}$$

with

$$(u \dot{+} v) \underset{\cdot}{+} w < u \dot{+} (v \underset{\cdot}{+} w) \iff \begin{cases} u = +\infty \text{ and } w = -\infty \; , \\ \text{or} \\ u = -\infty \text{ and } w = +\infty \text{ and } -\infty < v < +\infty \; . \end{cases} \tag{51d}$$

Finally, we have that

$$u \underset{\cdot}{+} (-v) \leq 0 \iff u \leq v \iff 0 \leq v \dot{+} (-u) \; , \tag{51e}$$

$$u \underset{\cdot}{+} (-v) \leq w \iff u \leq v \dot{+} w \iff u \underset{\cdot}{+} (-w) \leq v \; , \tag{51f}$$

$$w \leq v \dot{+} (-u) \iff u \underset{\cdot}{+} w \leq v \iff u \leq v \dot{+} (-w) \; . \tag{51g}$$

A.2 Background on Fenchel-Moreau conjugacies

We review general concepts and notations on Fenchel-Moreau conjugacies, then focus on the special case of the Fenchel conjugacy.

A.2.1 The general case

Let be given two sets $\mathbb{X}$ ("primal"), $\mathbb{Y}$ ("dual"), together with a *coupling* function

$$c : \mathbb{X} \times \mathbb{Y} \to \overline{\mathbb{R}} \; . \tag{52}$$

With any coupling, we associate *conjugacies* from $\overline{\mathbb{R}}^{\mathbb{X}}$ to $\overline{\mathbb{R}}^{\mathbb{Y}}$ and from $\overline{\mathbb{R}}^{\mathbb{Y}}$ to $\overline{\mathbb{R}}^{\mathbb{X}}$ as follows.

Definition 17 *The* c-Fenchel-Moreau conjugate *of a function* $f : \mathbb{X} \to \overline{\mathbb{R}}$*, with respect to the coupling* c*, is the function* $f^c : \mathbb{Y} \to \overline{\mathbb{R}}$ *defined by*

$$f^c(y) = \sup_{x \in \mathbb{X}} \Big(c(x,y) \underset{\cdot}{+} \big(-f(x) \big) \Big) \; , \;\; \forall y \in \mathbb{Y} \; . \tag{53a}$$

With the coupling c*, we associate the* reverse coupling c' *defined by*

$$c' : \mathbb{Y} \times \mathbb{X} \to \overline{\mathbb{R}} \; , \;\; c'(y,x) = c(x,y) \; , \;\; \forall (y,x) \in \mathbb{Y} \times \mathbb{X} \; . \tag{53b}$$

The c'-Fenchel-Moreau conjugate *of a function* $g : \mathbb{Y} \to \overline{\mathbb{R}}$*, with respect to the coupling* c'*, is the function* $g^{c'} : \mathbb{X} \to \overline{\mathbb{R}}$ *defined by*

$$g^{c'}(x) = \sup_{y \in \mathbb{Y}} \Big(c(x,y) \dotplus \big(- g(y) \big) \Big) \; , \;\; \forall x \in \mathbb{X} \; . \tag{53c}$$

The c-Fenchel-Moreau biconjugate *of a function* $f : \mathbb{X} \to \overline{\mathbb{R}}$*, with respect to the coupling* c*, is the function* $f^{cc'} : \mathbb{X} \to \overline{\mathbb{R}}$ *defined by*

$$f^{cc'}(x) = \big(f^{c} \big)^{c'}(x) = \sup_{y \in \mathbb{Y}} \Big(c(x,y) \dotplus \big(- f^{c}(y) \big) \Big) \; , \;\; \forall x \in \mathbb{X} \; . \tag{53d}$$

The biconjugate of a function $f : \mathbb{X} \to \overline{\mathbb{R}}$ satisfies

$$f^{cc'}(x) \leq f(x) \; , \;\; \forall x \in \mathbb{X} \; . \tag{53e}$$

A.2.2 The Fenchel conjugacy

When the sets $\mathbb{X}$ and $\mathbb{Y}$ are vector spaces equipped with a bilinear form $\langle \, , \rangle$, the corresponding conjugacy is the classical *Fenchel conjugacy.* For any functions $f : \mathbb{X} \to \overline{\mathbb{R}}$ and $g : \mathbb{Y} \to \overline{\mathbb{R}}$, we denote[5]

$$f^{\star}(y) = \sup_{x \in \mathbb{X}} \Big(\langle x \, , y \rangle \dotplus \big(- f(x) \big) \Big) \; , \;\; \forall y \in \mathbb{Y} \; , \tag{54a}$$

$$g^{\star'}(x) = \sup_{y \in \mathbb{Y}} \Big(\langle x \, , y \rangle \dotplus \big(- g(y) \big) \Big) \; , \;\; \forall x \in \mathbb{X} \; , \tag{54b}$$

$$f^{\star\star'}(x) = \sup_{y \in \mathbb{Y}} \Big(\langle x \, , y \rangle \dotplus \big(- f^{\star}(y) \big) \Big) \; , \;\; \forall x \in \mathbb{X} \; . \tag{54c}$$

