



HAL
open science

Generalisation of Farkas' lemma beyond closedness: a constructive approach via Fenchel-Rockafellar duality

Camille Pouchol, Emmanuel Trélat, Christophe Zhang

► To cite this version:

Camille Pouchol, Emmanuel Trélat, Christophe Zhang. Generalisation of Farkas' lemma beyond closedness: a constructive approach via Fenchel-Rockafellar duality. 2026. ⟨hal-05548185⟩

HAL Id: hal-05548185

<https://hal.science/hal-05548185v1>

Preprint submitted on 11 Mar 2026

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.



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License

Generalisation of Farkas' lemma beyond closedness: a constructive approach via Fenchel-Rockafellar duality

Camille Pouchol^a, Emmanuel Trélat^{b,c}, and Christophe Zhang^{b,c}

^aUniversité Paris Cité, CNRS, MAP5, F-75006 Paris, France.

^bSorbonne Université, Université Paris Cité, CNRS, Laboratoire Jacques-Louis Lions, 75005 Paris, France.

^cCAGE, INRIA, Paris, France.

Abstract

Farkas' lemma is an ubiquitous tool in optimisation, as it provides necessary and sufficient conditions to have $b \in A(P)$, where P is a closed convex cone, A is a (continuous) linear mapping and b is a fixed vector. The standard underlying hypothesis is the closedness of $A(P)$, which is not always satisfied and can be difficult to check. We devise a new method to generalise Farkas' lemma, based on a primal-dual pair of optimisation problems and Fenchel-Rockafellar duality theory. We work under the sole hypothesis that P be generated by a closed bounded convex set. This hypothesis is weaker than in previous generalisations of Farkas' lemma, which almost all require that $A(P)$ be closed, or, in few cases, that only P be closed. In our case, P (and a fortiori $A(P)$) is not necessarily closed; we uncover necessary and sufficient conditions both for $b \in A(P)$ and $b \in \overline{A(P)}$. For a given $\varepsilon \geq 0$, we exhibit constructive characterisations of $x \in P$ such that $\|Ax - b\| \leq \varepsilon$ when it exists, by means of optimality conditions. For $\varepsilon = 0$, these strongly rely on whether the dual problem admits a solution, and we discuss conditions under which it does. Finally, we also explain how, upon relaxation, we may apply our method to a nonconvex cone.

Contents

Contents	1
1 Introduction and main results	2
1.1 Farkas' lemma(s): necessary and sufficient conditions for solvability	3
1.2 Main contributions	3
1.3 Variants of Farkas' lemma: necessary and sufficient conditions for solvability	4
1.3.1 Closed convex cone	4
1.3.2 Cone generated by a bounded closed convex set	5
1.4 Constructive characterisations	6
1.4.1 Closed convex cone	7
1.4.2 Cone generated by a bounded closed convex set	8
1.4.3 General cone	9
1.5 Remarks and perspectives	10
2 Mathematical framework	11
2.1 Notations	11
2.1.1 Functions	11
2.1.2 Convex sets	11

2.1.3	Cones	12
2.2	Fenchel-Rockafellar duality and constructive characterisation of solutions	12
2.2.1	Primal and dual optimisation problems	12
2.2.2	Optimality conditions	13
3	Proofs	14
3.1	Approximate solvability	14
3.1.1	Necessary and sufficient condition for solvability	14
3.1.2	Constructive characterisation and uniqueness	15
3.2	Exact solvability	16
3.3	Closed convex case	17
3.4	Optimality conditions and constructiveness	18
3.4.1	Cone generated by a bounded closed convex set	18
3.4.2	General cone	19
4	Dual attainment and shared normal vectors	19
4.1	Conditions for the dual problem to attain its infimum	19
4.2	Geometric interpretation: shared normal vectors	20
4.3	On the existence of shared normal vectors	22
4.4	On the regularity condition	24
A	Elementary results in convex analysis	25
	References	27

1 Introduction and main results

Let X and Y be two (real) Hilbert spaces, $A : X \rightarrow Y$ a continuous bounded operator (an element of $\mathcal{L}(X, Y)$), $b \in Y$ and $P \subset X$ a cone containing 0. We are interested in the following problem, which is a variant of Farkas' celebrated lemma ([15]): do we have $b \in A(P)$, that is, is the vector b in the (somewhat) implicit cone $A(P)$?

This problem can also be formulated as a constrained linear equation:

$$\text{Solving the linear equation } Ax = b \text{ under the constraint } x \in P. \tag{1}$$

We let $A(P) = \{Ax, x \in P\}$.

We consider both

- (i) the **approximate solvability** problem,

$$\text{For any given } \varepsilon > 0, \text{ find } x_\varepsilon \in P \text{ such that } \|Ax_\varepsilon - b\|_Y \leq \varepsilon, \tag{2}$$

that is, determining whether $b \in \overline{A(P)}$.

- (ii) the **exact solvability** problem,

$$\text{Find } x \in P \text{ such that } Ax = b, \tag{3}$$

that is, determining whether $b \in A(P)$.

1.1 Farkas' lemma(s): necessary and sufficient conditions for solvability

The original version of Farkas' lemma [15] is concerned with the exact solvability problem within the cone \mathbb{R}_+^m in \mathbb{R}^m . It provides a necessary and sufficient condition (often formulated as an alternative):

Lemma 1.1. *Let $f, v_1, \dots, v_m \in \mathbb{R}^n$. Then, there exist $(f_1, \dots, f_m) \in \mathbb{R}_+^m$ such that*

$$f = \sum_{i=1}^m f_i v_i$$

if and only if

$$\{d \in \mathbb{R}^n, \langle v_i, d \rangle \leq 0, \forall i \in \{1, \dots, m\}\} \subset \{d \in \mathbb{R}^n, \langle d, f \rangle \leq 0\}.$$

For a survey on the rich history of Farkas' lemma and its generalisations, we refer to [14]. The most standard and used ones writes as the following alternative, with $P^\circ = \{x \in \mathbb{R}^n, \forall y \in P, \langle x, y \rangle \leq 0\}$ standing for the polar cone of P .

Lemma 1.2. *[Farkas' lemma] Let $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ and $P \subset \mathbb{R}^n$ be a closed convex cone such that $A(P)$ is closed. Then, given $b \in \mathbb{R}^m$, the following alternative holds:*

1. *the exact solvability problem admits a solution,*
2. *there exists $y \in \mathbb{R}^m$ such that $A^*y \in P^\circ$ and $\langle b, y \rangle > 0$.*

Farkas' lemma has been generalised in several directions: in particular, it can be stated in very abstract general infinite-dimensional settings [13]. However, the hypothesis that $A(P)$ is closed remains a cornerstone to all these results. In fact, in the above cited work, the authors prove that the Farkas alternative holds if and only if $A(P)$ is closed. Under this key assumption, let us note that distinguishing between exact and approximate solvability as we did above is irrelevant.

Up to our knowledge, only the two works [18, 19] formulate a generalisation without assuming the closedness of $A(P)$. Let us give one such result, in the particular case of reflexive Banach spaces; here we translate it into the Hilbertian setting we are interested in.

Theorem 1.1. *[Hernandez-Lerma, Lasserre, 1997] Assume that $P \subset X$ is a closed convex cone. Then, the following are equivalent for $b \in Y$:*

- (a) *The equation $Ax = b$ has a solution $x \in P$;*
- (b) *There exists $M > 0$ such that, for all $y \in Y$, $p \in P^\circ$,*

$$\langle b, y \rangle_Y + M \|A^*y + p\|_X \geq 0.$$

Given the equivalence proved in [13], foregoing the restrictive hypothesis that $A(P)$ be closed requires a stronger characterisation. Accordingly, the above theorem gives a quantified necessary sufficient condition. The underlying idea is to apply the "usual" Farkas lemma to an augmented system. The corresponding proof is rather intricate, it is not constructive, and the resulting condition, though quantified, is difficult to put in use since the inequality hinges on two variables.

1.2 Main contributions

This article presents a generalisation of Farkas' lemma, close in spirit to Theorem 1.1, which we state under hypotheses of increasing generality on the cones involved. As in [18, 19], we forego the hypothesis that $A(P)$ is closed, hence our formulation of the two distinct problems (2) and (3).

Our main contributions are then as follows:

- we recover the results of [18, 19] by means of an arguably simpler proof,

- not only do we forego of the assumption that $A(P)$ is closed (which is the usual hypothesis in Farkas' lemma, see Lemma 1.2 and [13]), we also weaken the underlying hypothesis that P itself is closed: our approach only requires that P can be generated by a bounded closed convex set containing 0,
- contrarily to other proofs, our proof is constructive under general hypotheses, in a sense we shall make more precise later on,
- the overall technique lends itself to relaxation, i.e., one can consider the approximate or exact solvability problem in nonconvex cones, and apply our strategy on well-chosen convex cones derived from them.

Algorithmic viewpoint. A key feature of our approach is that (approximate or exact) Farkas solutions can be *constructed* by solving an explicit convex optimisation problem. For a tolerance $\varepsilon \geq 0$, we minimise an **unconstrained** convex functional over Y (the dual problem \mathcal{D}_ε). Then, under suitable conditions, we may recover a primal variable x_ε through explicit first-order optimality conditions (projection onto a cone in the closed-cone case, or a support-function subgradient formula in the general case). This yields a practical primal–dual procedure to build either a solution $x \in P$ with $Ax = b$, or approximate solutions $x_\varepsilon \in P$ with $\|Ax_\varepsilon - b\|_Y \leq \varepsilon$, when they exist.

Related literature. Farkas-type statements and theorems of the alternative can be derived from convex duality in various settings (see, e.g., Farkas-type results for constrained optimisation problems, obtained via conjugate functions and conic duality in [9, 7, 8]).

In this article however, we address the classical problem of solving a linear equation on a (convex) cone, by formulating a well-chosen auxiliary optimisation problem. Problems of this form are already present in the works [24, 25]: given $c \in X$ fixed, $x \mapsto \langle c, x \rangle_X$ is maximised under the constraints $Ax = b$ and $x \in P$. A duality framework is then established to prove Farkas-type results on the solvability of the constrained equation $Ax = b$, $x \in P$.

In contrast to the above, we provide a unified *constructive* approach based on Fenchel–Rockafellar duality in a Hilbertian framework. We propose a different auxiliary optimisation problem (given by $(\mathcal{P}_\varepsilon)$), for which our approach yields explicit primal minimisers (which are automatically solutions of the solvability problems (2)–(3)) from dual minimisers. Our results cover cones of the form $P_r = \text{cone}(K_r)$ where the bounded convex generator K_r is closed but P_r (and thus $A(P_r)$) may fail to be closed.

Let us now introduce our results in full detail. For ease of readability, use and comparison with other results, we proceed in increasing order of generality.

1.3 Variants of Farkas' lemma: necessary and sufficient conditions for solvability

Let us first state necessary and sufficient conditions for the existence of solutions. Throughout the paper, X and Y are real Hilbert spaces, with inner products $\langle \cdot, \cdot \rangle_X$ and $\langle \cdot, \cdot \rangle_Y$ (with corresponding norms $\|\cdot\|_X$ and $\|\cdot\|_Y$), respectively. We use the Riesz representation theorem so that the adjoint A^* of a bounded linear operator $A \in \mathcal{L}(X, Y)$ is the Hilbert adjoint $A^* \in \mathcal{L}(Y, X)$ defined by $\langle Ax, y \rangle_Y = \langle x, A^*y \rangle_X$ for all $(x, y) \in X \times Y$.

1.3.1 Closed convex cone

Here, we assume that P is **closed and convex**, and for $x \in X$, we let x_+ be its projection onto P . Note that in this first setting, we already forego the usual hypothesis that $A(P)$ is closed, but retain the classical assumption of closedness for P , as in [18, 19].

Theorem 1.2. *Let $A \in \mathcal{L}(X, Y)$, $b \in Y$ and $P \subset X$ a closed convex cone. The following dual characterisations hold for any $b \in Y$:*

$$(i) \quad b \in \overline{A(P)} \iff \forall y \in Y, (A^*y)_+ = 0 \implies \langle b, y \rangle_Y \leq 0.$$

$$(ii) \quad b \in A(P) \iff \exists C > 0, \forall y \in Y, \langle b, y \rangle_Y \leq C \|(A^*y)_+\|_X.$$

Remark 1.1. If $A(P)$ is closed, (i) and (ii) are equivalent and the first solvability condition associated to (i) is clearly equivalent to the negation of item 2. of the standard Farkas' Lemma 1.2, because $z_+ = 0$ if and only if $z \in P^\circ$.

The second solvability condition, given in (ii), is a more compact and workable form of that uncovered in [18, 19]. Indeed, one can check that (ii) and Theorem 1.1[(b)] are actually equivalent as follows: in a Hilbert space and for a closed convex cone P , Moreau's decomposition $z = z_+ + z_-$ with $z_+ \in P$, $z_- \in P^\circ$ and $\langle z_+, z_- \rangle = 0$ yields

$$\inf_{p \in P^\circ} \|z - p\| = \|z_+\|.$$

Applying this identity to $z = A^*y$, the distance-to- P° term that appears in Theorem 1.1[(b)] can be rewritten as $\|(A^*y)_+\|$, leading to (ii).

Remark 1.2. Taking $P = X$, we recover the well-known fact that

$$b \in \overline{\text{ran}(A)} \iff (A^*y = 0 \implies \langle b, y \rangle_Y \leq 0) \iff (A^*y = 0 \implies \langle b, y \rangle_Y = 0) \iff b \perp \ker(A^*),$$

i.e., $\ker(A^*)^\perp = \overline{\text{ran}(A)}$.

