THE S-PROCEDURE VIA DUAL CONE CALCULUS

RAPHAEL HAUSER*

Abstract. Given a quadratic function h that satisfies a Slater condition, Yakubovich's S-Procedure (or S-Lemma) gives a characterization of all other quadratic functions that are copositive with h in a form that is amenable to numerical computations. In this paper we present a deep-rooted connection between the S-Procedure and the dual cone calculus formula $(K_1 \cap K_2)^* = K_1^* + K_2^*$, which holds for closed convex cones in $\mathbb{R}^2$. To establish the link with the S-Procedure, we generalize the dual cone calculus formula to a situation where K_1 is nonclosed, nonconvex and nonconic but exhibits sufficient mathematical resemblance to a closed convex cone. As a result, we obtain a new proof of the S-Lemma and an extension to Hilbert space kernels.

AMS subject classifications. Primary 90C20, 90C22. Secondary 49M20.

Key words. S-Lemma, S-Procedure, optimal control, robust optimization, convex separation.

1. Introduction.

Yakubovich's *S-Lemma* [9], also called *S-Procedure*, is a well-known result from robust control theory that characterizes all quadratic functions that are copositive with a given other quadratic function. A function g is called *copositive with* h if $h(x) \geq 0$ implies $g(x) \geq 0$.

THEOREM 1.1 (S-Lemma, [9]). *Let $g, h : \mathbb{R}^n \to \mathbb{R}$ be quadratic functions such that $h(x_0) > 0$ at some point $x_0 \in \mathbb{R}^n$. Then g is copositive with h if and only if there exists $\xi \geq 0$ such that $g(x) - \xi h(x) \geq 0$ for all $x \in \mathbb{R}^n$.*

Note that g and h are neither assumed to be convex nor homogeneous, and that the condition $g(x) - \xi h(x) \geq 0$ for all $x \in \mathbb{R}^n$ is easy to check, for a quadratic function $x \mapsto x^{\mathrm{T}}Qx + 2\ell^{\mathrm{T}}x + c$ can always be formulated so that the matrix Q is symmetric, and then the function is nonnegative everywhere on $\mathbb{R}^n$ if and only if the matrix $\left[\begin{smallmatrix} Q & \ell \\ \ell^{\mathrm{T}} & c \end{smallmatrix}\right]$ is positive semidefinite. The importance of this characterization is that it can be checked numerically.

Theorem 1.1 arose as a generalization of earlier results by Finsler [4], Hestenes & McShane [5] and Dines [3]. Megretsky & Treil [6] later extended the result further. The S-Lemma has suprisingly powerful consequences in robust optimization and control theory, as this result allows to replace certain nonconvex optimization problems by convex polynomial time solvable ones, and semi-infinite programming problems by optimization models with finitely many constraints. Indeed, Theorem 1.1 says that in an optimization problem in which the coefficients Q, ℓ, c of the polynomial g play the role of decision variables, the infinitely many constraints

$$g(x) \geq 0, \quad \forall x \in \mathbb{R}^n \text{ s.t. } h(x) \geq 0$$

can be replaced by a single matrix inequality

$$\begin{bmatrix} Q & \ell \\ \ell^{\mathrm{T}} & c \end{bmatrix} - \xi \begin{bmatrix} A & b \\ b^{\mathrm{T}} & d \end{bmatrix} \succeq 0,$$

*Mathematical Institute, University of Oxford, 24–29 St Giles', Oxford, OX1 3LB, United Kingdom, (hauser@maths.ox.ac.uk), and Pembroke College Oxford. The author was supported through grants GR/S34472 and EP/H02686X/1 from the Engineering and Physical Sciences Research Council of the UK.

where A, b, d are chosen such that $h(x) = x^{\mathrm{T}}Ax + 2b^{\mathrm{T}}x + d$, and where ξ is an auxiliary decision variable introduced by this lifting.

For a overviews of the history of the S-Lemma and its applications, see [7] and [2]. Three existing known approaches to proving Theorem 1.1 described in [7] are due to Yakubovich [9], Ben-Tal & Nemirovski [1] and Sturm & Zhang [8], and Yuan [10].