For any function $h : \mathbb{W} \to \overline{\mathbb{R}}$, its *epigraph* is $\mathrm{epi} h = \big\{ (w,t) \in \mathbb{W} \times \mathbb{R} \, \big| \, h(w) \leq t \big\}$, its *effective domain* is $\mathrm{dom} h = \big\{ w \in \mathbb{W} \, \big| \, h(w) < +\infty \big\}$. A function $h : \mathbb{W} \to \overline{\mathbb{R}}$ is said to be *proper* if it never takes the value $-\infty$ and that $\mathrm{dom} h \neq \emptyset$. When $\mathbb{W}$ is equipped with a topology, the function $h : \mathbb{W} \to \overline{\mathbb{R}}$ is said to be *lower semi continuous (lsc)* if its epigraph is closed, and is said to be *closed* if h is either *lower semi continuous (lsc)* and nowhere having the value $-\infty$, or is the constant function $-\infty$ [10, p. 15].

It is proved that, when the two vector spaces $\mathbb{X}$ and $\mathbb{Y}$ are *paired* in the sense of convex analysis[6], the Fenchel conjugacy induces a one-to-one correspondence between the closed convex functions on $\mathbb{X}$ and the closed convex functions on $\mathbb{Y}$ [10, Theorem 5]. Here, a function is said to be *convex* if its epigraph is convex. Notice that the set of closed convex functions is the set of proper convex functions united with the two constant functions $-\infty$ and $+\infty$.

[5] In convex analysis, one does not use the notation $\star'$, but simply the notation $\star$, as it is often the case that $\mathbb{X} = \mathbb{Y}$ in the Euclidian and Hilbertian cases.

[6] That is, $\mathbb{X}$ and $\mathbb{Y}$ are equipped with a bilinear form $\langle \, , \rangle$, and locally convex topologies that are compatible in the sense that the continuous linear forms on $\mathbb{X}$ are the functions $x \in \mathbb{X} \mapsto \langle x \, , y \rangle$, for all $y \in \mathbb{Y}$, and that the continuous linear forms on $\mathbb{Y}$ are the functions $y \in \mathbb{Y} \mapsto \langle x \, , y \rangle$, for all $x \in \mathbb{X}$.

A.3 One-sided linear couplings

Background on epi-composition. Let $\mathbb{W}$ and $\mathbb{X}$ be any two sets. The *epi-composition* operation combines a function $h : \mathbb{W} \to \overline{\mathbb{R}}$ with a mapping $\theta : \mathbb{W} \to \mathbb{X}$ to get a function $\inf\big[h \mid \theta\big] : \mathbb{X} \to \overline{\mathbb{R}}$ defined by [11, p. 27]

$$\inf\big[h \mid \theta\big](x) = \inf\big\{h(w) \,\big|\, w \in \mathbb{W} \;,\; \theta(w) = x\big\} \;,\; \forall x \in \mathbb{X} \;, \tag{55a}$$

with the convention that $\inf \emptyset = +\infty$ (and with the consequence that $\theta : \mathbb{W} \to \mathbb{X}$ need not be defined on all $\mathbb{W}$, but only on $\mathrm{dom}(h) = \big\{w \in \mathbb{W} \,\big|\, h(w) < +\infty\big\}$, the *effective domain* of h). The epi-composition has the following *invariance property*

$$h = f \circ \theta \;\text{ where }\; f : \mathbb{X} \to \overline{\mathbb{R}} \Rightarrow \inf\big[h \mid \theta\big] = f \dotplus \delta_{\theta(\mathbb{W})} \;, \tag{55b}$$

where δ_Z denotes the *characteristic function* of a set Z:

$$\delta_Z(z) = \begin{cases} 0 & \text{if } z \in Z \;, \\ +\infty & \text{if } z \notin Z \;. \end{cases} \tag{56}$$

Definition of one-sided linear couplings c_θ.