On the other hand, the necessary and sufficient condition in (ii) corresponds to the well-known condition of range inclusion in functional analysis (see [20, Lemma 2.4]): denoting by π_b the projection onto $\mathbb{R}b$, we have

$$\begin{aligned} \text{ran}(\pi_b) \subset \text{ran}(A) &\iff b \in \text{ran}(A) \iff \exists C > 0, \forall y \in Y, \langle b, y \rangle_Y \leq C \|A^*y\|_X \\ &\iff \exists C > 0, \forall y \in Y, |\langle b, y \rangle_Y| \leq C \|A^*y\|_X \\ &\iff \exists C > 0, \forall y \in Y, \|\pi_b y\| \leq C \|A^*y\|_X. \end{aligned}$$

Finally, we recover a well-known sufficient condition for the surjectivity of A : if there exists $C > 0$ such that for $y \in Y$, $\|A^*y\| \geq c\|y\|$, then for $b \in Y$,

$$|\langle b, y \rangle| \leq \|b\| \|y\| \leq \frac{\|b\|}{c} \|A^*y\|,$$

which implies $b \in \text{ran}(A)$ by Theorem 1.2 (ii).

1.3.2 Cone generated by a bounded closed convex set

Note that a mandatory hypothesis for all versions of Farkas' lemma we are aware of is that P is closed: this follows immediately from the usual assumption that $A(P)$ is closed [13]. In contrast, from this point on, we forego this hypothesis: indeed, the fundamental assumptions underlying our approach concern the choice of a generating set for the cone, rather than the cone itself.

We now assume that the cone P (denoted P_r in this case) is generated by a bounded closed convex set K_r containing 0, i.e., we have $\text{cone}(K_r) = P_r$. P_r is then still convex, but it need not be closed. On the other hand, if P_r is closed, setting $K_r = P_r \cap \overline{B(0, 1)}$ yields a bounded closed convex set containing 0 that does generate P_r . Hence, this hypothesis is indeed weaker than the closedness of P_r , and thus also weaker than the closedness of $A(P_r)$. As hinted at by the following remark, this significantly extends the class of possible convex cones for which one can formulate a Farkas lemma.

Remark 1.3. A sufficient condition for the closedness of $P_r = \text{cone}(K_r)$ is for 0 to be in the interior of K_r relative to P_r , see Lemma A.3.

As a counterexample (in \mathbb{R}^2) of a bounded closed convex set containing 0 which generates a nonclosed cone, note that

$$\text{cone}(\overline{B((1,0),1)}) = (\mathbb{R}_{++} \times \mathbb{R}) \cup \{(0,0)\},$$

and also that

$$\text{cone}(\overline{B((1,0),1)} \cap \overline{B((0,1),1)}) = (\mathbb{R}_{++})^2 \cup \{(0,0)\}.$$

Conversely, these two examples can be generalised to higher dimensions, and show that our approach can help handle exact and approximate resolution in useful non-closed cones, upon choosing appropriate generators.

We let $\sigma_{K_r} : x \mapsto \sup_{y \in K_r} \langle x, y \rangle_X$ be the support function of K_r .

Theorem 1.3. Let $A \in \mathcal{L}(X, Y)$, $b \in Y$, $K_r \subset X$ a bounded closed convex set containing 0, and $P_r = \text{cone}(K_r)$. The following dual characterisations hold for any $b \in Y$:

$$(i) \quad \begin{aligned} b \in \overline{A(P_r)} &\iff \forall y \in Y, \sigma_{K_r}(A^*y) = 0 \implies \langle b, y \rangle_Y \leq 0, \\ &\iff \forall y \in Y, A^*y \in P_r^\circ \implies \langle b, y \rangle_Y \leq 0. \end{aligned}$$

$$(ii) \quad b \in A(P_r) \iff \exists C > 0, \forall y \in Y, \langle b, y \rangle_Y \leq C \sigma_{K_r}(A^*y).$$

Remark 1.4. The dual characterisation (i) has two possible formulations due to the fact that $\sigma_{K_r}(z) = 0$ if and only if $z \in P_r^\circ$ (see Lemma 3.1). The second formulation illustrates the direct link between Theorem 1.3 and previous generalisations of Farkas' Lemma.

Remark 1.5. The two theorems 1.2 and 1.4 above given for a closed and convex cone P_r correspond to choosing $K_r = P_r \cap B(0,1)$ and applying Theorem 1.3.

1.4 Constructive characterisations

Our proof of the direct implications \implies in theorems 1.2 and 1.3 above is based on Fenchel-Rockafellar duality [21]. We denote by δ_C the indicator of a set $C \subset X$, in the convex-analytic sense, and we let $j_{K_r} = \inf\{\alpha > 0, x \in \alpha K_r\}$ be the gauge function of K_r .

For $\varepsilon \geq 0$, we introduce the auxiliary functional

$$\forall x \in X, \quad \mathcal{F}_\varepsilon(x) := \frac{1}{2}j_{K_r}^2(x) + \delta_{\{x \in X, \|Ax-b\|_Y \leq \varepsilon\}}(x).$$

Because $j_{K_r}(x) < +\infty$ if and only if $x \in P_r$, the expression of \mathcal{F}_ε is such that, for a given $\varepsilon \geq 0$

$$\forall x \in X, \quad \mathcal{F}_\varepsilon(x) < +\infty \iff \begin{cases} \|Ax - b\|_Y \leq \varepsilon \\ x \in P_r \end{cases}.$$

A natural way of finding such an x is to minimise \mathcal{F}_ε , we thus introduce the auxiliary optimisation problem

$$\inf_{x \in X} \mathcal{F}_\varepsilon(x). \tag{\mathcal{P}_\varepsilon}$$

We then have

$$x_\varepsilon^* \text{ is a minimiser for } (\mathcal{P}_\varepsilon) \implies \begin{cases} \|Ax_\varepsilon^* - b\|_Y \leq \varepsilon, \\ x_\varepsilon^* \in P_r. \end{cases} \tag{4}$$

Thus, solving the optimisation problem $(\mathcal{P}_\varepsilon)$ for all $\varepsilon > 0$ solves (2), and solving it for $\varepsilon = 0$ yields a solution to (3).

To solve this optimisation problem, we introduce its dual problem, following Fenchel-Rockafellar duality theory. For $\varepsilon \geq 0$, we show that it is given by

$$\inf_{y \in Y} J_\varepsilon(y), \quad J_\varepsilon(y) := \frac{1}{2} \sigma_{K_r}^2(A^*y) - \langle b, y \rangle_Y + \varepsilon \|y\|_Y. \quad (\mathcal{D}_\varepsilon)$$

Remark 1.6. *An important feature of this dual problem is that its cost function J_ε is unconstrained (finite everywhere), thanks to our choice of the auxiliary optimisation problem, which involves the function $\frac{1}{2} j_{K_r}^2$. This function has domain P , but its convex conjugate $\frac{1}{2} \sigma_{K_r}^2$ always has domain X .*

This is numerically relevant, and in contrast with the dual problems formulated in [24, 25], which are linearly constrained.

We establish strong duality between the primal and dual problems, namely:

$$\inf_{x \in X} \mathcal{F}_\varepsilon(x) = - \inf_{y \in Y} J_\varepsilon(y).$$

The optimality conditions derived for these primal dual pairs of optimisation problems then lead to constructive characterisations of solutions, when they exist.

Remark 1.7. *If $b \in A(P_r)$, the best constant $C = C(b)$ in Theorem 1.3 is given by (see Proposition 3.3):*

$$C(b) := \inf \{ j_{K_r}(x), Ax = b, x \in P_r \}.$$

If $b \in A(P_r)$, so is any positive multiple thereof, and $C(b)$ is clearly 1-positively homogeneous.

*By strong duality, $\frac{1}{2} C^2(b)$ also equals $-\inf \{ \frac{1}{2} \sigma_{K_r}^2(A^*y) - \langle b, y \rangle_Y, y \in Y \}$. As an infimum of continuous functions, $b \mapsto C(b)$ is upper-semicontinuous on $A(P_r)$.*

If $P_r = X$ and $K_r = \overline{B(0, 1)}$, $C(b)$ equals the minimal norm solution for the equation $Ax = b$. It also equals $\|A^\dagger b\|_X$, whenever the Moore-Penrose inverse A^\dagger is well-defined (that is, when $\text{Ran}(A)$ is closed).

1.4.1 Closed convex cone

As noted in Remark 1.5, the case of a closed convex cone P is a particular case of a cone P_r generated by a bounded closed convex set K_r . Taking $K_r := P \cap \overline{B(0, 1)}$, the primal problem (\mathcal{P}_ε) writes

$$\inf_{x \in X} \mathcal{F}_\varepsilon^{\text{cl}}(x), \quad \mathcal{F}_\varepsilon^{\text{cl}}(x) := \frac{1}{2} \|x\|_X^2 + \delta_P(x) + \delta_{\{x \in X, \|Ax - b\|_Y \leq \varepsilon\}}(x), \quad (\mathcal{P}_\varepsilon^{\text{cl}})$$

and its dual problem writes

$$\inf_{y \in Y} J_\varepsilon^{\text{cl}}(y), \quad J_\varepsilon^{\text{cl}}(y) := \frac{1}{2} \|(A^*y)_+\|_X^2 - \langle b, y \rangle_Y + \varepsilon \|y\|_Y, \quad (\mathcal{D}_\varepsilon^{\text{cl}})$$

where we recall that z_+ means the projection of a vector $z \in X$ onto P .

As described above, solving this primal-dual pair of optimisation problems leads to solving (2) and (3), thanks to (4), and yields constructive characterisations of solutions. Since these auxiliary problems involve elementary functions, the constructive characterisations we obtain are remarkably direct:

Theorem 1.4. *Under the hypotheses of Theorem 1.2, for any $b \in Y$,*

(i) *If $b \in \overline{A(P)}$, then for any $\varepsilon > 0$,*

- ($\mathcal{P}_\varepsilon^{\text{cl}}$) *admits a unique solution $x_\varepsilon^* \in P$, such that $\|Ax_\varepsilon^* - b\|_Y \leq \varepsilon$.*
- ($\mathcal{D}_\varepsilon^{\text{cl}}$) *admits a unique solution y_ε^* , and*

$$x_\varepsilon^* := (A^*y_\varepsilon^*)_+.$$

(ii) *If $b \in A(P)$, then*

- ($\mathcal{P}_0^{\text{cl}}$) *admits a unique solution $x^* \in P$, such that $Ax^* = b$.*
- **If furthermore ($\mathcal{D}_0^{\text{cl}}$) admits a solution y^* , then**

$$x^* := (A^*y^*)_+.$$

Least-norm (conic) pseudoinverse viewpoint. When $\varepsilon = 0$ and $b \in A(P)$, problem $(\mathcal{P}_0^{\text{cl}})$ computes the minimum-norm solution of the conic system $Ax = b$, $x \in P$. In this closed-cone setting, it is hence natural to define the cone-constrained least-norm pseudoinverse $A_P^\dagger(b) : A(P) \rightarrow P$ of A by

$$\forall b \in A(P), \quad \{A_P^\dagger(b)\} = \arg \min\{\|x\|_X, Ax = b, x \in P\}.$$

When $P = X$ and $\text{Ran}(A)$ is closed, this coincides with the usual Moore-Penrose pseudoinverse A^\dagger on $\text{Ran}(A)$. Thus, Theorem 1.4 does not merely prove solvability: it canonically selects, among all conic solutions, the least-norm one, and provides the explicit primal-dual representation $A_P^\dagger(b) = (A^*y^*(b))_+$, whenever a dual minimiser $y^*(b)$ exists. This also suggests a practical numerical procedure: solve the unconstrained dual problem $(\mathcal{D}_0^{\text{cl}})$ and recover the primal variable by projecting $A^*y^*(b)$ onto P .

1.4.2 Cone generated by a bounded closed convex set

In the more general case of a cone P_r generated by a bounded closed convex set K_r containing 0, the constructive characterisation is, accordingly, more involved.

Theorem 1.5. *Under the hypotheses of Theorem 1.3, for any $b \in Y$,*

(i) *If $b \in \overline{A(P_r)}$, then for any $\varepsilon > 0$,*

- $(\mathcal{P}_\varepsilon)$ *admits a solution $x_\varepsilon^* \in P_r$ such that $\|Ax_\varepsilon^* - b\|_Y \leq \varepsilon$,*
- $(\mathcal{D}_\varepsilon)$ *admits a unique solution y_ε^* , and*

$$x_\varepsilon^* \in \sigma_{K_r}(A^*y_\varepsilon^*) \partial\sigma_{K_r}(A^*y_\varepsilon^*). \quad (5)$$

(ii) *If $b \in A(P_r)$, then*

- (\mathcal{P}_0) *admits a solution $x^* \in P_r$ such that $Ax^* = b$,*
- *if furthermore (\mathcal{D}_0) admits a solution y^* , then*

$$x^* \in \sigma_{K_r}(A^*y^*) \partial\sigma_{K_r}(A^*y^*). \quad (6)$$

Remark 1.8. *For $\varepsilon > 0$, the dual variable y_ε^* can be interpreted as a normal-cone Lagrange multiplier associated with the feasibility constraint $Ax \in \overline{B(b, \varepsilon)}$. Indeed, from the saddle-point conditions (13) one has $y_\varepsilon^* \in -\partial\delta_{\overline{B(b, \varepsilon)}}(Ax_\varepsilon^*) = -N_{\overline{B(b, \varepsilon)}}(Ax_\varepsilon^*)$, where $N_{\overline{B(b, \varepsilon)}}(Ax_\varepsilon^*)$ is the normal cone to $\overline{B(b, \varepsilon)}$ at Ax_ε^* . In particular, if the constraint is active (and we will show it necessarily is when $\|b\| > \varepsilon$), i.e., $\|Ax_\varepsilon^* - b\|_Y = \varepsilon$, then $y_\varepsilon^* = \lambda_\varepsilon \frac{b - Ax_\varepsilon^*}{\|b - Ax_\varepsilon^*\|_Y}$ for some $\lambda_\varepsilon \geq 0$, which is the usual complementary-slackness relation for a (Hilbertian)-ball constraint.*

Note that the optimality conditions (5)–(6) are only necessary: if the subdifferentials involved contain more than one vector, there is *a priori* no way to determine which one the actual solution x_ε^* (or x^*) is. Thus, for the previous result to be more constructive, it is relevant to study tractable sufficient conditions ensuring that these subdifferentials are reduced to singletons (i.e., actually are gradients). In this case, the above inclusion then uniquely determines x_ε^* (or x^*).

