



HAL
open science

Constant Along Primal Rays Conjugacies and the 10 Pseudonorm

Jean-Philippe Chancelier, Michel de Lara

► **To cite this version:**

Jean-Philippe Chancelier, Michel de Lara. Constant Along Primal Rays Conjugacies and the 10 Pseudonorm. 2020. hal-02459673v1

HAL Id: hal-02459673

<https://hal.science/hal-02459673v1>

Preprint submitted on 29 Jan 2020 (v1), last revised 1 Jun 2021 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

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

$$\text{supp}(x) = \{j \in \{1, \dots, d\} \mid x_j \neq 0\} \subset \{1, \dots, d\}. \quad (1)$$

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

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

where $|K|$ denotes the cardinal of a subset $K \subset \{1, \dots, d\}$. 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:

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

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*

$$\ell_0^{\leq k} = \{x \in \mathbb{R}^d \mid \ell_0(x) \leq k\}, \quad \forall k \in \{0, 1, \dots, d\}, \quad (4a)$$

and the *level curves*

$$\ell_0^{=k} = \{x \in \mathbb{R}^d \mid \ell_0(x) = k\}, \quad \forall k \in \{0, 1, \dots, d\}. \quad (4b)$$

For any subset $K \subset \{1, \dots, d\}$, we denote the subspace of \mathbb{R}^d made of vectors whose components vanish outside of K by¹

$$\mathcal{R}_K = \mathbb{R}^K \times \{0\}^{-K} = \{x \in \mathbb{R}^d \mid x_j = 0, \forall j \notin K\} \subset \mathbb{R}^d, \quad (5)$$

where $\mathcal{R}_\emptyset = \{0\}$. We denote by $\pi_K : \mathbb{R}^d \rightarrow \mathcal{R}_K$ 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

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

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

$$\ell_0^{\leq k} = \{x \in \mathbb{R}^d \mid \ell_0(x) \leq k\} = \bigcup_{|K| \leq k} \mathcal{R}_K, \quad \forall k = 0, 1, \dots, d, \quad (7)$$

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

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 $\{\|\cdot\|_k\}_{k=1, \dots, d}$ 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 $l = 1, \dots, d$,*

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

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

$$\ell_0(x) \leq l \Rightarrow \|x\|_l = \|x\|_d. \quad (8b)$$

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

$$\ell_0(x) \geq \min \{k \in \{1, \dots, d\} \mid \|x\|_k = \|x\|_d\}. \quad (8c)$$

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

We say that a sequence $\{\|\cdot\|_k\}_{k=1,\dots,d}$ 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 $l = 1, \dots, d$,

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

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

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

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

$$\ell_0(x) = \min \{k \in \{1, \dots, d\} \mid \|x\|_k = \|x\|_d\}. \quad (9c)$$

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

$$\mathbb{S} = \{x \in \mathbb{R}^d \mid \|x\| = 1\}, \quad (10a)$$

$$\mathbb{B} = \{x \in \mathbb{R}^d \mid \|x\| \leq 1\}. \quad (10b)$$

Dual norms. We recall that the following expression

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

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

$$\mathbb{S}_\star = \{y \in \mathbb{R}^d \mid \|y\|_\star = 1\}, \quad (12a)$$

$$\mathbb{B}_\star = \{y \in \mathbb{R}^d \mid \|y\|_\star \leq 1\}. \quad (12b)$$

We have

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

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}° of the unit ball \mathbb{B} :

$$\mathbb{B}_\star = \mathbb{B}^\circ = \{y \in \mathbb{R}^d \mid \langle x, y \rangle \leq 1, \forall x \in \mathbb{B}\}. \quad (13b)$$

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

$$\mathbb{B}^{\circ\circ} = (\mathbb{B}^\circ)^\circ = \mathbb{B}, \quad (13c)$$

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

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

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

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

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,

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

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]

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

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:

$$\left(y \parallel_{\|\cdot\|} x \iff y \in N_{\mathbb{B}}\left(\frac{x}{\|x\|}\right) \iff x \in N_{\mathbb{B}_\star}\left(\frac{y}{\|y\|_\star}\right) \right), \quad \forall (x, y) \in \mathbb{R}^d \setminus \{0\} \times \mathbb{R}^d \setminus \{0\}. \quad (16)$$

Restriction norms.

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

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

$$\|x\|_K = \|x\|, \quad \forall x \in \mathcal{R}_K. \quad (17)$$

- the (K, \star) -norm $\|\cdot\|_{K, \star}$, on the subspace \mathcal{R}_K of \mathbb{R}^d , which is the norm $(\|\cdot\|_K)_\star$, 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)]

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

source norm $\ \cdot\ $	$\ \cdot\ _{(k)}^{\mathcal{R}}$	$\ \cdot\ _{(k),\star}^{\mathcal{R}}$
$\ \cdot\ _p$	(p, k) -support norm $\ x\ _{p,k}^{\text{sn}}$	top (k, q) -norm $\ y\ _{k,q}^{\text{tn}}$ $= \left(\sum_{j=1}^k y_{\nu(j)} ^q\right)^{1/q}, 1/p + 1/q = 1$
$\ \cdot\ _1$	$(1, k)$ -support norm ℓ_1 -norm $\ x\ _{1,k}^{\text{sn}} = \ x\ _1$	top (k, ∞) -norm ℓ_∞ -norm $\ y\ _{k,\infty}^{\text{tn}} = y_{\nu(1)} = \ y\ _\infty$
$\ \cdot\ _2$	$(2, k)$ -support norm	top $(k, 2)$ -norm $\ y\ _{k,2}^{\text{tn}} = \sqrt{\sum_{j=1}^k y_{\nu(j)} ^2}$
$\ \cdot\ _\infty$	(∞, k) -support norm	top $(k, 1)$ -norm $\ y\ _{k,1}^{\text{tn}} = \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 $k \in \{1, \dots, d\}$, we call coordinate- k norm the norm $\|\cdot\|_{(k)}^{\mathcal{R}}$ whose dual norm is the dual coordinate- k norm, denoted by $\|\cdot\|_{(k),\star}^{\mathcal{R}}$, with expression

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

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 $\sup_{K \subset \{1, \dots, d\}, |K| \leq k}$.

It is easily verified that $\|\cdot\|_{(k),\star}^{\mathcal{R}}$ indeed is a norm. We will adopt the convention $\|\cdot\|_{(0),\star}^{\mathcal{R}} = 0$ (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 $\{1, \dots, d\}$ such that $|y_{\nu(1)}| \geq |y_{\nu(2)}| \geq \dots \geq |y_{\nu(d)}|$. 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 $\|\cdot\|_{(k),\star}^{\mathcal{R}}$ in Definition 3 by

$$\mathbb{S}_{(k),\star}^{\mathcal{R}} = \{y \in \mathbb{R}^d \mid \|y\|_{(k),\star}^{\mathcal{R}} = 1\}, \quad k = 1, \dots, d, \quad (20a)$$

$$\mathbb{B}_{(k),\star}^{\mathcal{R}} = \{y \in \mathbb{R}^d \mid \|y\|_{(k),\star}^{\mathcal{R}} \leq 1\}, \quad k = 1, \dots, d. \quad (20b)$$

Proposition 4

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

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

- We have the equality

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

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

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

- The sequence $\left\{ \mathbb{B}_{(j),\star}^{\mathcal{R}} \right\}_{j=1, \dots, d}$ 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:

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

Proof.