Definition 18 *Let $\mathbb{X}$ and $\mathbb{Y}$ be two vector spaces equipped with a bilinear form $\langle\,,\rangle$. Let $\mathbb{W}$ be a set and*

$$\theta : \mathbb{W} \to \mathbb{X} \tag{57a}$$

be a mapping. We define the one-sided linear coupling c_θ *between $\mathbb{W}$ and $\mathbb{Y}$ by*

$$c_\theta : \mathbb{W} \times \mathbb{Y} \to \overline{\mathbb{R}} \;,\; c_\theta(w,y) = \langle \theta(w) \,, y\rangle \;,\; \forall w \in \mathbb{W} \;,\; \forall y \in \mathbb{Y} \;. \tag{57b}$$

Notice that, in a one-sided linear coupling, the second set posesses a linear structure (and is even paired with a vector space by means of a bilinear form), whereas the first set is not required to carry any structure.

c_θ-conjugates and biconjugates. Here are expressions for the conjugates and biconjugates of a function. We recall that, in convex analysis, $\sigma_X : \mathbb{Y} \to \overline{\mathbb{R}}$ denotes the *support function of a subset $X \subset \mathbb{X}$*:

$$\sigma_X(y) = \sup_{x \in X} \langle x \,, y\rangle \;,\; \forall y \in \mathbb{Y} \;. \tag{58}$$

Proposition 19 *For any function $g : \mathbb{Y} \to \overline{\mathbb{R}}$, the c'_θ-Fenchel-Moreau conjugate is given by*

$$g^{c'_\theta} = g^{\star'} \circ \theta \;. \tag{59a}$$

For any function $h : \mathbb{W} \to \overline{\mathbb{R}}$, the c_θ-Fenchel-Moreau conjugate is given by

$$h^{c_\theta} = \big(\inf\big[h \mid \theta\big]\big)^\star \;, \tag{59b}$$

where the epi-composition $\inf\big[h \mid \theta\big]$ *has been introduced in* (55a)*, and the c_θ-Fenchel-Moreau biconjugate is given by*

$$h^{c_\theta c_\theta{}'} = \big(h^{c_\theta}\big)^{\star'} \circ \theta = h^{c_\theta \star'} \circ \theta = \big(\inf\big[h \mid \theta\big]\big)^{\star\star'} \circ \theta \;. \tag{59c}$$

We observe that the c_θ-Fenchel-Moreau conjugate h^{c_θ} is a closed convex function on $\mathbb{R}^d$ (see §A.2.2). For any subset $W \subset \mathbb{W}$, the $(-c_\theta)$-Fenchel-Moreau conjugate of the characteristic function of W is given by

$$\delta_W^{-c_\theta} = \sigma_{-\theta(W)} \; , \;\; \forall W \subset \mathbb{W} \; . \tag{59d}$$

c_θ-convex functions. We recall that so-called *c_θ-convex functions* are all functions $h : \mathbb{R}^d \to \overline{\mathbb{R}}$ of the form $g^{c'_\theta}$, for any function $g : \mathbb{R}^d \to \overline{\mathbb{R}}$, or, equivalently, all functions of the form $h^{c_\theta c_\theta{}'}$, for any function $h : \mathbb{R}^d \to \overline{\mathbb{R}}$, or, equivalently, all functions that are equal to their c_θ-biconjugate ($h^{c_\theta c_\theta{}'} = h$) [13, 12, 8].

Proposition 20 *A function is c_θ-convex if and only if it is the composition of a closed convex function on $\mathbb{R}^d$ with the mapping θ in* (57a)*. More precisely, for any function $h : \mathbb{R}^d \to \overline{\mathbb{R}}$, we have the equivalences*

$$h \text{ is } c_\theta\text{-convex} \tag{60a}$$

$$\iff h = h^{c_\theta c_\theta{}'} \tag{60b}$$

$$\iff h = \underbrace{\big(h^{c_\theta}\big)^{\star'}}_{\text{closed convex function}} \circ\theta \tag{60c}$$

$$\iff \text{there exists a closed convex function } f : \mathbb{X} \to \overline{\mathbb{R}} \text{ such that } h = f \circ \theta \; . \tag{60d}$$

Proof. If $h^{c_\theta c_\theta{}'} = h$, then $h = \big(h^{c_\theta}\big)^{\star'} \circ \theta$ by (59c), where the function $\big(h^{c_\theta}\big)^{\star'}$ is closed convex.