In this paper we give a new proof of the S-Lemma that is based on a generalization of the dual cone calculus formula $(K_1 \cap K_2)^* = K_1^* + K_2^*$, which is known to hold true for closed convex cones $K_1, K_2 \subseteq \mathbb{R}^2$, to a situation where K_1 is nonclosed, nonconvex and nonconic but exhibits sufficient mathematical resemblance to a closed convex cone. For this purpose we introduce a weak notion of convexity, *homogenization-convexity*, the theory of which will be developed in Section 2. Our proof extends quite straighforwardly to an S-Lemma for Hilbert space kernels. The techniques we employ are elementary. The main ideas of the proof merely require linear algebra in two dimensions. The S-Lemma and its extension to Hilbert space kernels are then obtained by a lifting.

Among the existing proofs of the S-Lemma, Yakubovich's orginal proof is closest in spirit to the proof presented in this paper. Yakubovich employed a result of Dines [3], which shows that the joint range $\{(f(x), g(x)) : x \in \mathbb{R}^n\}$ of two homogeneous quadratic functions f, g on $\mathbb{R}^n$ is convex. Our own approach is based on showing that the projection of the set

$$\left\{ \left[\begin{smallmatrix} x \\ 1 \end{smallmatrix}\right] \left[\begin{smallmatrix} x \\ 1 \end{smallmatrix}\right]^{\mathrm{T}} : x \in \mathbb{R}^n \right\}$$

into a 2-dimensional subspace satisfies the weaker notion of homogenization-convexity. Once this is established, the S-Lemma follows from our generalized dual cone calculus formula.

1.1. Notation. The inner product on any Hilbert space V is denoted by $\langle \cdot, \cdot \rangle$. This inner product defines the canonical self-duality isomorphism on V and the canonical norm $\|\cdot\|$. The topological closure and boundary of a set $C \subseteq V$ under the induced topology are denoted by $\mathrm{clo}[C]$ and ∂C. The convex, conic and homogeneous hulls of C are denoted by

$$\begin{aligned} \mathrm{conv}(C) &:= \left\{ \sum_{i=1}^{n} \lambda_i x_i : n \in \mathbb{N},\, \lambda_i \geq 0,\, x_i \in C,\, \forall i,\, \sum_{i=1}^{n} \lambda_i = 1 \right\}, \\ \mathrm{cone}(C) &:= \left\{ \sum_{i=1}^{n} \lambda_i x_i : n \in \mathbb{N},\, \lambda_i \geq 0,\, x_i \in C,\, \forall i \right\}, \\ \mathrm{hom}(C) &:= \{\tau x : \tau \geq 0,\, x \in C\}. \end{aligned}$$

The relation between these three concepts is that $\mathrm{cone}(C) = \mathrm{hom}(\mathrm{conv}(C))$.

Definition 1.2. *For any $C \subseteq V$ we refer to the set* $\mathrm{clo}[\mathrm{hom}(C)]$ *as the* homogenization *of C.*

We denote the unit sphere in $(V, \langle \cdot, \cdot \rangle)$ by $\mathbb{S}(V)$ and the spherical projection by

$$\begin{aligned} \mathfrak{q} : V \setminus \{0\} &\to \mathbb{S}(V) \\ x &\mapsto \frac{x}{\|x\|}. \end{aligned}$$

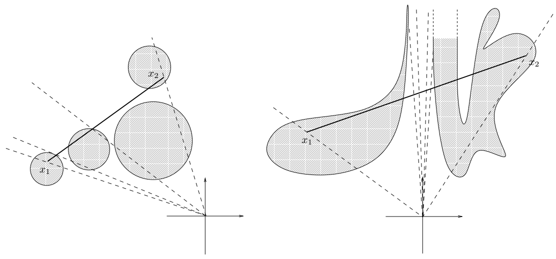

FIG. 2.1. *The shaded sets C are homogenization-convex. For every pair of points $x_1, x_2 \in C$, every point on the interval $[x_1, x_2]$ can be produced as a limit of positively scaled points in C. In the example on the right the two connected components of C go off to infinity in the same asymptotic direction.*

Note that the spherical projection is not defined at the origin of V. Nonetheless, by abuse of language, if $C \subseteq V$ we write $\mathfrak{q}(C)$ for $\mathfrak{q}(C \setminus \{0\})$. The set of *recession directions* of a set $C \subset V$ is given by

$$\mathrm{rec}(C) := \{s \in \mathbb{S}(V) : \forall \tau, \varepsilon > 0, \ \exists\, x \in C \text{ s.t. } \|s - \mathfrak{q}(x)\| < \varepsilon, \ \|x\| > \tau\}.$$

For any $x_1, x_2 \in V \setminus \{0\}$ we write

$$[x_1, x_2] = \{\xi x_2 + (1 - \xi)x_1 : \xi \in [0, 1]\}$$

for the straight-line segment between x_1 and x_2. For $y_1, y_2 \in \mathbb{S}(V)$, we write $[y_1, y_2] := \mathfrak{q}([x_1, x_2])$, where $x_1 \in \mathfrak{q}^{-1}(y_1)$ and $x_2 \in \mathfrak{q}^{-1}(y_2)$. It is easy to check that the definition of $[y_1, y_2]$ does not depend on the specific choice of x_1 and x_2. A subset $S \subset \mathbb{S}(V)$ is *spherically convex* if $[y_1, y_2] \subset S$ for all $y_1, y_2 \in S$.