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

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

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

First, we establish the (known) fact that $\overline{\ell_0^{\overline{=k}}} = \ell_0^{\leq k}$. The inclusion $\overline{\ell_0^{\overline{=k}}} \subset \ell_0^{\leq k}$ is easy because, on the one hand, $\ell_0^{\overline{=k}} \subset \ell_0^{\leq k}$ 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 $\ell_0^{\leq k} \subset \overline{\ell_0^{\overline{=k}}}$. For this purpose, we consider $x \in \ell_0^{\leq k}$. If $x \in \ell_0^{\overline{=k}}$, obviously $x \in \overline{\ell_0^{\overline{=k}}}$. Therefore, we suppose that $\ell_0(x) = l < k$. By definition of $\ell_0(x)$ in (2), there exists $L \subset \{1, \dots, d\}$ 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 $x^\epsilon \rightarrow x$ when $\epsilon \rightarrow 0$. This proves that $\ell_0^{\leq k} \subset \overline{\ell_0^{\overline{=k}}}$.

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

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

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

$$\mathbb{S}_{(k)}^{\mathcal{R}} = \{x \in \mathbb{R}^d \mid \|x\|_{(k)}^{\mathcal{R}} = 1\}, \quad (25a)$$

$$\mathbb{B}_{(k)}^{\mathcal{R}} = \{x \in \mathbb{R}^d \mid \|x\|_{(k)}^{\mathcal{R}} \leq 1\}. \quad (25b)$$

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

Proposition 5

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

$$\mathbb{B}_{(k)}^{\mathcal{R}} = \overline{\text{co}}\left(\bigcup_{|K| \leq k} (\mathcal{R}_K \cap \mathbb{S})\right), \quad (26)$$

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

- We have the equality

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

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

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

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

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

Proof.

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

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

and we conclude that $\mathbb{B}_{(k)}^{\mathcal{R}} = \overline{\text{co}}(\bigcup_{|K| \leq k} (\mathcal{R}_K \cap \mathbb{S}))$ 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 $\mathbb{B}_{(j)}^{\mathcal{R}} = (\mathbb{B}_{(j), \star}^{\mathcal{R}})^{\circ}$, the polar set of $\mathbb{B}_{(j), \star}^{\mathcal{R}}$.

This ends the proof. □

We recall that the normed space $(\mathbb{R}^d, \|\cdot\|)$ 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 $(\mathbb{R}^d, \|\cdot\|_p)$, 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 $\left\{\|\cdot\|_{(j)}^{\mathcal{R}}\right\}_{j=1, \dots, d}$ 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 $\left\{ \|\cdot\|_{(j)}^{\mathcal{R}} \right\}_{j=1,\dots,d}$ of coordinate- k norms is decreasingly graded with respect to the ℓ_0 pseudonorm, that is, for any $l = 1, \dots, d$,

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

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

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

Proof.

- We prove Item 1. As the sequence $\left\{ \|\cdot\|_{(j)}^{\mathcal{R}} \right\}_{j=1,\dots,d}$ 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 $k \in \{1, \dots, d\}$, we have²

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

Therefore, we have obtained (30a).

- We prove Item 2. As the sequence $\left\{ \|\cdot\|_{(j)}^{\mathcal{R}} \right\}_{j=1,\dots,d}$ 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).

²In 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 $(\mathbb{R}^d, \|\cdot\|)$ is strictly convex. Then, for any $x \in \mathbb{R}^d$ and for any $k \in \{1, \dots, d\}$, we have³

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

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

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

Therefore, we have obtained (30b).

This ends the proof. □

	$\left\{ \ \cdot\ _{(j)}^{\mathcal{R}} \right\}_{j=1, \dots, d}$	
	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 $\left\{ \|\cdot\|_{(j)}^{\mathcal{R}} \right\}_{j=1, \dots, d}$ be graded (second column), it suffices that $\|\cdot\|$ be any norm; to obtain that the sequence $\left\{ \|\cdot\|_{(j)}^{\mathcal{R}} \right\}_{j=1, \dots, d}$ 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 $\left\{ \|\cdot\|_{p,j}^{\text{sn}} \right\}_{j=1, \dots, d}$ 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):

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

³See Footnote 2.

We also have that the sequence $\{\|\cdot\|_{p,j}^{\text{sn}}\}_{j=1,\dots,d}$ 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):

$$\ell_0(x) \geq \min \left\{ k \in \{1, \dots, d\} \mid \|x\|_{\infty, k}^{\text{sn}} = \|x\|_{\infty} \right\}, \quad \forall x \in \mathbb{R}^d. \quad (31b)$$

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

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

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 $\mathbb{B}_{(k)}^{\mathcal{R}}$ in (25b) is the unit ball of the norm $\|\cdot\|_{(k)}^{\mathcal{R}}$.

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

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

$$\begin{aligned} \ell_0^{\leq k} \cap \mathbb{S} &= \overline{\text{co}}(\ell_0^{\leq k} \cap \mathbb{S}) \cap \mathbb{S} \\ &\text{(because } \ell_0^{\leq k} \cap \mathbb{S} \subset \mathbb{S} \text{ and is closed, and because the unit ball } \mathbb{B} \text{ is rotund)} \\ &= \overline{\text{co}}\left(\bigcup_{|K| \leq k} (\mathcal{R}_K \cap \mathbb{S})\right) \cap \mathbb{S} && \text{(by (7))} \\ &= \mathbb{B}_{(k)}^{\mathcal{R}} \cap \mathbb{S}. && \text{(by (26))} \end{aligned}$$

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 ζ , or *Capra*, between \mathbb{R}^d and \mathbb{R}^d by

$$\forall y \in \mathbb{R}^d, \quad \begin{cases} \zeta(x, y) &= \frac{\langle x, y \rangle}{\|x\|}, \quad \forall x \in \mathbb{R}^d \setminus \{0\}, \\ \zeta(0, y) &= 0. \end{cases} \quad (33)$$

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:

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

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*:

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

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 $g : \mathbb{R}^d \rightarrow \overline{\mathbb{R}}$, the ζ' -Fenchel-Moreau conjugate is given by

$$g^{\zeta'} = g^{\star'} \circ n. \quad (35a)$$

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

$$f^{\zeta} = (\inf [f | n])^{\star}, \quad (35b)$$

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

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

and the ζ -Fenchel-Moreau biconjugate is given by

$$f^{\zeta\zeta'} = (f^{\zeta})^{\star'} \circ n = (\inf [f | n])^{\star\star'} \circ n. \quad (35d)$$