If there exists a closed convex function $f : \mathbb{X} \to \overline{\mathbb{R}}$ such that $h = f \circ \theta$, then $\inf\big[h \mid \theta\big] = f \dot{+} \delta_{\theta(\mathbb{W})}$ by (55b), and therefore $h^{c_\theta c_\theta{}'} = \big(\inf\big[h \mid \theta\big]\big)^{\star\star'} \circ\theta = \big(f \dot{+} \delta_{\theta(\mathbb{W})}\big)^{\star\star'} \circ\theta$ by (59c). Now, as $f \dot{+} \delta_{\theta(\mathbb{W})} \geq f$, we get that $\big(f \dot{+} \delta_{\theta(\mathbb{W})}\big)^{\star\star'} \geq f^{\star\star'} = f$, where the last equality holds because the function $f : \mathbb{X} \to \overline{\mathbb{R}}$ is closed convex. As a consequence, we obtain that $h^{c_\theta c_\theta{}'} \geq f \circ \theta = h$. Now, by (53e), we always have the inequality $h^{c_\theta c_\theta{}'} \leq h$. Thus, we conclude that $h^{c_\theta c_\theta{}'} = h$.

This ends the proof. □

c_θ-subdifferential. Following the definition of the subdifferential of a function with respect to a duality in [1], we define the *c_θ-subdifferential* of the function $h : \mathbb{R}^d \to \overline{\mathbb{R}}$ at $w \in \mathbb{R}^d$ by

$$\partial_{c_\theta} h(w) = \big\{ y \in \mathbb{R}^d \,\big|\, c_\theta(w', y) \dot{+} \big(-h(w')\big) \leq c_\theta(w, y) \dot{+} \big(-h(w)\big) \; , \;\; \forall w' \in \mathbb{R}^d \big\} \tag{61a}$$

$$= \big\{ y \in \mathbb{R}^d \,\big|\, h^{c_\theta}(y) = c_\theta(w, y) \dot{+} \big(-h(w)\big) \big\} \tag{61b}$$

$$= \big\{ y \in \mathbb{R}^d \,\big|\, \big(\inf\big[h \mid \theta\big]\big)^{\star}(y) = \langle \theta(w) \,, y \rangle \dot{+} \big(-h(w)\big) \big\} \; . \tag{61c}$$

References

[1] M. Akian, S. Gaubert, and V. Kolokoltsov. Invertibility of functional Galois connections. *Comptes Rendus Mathematique*, 335(11):883 – 888, 2002.

[2] C. D. Aliprantis and K. C. Border. *Infinite dimensional analysis.* Springer-Verlag, Berlin, second edition, 1999.

[3] H. H. Bauschke and P. L. Combettes. *Convex analysis and monotone operator theory in Hilbert spaces.* CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC. Springer-Verlag, second edition, 2017.

[4] J.-P. Chancelier and M. De Lara. Hidden convexity in the l_0 pseudonorm, June 2019. Preprint hal-02146454, arXiv:1906.04038.

[5] J.-P. Chancelier and M. De Lara. Orthant-strictly monotonic norms, graded sequences and generalized top-k and k-support norms for sparse optimization, 2020. Preprint hal-02456433, arXiv:2001.10441.

[6] J.-P. Chancelier and M. De Lara. Variational formulations for the l_0 pseudonorm and applications to sparse optimization, 2020. Preprint.

[7] J. B. Hiriart-Ururty and C. Lemaréchal. *Convex Analysis and Minimization Algorithms I.* Springer-Verlag, Berlin, 1993.

[8] J. E. Martínez-Legaz. Generalized convex duality and its economic applications. In S. S. Hadjisavvas N., Komlósi S., editor, *Handbook of Generalized Convexity and Generalized Monotonicity. Nonconvex Optimization and Its Applications*, volume 76, pages 237–292. Springer-Verlag, 2005.

[9] J. J. Moreau. Inf-convolution, sous-additivité, convexité des fonctions numériques. *J. Math. Pures Appl. (9)*, 49:109–154, 1970.

[10] R. T. Rockafellar. *Conjugate Duality and Optimization.* CBMS-NSF Regional Conference Series in Applied Mathematics. Society for Industrial and Applied Mathematics, 1974.

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

[12] A. Rubinov. *Abstract convexity and global optimization*, volume 44 of *Nonconvex Optimization and its Applications.* Kluwer Academic Publishers, Dordrecht, 2000.

[13] I. Singer. *Abstract Convex Analysis.* Canadian Mathematical Society Series of Monographs and Advanced Texts. John Wiley & Sons, Inc., New York, 1997.

[14] C. Zalinescu. *Convex Analysis in General Vector Spaces.* World Scientific, 2002.