2. Homogenization-Convexity.

Our approach to the S-Lemma hinges on a weak notion of convexity that we shall now define.

DEFINITION 2.1. *A set $C \subseteq \mathbb{R}^2$ is homogenization-convex if the homogenization* $\mathrm{clo}[\mathrm{hom}(C)]$ *of C is a convex subset of* $\mathbb{R}^2$.

A few alternative characterizations provide further insight:

LEMMA 2.2. *The following conditions on a set $C \subseteq \mathbb{R}^2$ are equivalent:*

i) *C is homogenization-convex.*
ii) $\mathrm{clo}[\mathfrak{q}(C)]$ *is spherically convex in* $\mathbb{S}(V)$
iii) $\mathrm{clo}\big[\mathrm{hom}(C)\big] = \mathrm{clo}\big[\mathrm{cone}(C)\big]$.
iv) $\mathfrak{q}([x_1, x_2]) \subset \mathrm{clo}[\mathfrak{q}(C)]$ *for all $x_1, x_2 \in C$,*

Proof. i)⇔ ii) ⇔ iii) follow immediately from $\mathfrak{q}^{-1}(\mathrm{clo}[\mathfrak{q}(C)]) = \mathrm{clo}[\mathrm{hom}(C)] \setminus \{0\}$ and from the characterization of $\mathrm{cone}(C)$ as the smallest convex set K such that $C \subseteq K$ and $\mathrm{hom}(K) = K$. ii)⇒ iv) follows from the definition of spherical convexity. iv) ⇒ ii): Let $y_1, y_2 \in \mathrm{clo}[\mathfrak{q}(C)]$ and $x_i \in \mathfrak{q}^{-1}(y_i)$ $(i = 1, 2)$. If $x_1 \sim \pm x_2$, then $[y_1, y_2] = \{y_1, y_2\} \subset \mathrm{clo}[\mathfrak{q}(C)]$. Otherwise, x_1 and x_2 are linearly independent, and for all $\lambda \in [0, 1]$, $\mathfrak{q}(\lambda y_1 + (1 - \lambda)y_2) = \lim_{n\to\infty} \mathfrak{q}\left(\lambda x_1^n + (1 - \lambda)x_2^n\right) \in \mathrm{clo}\left[\mathfrak{q}(C)\right]$ for some sequences $(x_i^n)_{\mathbb{N}} \subset C$ for which $\mathfrak{q}(x_i^n) \to y_i$, $(i = 1, 2)$. □

It follows from Lemma 2.2 iii) that if C is convex then C is homogenization-convex. The examples of Figure 2.1 illustrate that the reverse relationship is not true. See also Figure 2.2 for examples of sets that are *not* homogenization-convex. The following example is relevant to our proof of the S-Lemma:

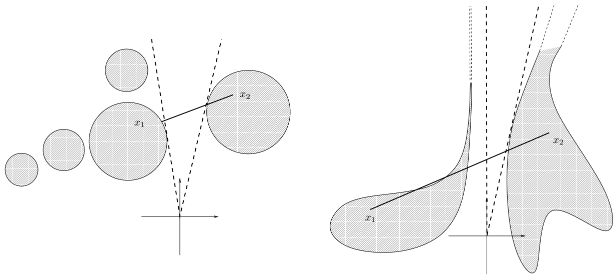

FIG. 2.2. *The shaded sets C are not homogenization-convex. The directions going through the part of $[x_1, x_2]$ cut out by the bold-face dotted lines is not covered by any point in* $\text{clo}[\mathfrak{q}(C)]$. *In the example on the right the two connected components of C go off to infinity in different asymptotic directions.*

EXAMPLE 2.3. *Let $x(t) = a_0 + ta_1 + t^2 a_2$ and $y(t) = b_0 + tb_1 + t^2 b_2$ for some real coefficients a_i, b_i $(i = 0, 1, 2)$. Then $C := \{[\,x(t)\ y(t)\,]^{\mathrm{T}} : t \in \mathbb{R}\}$ is homogenization-convex.*