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

Capra-convex functions. We recall that so-called $\dot{\varsigma}$ -convex functions are all functions of the form $g^{\dot{\varsigma}'}$, for any $g : \mathbb{R}^d \rightarrow \overline{\mathbb{R}}$, or, equivalently, all functions of the form $f^{\dot{\varsigma}\dot{\varsigma}'}$, for any $f : \mathbb{R}^d \rightarrow \overline{\mathbb{R}}$, or, equivalently, all functions that are equal to their $\dot{\varsigma}$ -biconjugate ($f^{\dot{\varsigma}\dot{\varsigma}'} = f$) [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 $\dot{\varsigma}$ -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 $h : \mathbb{R}^d \rightarrow \overline{\mathbb{R}}$, we have the equivalences*

h is $\dot{\varsigma}$ -convex

$$\Leftrightarrow h = h^{\dot{\varsigma}\dot{\varsigma}'}$$

$$\Leftrightarrow h = (h^{\dot{\varsigma}})^{\star'} \circ n \text{ (where } (h^{\dot{\varsigma}})^{\star'} \text{ is a closed convex function)}$$

$$\Leftrightarrow \text{there exists a closed convex function } f : \mathbb{R}^d \rightarrow \overline{\mathbb{R}} \text{ such that } h = f \circ n .$$

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 $f : \mathbb{R}^d \rightarrow \overline{\mathbb{R}}$ at $x \in \mathbb{R}^d$ has the following expressions

$$\partial_{\dot{\varsigma}} f(x) = \{y \in \mathbb{R}^d \mid \dot{\varsigma}(x', y) \dot{+} (-f(x')) \leq \dot{\varsigma}(x, y) \dot{+} (-f(x)), \forall x' \in \mathbb{R}^d\} \quad (37a)$$

$$= \{y \in \mathbb{R}^d \mid f^{\dot{\varsigma}}(y) = \dot{\varsigma}(x, y) \dot{+} (-f(x))\} \quad (37b)$$

$$= \{y \in \mathbb{R}^d \mid (\inf [f \mid n])^{\star}(y) = \langle n(x), y \rangle \dot{+} (-f(x))\}, \quad (37c)$$

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

$$\partial_{\dot{\varsigma}} f(0) = \{y \in \mathbb{R}^d \mid (\inf [f \mid n])^{\star}(y) = -f(0)\} \quad (37d)$$

$$\partial_{\dot{\varsigma}} f(x) = \{y \in \mathbb{R}^d \mid (\inf [f \mid n])^{\star}(y) = \frac{\langle x, y \rangle}{\|x\|} \dot{+} (-f(x)), \forall x \in \mathbb{R}^d \setminus \{0\}\}. \quad (37e)$$

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 $\delta_{\ell_0^{\leq k}}^{\star} = +\infty$, for all $k = 0, 1, \dots, d$ — 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 $\left\{ \|\cdot\|_{(j),\star}^{\mathcal{R}} \right\}_{j=1,\dots,d}$ of dual coordinate- k norms in Definition 3, and associated coupling \dagger in (33).*

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

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

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

Proof. We prove (38):

$$(\varphi \circ \ell_0)^{\dagger} = \left(\inf_{j=0,1,\dots,d} [\delta_{\ell_0^{\leq j}} \dagger \varphi(j)] \right)^{\dagger}$$

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

$$\begin{aligned} &= \sup_{j=0,1,\dots,d} [\delta_{\ell_0^{\leq j}} \dagger \varphi(j)]^{\dagger} \quad (\text{as conjugacies, being dualities, turn infima into suprema}) \\ &= \sup_{j=0,1,\dots,d} [\delta_{\ell_0^{\leq j}}^{\dagger} \dagger (-\varphi(j))] \quad (\text{by property of conjugacies}) \\ &= \sup_{j=0,1,\dots,d} [\sigma_{n(\ell_0^{\leq j})} \dagger (-\varphi(j))] \quad (\text{as } \delta_{\ell_0^{\leq j}}^{\dagger} = \sigma_{n(\ell_0^{\leq j})} \text{ by (59d)}) \\ &= \sup_{j=0,1,\dots,d} \left\{ \sup \{0, \sigma_{\ell_0^{\leq j} \cap \mathbb{S}}\} \dagger (-\varphi(j)) \right\} \end{aligned}$$

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

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

This ends the proof. \square

With φ the identity function on $\{0, 1, \dots, d\}$, we find the Capra-conjugate of the ℓ_0 pseudonorm. With the functions $\varphi = \delta_{\{0,1,\dots,k\}}$ (for any $k \in \{0, 1, \dots, d\}$), 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 $k = 0, 1, \dots, d$, 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 $\{\|\cdot\|_{(j)}^{\mathcal{R}}\}_{j=1,\dots,d}$ of coordinate- k norms and sequence $\{\|\cdot\|_{\star(j)}^{\mathcal{R}}\}_{j=1,\dots,d}$ of dual coordinate- k norms, as in Definition 3, and with associated Capra coupling $\dot{\zeta}$ in (33).*

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

$$(\varphi \circ \ell_0)^{\dot{\zeta}\dot{\zeta}'}(x) = ((\varphi \circ \ell_0)^{\dot{\zeta}})^{\star'}\left(\frac{x}{\|x\|}\right), \quad \forall x \in \mathbb{R}^d \setminus \{0\}, \quad (39a)$$

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

$$((\varphi \circ \ell_0)^{\dot{\zeta}})^{\star'} = \left(\sup_{j=0,1,\dots,d} [\|\cdot\|_{(j),\star}^{\mathcal{R}} - \varphi(j)] \right)^{\star'}, \quad (39b)$$

and also has the following four expressions as a Fenchel biconjugate

$$= \left(\inf_{j=0,1,\dots,d} [\delta_{\mathbb{B}_{(j)}^{\mathcal{R}}} \dot{+} \varphi(j)] \right)^{\star\star'}, \quad (39c)$$

hence the function $((\varphi \circ \ell_0)^{\dot{\zeta}})^{\star'}$ is the largest closed convex function below the integer valued function $\inf_{j=0,1,\dots,d} [\delta_{\mathbb{B}_{(j)}^{\mathcal{R}}} \dot{+} \varphi(j)]$, which is such that $x \in \mathbb{B}_{(j)}^{\mathcal{R}} \setminus \mathbb{B}_{(j-1)}^{\mathcal{R}} \mapsto \varphi(j)$ for $l = 1, \dots, d$, 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 $\mathbb{B}_{(1)}^{\mathcal{R}} \subset \dots \subset \mathbb{B}_{(j-1)}^{\mathcal{R}} \subset \mathbb{B}_{(j)}^{\mathcal{R}} \subset \dots \subset \mathbb{B}_{(d)}^{\mathcal{R}} = \mathbb{B}$ by (24)), that is,

$$= \left(x \mapsto \inf \{ \varphi(j) \mid x \in \mathbb{B}_{(j)}^{\mathcal{R}}, j \in \{0, 1, \dots, d\} \} \right)^{\star\star'}, \quad (39d)$$