Singular vectors. Let us introduce the set of *singular vectors* $\text{sing}(K_r)$ of K_r , that is, the set of vectors $z \in X$ such that the above subdifferential $\partial\sigma_{K_r}(z)$ is not reduced to a singleton. Equivalently, this corresponds to σ_{K_r} failing to be differentiable at z . For instance, if K_r is a polytope, singular vectors are exactly those normals that expose a face of dimension ≥ 1 (an edge, a facet, etc).

We are thus led to the hypothesis

$$\forall y \in Y, \quad A^*y \in \text{sing}(K_r)^c \cup P_r^\circ, \quad (\text{H})$$

which can also be written as $\text{Ran}(A^*) \subset \text{sing}(K_r)^c \cup P_r^\circ$.

Assumption H can be seen as a non-degeneracy condition ensuring that the support-function subgradient used in (5)–(6) for the reconstruction of x_ε^* is single-valued along $\text{Ran}(A^*)$.

Proposition 1.1. Under (H), conditions (5) and (6) become, respectively,

$$\{x_\varepsilon^*\} = \sigma_{K_r}(A^*y_\varepsilon^*) \partial\sigma_{K_r}(A^*y_\varepsilon^*), \quad (7)$$

and

$$\{x^*\} = \sigma_{K_r}(A^*y^*) \partial\sigma_{K_r}(A^*y^*). \quad (8)$$

That is, these optimality conditions are necessary and sufficient.

Of course, assumption (H) is a much stronger hypothesis than what is actually needed: we only need the inclusion underlying (H) to hold at a minimiser y^* of J_ε .

Remark 1.9. A typical example where (H) holds is given in the previous section, with P_r a convex and closed cone. As noted in Remark 1.5, the corresponding generator set K_r is the intersection of P_r with the unit ball. In this case, we have $\text{sing}(K_r) \subset P_r^\circ$. Indeed, using Lemmas 3.4 and A.4, for any $v \in X$, if $v_+ \neq 0$, then $\sigma_{K_r}(v) = \|v_+\|_X$ and $\partial\sigma_{K_r}(v) = v_+/\|v_+\|_X$. Hence, $v \in \text{sing}(K_r)$ implies $v_+ = 0$, i.e., $v \in P_r^\circ$.

As a result,

$$\text{sing}(K_r)^c \cup P_r^\circ = X,$$

and (H) holds.

This illustrates the importance of the choice of a generator for the convex cone, when this choice is available, and a generic example thereof.

1.4.3 General cone

We now consider the case where P might also not be convex: we let P be any cone, and let K be a bounded set containing 0 that generates P , i.e., such that $\text{cone}(K) = P$. We **relax the constraints** by defining K_r to be the closed convex hull of K , namely $K_r := \overline{\text{conv}}(K)$ and $P_r = \text{cone}(K_r)$, a convex cone that is not necessarily closed. The diagram below summarises the overall relaxation approach.

$$\begin{array}{ccc} P & \xrightarrow{\subset} & P_r \\ \text{cone} \uparrow & & \uparrow \text{cone} \\ K & \xrightarrow{\overline{\text{conv}}} & K_r \end{array}$$

We now apply the previous approach to the cone P_r and corresponding generating set K_r . The key question is now whether the constructive characterisations (5)–(6) lead to elements in the original constraint cone P , and not merely in P_r . This motivates the following assumption.

$$\text{ext}(K_r) \subset K, \quad (\text{E})$$

where $\text{ext}(K_r)$ refers to extreme vectors of K_r , i.e., that vectors of K_r that cannot be written as a nontrivial convex combination of vectors in K_r .

Assumption (E) will often be satisfied in practice. Indeed, since K is assumed to be bounded, so is the closed convex set K_r ; hence the latter set is weakly compact. As a result, an application of Milman's theorem (see [23][Theorem 3.24]) with either the strong or the weak topology shows that (E) holds true as soon as either one of the following hypotheses holds

- K is weakly closed,
- K_r is strongly compact and K is strongly closed.

Theorem 1.6. *There hold*

- (i) *If $b \in \overline{A(P_r)}$, then for any $\varepsilon > 0$,*

- $(\mathcal{P}_\varepsilon)$ admits a solution $x_\varepsilon^* \in P_r$ such that $\|Ax_\varepsilon^* - b\|_Y \leq \varepsilon$,
- $(\mathcal{D}_\varepsilon)$ admits a unique solution y_ε^* ,
- under (H)-(E), x_ε^* is the unique solution to $(\mathcal{P}_\varepsilon)$, $x_\varepsilon^* \in P$ and

$$\{x_\varepsilon^*\} = \sigma_{K_r}(A^*y_\varepsilon^*) \partial\sigma_{K_r}(A^*y_\varepsilon^*).$$

(ii) If $b \in A(P_r)$,

- (\mathcal{P}_0) admits a solution $x^* \in P_r$ such that $Ax^* = b$,
- if furthermore (\mathcal{D}_0) admits a solution y^* , then under (H)-(E), $x^* \in P$ and

$$\{x^*\} = \sigma_{K_r}(A^*y^*) \partial\sigma_{K_r}(A^*y^*).$$

Notice that in the above theorem, (i) implies that, as soon as $b \in \overline{A(P_r)}$, then $b \in \overline{A(P)}$. Indeed, the existence of solutions to the auxiliary problem $(\mathcal{P}_\varepsilon)$ and its dual $(\mathcal{D}_\varepsilon)$ holds independently of $b \in Y$. We have thus obtained the following corollary:

Corollary 1.1. *Under (H)-(E), there holds $\overline{A(P_r)} = \overline{A(P)}$.*

Remark 1.10. *This type of statement is common in optimal control theory, in the form of so-called bang-bang principles: in many optimal control problems, the set of reachable states using only controls which saturate the constraints (bang-bang controls), is equal to the set of reachable states using all admissible controls [17]. Controls which saturate the constraints are typically extreme points of the control constraint set.*

1.5 Remarks and perspectives

Uniqueness. In all our results regarding the approximate solvability problem, the minimiser of the dual functional is unique. In contrast, that is no longer the case for the exact solvability problem, even in simple cases.

Consider $X = Y = \mathbb{R}^2$, $P_r = \mathbb{R}_+^2$, $K_r = P_r \cap \overline{B(0, 1)}$, $A = \text{Id}$, $b = (1, 0)$. Clearly, $b \in A(P_r)$. The dual functional J_0 then writes

$$J_0(y) = J_0(y_1, y_2) = \frac{1}{2}\|y_+\|_2^2 - y_1.$$

One easily checks that J reaches its infimum value $-\frac{1}{2}$ on the whole half-line $\{(1, a), a \leq 0\}$. Note however that the constrained linear equation $Ax = b$ with $x \in P$ has a unique solution.

Generalisations. The Hilbert space setting we have chosen is sufficiently general to be suitable for applications requiring infinite-dimensional spaces (e.g., linear control theory), yet sufficiently simple to be tractable.

We are however fairly confident that the core of our results can be generalised to reflexive Banach spaces, or at least to uniformly convex Banach spaces, upon using proper generalisations of the relevant tools (notably, Moreau's decomposition in Banach spaces [12]).

Constructiveness. In the second parts of Theorems 1.4, 1.5, 1.6, the dual problem for the exact solvability problem may or may not admit a minimiser. This issue, which plays a central role in the constructiveness of our method, is discussed in further detail in Section 4, where we introduce the notion of *weak constructiveness*. We will focus on the existence of an identity characterising the solution of the primal problem (when it exists), notwithstanding hypotheses of the type (H)-(E) guaranteeing that this identity characterises x^* uniquely. We obtain some partial results and counterexamples which raise important questions on the general issue of constructive characterisations for solutions to constrained equations.

Intrinsic condition for the exact solvability problem. In the case of a closed cone, the results we provide in Theorems 1.2 and 1.4 are intrinsic to the cone. This is also the case for the necessary and sufficient condition we provide in Theorem 1.3 for the approximate solvability problem.

However, the rest of our results seem to depend strongly on the generator K_r of our cone. It would be interesting to obtain an intrinsic version of these results, *e.g.* an analog to the identity $x^* = (A^*y^*)_+$, which would hold for a convex (and not necessarily closed) cone.

Class of cones under consideration. In the general case, we consider convex cones that are generated by a bounded closed convex set containing 0. It would be relevant to describe this class of cones in a less indirect way. Going further, and considering a somewhat dual question, can we describe the set of bounded closed convex generators for such a cone?

This is related to the previous paragraph concerning a possible intrinsic characterisation of exact solvability. Indeed, the necessary and sufficient condition for exact solvability, and the constructive characterisation of the solution x^* , seem to depend on the generator K of the cone. Having a more detailed description of the class of possible generators, and more knowledge on the properties of the cone P , could allow us to consider either an “optimal” generator to write these results, or a “generic” one. For example, in the case where the cone is closed, the generic generator seems to be the intersection of the cone with the closed unit ball.

Numerical approximations of solutions. Because of its constructive nature and its reliance upon the Fenchel-Rockafellar Theorem, our approach lends itself to using algorithms that find saddle-points of the corresponding Lagrangian, such as the Chambolle-Pock algorithm [10].

For applications to control theory or inverse problems associated to ODEs and/or PDEs, at least one of the Hilbert spaces is infinite-dimensional. This poses the question of how to appropriately discretise the involved functionals, especially in view of preserving the primal-dual structure at the discrete level.

2 Mathematical framework

2.1 Notations

For the readers’ convenience, let us define the convex analytic tools used in the present work, in some generic (real) Hilbert space H .

2.1.1 Functions

Let $f : H \rightarrow \mathbb{R} \cup \{+\infty\}$.

We let $\text{dom}(f) := \{x \in H, f(x) < +\infty\}$ be its **domain**; f is said to be **proper** when $\text{dom}(f) \neq \emptyset$. We define its **convex conjugate** by

$$\forall x \in H, \quad f^*(x) := \sup_{y \in H} (\langle x, y \rangle - f(y)).$$

We let $\Gamma_0(H)$ be the set of all proper, convex and lower semicontinuous functions on H and recall Fenchel-Moreau’s identity: if $f \in \Gamma_0(H)$ then $f^* \in \Gamma_0(H)$ and $(f^*)^* = f$.

Finally, for a given $f \in \Gamma_0(H)$, and for $x \in H$, we define the **(convex) subdifferential** of f at x through

$$\partial f(x) := \{p \in H, \forall y \in H, \quad f(y) \geq f(x) + \langle p, y - x \rangle\}.$$

2.1.2 Convex sets

Let $C \subset H$ be a convex set.

We let δ_C be its **indicator** function, namely $\delta_C(x) = 0$ if $x \in C$ and $+\infty$ otherwise, and $\sigma_C = \delta_C^*$ be its **support** function; in other words, $\sigma_C(x) = \sup_{y \in C} \langle x, y \rangle$. We also define its **gauge** function j_C , defined for $x \in H$ by $j_C(x) = \inf\{\alpha > 0, x \in \alpha C\}$.

The set $\text{ext}(C)$ of **extreme** vectors of C refers to vectors $x \in C$ that cannot be written as a non-trivial convex combination of two vectors in C .

If C is closed and bounded, $\sigma_C(x) < +\infty$ for all $x \in C$ and the supremum defining $\sigma_C(x)$ is a maximum. In this case, recall that $\text{sing}(C)$ denotes the set of **singular vectors** to C , namely those vectors $x \in H$ for which the set of maximisers associated to the maximum defining $\sigma_C(x)$ is not reduced to a singleton.

Finally, for $x \in H$, we let $N_C(x) = \{p \in H, \forall y \in C, \langle p, x \rangle \geq \langle p, y \rangle\}$ be the **normal cone** to C at x .

For an arbitrary set $K \subset H$, we let $\overline{\text{conv}}(K)$ be its **closed convex hull**, namely the smallest closed and convex set containing K .

2.1.3 Cones

Let $P \subset H$ be a cone.

We let $P^\circ = \{x \in H, \forall y \in P, \langle x, y \rangle \leq 0\}$ be the **polar cone** of P . If P is closed and convex, given $x \in H$ we let x_+ be its projection onto P and x_- be its projection onto P° . Moreau's decomposition theorem yields $x = x_+ + x_-$ with $\langle x_+, x_- \rangle = 0$.