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

$$= \left(\inf_{j=0,1,\dots,d} [\delta_{\mathbb{S}_{(j)}^{\mathcal{R}}} \dot{+} \varphi(j)] \right)^{\star\star'}, \quad (39e)$$

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

$$= \left(x \mapsto \inf \{ \varphi(j) \mid x \in \mathbb{S}_{(j)}^{\mathcal{R}}, j \in \{0, 1, \dots, d\} \} \right)^{\star\star'}, \quad (39f)$$

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

(39g)

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

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

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

3. For any function $\varphi : \{0, 1, \dots, d\} \rightarrow \mathbb{R}_+$, that is, with nonnegative finite values, and such that $\varphi(0) = 0$, the function $((\varphi \circ \ell_0)^{\dot{\zeta}})^{\star'}$ is proper convex lsc and has the following two variational expressions⁴

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

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

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

$$(\varphi \circ \ell_0)^{\dot{\zeta}\dot{\zeta}'}(x) = \frac{1}{\|x\|} \min_{\substack{z^{(1)} \in \mathbb{R}^d, \dots, z^{(d)} \in \mathbb{R}^d \\ \sum_{j=1}^d \|z^{(j)}\|_{(j)}^{\mathcal{R}} \leq \|x\| \\ \sum_{j=1}^d z^{(j)} = x}} \sum_{j=1}^d \|z^{(j)}\|_{(j)}^{\mathcal{R}} \varphi(j), \quad \forall x \in \mathbb{R}^d \setminus \{0\}. \quad (40)$$

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

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

1. Let the function $\varphi : \{0, 1, \dots, d\} \rightarrow \overline{\mathbb{R}}$ 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

$$\begin{aligned}
((\varphi \circ \ell_0)^{\dot{\zeta}})^{\star'} &= \left(\sup_{j=0,1,\dots,d} [\|\cdot\|_{\mathbb{B}(j),\star}^{\mathcal{R}} - \varphi(j)] \right)^{\star'} && \text{(by (38))} \\
&= \left(\sup_{j=0,1,\dots,d} [\sigma_{\mathbb{B}(j)}^{\mathcal{R}} - \varphi(j)] \right)^{\star'} \\
\text{(by (13a) as } \mathbb{B}_{(j)}^{\mathcal{R}} \text{ is the unit ball of the norm } \|\cdot\|_{\mathbb{B}(j)}^{\mathcal{R}} \text{ by (25b), and with the convention } \mathbb{B}_{(0)}^{\mathcal{R}} = \{0\}) & \\
&= \left(\sup_{j=0,1,\dots,d} [\delta_{\mathbb{B}(j)}^{\star\mathcal{R}} - \varphi(j)] \right)^{\star'} && \text{(because } \delta_{\mathbb{B}(j)}^{\star\mathcal{R}} = \sigma_{\mathbb{B}(j)}^{\mathcal{R}} \text{)} \\
&= \left(\sup_{j=0,1,\dots,d} (\delta_{\mathbb{B}(j)}^{\mathcal{R}} + \varphi(j))^{\star} \right)^{\star'} && \text{(by property of conjugacies)} \\
&= \left(\left(\inf_{j=0,1,\dots,d} [\delta_{\mathbb{B}(j)}^{\mathcal{R}} + \varphi(j)] \right)^{\star} \right)^{\star'} \\
&\quad \text{(as conjugacies, being dualities, turn infima into suprema)} \\
&= \left(\inf_{j=0,1,\dots,d} [\delta_{\mathbb{B}(j)}^{\mathcal{R}} + \varphi(j)] \right)^{\star\star'} && \text{(by (54c))}
\end{aligned}$$

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 $\mathbb{B}_{(j)}^{\mathcal{R}}$ by the spheres $\mathbb{S}_{(j)}^{\mathcal{R}}$, since $\|\cdot\|_{\mathbb{B}(j),\star}^{\mathcal{R}} = \sigma_{\mathbb{S}(j)}^{\mathcal{R}} = \delta_{\mathbb{S}(j)}^{\star\mathcal{R}}$. Thus, we obtain (39e) and (39f).

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

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

$$\bigcap_{j=0,1,\dots,n} \text{dom } f_j \neq \emptyset \Rightarrow \left(\sup_{j=0,1,\dots,n} f_j \right)^{\star} = \min_{(\lambda_0, \lambda_1, \dots, \lambda_n) \in \Delta_{n+1}} \left(\sum_{j=0}^n \lambda_j f_j \right)^{\star}, \quad (41)$$

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

We obtain

$$((\varphi \circ \ell_0)^{\dot{\zeta}})^{\star'} = \left(\sup_{j=0,1,\dots,d} \left[\|\cdot\|_{\mathbb{B}_{(j)}^{\mathcal{R}}} - \varphi(j) \right] \right)^{\star'} \quad (\text{by (38)})$$

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

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

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

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

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

as, for all $j = 1, \dots, d$, $\lambda_j \sigma_{\mathbb{B}_{(j)}^{\mathcal{R}}} = \sigma_{\lambda_j \mathbb{B}_{(j)}^{\mathcal{R}}}$ 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]

$$= \min_{(\lambda_0, \lambda_1, \dots, \lambda_d) \in \Delta_{d+1}} \left(\sigma_{\sum_{j=1}^d \lambda_j \mathbb{B}_{(j)}^{\mathcal{R}}} - \sum_{j=0}^d \lambda_j \varphi(j) \right)^{\star'} \quad (\text{thanks to the convention } \mathbb{B}_{(0)}^{\mathcal{R}} = \{0\})$$

$$= \min_{(\lambda_0, \lambda_1, \dots, \lambda_d) \in \Delta_{d+1}} \left(\left(\sigma_{\sum_{j=1}^d \lambda_j \mathbb{B}_{(j)}^{\mathcal{R}}} \right)^{\star'} + \sum_{j=0}^d \lambda_j \varphi(j) \right) \quad (\text{by property of conjugacies})$$

$$= \min_{(\lambda_0, \lambda_1, \dots, \lambda_d) \in \Delta_{d+1}} \left(\delta_{\sum_{j=1}^d \lambda_j \mathbb{B}_{(j)}^{\mathcal{R}}} + \sum_{j=0}^d \lambda_j \varphi(j) \right) \quad (\text{because } \sum_{j=1}^d \lambda_j \mathbb{B}_{(j)}^{\mathcal{R}} \text{ is a closed convex set.})$$

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

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

which is (39h).

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

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

$$\begin{aligned} ((\varphi \circ \ell_0)^{\dot{\zeta}})^{\star'}(x) &= \min_{\substack{(\lambda_0, \lambda_1, \dots, \lambda_d) \in \Delta_{d+1} \\ x \in \sum_{j=1}^d \lambda_j \mathbb{B}_{(j)}^{\mathcal{R}}}} \sum_{j=1}^d \lambda_j \varphi(j) && \text{(because } \varphi(0) = 0) \\ &= \min_{\substack{z^{(1)} \in \mathbb{B}_{(1)}^{\mathcal{R}}, \dots, z^{(d)} \in \mathbb{B}_{(d)}^{\mathcal{R}} \\ \lambda_1 \geq 0, \dots, \lambda_d \geq 0 \\ \sum_{j=1}^d \lambda_j \leq 1 \\ \sum_{j=1}^d \lambda_j z^{(j)} = x}} \sum_{j=1}^d \lambda_j \varphi(j) \end{aligned}$$

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

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

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

$$\begin{aligned} &= \min_{\substack{z^{(1)} \in \mathbb{R}^d, \dots, z^{(d)} \in \mathbb{R}^d \\ \sum_{j=1}^d \|\|z^{(j)}\|\|_{(j)}^{\mathcal{R}} \leq 1 \\ \sum_{j=1}^d z^{(j)} = x}} \sum_{j=1}^d \varphi(j) \|\|z^{(j)}\|\|_{(j)}^{\mathcal{R}}, \end{aligned}$$

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

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

$$(\varphi \circ \ell_0)^{\dot{\zeta}\dot{\zeta}'}(x) = \frac{1}{\|\|x\|\|} \min_{\substack{z^{(1)} \in \mathbb{R}^d, \dots, z^{(d)} \in \mathbb{R}^d \\ \sum_{j=1}^d \|\|z^{(j)}\|\|_{(j)}^{\mathcal{R}} \leq \|\|x\|\| \\ \sum_{j=1}^d z^{(j)} = x}} \sum_{j=1}^d \varphi(j) \|\|z^{(j)}\|\|_{(j)}^{\mathcal{R}}, \quad \forall x \in \mathbb{R}^d \setminus \{0\},$$

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

This ends the proof. \square

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

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

The following statements are equivalent.

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

$$\delta_{\ell_0^{\leq k}}^{\zeta\zeta'} = \delta_{\ell_0^{\leq k}}, \quad k = 0, 1, \dots, d. \quad (42)$$

Proof.

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

$$\begin{aligned} \delta_{\ell_0^{\leq k}}^{\zeta\zeta'} &= \left(\inf_{j=0,1,\dots,d} \left[\delta_{\mathbb{B}_{(j)}^{\mathcal{R}}} \dot{+} \delta_{\{0,1,\dots,k\}}(j) \right] \right)^{\star\star'} \circ n \quad (\text{by (39c) with the functions } \varphi = \delta_{\{0,1,\dots,k\}}) \\ &= \left(\inf_{j=0,1,\dots,k} \delta_{\mathbb{B}_{(j)}^{\mathcal{R}}} \right)^{\star\star'} \circ n \\ &= \left(\delta_{\mathbb{B}_{(k)}^{\mathcal{R}}} \right)^{\star\star'} \circ n \\ &\quad (\text{by the inclusions } \mathbb{B}_{(1)}^{\mathcal{R}} \subset \dots \subset \mathbb{B}_{(k)}^{\mathcal{R}} \text{ in (29) and by the convention } \mathbb{B}_{(0)}^{\mathcal{R}} = \{0\}) \\ &= \delta_{\mathbb{B}_{(k)}^{\mathcal{R}}} \circ n \quad (\text{because the unit ball } \mathbb{B}_{(k)}^{\mathcal{R}} \text{ is closed and convex}) \\ &= \delta_{n^{-1}(\mathbb{B}_{(k)}^{\mathcal{R}})} \end{aligned}$$

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

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

Therefore, we have

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

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

This ends the proof. \square

Notice that, by Item 2 in Proposition 6, it suffices that the normed space $(\mathbb{R}^d, \|\cdot\|)$ be strictly convex to obtain that the characteristic functions $\delta_{\ell_0^{\leq k}}$ are ζ -convex, for all $k = 0, 1, \dots, d$. 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 ζ -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 $k = 0, 1, \dots, d$ (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 $x \in \mathbb{R}^d \setminus \{0\}$ (indeed, this is a consequence of $\ell_0^{\star\star}(x) = 0 \neq \ell_0(x)$ when $x \in \mathbb{R}^d \setminus \{0\}$). 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 $\left\{ \|\cdot\|_{\star(j)}^{\mathcal{R}} \right\}_{j=1, \dots, d}$ of dual coordinate- k norms, as in Definition 3, and associated coupling ζ in (33).*

Let a function $\varphi : \{0, 1, \dots, d\} \rightarrow \overline{\mathbb{R}}$ 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*

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

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

- *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 $\partial_{\zeta}(\varphi \circ \ell_0)(x) = \mathbb{R}^d$,*
- *if $l = \ell_0(x) \geq 1$ and $\varphi(l) = +\infty$ and there exists $j \in \{0, 1, \dots, d\}$ such that $\varphi(j) \neq +\infty$, then $\partial_{\zeta}(\varphi \circ \ell_0)(x) = \emptyset$,*

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

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

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

Proof. We have

$$\begin{aligned} y \in \partial_{\dot{\zeta}}(\varphi \circ \ell_0)(x) &\iff (\varphi \circ \ell_0)^{\dot{\zeta}}(y) = \dot{\zeta}(x, y) \dagger (-\varphi \circ \ell_0)(x) \\ &\quad \text{(by definition (37b) of the Capra-subdifferential)} \\ &\iff \sup_{j=0,1,\dots,d} [\|y\|_{(j),\star}^{\mathcal{R}} - \varphi(j)] = \dot{\zeta}(x, y) \dagger (-\varphi \circ \ell_0)(x) \\ &\quad \text{(as } (\varphi \circ \ell_0)^{\dot{\zeta}}(y) = \sup_{j=0,1,\dots,d} [\|y\|_{(j),\star}^{\mathcal{R}} - \varphi(j)] \text{ by (38))} \\ &\iff \left(x = 0 \text{ and } \sup_{j=0,1,\dots,d} [\|y\|_{(j),\star}^{\mathcal{R}} - \varphi(j)] = -\varphi(0) \right) \\ \text{or } &\left(x \neq 0 \text{ and } \sup_{j=0,1,\dots,d} [\|y\|_{(j),\star}^{\mathcal{R}} - \varphi(j)] = \frac{\langle x, y \rangle}{\|x\|} - \varphi(\ell_0(x)) \right) \\ &\quad \text{(by definition (33) of } \dot{\zeta}(x, y) \text{)} \end{aligned}$$

Therefore, on the one hand, we obtain that

$$\begin{aligned} y \in \partial_{\dot{\zeta}}(\varphi \circ \ell_0)(0) &\iff \|y\|_{(j),\star}^{\mathcal{R}} - \varphi(j) \leq -\varphi(0), \quad \forall j = 1, \dots, d \quad \text{(as } \|y\|_{(0),\star}^{\mathcal{R}} = 0 \text{ by convention)} \\ &\iff \|y\|_{(j),\star}^{\mathcal{R}} \leq \varphi(j) \dagger (-\varphi(0)), \quad \forall j = 1, \dots, d \quad \text{(using (51f))} \\ &\iff y \in \bigcap_{j=1,\dots,d} [\varphi(j) \dagger (-\varphi(0))] \mathbb{B}_{(j),\star}^{\mathcal{R}}, \end{aligned}$$