Finally, we highlight two useful identities (see [1, Proposition 6.26] and [1, Proposition 6.34] respectively): for two nonempty convex cones P_1 and P_2 , there holds

$$P_1^\circ \cap P_2^\circ = (P_1 + P_2)^\circ, \quad (\overline{P_1} \cap \overline{P_2})^\circ = \overline{P_1^\circ + P_2^\circ}. \quad (9)$$

2.2 Fenchel-Rockafellar duality and constructive characterisation of solutions

Let $P_r = \text{cone}(K_r)$ be a cone, where K_r is a closed bounded convex set containing 0. Hence P_r is a convex cone containing 0, and is not necessarily closed. It is straightforward but essential to note the identity

$$\text{dom}(j_{K_r}) = P_r. \quad (10)$$

The purpose of this section is to define the (primal) optimisation problem underlying our results, derive its dual problem, establish strong duality and make the corresponding saddle-point conditions explicit.

2.2.1 Primal and dual optimisation problems

For any $\varepsilon \geq 0$, we consider the *primal optimisation problem*,

$$\pi_\varepsilon = \inf_{x \in X} \frac{1}{2} j_{K_r}^2(x) + \delta_{\{x \in X, \|Ax - b\|_Y \leq \varepsilon\}}(x) = \inf_{x \in X} \frac{1}{2} j_{K_r}^2(x) + \delta_{\overline{B(b, \varepsilon)}}(Ax). \quad (11)$$

For a given $\varepsilon \geq 0$ and in view of (10), note that the functional underlying the primal problem is not identically equal to $+\infty$ if and only if there exists $x \in X$ such that $x \in P_r$ and $\|Ax - b\|_Y \leq \varepsilon$. In other words, $\pi_\varepsilon < +\infty$ if and only if there exists $x \in X$ such that $x \in P_r$ and $\|Ax - b\|_Y \leq \varepsilon$. Consequently, the case $\varepsilon = 0$ amounts to exact solvability.

Proposition 2.1. *For any $\varepsilon \geq 0$, the following strong duality holds*

$$\pi_\varepsilon = \inf_{x \in X} \frac{1}{2} j_{K_r}^2(x) + \delta_{\overline{B(b, \varepsilon)}}(Ax) = - \inf_{y \in Y} J_\varepsilon(y), \quad (12)$$

where the dual functional J_ε is given by

$$\forall y \in Y, \quad J_\varepsilon(y) := \frac{1}{2} \sigma_{K_r}^2(A^*y) - \langle b, y \rangle_Y + \varepsilon \|y\|_Y.$$

Furthermore, if π_ε is finite, it is attained.

Since K_r is bounded, σ_{K_r} is Lipschitz continuous. The function J_ε is thus continuous and, in particular, takes finite values on the whole of Y .

Proof. The primal functional can be written in the form $F(x) + G(Ax)$, where $F = \frac{1}{2}j_{K_r}^2$, $G = \delta_{\overline{B(b,\varepsilon)}}$. Clearly, $G \in \Gamma_0(Y)$, while by Lemma A.2 (ii), $F \in \Gamma_0(X)$. Its dual functional, defined on Y , is then given by

$$y \mapsto F^*(A^*y) + G^*(-y)$$

It is standard that $G^* = \delta_{\overline{B(b,\varepsilon)}}^* = \langle b, \cdot \rangle_Y + \varepsilon \|\cdot\|_Y$, and Lemma A.2 (ii) shows that $F^* = (\frac{1}{2}j_{K_r}^2)^* = \frac{1}{2}\sigma_{K_r}^2$.

We now apply the Fenchel-Rockafellar Theorem [22]: if there exists $y \in Y$ such that $G^*(-y) < +\infty$ and F^* is continuous at A^*y , then strong duality holds and the infimum of the primal optimisation problem is attained if finite. Here, G^* takes finite values on Y , while F^* is continuous on X as already mentioned. Hence any $y \in Y$ can be used to apply the Fenchel-Rockafellar theorem, and the proof is finished. \square

2.2.2 Optimality conditions

What follows is standard folklore related to (strong) duality associated to the Fenchel-Rockafellar Theorem [22].

A couple $(x^*, y^*) \in X \times Y$ is a primal-dual optimal pair (i.e., x^* minimises the primal functional, and y^* minimises the dual functional) if and only if $(x^*, -y^*)$ is a saddle-point of the Lagrangian

$$(x, z) \in X \times Y \mapsto \langle z, Ax \rangle_Y + F(x) - G^*(z) = \langle z, Ax \rangle_Y + \frac{1}{2}j_{K_r}^2(x) - \langle b, z \rangle_Y - \varepsilon \|z\|_Y.$$

This is also equivalent to the two following first order optimality conditions

$$\begin{cases} x^* \in \partial F^*(A^*y^*) = \partial(\frac{1}{2}\sigma_{K_r}^2)(A^*y^*) = \sigma_{K_r}(A^*y^*) \partial \sigma_{K_r}(A^*y^*) \\ y^* \in -\partial G(Ax^*) \end{cases}, \quad (13)$$

where the identity used above in the first line is due to the chain rule for subdifferentials, see for instance [11][Theorem 2.3.9, (ii)].

By the results and the remarks of the previous subsection, given that $\pi_\varepsilon = -\inf_{y \in Y} J_\varepsilon(y)$ for all $\varepsilon \geq 0$, we have

$$\begin{aligned} \forall \varepsilon > 0, \inf_{y \in Y} J_\varepsilon(y) > -\infty &\implies b \in \overline{A(P_r)}, \\ \inf_{y \in Y} J_0(y) > -\infty &\implies b \in A(P_r). \end{aligned} \quad (14)$$

Combining (14) with the optimality conditions (13), we immediately obtain:

Lemma 2.1. *Let $\varepsilon \geq 0$. If J_ε admits a minimiser y_ε^* , then there exists*

$$x_\varepsilon^* \in \sigma_{K_r}(A^*y_\varepsilon^*) \partial \sigma_{K_r}(A^*y_\varepsilon^*),$$

which is a minimiser for $(\mathcal{P}_\varepsilon)$. In particular, $x_\varepsilon^ \in P_r$ and $\|Ax_\varepsilon^* - b\|_Y \leq \varepsilon$.*

Corollary 2.1. *If (x^*, y^*) is a primal-dual optimal pair such that $\sigma_{K_r}(A^*y^*) \partial \sigma_{K_r}(A^*y^*)$ is reduced to a singleton, then this singleton equals x^* .*

The proofs of our main results will rely greatly on the above observations, and a careful examination of the relationship between the Farkas problems (2) and (3), and the optimisation problems $(\mathcal{P}_\varepsilon)$ and $(\mathcal{D}_\varepsilon)$ for $\varepsilon \geq 0$.

3 Proofs

Outline of the proofs. We first focus on cones generated by a bounded closed convex set containing 0, *i.e.*, Theorems 1.3 and 1.5.

We tackle approximate solvability, then exact solvability. Section 3.1 provides:

- proof of Theorem 1.3 (i) - necessary and sufficient conditions for approximate solvability;
- proof of the uniqueness of dual minimisers (Lemma 3.2);
- proof of Theorem 1.5 (i) - constructive characterisation and uniqueness of solutions in the approximate case.

In Section 3.2:

- we prove Theorem 1.3 (ii) - a necessary and sufficient condition for exact solvability;
- we deduce Theorem 1.5 (ii) regarding the characterisation of solutions when they exist.

Section 3.3 derives Theorems 1.2 and 1.4, which are particular instances of the above theorems, with a closed cone.

Finally, in Section 3.4, we focus on situations where the solution can be completely characterised:

- in Subsection 3.4.1 we address the case of convex cones, proving Proposition 1.1;
- in Subsection 3.4.2 we turn to general cones using relaxation, and prove Theorem 1.6.

Within this section, we shall often use the trivial fact that σ_{K_r} is 1-positively homogeneous.

3.1 Approximate solvability

3.1.1 Necessary and sufficient condition for solvability

We begin by proving Theorem 1.3, (i). First, we check that:

Lemma 3.1. *For $y \in Y$, $\sigma_{K_r}(A^*y) = 0$ if and only if $A^*y \in P_r^\circ$.*

Proof. Let us first remark that $\sigma_{K_r}(A^*y) = 0$ if and only if $\langle A^*y, v \rangle_X \leq 0$ for all $v \in K_r$. Given that $P_r = \text{cone}(K_r)$, it is equivalent to $\langle A^*y, v \rangle_X \leq 0$ for all $v \in P_r$, *i.e.*, $A^*y \in P_r^\circ$. \square

We now have two equivalent formulations of the solvability condition in Theorem 1.3, (i), one that is intrinsic to P_r , and one that depends on the choice of generator K_r . We now proceed with the proof, using the latter.

Lemma 3.2. *The following holds:*

$$\left(\forall y \in Y, \sigma_{K_r}(A^*y) = 0 \implies \langle b, y \rangle_Y \leq 0 \right) \implies \forall \varepsilon > 0, \inf_{y \in Y} J_\varepsilon(y) = \min_{y \in Y} J_\varepsilon(y) > -\infty.$$

Proof. Assume that, for all $y \in Y$, $\sigma_{K_r}(A^*y) = 0 \implies \langle b, y \rangle_Y \leq 0$.

Since the functionals $J_\varepsilon, \varepsilon > 0$ are in $\Gamma_0(Y)$, the result will be proved if we establish that J_ε is coercive for all $\varepsilon > 0$.

We let $\varepsilon > 0$ be fixed. By contradiction, if J_ε is not coercive there exists a sequence $(y_n) \in Y^\mathbb{N}$ such that $\|y_n\|_X \rightarrow \infty$ with $(J(y_n))$ upper bounded, say $J(y_n) \leq M$ for all $n \in \mathbb{N}$. Denoting $z_n := \frac{y_n}{\|y_n\|_Y}$ and since $\sigma_{K_r}^2$ is 2-positively homogeneous, we have

$$\frac{1}{2} \|y_n\|_Y^2 \sigma_{K_r}^2(A^*z_n) - \|y_n\|_Y \langle b, z_n \rangle_Y + \varepsilon \|y_n\|_Y = J_\varepsilon(y_n) \leq M.$$

Consequently, if $\liminf_{n \rightarrow \infty} \sigma_{K_r}^2(A^*z_n) > 0$, then the left-hand side above diverges to $+\infty$ which is not possible. Let us now treat the remaining case where $\liminf_{n \rightarrow \infty} \sigma_{K_r}^2(A^*z_n) = 0$. We may upon extraction of a subsequence assume both that $\lim_{n \rightarrow \infty} \sigma_{K_r}^2(A^*z_n) = 0$, and owing to the boundedness of (z_n) , that $z_n \rightharpoonup z$ weakly in Y for some $z \in Y$. Since $A^* \in L(Y, X)$, we have $A^*z_n \rightharpoonup A^*z$ weakly in X .

The function $\sigma_{K_r}^2$ is convex and strongly lower semicontinuous on E , hence it is (sequentially) weakly lower semicontinuous, which implies that $\sigma_{K_r}^2(A^*z) \leq \liminf_{n \rightarrow \infty} \sigma_{K_r}^2(A^*z_n) = 0$ and enforces $\sigma_{K_r}^2(A^*z) = 0$. By our assumption, this implies $\langle b, z \rangle_Y \leq 0$. In this case, we have

$$\|y_n\|_Y (-\langle b, z_n \rangle_Y + \varepsilon) \leq J_\varepsilon(y_n) \leq M,$$

and at the limit we again reach a contradiction. \square

Let us now prove Theorem 1.3 (i), namely the following Proposition.

Proposition 3.1. *We have the equivalence*

$$b \in \overline{A(P_r)} \iff \forall y \in Y, \sigma_{K_r}(A^*y) = 0 \implies \langle b, y \rangle_Y \leq 0.$$

Proof. Assume that $b \in \overline{A(P_r)}$: we may find $(w_n) \in P_r^{\mathbb{N}}$ such that the sequence (Aw_n) converges to b . By contradiction, and using Lemma 3.1, assume that there exists $y \in Y$ such that $A^*y \in P_r^{\circ}$ and $\langle b, y \rangle_Y > 0$. Then we should have

$$0 \geq \langle A^*y, w_n \rangle_X = \langle y, Aw_n \rangle_Y \longrightarrow \langle y, b \rangle_Y > 0,$$

which is impossible.

We now tackle the converse assertion, assuming that, for all $y \in Y$, $\sigma_{K_r}(A^*y) = 0 \implies \langle b, y \rangle_Y \leq 0$. By Lemma 3.2, this implies that $\inf_{y \in Y} J_\varepsilon(y) > -\infty$ for all $\varepsilon > 0$. Combined with (14), this concludes the proof that $b \in \overline{A(P_r)}$. \square

We can sum up the relationship of the auxiliary optimisation problems $(\mathcal{P}_\varepsilon)$ – $(\mathcal{D}_\varepsilon)$ to the approximate Farkas problem in this way:

Corollary 3.1. *The following propositions are equivalent:*

- (i) $b \in \overline{A(P_r)}$;
- (ii) For all $\varepsilon > 0$, $\inf_{y \in Y} J_\varepsilon(y) > -\infty$;
- (iii) For all $\varepsilon > 0$, $(\mathcal{P}_\varepsilon)$ admits a solution x_ε^* .

Proof. The implication (i) \implies (ii) is a consequence of Proposition 3.1 and Lemma 3.2.

The implication (ii) \implies (iii) is a direct consequence of the strong duality given in Proposition 2.1: $\pi_\varepsilon = -\inf_Y J_\varepsilon < +\infty$ and is hence attained at a minimiser x_ε^* .