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

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

$$y \in \partial_{\dot{\zeta}}(\varphi \circ \ell_0)(x) \iff \sup_{j=0,1,\dots,d} [\|y\|_{(j),\star}^{\mathcal{R}} - \varphi(j)] = \frac{\langle x, y \rangle}{\|x\|} - \varphi(\ell_0(x)). \quad (45a)$$

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

$$\begin{aligned} y \in \partial_{\dot{\zeta}}(\varphi \circ \ell_0)(x) &\iff \sup_{j=0,1,\dots,d} [\|y\|_{(j),\star}^{\mathcal{R}} - \varphi(j)] = \frac{\langle x, y \rangle}{\|x\|} - \varphi(l) \quad \text{(by (45a) with } \ell_0(x) = l \text{)} \\ &\iff \|y\|_{(l),\star}^{\mathcal{R}} - \varphi(l) \leq \sup_{j=0,1,\dots,d} [\|y\|_{(j),\star}^{\mathcal{R}} - \varphi(j)] = \frac{\langle x, y \rangle}{\|x\|} - \varphi(l) \\ &\iff \|y_L\|_{L,\star} - \varphi(l) \leq \|y\|_{(l),\star}^{\mathcal{R}} - \varphi(l) \leq \sup_{j=0,1,\dots,d} [\|y\|_{(j),\star}^{\mathcal{R}} - \varphi(j)] = \frac{\langle x, y \rangle}{\|x\|} - \varphi(l) \end{aligned}$$

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

$$\begin{aligned} \Leftrightarrow \|y_L\|_{L,\star} - \varphi(l) &\leq \|y\|_{(l),\star}^{\mathcal{R}} - \varphi(l) \leq \sup_{j=0,1,\dots,d} [\|y\|_{(j),\star}^{\mathcal{R}} - \varphi(j)] = \frac{\langle x, y \rangle}{\|x\|} - \varphi(l) \leq \|y_L\|_{L,\star} - \varphi(l) \\ &\quad (\text{as we have } \frac{\langle x, y \rangle}{\|x\|} = \frac{\langle x_L, y_L \rangle}{\|x_L\|} \leq \|y_L\|_{L,\star} \text{ since } x = x_L \text{ and by (14a)}) \end{aligned}$$

$$\begin{aligned} \Leftrightarrow \|y_L\|_{L,\star} - \varphi(l) &= \|y\|_{(l),\star}^{\mathcal{R}} - \varphi(l) = \sup_{j=0,1,\dots,d} [\|y\|_{(j),\star}^{\mathcal{R}} - \varphi(j)] = \frac{\langle x, y \rangle}{\|x\|} - \varphi(l) \\ &\quad (\text{as all terms in the inequalities are necessarily equal}) \end{aligned}$$

$$\Leftrightarrow \begin{cases} \text{either } \varphi(l) = -\infty \\ \text{or } (\varphi(l) = +\infty \text{ and } \varphi(j) = +\infty, \forall j = 0, 1, \dots, d) \\ \text{or } \left(-\infty < \varphi(l) < +\infty \text{ and} \right. \\ \quad \left. \|y_L\|_{L,\star} = \|y\|_{(l),\star}^{\mathcal{R}} = \frac{\langle x, y \rangle}{\|x\|} \text{ and } \|y\|_{(l),\star}^{\mathcal{R}} - \varphi(l) = \sup_{j=0,1,\dots,d} [\|y\|_{(j),\star}^{\mathcal{R}} - \varphi(j)] \right) . \end{cases}$$

Let us make a brief insert and notice that

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

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

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

$$\begin{aligned} y \in \partial_{\mathcal{C}}(\varphi \circ \ell_0)(x) &\Leftrightarrow \|y_L\|_{L,\star} = \|y\|_{(l),\star}^{\mathcal{R}} = \frac{\langle x, y \rangle}{\|x\|} \text{ and } l \in \operatorname{argmax}_{j=0,1,\dots,d} [\|y\|_{(j),\star}^{\mathcal{R}} - \varphi(j)] \\ &\Leftrightarrow \|y_L\|_{L,\star} = \|y\|_{(l),\star}^{\mathcal{R}} \text{ and } \langle x, y \rangle = \|x\| \times \|y\|_{(l),\star}^{\mathcal{R}} \text{ and } l \in \operatorname{argmax}_{j=0,1,\dots,d} [\|y\|_{(j),\star}^{\mathcal{R}} - \varphi(j)] \\ &\Leftrightarrow \langle x, y \rangle = \|x\| \times \|y\|_{(l),\star}^{\mathcal{R}} \text{ and } l \in \operatorname{argmax}_{j=0,1,\dots,d} [\|y\|_{(j),\star}^{\mathcal{R}} - \varphi(j)] \\ &\quad (\text{as just established in the insert}) \end{aligned}$$

$$\begin{aligned} \Leftrightarrow \langle x, y \rangle &= \|x\|_{(l)}^{\mathcal{R}} \times \|y\|_{(l),\star}^{\mathcal{R}} \text{ and } l \in \operatorname{argmax}_{j=0,1,\dots,d} [\|y\|_{(j),\star}^{\mathcal{R}} - \varphi(j)] \\ &\quad (\text{as } \ell_0(x) = l \Rightarrow \|x\| = \|x\|_{(l)}^{\mathcal{R}} \text{ by (30a)}) \end{aligned}$$

$$\Leftrightarrow y \in N_{\mathbb{B}_{(l)}^{\mathcal{R}}}\left(\frac{x}{\|x\|_{(l)}^{\mathcal{R}}}\right) \text{ and } l \in \operatorname{argmax}_{j=0,1,\dots,d} [\|y\|_{(j),\star}^{\mathcal{R}} - \varphi(j)] \quad (\text{by the equivalence in (16)})$$

This ends the proof. \square

With φ the identity function on $\{0, 1, \dots, d\}$, we find the Capra-subdifferential of the l_0 pseudonorm. With the functions $\varphi = \delta_{\{0,1,\dots,k\}}$ (for any $k \in \{0, 1, \dots, d\}$), we find the Capra-subdifferentials of the characteristic functions $\delta_{l_0^{\leq k}}$ 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 $\varphi : \{0, 1, \dots, d\} \rightarrow [0, +\infty[$, such that $\varphi(j) > \varphi(0) = 0$ for all $j = 1, \dots, d$, there exists a norm $\|\cdot\|_{(\varphi)}^{\mathcal{R}}$ characterized

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

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

or, equivalently,

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

- or by the inf-convolution

$$\|\cdot\|_{(\varphi)}^{\mathcal{R}} = \bigsqcap_{j=1,\dots,d} \left(\varphi(j)\|\cdot\|_{(j)}^{\mathcal{R}} \right), \quad (46c)$$

that is,

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

Proof.

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

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

• We have

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

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

$$= \square_{j=1,\dots,d} \delta_{\varphi(j)\mathbb{B}_{(j),\star}^{\mathcal{R}}}^{\star}$$

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

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

This ends the proof. \square

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 $\varphi : \{0, 1, \dots, d\} \rightarrow [0, +\infty[$, such that $\varphi(j) > \varphi(0) = 0$ for all $j = 1, \dots, d$, we have the inequalities

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

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

Proof. We consider the coupling $\dot{\varphi}$ in (33).

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

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

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

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

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

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

For any function $\varphi : \{0, 1, \dots, d\} \rightarrow [0, +\infty[$, such that $\varphi(j) > \varphi(0) = 0$ for all $j = 1, \dots, d$, 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 $(\varphi(j))^q \geq j$, for all $j = 1, \dots, d$. 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 $j = 1, \dots, d$, the inequality (47) yields that $\frac{\|\|x\|\|_1}{\|\|x\|\|_p} \leq (\ell_0(x))^{1/q}$, 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
$\delta_{\ell_0^{\leq k}}^{(-\star)} = +\infty$	$\delta_{\ell_0^{\leq k}}^{-\dot{\mathcal{C}}} = \ \cdot\ _{(k),\star}^{\mathcal{R}}$
$\delta_{\ell_0^{\leq k}}^{\star\star'} = -\infty$	$\delta_{\ell_0^{\leq k}}^{\dot{\mathcal{C}}\dot{\mathcal{C}}'} = \delta_{\{x \in \mathbb{R}^d \mid \ x\ _{(k)}^{\mathcal{R}} = \ x\ \}}$
$\partial \delta_{\ell_0^{\leq k}}(x) = \emptyset$ $\forall x \in \mathbb{R}^d$	$\partial_{\dot{\mathcal{C}}} \delta_{\ell_0^{\leq k}}(x) = \begin{cases} \emptyset & \text{if } \ell_0(x) = k+1, \dots, d, \\ N_{\mathbb{B}_{(k)}^{\mathcal{R}}}(\frac{x}{\ x\ _{(k)}^{\mathcal{R}}}) & \text{if } \ell_0(x) = 1, \dots, k, \\ \{0\} & \text{if } \ell_0(x) = 0 \end{cases}$ $\forall x \in \mathbb{R}^d$
$\ell_0^{\star} = \delta_{\{0\}}$	$\ell_0^{\dot{\mathcal{C}}} = \sup_{j=0,1,\dots,d} [\ \cdot\ _{(j),\star}^{\mathcal{R}} - j]$
$\ell_0^{\star\star'} = 0$	$\ell_0^{\dot{\mathcal{C}}\dot{\mathcal{C}}'}(x) = \frac{1}{\ x\ } \min_{\substack{z^{(1)} \in \mathbb{R}^d, \dots, z^{(d)} \in \mathbb{R}^d \\ \sum_{j=1}^d \ z^{(j)}\ _{(j)}^{\mathcal{R}} \leq \ x\ \\ \sum_{j=1}^d z^{(j)} = x}} \sum_{j=1}^d j \ z^{(j)}\ _{(j)}^{\mathcal{R}}, \quad \forall x \in \mathbb{R}^d \setminus \{0\}$ $\ell_0^{\dot{\mathcal{C}}\dot{\mathcal{C}}'}(0) = 0$
$\partial \ell_0(0) = \{0\}$	$\partial_{\dot{\mathcal{C}}} \ell_0(0) = \bigcap_{j=1,\dots,d} j \mathbb{B}_{(j),\star}^{\mathcal{R}} = \mathbb{B}_{(\text{Id}),\star}^{\mathcal{R}}$
$\partial \ell_0(x) = \emptyset$ $\forall x \in \mathbb{R}^d \setminus \{0\}$	$y \in \partial_{\dot{\mathcal{C}}} \ell_0(x) \iff \begin{cases} y \in N_{\mathbb{B}_{(l)}^{\mathcal{R}}}(\frac{x}{\ x\ _{(l)}^{\mathcal{R}}}) \\ \text{and } l \in \operatorname{argmax}_{j=0,1,\dots,d} [\ y\ _{(j),\star}^{\mathcal{R}} - j] \end{cases}$ $\forall x \in \mathbb{R}^d \setminus \{0\}, \text{ where } l = \ell_0(x) \geq 1$

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 $k = 0, 1, \dots, d$

Moreau lower addition

The Moreau *lower addition* extends the usual addition with

$$(+\infty) \dagger (-\infty) = (-\infty) \dagger (+\infty) = -\infty . \quad (49a)$$

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

$$u \leq u' , \quad v \leq v' \Rightarrow u \dagger v \leq u' \dagger v' , \quad (49b)$$

$$(-u) \dagger (-v) \leq -(u \dagger v) , \quad (49c)$$

$$(-u) \dagger u \leq 0 , \quad (49d)$$

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

$$\sup_{a \in \mathbb{A}} f(a) \dagger \sup_{b \in \mathbb{B}} g(b) = \sup_{a \in \mathbb{A}, b \in \mathbb{B}} (f(a) \dagger g(b)) , \quad (49e)$$

$$\inf_{a \in \mathbb{A}} f(a) \dagger \inf_{b \in \mathbb{B}} g(b) \leq \inf_{a \in \mathbb{A}, b \in \mathbb{B}} (f(a) \dagger g(b)) , \quad (49f)$$

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

Moreau upper addition

The Moreau *upper addition* extends the usual addition with

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

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

$$u \leq u' , \quad v \leq v' \Rightarrow u \dot{\dagger} v \leq u' \dot{\dagger} v' , \quad (50b)$$

$$(-u) \dot{\dagger} (-v) \geq -(u \dot{\dagger} v) , \quad (50c)$$

$$(-u) \dot{\dagger} u \geq 0 , \quad (50d)$$

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

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

$$\sup_{a \in \mathbb{A}} f(a) \dot{\dagger} \sup_{b \in \mathbb{B}} g(b) \geq \sup_{a \in \mathbb{A}, b \in \mathbb{B}} (f(a) \dot{\dagger} g(b)) , \quad (50f)$$

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

Joint properties of the Moreau lower and upper addition

We obviously have that

$$u \dot{+} v \leq u \dot{+} v . \quad (51a)$$

The Moreau lower and upper additions are related by

$$-(u \dot{+} v) = (-u) \dot{+} (-v) , \quad -(u \dot{+} v) = (-u) \dot{+} (-v) . \quad (51b)$$

They satisfy the inequality

$$(u \dot{+} v) \dot{+} w \leq u \dot{+} (v \dot{+} w) . \quad (51c)$$

with

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

Finally, we have that

$$u \dot{+} (-v) \leq 0 \iff u \leq v \iff 0 \leq v \dot{+} (-u) , \quad (51e)$$

$$u \dot{+} (-v) \leq w \iff u \leq v \dot{+} w \iff u \dot{+} (-w) \leq v , \quad (51f)$$

$$w \leq v \dot{+} (-u) \iff u \dot{+} w \leq v \iff u \leq v \dot{+} (-w) . \quad (51g)$$