Finally, recalling the implication (4) given in the introduction, (iii) implies that for all $\varepsilon > 0$ there exists $x_\varepsilon^* \in P_r$ such that $\|Ax_\varepsilon^* - b\|_Y \leq \varepsilon$, i.e., $b \in \overline{A(P_r)}$. \square

3.1.2 Constructive characterisation and uniqueness

We now turn to the proof of Theorem 1.5 (i). The previous section already provides the existence of a solution x_ε^* to $(\mathcal{P}_\varepsilon)$, and a minimiser for the dual functional J_ε , under the assumption that $b \in \overline{A(P_r)}$.

We now prove that this dual minimiser is always unique:

Proposition 3.2. *If $b \in \overline{A(P_r)}$, then J_ε admits a unique minimiser for all $\varepsilon > 0$.*

Proof. Since $b \in \overline{A(P_r)}$, Lemma 3.2 shows that for all $\varepsilon > 0$, J_ε has a finite infimum, which is in fact a minimum.

Let $\varepsilon > 0$ be fixed, and denote by X^* the set of minimisers for the primal problem $(\mathcal{P}_\varepsilon)$, and by Y^* the set of minimisers for the dual problem $(\mathcal{D}_\varepsilon)$. Let $(x^*, y^*) \in X^* \times Y^*$ be a primal-dual optimal pair, which exists by Corollary 3.1.

If $\|b\|_Y \leq \varepsilon$, $x^* = 0$ is an optimal primal variable since $\|A0 - b\|_Y = \|b\|_Y \leq \varepsilon$. But $j_{K_r}(x) = 0$ if and only if $x = 0$ by Lemma A.1(i), hence 0 is the unique optimal primal variable. In particular, $AX^* = \{0\}$.

If $\|b\|_Y > \varepsilon$, $0 \notin X^*$ so $\pi_\varepsilon > 0$. By (12), $\min J_\varepsilon < 0$ and thus $0 \notin Y^*$. Then, the relation $y^* \in -\partial\delta_{\overline{B(b, \varepsilon)}}(Ax^*)$ of (13), implies that Ax^* must lie at the boundary of $\overline{B(b, \varepsilon)}$. Since X^* is convex, so is AX^* . Hence, AX^* is a convex subset of the boundary of $\overline{B(b, \varepsilon)}$, *i.e.*, the sphere $\{y \in Y, \|y - b\|_Y = \varepsilon\}$. The closed ball being strictly convex in the Hilbert space Y , any convex subset of its boundary must be a singleton. Thus, there exists some $b^* \in \overline{B(b, \varepsilon)}$ with $\|b^* - b\|_Y = \varepsilon$ such that

$$AX^* = \{b^*\}. \quad (15)$$

Thus, in any case, the set AX^* is always reduced to a single point, which we can denote b^* in general.

We now return to the inclusion $y^* \in -\partial\delta_{\overline{B(b, \varepsilon)}}(Ax^*) = -\partial\delta_{\overline{B(b, \varepsilon)}}(b^*)$. First assume $\|b\|_Y \leq \varepsilon$, then $b^* = 0$ and $y^* \in -\partial\delta_{\overline{B(b, \varepsilon)}}(0)$. If $\|b\|_Y < \varepsilon$, $0 \in B(b, \varepsilon)$, and thus the latter subdifferential reduces to $\{0\}$, leading to $y^* = 0$. Otherwise, $\|b\|_Y = \varepsilon$, $0 \in \partial\overline{B(b, \varepsilon)}$ and we find

$$y^* \in \left\{ \lambda \frac{b}{\varepsilon}, \lambda \geq 0 \right\} = \{\lambda b, \lambda \geq 0\}.$$

Restricting the function J_ε to the above half-line, using the homogeneities of each of its terms, and the fact that $\|b\|_Y = \varepsilon$, we get

$$\gamma_0(\lambda) := J_\varepsilon(\lambda b) = a_0 \lambda^2, \quad \lambda \geq 0,$$

where $a_0 = \frac{1}{2} \sigma_{K_r}^2(A^*b)$. It is clear that 0 is the unique minimiser of γ_0 . In other words, we have proved $y^* = 0$ whenever $\|b\|_Y \leq \varepsilon$.

Now assume $\|b\|_Y > \varepsilon$. In this case, note that $y^* \neq 0$, since otherwise the dual problem would admit the minimum 0, hence the primal optimisation problem would also be of minimum 0. If so, $x = 0$ is an optimal primal variable, which would mean that $\|A0 - b\|_Y = \|b\|_Y \leq \varepsilon$, a contradiction.

To prove the uniqueness of $y^* \neq 0$, we argue as follows. Since b^* lies at the boundary of $\overline{B(b, \varepsilon)}$, we have

$$y^* \in \left\{ \lambda \left(\frac{b - b^*}{\varepsilon} \right), \lambda \geq 0 \right\} = \{\lambda(b - b^*), \lambda \geq 0\}.$$

Restricting J_ε to the above half-line as previously, we find

$$\gamma(\lambda) := J_\varepsilon(\lambda(b - b^*)) = a_1 \lambda^2 + a_2 \lambda, \quad \lambda \geq 0,$$

where, using $\|b - b^*\|_Y = \varepsilon$ and the homogeneities involved $a_1 = \frac{1}{2} \sigma_{K_r}^2(A^*(b - b^*))$ and $a_2 = -\langle b, b - b^* \rangle_Y + \varepsilon^2$. Because J_ε is bounded below and has negative minimum, we must have $a_1 > 0$ and $a_2 < 0$.

Consequently, γ has a unique minimiser $\lambda^* := -\frac{a_2}{2a_1} > 0$. Hence, $y^* = \lambda^*(b - b^*)$, and the dual optimal variable is unique. \square

Combined with Lemma 2.1, this proves Theorem 1.5 (i) regarding the characterisation of minimisers for $(\mathcal{P}_\varepsilon)$.

3.2 Exact solvability

Since we are interested in exact solvability, we set $\varepsilon = 0$ and will be using the corresponding dual functional J_0 . We here prove Theorem 1.3 (ii), and then Theorem 1.5 (ii).

Lemma 3.3. *The following holds:*

$$\exists C > 0, \forall y \in Y, \langle b, y \rangle_Y \leq C \sigma_{K_r}(A^*y) \implies \inf_{y \in Y} J_0(y) > -\infty.$$

Proof. A lower bound for J_0 is readily derived:

$$\forall y \in Y, \quad J_0(y) = \frac{1}{2} \sigma_{K_r}^2(A^*y) - \langle b, y \rangle_Y \geq \frac{1}{2} \sigma_{K_r}^2(A^*y) - C \sigma_{K_r}(A^*y).$$

Since the function $z \mapsto \frac{1}{2}z^2 - z$ is lower bounded on \mathbb{R} , so is J_0 . \square

We can now prove Theorem 1.3 (ii), which we restate in the following proposition:

Proposition 3.3. *We have the equivalence*

$$b \in A(P_r) \iff \exists C > 0, \forall y \in Y, \langle b, y \rangle_Y \leq C \sigma_{K_r}(A^*y).$$

If $b \in A(P_r)$, the best constant C is given by $C = \sqrt{2\pi_0}$.

Proof. Assume that $b \in A(P_r)$. Let $x \in P_r$ be such that $Ax = b$. Since $x \in P_r$, we may write $x = j_{K_r}(x)v$ with $v \in K_r$, by Lemma A.1(iv). For all $y \in Y$, we write

$$\langle b, y \rangle_Y = \langle Ax, y \rangle_Y = \langle x, A^*y \rangle_Y = j_{K_r}(x) \langle v, A^*y \rangle \leq C \sigma_{K_r}(A^*y),$$

with $C = j_{K_r}(x)$. By minimising over all possible $x \in P_r$ such that $Ax = b$, we may in fact set $C = \sqrt{2\pi_0}$.

We now assume that the inequality holds. In this case, Lemma 3.3 shows that J_0 has a finite infimum, and we may conclude that $b \in A(P_r)$ by (14). \square

We can now state the analog of Corollary 3.1 for exact solvability problems:

Corollary 3.2. *The following propositions are equivalent:*

- (i) $b \in A(P_r)$;
- (ii) $\inf_{y \in Y} J_0(y) > -\infty$;
- (iii) *The primal problem (\mathcal{P}_0) admits a solution x^* .*

Proof. The implication (i) \implies (ii) is a consequence of Proposition 3.3 and Lemma 3.3.

The implication (ii) \implies (iii) is a direct consequence of the strong duality given in Proposition 2.1: $\pi_0 = -\inf_Y J_0 < +\infty$ and is hence attained at a minimiser x^* .

Finally, recalling the implication (4) given in the introduction, (iii) implies that $Ax^* = b$, i.e., $b \in A(P_r)$. \square

To prove Theorem 1.5 (ii), we now let $b \in A(P_r)$. Note that, in contrast to approximate solvability problems, the associated dual functional J_0 does not always admit a minimiser, even though its infimum is finite. As a result, even though a minimiser x^* always exists for (\mathcal{P}_0) , the constructive characterisation (6) may not be available. However, under the hypothesis that J_0 does admit a minimiser, we can apply Lemma 2.1 with $\varepsilon = 0$ and conclude.

3.3 Closed convex case

We now turn to the particular case of closed convex cones, described by Theorems 1.2 and 1.4. Let P_r be a closed and convex cone containing 0. Then, noting that $\text{cone}(P_r \cap \overline{B(0,1)}) = P_r$, we have a natural generating set $P_r \cap \overline{B(0,1)}$ which is a closed, convex and bounded set containing 0.

Lemma 3.4. *Let $K_r := P_r \cap \overline{B(0,1)}$. Then for all $x \in X$, there holds*

$$\sigma_{K_r}(x) = \|x_+\|_X \quad \text{and} \quad \sigma_{K_r}(x) \partial(\sigma_{K_r})(x) = \{x_+\}.$$

Proof. Let us start with the case where $x_+ = 0$. In this case $x \in P_r^\circ$, hence $\sigma_{K_r}(x) \leq 0$, and since K_r contains 0, we have $\sigma_{K_r}(x) = 0 = \|x_+\|_X$. Hence, both identities hold.

Now assume that $x_+ \neq 0$. For all $v \in K_r$, i.e., $v \in P_r$ with $\|v\|_X \leq 1$,

$$\langle v, x \rangle_X = \langle v, x_+ \rangle_X + \langle v, x_- \rangle_X \leq \langle v, x_+ \rangle_X \leq \|v\|_X \|x_+\|_X \leq \|x_+\|_X,$$

where equality is attained with the choice $v = \frac{x_+}{\|x_+\|_X}$ since $\langle x_+, x_- \rangle_X = 0$. Hence, $\sigma_{K_r}(x) = \|x_+\|_X$.

Let us show $\frac{x_+}{\|x_+\|_X}$ is the unique choice to reach equality. Indeed, if $v \in P_r$, $\|v\|_X \leq 1$ is such that $\langle v, x \rangle_X = \|x_+\|_X$, then

$$\|x_+\|_X = \langle x_+, v \rangle_X \leq \|x_+\|_X \|v\|_X \leq \|x_+\|_X.$$

All these inequalities must be equalities, thus enforcing $\|v\|_X = 1$ on the one hand, and nonnegative linear dependence of v and x_+ on the other hand, which amounts to $v = \frac{x_+}{\|x_+\|_X}$. We have proved $\partial(\sigma_{K_r})(x) = \{\frac{x_+}{\|x_+\|_X}\}$, and hence $\sigma_{K_r}(x) \partial(\sigma_{K_r})(x) = \{x_+\}$. \square

Lemma 3.5. *Let $K_r := P_r \cap \overline{B(0, 1)}$. If $b \in \overline{A(P_r)}$, then for every $\varepsilon > 0$, $(\mathcal{P}_\varepsilon^{\text{cl}})$ admits a unique solution. If $b \in A(P_r)$, then $(\mathcal{P}_0^{\text{cl}})$ admits a unique solution.*

Proof. In both cases, we know that the primal problems reach their infima, which we denote $\frac{1}{2}(\lambda_\varepsilon^*)^2$ (resp. $\frac{1}{2}(\lambda^*)^2$). Since the objective functionals are convex, the sets of minimisers are convex, and are subsets of $\lambda_\varepsilon^* S^1$ (resp. $\lambda^* S^1$) where S^1 is the unit sphere of X . The latter set being strictly convex, both sets of minimisers are reduced to a singleton. \square

Lemmas 3.4 and 3.5 combined with Theorem 1.3 (resp. Theorem 1.5) proves Theorem 1.2 (resp. Theorem 1.4).

3.4 Optimality conditions and constructiveness

Recall that, for a given $\varepsilon \geq 0$, a primal-dual optimal pair (x^*, y^*) satisfies $x^* \in \sigma_{K_r}(A^* y^*) \partial \sigma_{K_r}(A^* y^*)$. Most of what follows is dedicated to analysing this very inclusion in various cases. This will allow us to prove Proposition 1.1 and Theorem 1.6.

3.4.1 Cone generated by a bounded closed convex set

Here, we assume that $P_r = \text{cone}(K_r)$ with K_r a closed convex bounded set containing 0. As a result, P_r is convex but not necessarily closed.

In this general setting, the set $\sigma_{K_r}(x) \partial(\sigma_{K_r})(x)$ is not necessarily reduced to a singleton for a given $x \in X$.