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

$$c : \mathbb{X} \times \mathbb{Y} \rightarrow \overline{\mathbb{R}} . \quad (52)$$

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 $f : \mathbb{X} \rightarrow \overline{\mathbb{R}}$, with respect to the coupling c , is the function $f^c : \mathbb{Y} \rightarrow \overline{\mathbb{R}}$ defined by*

$$f^c(y) = \sup_{x \in \mathbb{X}} \left(c(x, y) \dot{+} (-f(x)) \right) , \quad \forall y \in \mathbb{Y} . \quad (53a)$$

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

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

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

$$g^{c'}(x) = \sup_{y \in \mathbb{Y}} \left(c(x, y) \dagger (-g(y)) \right), \quad \forall x \in \mathbb{X}. \quad (53c)$$

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

$$f^{cc'}(x) = (f^c)^{c'}(x) = \sup_{y \in \mathbb{Y}} \left(c(x, y) \dagger (-f^c(y)) \right), \quad \forall x \in \mathbb{X}. \quad (53d)$$

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

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

A.2.2 The Fenchel conjugacy

When the sets \mathbb{X} and \mathbb{Y} are vector spaces equipped with a bilinear form $\langle \cdot, \cdot \rangle$, the corresponding conjugacy is the classical *Fenchel conjugacy*. For any functions $f : \mathbb{X} \rightarrow \overline{\mathbb{R}}$ and $g : \mathbb{Y} \rightarrow \overline{\mathbb{R}}$, we denote⁵

$$f^*(y) = \sup_{x \in \mathbb{X}} \left(\langle x, y \rangle \dagger (-f(x)) \right), \quad \forall y \in \mathbb{Y}, \quad (54a)$$

$$g^{*'}(x) = \sup_{y \in \mathbb{Y}} \left(\langle x, y \rangle \dagger (-g(y)) \right), \quad \forall x \in \mathbb{X}, \quad (54b)$$

$$f^{**'}(x) = \sup_{y \in \mathbb{Y}} \left(\langle x, y \rangle \dagger (-f^*(y)) \right), \quad \forall x \in \mathbb{X}. \quad (54c)$$

For any function $h : \mathbb{W} \rightarrow \overline{\mathbb{R}}$, its *epigraph* is $\text{epih} = \{(w, t) \in \mathbb{W} \times \mathbb{R} \mid h(w) \leq t\}$, its *effective domain* is $\text{dom}h = \{w \in \mathbb{W} \mid h(w) < +\infty\}$. A function $h : \mathbb{W} \rightarrow \overline{\mathbb{R}}$ is said to be *proper* if it never takes the value $-\infty$ and that $\text{dom}h \neq \emptyset$. When \mathbb{W} is equipped with a topology, the function $h : \mathbb{W} \rightarrow \overline{\mathbb{R}}$ 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 analysis⁶, 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 $+\infty$.

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

⁶That is, \mathbb{X} and \mathbb{Y} are equipped with a bilinear form $\langle \cdot, \cdot \rangle$, 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 $h : \mathbb{W} \rightarrow \overline{\mathbb{R}}$ with a mapping $\theta : \mathbb{W} \rightarrow \mathbb{X}$ to get a function $\inf [h \mid \theta] : \mathbb{X} \rightarrow \overline{\mathbb{R}}$ defined by [11, p. 27]

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

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

$$h = f \circ \theta \text{ where } f : \mathbb{X} \rightarrow \overline{\mathbb{R}} \Rightarrow \inf [h \mid \theta] = f \dot{+} \delta_{\theta(\mathbb{W})}, \quad (55b)$$

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

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

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 $\langle \cdot, \cdot \rangle$. Let \mathbb{W} be a set and

$$\theta : \mathbb{W} \rightarrow \mathbb{X} \quad (57a)$$

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

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

Notice that, in a one-sided linear coupling, the second set possesses 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, $\sigma_X : \mathbb{Y} \rightarrow \overline{\mathbb{R}}$ denotes the *support function of a subset* $X \subset \mathbb{X}$:

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

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

$$g^{c'_\theta} = g^{\star'} \circ \theta. \quad (59a)$$

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

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

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

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

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

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

c_θ -convex functions. We recall that so-called c_θ -convex functions are all functions $h : \mathbb{R}^d \rightarrow \overline{\mathbb{R}}$ of the form g^{c_θ} , for any function $g : \mathbb{R}^d \rightarrow \overline{\mathbb{R}}$, or, equivalently, all functions of the form $h^{c_\theta c_{\theta'}}$, for any function $h : \mathbb{R}^d \rightarrow \overline{\mathbb{R}}$, 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 $h : \mathbb{R}^d \rightarrow \overline{\mathbb{R}}$, we have the equivalences*

$$h \text{ is } c_\theta\text{-convex} \quad (60a)$$

$$\iff h = h^{c_\theta c_{\theta'}} \quad (60b)$$

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

$$\iff \text{there exists a closed convex function } f : \mathbb{X} \rightarrow \overline{\mathbb{R}} \text{ such that } h = f \circ \theta. \quad (60d)$$

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

If there exists a closed convex function $f : \mathbb{X} \rightarrow \overline{\mathbb{R}}$ such that $h = f \circ \theta$, then $\inf [h \mid \theta] = f \dagger \delta_{\theta(\mathbb{W})}$ by (55b), and therefore $h^{c_\theta c_{\theta'}} = (\inf [h \mid \theta])^{\star'} \circ \theta = (f \dagger \delta_{\theta(\mathbb{W})})^{\star'} \circ \theta$ by (59c). Now, as $f \dagger \delta_{\theta(\mathbb{W})} \geq f$, we get that $(f \dagger \delta_{\theta(\mathbb{W})})^{\star'} \geq f^{\star\star'} = f$, where the last equality holds because the function $f : \mathbb{X} \rightarrow \overline{\mathbb{R}}$ 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 $h : \mathbb{R}^d \rightarrow \overline{\mathbb{R}}$ at $w \in \mathbb{R}^d$ by

$$\partial_{c_\theta} h(w) = \{y \in \mathbb{R}^d \mid c_\theta(w', y) \dagger (-h(w')) \leq c_\theta(w, y) \dagger (-h(w)), \quad \forall w' \in \mathbb{R}^d\} \quad (61a)$$

$$= \{y \in \mathbb{R}^d \mid h^{c_\theta}(y) = c_\theta(w, y) \dagger (-h(w))\} \quad (61b)$$

$$= \{y \in \mathbb{R}^d \mid (\inf [h \mid \theta])^{\star}(y) = \langle \theta(w), y \rangle \dagger (-h(w))\}. \quad (61c)$$

References

- [1] M. Akian, S. Gaubert, and V. Kolokoltsov. Invertibility of functional Galois connections. *Comptes Rendus Mathématique*, 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-Uruty 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.