Lemma 3.6. *Assume that (H) holds. Then,*

$$\forall y \in Y, \quad \sigma_{K_r}(A^* y) \partial \sigma_{K_r}(A^* y) \text{ is reduced to a singleton.}$$

Proof. Let $x \in X$. Let us first remark that the latter set is never empty, because

$$\partial \sigma_{K_r}(x) = \left\{ v \in K_r, \langle x, v \rangle_X = \sup_{v \in K_r} \langle x, v \rangle_X \right\},$$

and the weakly continuous function $v \mapsto \langle x, v \rangle_X$ is bounded and attains its supremum on the weakly compact set K_r .

Since $\sigma_{K_r}(x) = 0$ if and only if $x \in P_r^\circ$, and by definition of the cone of singular vectors, we end up with the equivalence

$$\sigma_{K_r}(x) \partial \sigma_{K_r}(x) \text{ is reduced to a singleton} \iff x \in P_r^\circ \cup \text{sing}(K_r)^c.$$

Under (H), we obtain the result by applying the above to $x = A^* y$ for $y \in Y$. \square

This Lemma then immediately yields Proposition 1.1.

3.4.2 General cone

Here, we assume that P is any cone containing 0, K is a bounded set containing 0 such that $\text{cone}(K) = P$. Then, we let $K_r := \overline{\text{conv}}(K)$ and $P_r = \text{cone}(K_r)$. Recall that the assumption (H) refers to the inclusion $\text{Ran}(A^*) \subset P_r^\circ \cup \text{sing}(K_r)^c$, while (E) refers to the assumption $\text{ext}(K_r) \subset K$.

Lemma 3.7. *Assume that (H)-(E) holds. Then,*

$$\forall y \in Y, \quad \sigma_{K_r}(A^*y) \partial \sigma_{K_r}(A^*y) \text{ is reduced to a single element in } P.$$

Proof. By Lemma 3.6, we know that $\sigma_{K_r}(A^*y) \partial \sigma_{K_r}(A^*y)$ is reduced to a singleton. If $\sigma_{K_r}(A^*y) = 0$, this singleton equals 0, which is an element of P . If $\sigma_{K_r}(A^*y) \neq 0$, then this singleton is a multiple of a unique vector v in K_r maximising the function $v \mapsto \langle A^*y, v \rangle_X$ over K_r . The set of maximisers of a convex function over a convex set must contain at least an extreme point. Hence, v is in fact in $\text{ext}(K_r)$, and by (E), it is an element of K . Consequently, $\sigma_{K_r}(A^*y) \partial(\sigma_{K_r})(A^*y)$ contains the single element $\sigma_{K_r}(A^*y) v$, which lies in the cone $\text{cone}(K) = P$. \square

This lemma, combined with Corollary 2.1 and Theorems 1.4 and 1.5 respectively, completes the proofs of Theorem 1.6 (i) and (ii) respectively.

4 Dual attainment and shared normal vectors

In this Section, we discuss further the constructive character of our approach. Constructiveness is to be understood here in the sense that an exact (or approximate) solution x^* to (1) can be characterised by

$$x^* \in \mathcal{F}(y^*),$$

where $y^* \in Y$, \mathcal{F} is a set-valued mapping on Y , and both y^* and \mathcal{F} can be approximated numerically. The situation is even more favorable if the identity above completely characterises x^* - *i.e.*, $\mathcal{F}(y^*)$ is reduced to a singleton.

Given Theorems 1.4, 1.5 and 1.6, we will consider the more specific terminology of a **weakly constructive solution** for situations where *the dual problem admits a minimiser*, which leads to a constructive characterisation by Theorems 1.4, 1.5 and 1.6. The more amenable situation where the saddle-point relation uniquely defines x^* corresponds to what we call a **constructive solution**.

Let us first remark that, by Theorem 1.6, our method is always weakly constructive for approximate resolution. We thus focus this section's study to the problem of exact resolution.

4.1 Conditions for the dual problem to attain its infimum

Recall that Theorems 1.2 and 1.3 rely on an unusual application of Fenchel-Rockafellar's duality theorem, in which **regularity conditions on the dual problem** are assumed, leading to the solution of the primal problem being reached at a minimiser x^* if the infimum value of the primal problem is finite. This approach does not, however, tell us whether the dual problem reaches its infimum, as remarked in Theorem 1.4.

Still, it is clear from Theorems 1.4 and 1.6 that if the dual problem admits a minimiser y^* , then our method is weakly constructive through the saddle-point condition

$$x^* \in \sigma_K(A^*y^*) \partial \sigma_K(A^*y^*).$$

It is therefore natural to explore sufficient (potentially necessary) conditions for which the dual functional reaches its infimum. There is extensive literature on this matter. Let us mention, in close relation to our work, the articles [24, 25], which study the constrained linear equation (1) by considering a linear auxiliary optimisation problem and its dual. In these works, the author derives sufficient conditions for the dual to reach its infimum. In particular, one of these sufficient conditions is for (1) to have a solution, along with a generalised Slater condition. In a more general framework, [6] provides a necessary

and sufficient condition, but it requires knowing the infimum of the primal problem and hence usually is intractable from the practical point of view.

Various sufficient conditions have been proposed, relying on appropriate strong duality: under a *regularity condition on the primal problem*, the primal and dual problems have the same value and the dual problem reaches its infimum if it is finite. The idea is thus to formulate weaker regularity conditions on the primal problem (using weakened notions of the relative interior) that still ensure strong duality see [16, 4, 5, 2, 3]. Relying on the notions of quasi-relative interior (and strong quasi-relative interior, respectively),¹ one such rather weak condition is $0 \in \text{sqri}(P - \{x \in X, Ax = b\})$ [5]. Some relevant settings covered by all these works are:

- the case where $P = X$,
- when X is of finite dimension, the case where P is a subspace of X ,
- when Y is of finite dimension, the case where $b \in A \text{qri}(P)$ provided that $\text{qri}(P) \neq \emptyset$.

The first two conditions are somewhat restrictive, while the last condition is difficult to check in practice, since it requires knowledge about the set of solutions to the affine equation $Ax = b$ in relation to the cone P_r , which is precisely what we hope to study.

Another possible way to proceed would be to exploit the saddle-point conditions directly:

Lemma 4.1. *Let K be a bounded closed convex set containing 0, and $P = \text{cone}(K)$. Suppose that $b \in A(P)$, so that the primal problem admits a minimiser x^* . Then, $y^* \in Y$ is a minimiser to the dual problem if and only if*

$$x^* \in \sigma_K(A^*y^*) \partial \sigma_K(A^*y^*).$$

Proof. Given that x^* is a minimiser for the primal problem, y^* is a minimiser if and only if (x^*, y^*) fulfill the saddle-point conditions (13). Given that, by hypothesis, $Ax^* = b$, we have $\partial \delta_{\{b\}}(Ax^*) = \partial \delta_{\{b\}}(b) = Y$. Hence, the second saddle-point condition is always satisfied, and the first saddle-point condition is a necessary and sufficient condition for y^* to be a minimiser for the dual problem. \square

A sufficient condition for constructiveness immediately follows:

Lemma 4.2. *Let P be a closed convex cone, $K := P \cap \overline{B(0, 1)}$.*

If $P \subset \text{ran}(A^)$, then for every $b \in A(P)$, the solution to (3) is constructive.*

Proof. Let $b \in A(P)$, and denote x^* a solution to (3). By hypothesis, there exists $y^* \in Y$ such that $x^* = A^*y^*$. Then, $x^* = A^*y^* = (A^*y^*)_+ = \sigma_K(A^*y^*) \partial \sigma_K(A^*y^*)$ by Lemma 3.4. \square

For the remainder of this section, we shall endeavour to study the problem of minimising the dual functional, by finding vectors that satisfy the saddle-point condition. In the next subsection, we shall see how this can be understood through a geometric interpretation.

4.2 Geometric interpretation: shared normal vectors

In this section, unless otherwise stated, we operate under the following assumptions:

Assumption 4.1.

- $K \subset X$ is a bounded, closed and convex set containing 0. We define, as usual, $P = \text{cone}(K)$. Recall that P is not necessarily closed.
- $b \neq 0$.

¹For a subset C of a Hilbert space H , the quasi-relative interior $\text{qri}(C)$ is the set of vectors $x \in H$ such that $\overline{\text{cone}(C - x)}$ is a subspace of X , while the strong quasi-relative interior $\text{sqri}(C)$ is the set of vectors $x \in H$ such that $\text{cone}(C - x)$ is a closed subspace of X .

- the primal problem (\mathcal{P}_0) has a solution $x^* \neq 0$, i.e., $Ax^* = b$ and

$$\pi_0 = \inf_{Ax=b} \frac{1}{2}j_K^2(x) = \frac{1}{2}j_K^2(x^*) =: \frac{1}{2}\lambda^*.$$

Under the two first assumptions, and if $b \in A(P)$, then the last assumption is automatically satisfied as per Corollary 3.2.

Note that we thus have $A^{-1}(b) = x^* + \ker(A)$. Then, clearly, for any $x \in A^{-1}(b)$, we have $N_{A^{-1}(b)}(x) = (\text{Ker}(A))^\perp = \overline{\text{ran}(A^*)}$.

On the other hand, recall that by property of the support function, for any convex set $C \subset X$, and for any $v \in X$,

$$\partial\sigma_C(v) = \arg \max_{x \in C} \langle x, v \rangle_X.$$

Moreover, by definition of the normal cone, we have

$$v \in N_C(x) \iff x \in \partial\sigma_C(v). \quad (16)$$

We now establish a proposition that gives us a necessary, and sometimes sufficient, condition for the dual problem to reach its infimum. It relies on the notion of shared normal vectors.

Definition 4.1 (Shared normal vector). *Let $C_1, C_2 \subset X$ be closed convex sets and $x \in C_1 \cap C_2$. Given a vector $v \in X$, we say that C_1 and C_2 share the normal vector v at x if $v \in N_{C_1}(x) \cap N_{C_2}(x)$, i.e., the two sets admit a common supporting hyperplane at the contact point x with normal v .*

Proposition 4.1. *Under Assumption 4.1, the properties*

- (i) *the dual problem reaches its infimum at $y^* \in Y$,*
- (ii) *the convex sets $A^{-1}(b)$ and λ^*K share a nonzero normal vector $v \in X$ at their contact point x^* , such that the regularity condition holds:*

$$\langle v, x^* \rangle = \sigma_{\lambda^*K}(v) > 0, \quad (17)$$

satisfy (i) \implies (ii). Moreover, if $\text{ran}(A^)$ is closed, (i) \iff (ii).*

Proof. We first prove (i) \implies (ii). Assuming (i), recall the optimality condition (13)

$$x^* \in \sigma_K(A^*y^*) \partial\sigma_K(A^*y^*).$$

Strong duality (Proposition 2.1) implies

$$\frac{1}{2}(\lambda^*)^2 = \pi_0 = -\min_Y J_0 = -\frac{1}{2}\sigma_K(A^*y^*)^2 + \langle b, y^* \rangle = -\frac{1}{2}\sigma_K(A^*y^*)^2 + \langle Ax^*, y^* \rangle.$$

The optimality condition above clearly implies that $\langle x^*, A^*y^* \rangle = \sigma_K(A^*y^*)^2$ so that we have

$$\frac{1}{2}(\lambda^*)^2 = \frac{1}{2}\sigma_K(A^*y^*)^2,$$

i.e., since $\sigma_K(A^*y^*) \geq 0$ by property of K ,

$$0 < \lambda^* = \sigma_K(A^*y^*).$$

In particular, this yields the regularity condition (17).

We can then rewrite the optimality condition as

$$x^* \in \partial\sigma_{\lambda^*K}(A^*y^*),$$

which, thanks to (16), implies that $A^*y^* \in N_{\lambda^*K}(x^*)$. On the other hand, since $A^*y^* \in \text{ran}(A^*) \subset (\ker(A))^\perp = N_{A^{-1}b}(x^*)$, this proves (ii).

To prove the converse under the hypothesis that $\text{ran}(A^*)$ is closed, first note that $(\ker(A))^\perp = \overline{\text{ran}(A^*)} = \text{ran}(A^*)$. Thus, if λ^*K and $A^{-1}(b)$ share a normal vector v , then there exists $y \in Y$ such that $v = A^*y$, and we can write, by definition of $N_{\lambda^*K}(x^*)$,

$$x^* \in \partial\sigma_{\lambda^*K}(A^*y) = \lambda^*\partial\sigma_K(A^*y).$$

Moreover, by hypothesis, $\sigma_K(A^*y) = \frac{1}{\lambda^*}\langle A^*y, x^* \rangle_X > 0$. We can then define

$$y^* := \sqrt{\frac{\lambda^*}{\sigma_K(A^*y)}}y.$$

Then,

$$x^* \in \lambda^*\partial\sigma_K(A^*y) = \lambda^*\sqrt{\frac{\sigma_K(A^*y)}{\lambda^*}}\partial\sigma_K(A^*y^*) = \sigma_K(A^*y^*)\partial\sigma_K(A^*y^*).$$

Using the notations introduced when discussing Fenchel-Rockafellar duality, namely $F = \frac{1}{2}j_K^2$ and $G = \delta_{\{b\}}$, the above rewrites

$$x^* \in \partial(\frac{1}{2}\sigma_K^2)(A^*y^*) = \partial F^*(A^*y^*).$$

Hence

$$Ax^* = b \in A\partial F^*(A^*y^*),$$

which is equivalent to

$$0 \in A\partial F^*(A^*y^*) - \partial G^*(y^*) = \partial(F^* \circ A^* + G^*(-))(y^*) = \partial J_0(y^*).$$

Note that the above computation on subdifferentials is guaranteed by the sufficient stability condition used to establish strong duality [22], namely

$$A^* \text{dom } G^*(-) \cap \text{cont}(F^*) \neq \emptyset.$$

Hence, the dual problem reaches its infimum at y^* . □

4.3 On the existence of shared normal vectors

Given the above section, we now focus on the existence of shared normal vectors between the sets λ^*K and $A^{-1}b$. We begin with a counterexample in infinite dimension.

A counterexample in infinite dimension: convex cone (not closed), no shared normal vectors. We here here heavily draw from [1, Example 6.11 iv].

Let $X = \ell^2(\mathbb{N})$, let $\{e_n\}_{n \in \mathbb{N}}$ be a Hilbert basis of X . We define the following convex, closed, bounded set:

$$C := \left\{ x = \sum_{n \in \mathbb{N}} x_n e_n, \quad -2^{-n} \leq x_n \leq 4^{-n}, \quad \forall n \in \mathbb{N}. \right\}$$

This set generates a dense cone (see [1]), but not the whole space. Define

$$x^* := \sum_{n \in \mathbb{N}} 2^{-n} e_n.$$

Then, $-x^* \in C$ but $x^* \notin C$. Define the following generating set

$$K := C + x^* = \left\{ x = \sum_{n \in \mathbb{N}} x_n e_n, \quad 0 \leq x_n \leq 4^{-n} + 2^{-n}, \quad \forall n \in \mathbb{N} \right\},$$

which is clearly convex, bounded, closed, and contains 0. Moreover, $x^* \in K$, $j_K(x^*) = 1$.

Let $\mathcal{I} \subset \mathbb{N}$ such that $\text{card}(\mathcal{I}) < +\infty$. Then, for any $A \in L(X, Y)$ such that $\ker(A) = \text{Span}\{e_n, n \in \mathcal{I}\}$,

$$\inf_{x \in x^* + \ker(A)} \frac{1}{2} j_K^2(x) = \frac{1}{2} j_K^2(x^*) = \frac{1}{2}.$$

Hence, K , A , and x^* satisfy Assumption 4.1 with $\lambda^* = 1$.

Now, by density of $\text{cone}(C)$,

$$(K - x^*)^\circ = C^\circ = \{0\},$$

i.e.,

$$N_K(x^*) = \{0\}.$$

Hence,

$$N_K(x^*) \cap N_{A^{-1}(b)}(x^*) = \{0\},$$

i.e., there are no nontrivial shared normal vectors, even though x^* minimises $\frac{1}{2} j_K^2$ on $A^{-1}(b)$. By Proposition 4.1, the associated dual problem does not have a minimiser even though its infimum is finite.

An important partial result on the existence of shared normal vectors. The above counterexample precludes any general result on shared normal vectors in infinite dimension. However, in finite dimension, we are able to prove a positive result:

Lemma 4.3. *Suppose that X has finite dimension. Then, under Assumption 4.1, λ^*K and $A^{-1}b$ share a nonzero normal vector in x^* .*

Proof. We proceed by contradiction. Suppose that λ^*K and $A^{-1}b$ do not share a nonzero normal vector in x^* :

$$N_{A^{-1}(b)}(x^*) \cap N_{\lambda^*K}(x^*) = \{0\},$$

i.e.,

$$(\lambda^*K - x^*)^\circ \cap \text{ran}(A^*) = \{0\}.$$

Taking the cones generated by the sets involved above, we arrive at

$$\text{cone}((\lambda^*K - x^*)^\circ) \cap \text{ran}(A^*) = \{0\}.$$

By the elementary result (9), we then have

$$(\text{cone}(\lambda^*K - x^*) + \ker(A))^\circ = \{0\},$$

i.e.,

$$\overline{\text{cone}(\lambda^*K - x^*) + \ker(A)} = X.$$

In finite dimension, it is standard that a dense convex set is the whole space. Indeed, if it were not the whole space, it would be included in a closed half-space [21][Corollary 11.5.2], and such a half-space cannot be dense.

Hence, we obtain that for any $\nu > 0$, $\nu x^* \in \text{cone}(\lambda^*K - x^*) + \ker(A)$. That is, there exist $\rho \geq 0$, $k \in K$, $v \in \ker(A)$ such that

$$\nu x^* = \rho(\lambda^*k - x^*) + v,$$

i.e.,

$$(\nu + \rho)x^* = \rho\lambda^*k + v,$$

i.e.,

$$x^* + v' = \frac{\rho}{\nu + \rho} \lambda^*k, \quad v' \in \ker(A),$$

which implies

$$A(x^* + v') = b, \quad j_K(x^* + v') = \frac{\rho}{\rho + \nu} \lambda^* < \lambda^* = \inf_{Ax=b} j_K(x),$$

we thus arrive at a contradiction. □

4.4 On the regularity condition

Notice that, in Proposition 4.1, the condition $\langle v, x^* \rangle > 0$ for the shared normal vector v might not seem very natural in the context of a geometric interpretation. We now show a simple counterexample, illustrating how the unfavorable situation $\langle v, x^* \rangle_X = 0$ can in fact occur even in simple settings.

Shared normal vectors, no minimiser for the dual functional (closed cone) Take $X = \mathbb{R}^3$, $Y = \mathbb{R}^2$, $\alpha > 0$, and let P be the convex closed cone given by

$$P := \{(x_1, x_2, x_3) \in \mathbb{R}^3, \quad x_3 \geq 0, \quad x_1^2 + x_2^2 \leq \alpha^2 x_3^2\},$$

and denote as usual $K := P \cap \overline{B(0, 1)}$. One easily checks that

$$P^\circ = \{(x_1, x_2, x_3) \in \mathbb{R}^3, \quad x_3 \leq 0, \quad x_1^2 + x_2^2 \leq \frac{1}{\alpha^2} x_3^2\}.$$

Now define

$$A := \begin{pmatrix} 1 & 0 & -\alpha \\ \alpha & 1 + \alpha^2 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

One easily checks that

$$x^* := \left(\frac{\alpha}{1+\alpha^2}, 0, \frac{1}{1+\alpha^2}\right) \in P, \quad Ax^* = b.$$

Moreover, given the set of solutions

$$A^{-1}(b) = x^* + \ker(A) = \left\{ \left(\frac{\alpha}{1+\alpha^2} - \alpha s, s, \frac{1}{1+\alpha^2} - s\right), s \in \mathbb{R} \right\},$$

we have

$$A^{-1}(b) \cap P = \{x^*\}.$$

Hence, x^* is also a (in fact, the) solution to the primal problem (11).

Let us here denote $\lambda^* = \|x^*\|_X$. We have on the one hand

$$\text{ran}(A^*) = \text{Span}\{(1, 0, -\alpha); (\alpha, 1 + \alpha^2, 1)\},$$

and on the other hand, using Lemma A.4,

$$N_{\lambda^* K}(x^*) = \text{cone}(\{x^*\}) + P^\circ \cap \{x^*\}^\perp.$$

Now, it is clear that

$$P^\circ \cap \{x^*\}^\perp = \text{cone}(\{(1, 0, -\alpha)\}) \subset \text{ran}(A^*).$$

Moreover, $x^* \notin \text{ran}(A^*)$, so that finally

$$N_{\lambda^* K}(x^*) \cap \text{ran}(A^*) = P^\circ \cap \{x^*\}^\perp.$$

In particular, all shared normal vectors of $A^{-1}(b)$ and $\lambda^* K$ are normal to x^* , which by Proposition 4.1 implies that the dual problem associated to b does not have a minimiser, even though its infimum is finite.

A sufficient condition. In the above counterexample, $\lambda^* K$ and $\text{ran}(A^*)$ share nontrivial normal vectors, but they are all orthogonal to x^* . We now give a sufficient condition under which this type of situation cannot arise. Its proof is based on a separation result between two well-chosen convex sets.

Proposition 4.2. *We work under Assumption 4.1, and assume furthermore that P is closed, that $K = P \cap \overline{B(0, 1)}$, and that $\text{ran}(A^*)$ is closed.*

If for all $x \in P$, $\text{Span}(\{x\}) + P$ is closed, then the dual problem always admits a minimiser.

Proof. Since $\text{ran}(A^*)$ is closed, by Proposition 4.1, the dual problem admits a minimiser if and only if there exists a vector $v \in N_{\lambda^*K}(x^*) \cap \text{ran}(A^*)$ such that $\langle v, x^* \rangle_X > 0$. Proceed by contradiction: suppose

$$\text{ran}(A^*) \cap \{v \in N_{\lambda^*K}(x^*), \langle v, x^* \rangle > 0\} = \emptyset.$$

Since, for any $v \in N_{\lambda^*K}(x^*)$, $\langle v, x^* \rangle_X \geq 0$, this also writes

$$N_{\lambda^*K}(x^*) \cap \text{ran}(A^*) = N_{\lambda^*K}(x^*) \cap N_{A^{-1}(b)}(x^*) \subset \{x^*\}^\perp. \quad (18)$$

Recall now that, by Lemma A.4, since $\|x^*\|_X = \lambda^*$,

$$N_{\lambda^*K}(x^*) = \text{cone}(\{x^*\}) + P^\circ \cap \{x^*\}^\perp$$

so

$$\{v \in N_{\lambda^*K}(x^*), \langle v, x^* \rangle > 0\} = \{\rho x^* + v, \rho > 0, v \in P^\circ \cap \{x^*\}^\perp\},$$

and, by linearity, our assumption implies

$$\text{ran}(A^*) \cap (x^* + \overline{B(0,1)} \cap P^\circ \cap \{x^*\}^\perp) = \emptyset.$$

Now since the sets in the above intersection are convex and closed, and $x^* + \overline{B(0,1)} \cap P^\circ \cap \{x^*\}^\perp$ is (weakly) compact, we can use a strict separation theorem: there exists $h \in X$, $h \neq 0$ and $\alpha > 0$ such that

$$\forall v \in \overline{B(0,1)} \cap P^\circ \cap \{x^*\}^\perp, \forall y \in Y, \quad \langle h, x^* + v \rangle_X \leq \langle h, A^*y \rangle_X - \alpha. \quad (19)$$

For the above inequality to hold, one must have $h \in \text{ran}(A^*)^\perp = \ker(A)$, which leads to

$$\forall v \in \overline{B(0,1)} \cap P^\circ \cap \{x^*\}^\perp, \quad \langle h, x^* + v \rangle_X \leq -\alpha.$$

Since $\text{cone}(x^* + \overline{B(0,1)} \cap P^\circ \cap \{x^*\}^\perp) \subset N_{\lambda^*K}(x^*)$, this implies that $h \in N_{\lambda^*K}(x^*)^\circ$.

Now, using in order Lemma A.4 and the identities (9) (since P and $\{x^*\}^\perp$ are closed cones),

$$N_{\lambda^*K}(x^*)^\circ = (\text{cone}(\{x^*\}) + P^\circ \cap \{x^*\}^\perp)^\circ = \text{cone}(\{x^*\})^\circ \cap (P^\circ \cap \{x^*\}^\perp)^\circ = \text{cone}(\{x^*\})^\circ \cap \overline{\text{Span}(\{x^*\}) + P}.$$

By hypothesis, $\text{Span}(\{x^*\}) + P$ is closed, so we obtain

$$h \in \ker(A) \cap \text{cone}(\{x^*\})^\circ \cap (\text{Span}(\{x^*\}) + P).$$

Now, without loss of generality, we write $h = \rho x^* + x_0$, with $\rho \in [-1, 1]$ and $x_0 \in P$. Then, for $s \in [0, 1]$, $1 + s\rho \geq 0$ and

$$x^* + sh = (1 + s\rho)x^* + sx_0 \in P \cap A^{-1}(b),$$

since $P + P \subset P$, by definition of a convex cone. Moreover, considering $f : s \mapsto \|x^* + sh\|_X^2$, we find that $f'(0) = 2\langle x^*, h \rangle_X \leq -2\alpha < 0$, from (19). Hence, for s small enough, $\|x^* + sh\|_X^2 \leq \|x^*\|_X^2$, and $x^* + sh \in P \cap A^{-1}(b)$, which contradicts the definition of x^* . \square

A Elementary results in convex analysis

We here gather several basic results used throughout the work. We do not claim any originality, but provide them here for completeness and readability.

Throughout this subsection, we let H be a Hilbert space and C be a non-empty, bounded, closed and convex set containing 0.

Lemma A.1. *For any $u \in H$, one has*

$$(i) \quad j_C(u) = 0 \iff u = 0,$$

- (ii) for any $\alpha > j_C(u)$, $u \in \alpha C$,
- (iii) for any $\alpha > 0$, $j_C(u) \leq \alpha \iff u \in \alpha C$,
- (iv) $u \in \text{cone}(C) \iff u \in j_C(u)C$.

Proof. If $j_C(u) = 0$, then we may find (ε_n) with $\varepsilon_n > 0$, $\varepsilon_n \rightarrow 0$ and $(c_n) \in C^{\mathbb{N}}$ such that for all $n \in \mathbb{N}$, $u = \varepsilon_n c_n$. Since C is bounded, we uncover $u = 0$ at the limit $n \rightarrow +\infty$. The converse assumption is trivial since $0 \in C$, so (i) is proved.

Since $\alpha > j_C(u)$, we may find $j_C(u) < \beta \leq \alpha$ such that $u \in \beta C$. Hence there exists $c \in C$ such that $u = \beta c = \alpha(\frac{\beta}{\alpha}c) = \alpha(\frac{\beta}{\alpha}c + (1 - \frac{\beta}{\alpha})0)$. The vector $\frac{\beta}{\alpha}c$ appears as the convex combination of $c \in C$ and $0 \in C$, which shows that $u \in \alpha C$ and proves (ii).

Let $\alpha > 0$ and $u \in H$. If $u \in \alpha C$, then $j_C(u) \leq \alpha$ by definition. Conversely, assume $j_C(u) \leq \alpha$. We may find a sequence (α_n) with $\alpha_n > \alpha$ converging to α . Since $\alpha_n > j_C(u)$, (ii) ensures the existence of $c_n \in C$ such that $u = \alpha_n c_n \iff c_n = \alpha_n^{-1}u$. Passing to the limit $n \rightarrow +\infty$ (since $\alpha > 0$), we find that (c_n) converges to $c := \alpha^{-1}u$. By closedness of C , $c \in C$, showing that $u \in \alpha C$ and finishing the proof of (iii).

Let $u \in H$. If $u = 0$, the equivalence is clear: $0 \in \text{cone}(C)$ by the assumption $0 \in C$, and since $j_C(u) = 0$, any $v \in C$ is such that $0 = u = j_C(u)v = 0$. Now assume that $u \neq 0$. By (i), we know that $j_C(u) \neq 0$, hence if $u = j_C(u)v$ with $v \in C$, we have $u \in \text{cone}(C)$. Conversely, if $u \in \text{cone}(C)$, we use (iii) with $\alpha = j_C(u) > 0$, showing that $u \in \alpha C = j_C(u)C$, as wanted. \square

Lemma A.2. *There hold*

- (i) $j_C \in \Gamma_0(H)$,
- (ii) $\frac{1}{2}\sigma_C^2 \in \Gamma_0(H)$, $\frac{1}{2}j_C^2 \in \Gamma_0(H)$ and $(\frac{1}{2}j_C^2)^* = \frac{1}{2}\sigma_C^2$.

Proof. Since $j_C \geq 0$ and $j_C(0) = 0$, j_C is proper. The convexity of j_C is standard, see for instance [1] (Example 8.27). Let us prove that j_C is lower semicontinuous. Given $\alpha \geq 0$, we shall establish that $\{x \in H, j_C(x) \leq \alpha\}$, is closed. let (x_n) be a sequence taking values in the previous set, converging to $x \in H$. By Lemma A.1 (iv), we may write $x_n = j_C(x_n)v_n$, with $(v_n) \in C^{\mathbb{N}}$. Upon extracting a subsequence, we may assume that (v_n) converges weakly to some $v \in C$, by sequential weak compactness of C , and that the bounded sequence $(j_C(x_n))$ converges to some $\lambda \geq 0$ with $\lambda \leq \alpha$. Hence, (x_n) converges weakly to λv , which by uniqueness of the weak limit entails $x = \lambda v$, showing that $j_C(x) = \lambda \leq \alpha$, as wanted.

Let us now address (ii). We rather show that $(\frac{1}{2}\sigma_C^2)^* = \frac{1}{2}j_C^2$, which will prove the identity by Fenchel-Moreau's theorem. Since $0 \in C$, $\sigma_C \geq 0$. If $f \in \Gamma_0(H)$, $f \geq 0$, then $f^2 \in \Gamma_0(H)$. Hence $\frac{1}{2}\sigma_C^2 \in \Gamma_0(H)$ and $\frac{1}{2}j_C^2 \in \Gamma_0(H)$.

Now let us prove the equality. For $x = 0$, it holds true since $j_C(0) = 0$ and $(\frac{1}{2}\sigma_C^2)^*(0) = -\frac{1}{2} \inf_{y \in H} \sigma_C^2(y) = 0$.

We now assume that $x \neq 0$. For $f \in \Gamma_0(H)$ and $g \in \Gamma_0(\mathbb{R})$ non-decreasing, we recall the composition formula

$$(g \circ f)^*(x) = \inf_{\alpha \geq 0} \left(g^*(\alpha) + \alpha f^*\left(\frac{x}{\alpha}\right) \right),$$

with the convention that for $\alpha = 0$, $0 f^*\left(\frac{x}{\alpha}\right) = \sigma_{\text{dom}(f)}(y)$. Hence for $x \in H \setminus \{0\}$, one finds

$$\begin{aligned} \left(\frac{1}{2}\sigma_C^2\right)^*(x) &= \inf_{\alpha > 0} \left(\frac{1}{2}\alpha^2 + \alpha \delta_C\left(\frac{x}{\alpha}\right) \right) = \inf_{\alpha > 0} \left(\frac{1}{2}\alpha^2 + \alpha \delta_C\left(\frac{x}{\alpha}\right) \right) \\ &= \frac{1}{2} \left(\inf_{\alpha > 0, x \in \alpha C} \alpha \right)^2 = \frac{1}{2}j_C^2(x), \end{aligned}$$

where we discarded $\alpha = 0$ since

$$0 \delta_C\left(\frac{x}{0}\right) = \sigma_{\text{dom}(\sigma_C)}(x) = \sigma_H(x) = \delta_{\{0\}}(x) = +\infty,$$

using that $\text{dom}(\sigma_C) = H$, by boundedness of C . \square

Lemma A.3. *If 0 is in the interior of C relative to the cone it generates, i.e., if*

$$\exists \delta > 0, \quad \text{cone}(C) \cap \overline{B(0, \delta)} \subset C,$$

then $\text{cone}(C)$ is closed.

Proof. Let $(p_n) \in \text{cone}(C)^\mathbb{N}$ be a convergent subsequence, of limit $p \in H$. By Lemma A.1(iv), we may write $p_n = \lambda_n c_n$ with $(c_n) \in C^\mathbb{N}$ and $\lambda_n = j_C(p_n)$. If $p = 0$, there is nothing to prove, so we may assume $p \neq 0$. Since (c_n) is bounded by hypothesis, and (p_n) converges to $p \neq 0$, $\liminf \lambda_n > 0$. Hence, if (λ_n) is upper bounded, then upon extraction we may write $\lambda_n \rightarrow \lambda$ with $\lambda \neq 0$. In this case, we find that (c_n) converges to $c := \frac{p}{\lambda}$. Since C is closed, we have $c \in C$ and hence $p = \lambda c \in \text{cone}(C)$.

To conclude, we only need to show that (λ_n) cannot have a diverging subsequence. By contradiction, assume that it is the case: upon extraction, we may assume that $\lambda_n \rightarrow +\infty$ as $n \rightarrow +\infty$. Then we may form the sequence $w_n = \lambda_n^{-1/2} p_n = \lambda_n^{1/2} c_n$. This sequence satisfies $w_n \rightarrow 0$ as $n \rightarrow +\infty$ as well as $(w_n) \in \text{cone}(C)^\mathbb{N}$. We also compute $j_C(w_n) = \lambda_n^{-1/2} j_C(p_n) = \lambda_n^{1/2} \rightarrow +\infty$ as $n \rightarrow +\infty$. Hence for n large enough we have both $w_n \in \text{cone}(C) \cap \overline{B(0, \delta)}$ and $w_n \notin C$, contradicting the assumption that 0 is the interior of C relative to $\text{cone}(C)$. \square

Finally, we prove a technical Lemma which will help us characterise various situations with a **closed** cone P .

Lemma A.4. *Let P be a closed convex cone, and $K := P \cap \overline{B(0, 1)}$. For $\lambda > 0$ and $x \in \lambda K$, we have the following characterisation for the set of normal vectors $N_{\lambda K}(x)$:*

$$N_{\lambda K}(x) = \begin{cases} \text{cone}(\{x\}) + P^\circ \cap \{x\}^\perp & \text{if } j_K(x) = \lambda, \\ P^\circ \cap \{x\}^\perp & \text{otherwise.} \end{cases}$$

Proof. Recall that, using the definition of σ_K and Lemma 3.4,

$$\begin{aligned} v \in N_{\lambda K}(x) &\iff x \in \partial \sigma_{\lambda K}(v) \\ &\iff \lambda \sigma_K(v) = \langle x, v \rangle_X \\ &\iff \lambda \|v_+\|_X = \langle v_+, x \rangle_X + \langle v_-, x \rangle_X. \end{aligned}$$

Now, for $v \in X$, $\langle v_-, x \rangle_X \leq 0$ because $x \in P$, $v_- \in P^\circ$. Moreover, if $x \in \partial \sigma_{\lambda K}(v)$, then $\|x\|_X \leq \lambda$ since $K \subset \overline{B(0, 1)}$, and

$$\lambda \|v_+\|_X = \sigma_{\lambda K}(v) = \langle v, x \rangle_X = \langle v_+, x \rangle_X + \langle v_-, x \rangle_X \leq \lambda \|v_+\|_X + \langle v_-, x \rangle_X \leq \lambda \|v_+\|_X.$$

Hence, the inequalities above are equalities, which yields

$$v \in N_{\lambda K}(x) \iff \langle v_-, x \rangle_X = 0, \quad \lambda \|v_+\|_X = \langle v_+, x \rangle_X.$$

Now, $j_K(x) \leq \lambda$. If this inequality is strict, then $\lambda \|v_+\|_X = \langle v_+, x \rangle_X \iff v_+ = 0$. Otherwise, $\|x\|_X = j_K(x) = \lambda$ and, since $v_+ \in P$, $\lambda \|v_+\|_X = \langle v_+, x \rangle_X \iff \exists \rho \geq 0, v = \rho x$, which concludes the proof. \square

References

- [1] Heinz Bauschke and Patrick Louis Combettes. Convex analysis and monotone operator theory in Hilbert spaces, 2011. *CMS books in mathematics*, 10:978–1.
- [2] Jonathan M Borwein and Rafal Goebel. Notions of relative interior in Banach spaces. *Journal of Mathematical Sciences*, 115(4), 2003.

- [3] Jonathan M Borwein, Jon D Vanderwerff, et al. *Convex functions: constructions, characterizations and counterexamples*, volume 172. Cambridge University Press Cambridge, 2010.
- [4] Radu Ioan Boț and Ernő Robert Csetnek. A comparison of some recent regularity conditions for Fenchel duality. *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, pages 111–130, 2011.
- [5] Radu Ioan Boț and Ernő Robert Csetnek. Regularity conditions via generalized interiority notions in convex optimization: new achievements and their relation to some classical statements. *Optimization*, 61(1):35–65, 2012.
- [6] Radu Ioan Boț, Ernő Robert Csetnek, and Andreea Moldovan. Revisiting some duality theorems via the quasirelative interior in convex optimization. *Journal of Optimization Theory and Applications*, 139:67–84, 2008.
- [7] Radu Ioan Boț, Ioan Bogdan Hodrea, and Gert Wanka. Farkas-type results for inequality systems with composed convex functions via conjugate duality. *Journal of Mathematical Analysis and Applications*, 322(1):316–328, 2006.
- [8] Radu Ioan Boț, Ioan Bogdan Hodrea, and Gert Wanka. Some new Farkas-type results for inequality systems with DC functions. *Journal of Global Optimization*, 39(4):595–608, 2007.
- [9] Radu Ioan Boț and Gert Wanka. Farkas-type results with conjugate functions. *SIAM Journal on Optimization*, 15(2):540–554, 2005.
- [10] Antonin Chambolle and Thomas Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal of Mathematical Imaging and Vision*, 40, 05 2011.
- [11] Frank H Clarke. *Optimization and nonsmooth analysis*. SIAM, 1990.
- [12] Patrick L Combettes and Noli N Reyes. Moreau’s decomposition in Banach spaces. *Mathematical Programming*, 139(1):103–114, 2013.
- [13] Bruce Desmond Craven and Jaromir Joseph Koliha. Generalizations of Farkas’ theorem. *SIAM Journal on Mathematical Analysis*, 8(6):983–997, 1977.
- [14] Nguyen Dinh and Vintiya Jeyakumar. Farkas’ lemma: three decades of generalizations for mathematical optimization. *Top*, 22(1):1–22, 2014.
- [15] Julius Farkas. Theorie der einfachen Ungleichungen. *Journal für die reine und angewandte Mathematik (Crelles Journal)*, 1902(124):1–27, 1902.
- [16] Muddappa Seetharama Gowda and Marc Teboulle. A comparison of constraint qualifications in infinite-dimensional convex programming. *SIAM Journal on Control and Optimization*, 28(4):925–935, 1990.
- [17] Henry Hermes and Joseph P. LaSalle. *Functional Analysis and Time Optimal Control*, volume 56 of *Mathematics in Science and Engineering*. Academic Press, New York, 1969.
- [18] Onesimo Hernandez-Lerma and Jean Bernard Lasserre. Cone-constrained linear equations in Banach spaces. *Journal of Convex Analysis*, 4:149–164, 1997.
- [19] Jean Bernard Lasserre. A Farkas lemma without a standard closure condition. *SIAM journal on control and optimization*, 35(1):265–272, 1997.
- [20] Xungjing Li and Jiongmin Yong. *Optimal control theory for infinite dimensional systems*. Springer Science & Business Media, 2012.

- [21] R. Tyrrell Rockafellar. *Convex Analysis*, volume 28 of *Princeton Mathematical Series*. Princeton University Press, Princeton, NJ, 1970.
- [22] Ralph Rockafellar. Duality and stability in extremum problems involving convex functions. *Pacific Journal of Mathematics*, 21(1):167–187, 1967.
- [23] Walter Rudin. *Functional Analysis*. International series in pure and applied mathematics. McGraw-Hill, 1991.
- [24] Constantin Zălinescu. On zero duality gap and the Farkas lemma for conic programming. *Mathematics of Operations Research*, 33(4):991–1001, 2008.
- [25] Constantin Zălinescu. On duality gap in linear conic problems. *Optimization Letters*, 6(3):393–402, 2012.