Proof. When (a_2, a_1) and (b_2, b_1) are linearly dependent, then there exist $\eta_1, \eta_2 \in \mathbb{R}$, not both zero, such that $\eta_1 a_i + \eta_2 b_i = 0$ $(i = 1, 2)$, and then C is a subset of the line $\{z \in \mathbb{R}^2 : \eta_1 z_1 + \eta_2 z_2 = \eta_1 a_0 + \eta_2 b_0\}$. Since C is connected by arcs, it must be an interval, hence convex. This implies that C is homogenization-convex. In the case where (a_2, a_1) and (b_2, b_1) are linearly independent, there exist $\xi_1, \xi_2 \in \mathbb{R}$ such that $\xi_1(a_2, a_1) + \xi_2(b_2, b_1) = (0, 1)$, so that $\xi_1 x(t) + \xi_2 y(t) = t + c$ for some $c \in \mathbb{R}$. Furthermore, we may assume without loss of generality that $b_2 \neq 0$. The set of loci $\{(x, y) : x = x(t),\, y = y(t),\, t \in \mathbb{R}\}$ is then characterised by the equation

$$y = b_2 \left(\xi_1 x + \xi_2 y - c\right)^2 + b_1 \left(\xi_1 x + \xi_2 y - c\right) + b_0.$$

This is the general equation of a parabola. Hence, $C = \partial K$, where K is the set of points enclosed by the parabola. K being a convex set with unique recession direction (a_2, b_2), the homogenization-convexity of C is a special case of Example 2.5 below. ◻

EXAMPLE 2.4. *Let $C = \partial K$ where K is a closed convex subset of $\mathbb{R}^2$ with complement $K^c = \mathbb{R}^2 \setminus K$ and such that $0 \in \text{int}[K^c]$. Then C is homogenization-convex.*

Proof. Consider the map

$$\begin{aligned} \sigma : \mathbb{R}^2 &\to \mathbb{R} \cup \{+\infty\} \\ v &\mapsto \inf \{\tau \geq 0 : \tau v \in K\}, \end{aligned}$$

defined for all $v \in K$, where $\inf \emptyset := +\infty$ as usual. Choose arbitrary points $x_1, x_2 \in C$. If x_1, x_2 are linearly dependent, then $\mathfrak{q}([x_1, x_2]) \subseteq \{\mathfrak{q}(x_1), \mathfrak{q}(x_2)\} \subseteq \text{clo}[\mathfrak{q}(C)]$. Else x_1, x_2 are linearly independent, and for any point $x \in [x_1, x_2]$, we have $x \neq 0$, so that $\mathfrak{q}(x)$ is well defined. Since $x \in K$, we have $\sigma(x) \leq 1$, and since $0 \in \text{int}[K^c]$, $\sigma(x) > 0$. Furthermore, $\sigma(x)x \in \partial K = C$, so that $x = \sigma(x)^{-1}(\sigma(x)x) \in \text{hom}(C)$. Since x was chosen arbitrarily, this shows that $\mathfrak{q}([x_1, x_2]) \subset \text{clo}[\mathfrak{q}(C)]$, and the claim follows from Lemma 2.2 iv). ◻

EXAMPLE 2.5. *Let $C = \partial K$ where K is a closed convex subset of $\mathbb{R}^2$ with at most one recession direction. Then C is homogenization-convex.*

Proof. We may assume without loss of generality that $0 \in K$, for otherwise our claim is true by virtue of Example 2.4. Consider the map

$$\varsigma(v) := \sup\{\tau \geq 0 : \tau v \in K\},$$

defined for all $v \in K$. Then $\varsigma(\cdot)$ takes finite values on $K \setminus \left(\mathfrak{q}^{-1}(\mathrm{rec}(K)) \cup \{0\}\right)$. Since $\varsigma(v)v \in \partial K = C$ when $\varsigma(v)$ is finite, it follows that

$$\mathrm{hom}(K) \setminus \mathfrak{q}^{-1}(\mathrm{rec}(K)) \subseteq \mathrm{hom}(C) \subseteq \mathrm{hom}(K). \tag{2.1}$$

By assumption, $\mathrm{rec}(K)$ is either empty or a singleton. If $\dim(K) = 1$, then $C = K$. Otherwise, taking closures in (2.1) reveals that $\mathrm{clo}[\mathrm{hom}(K)] = \mathrm{clo}[\mathrm{hom}(C)]$, and by convexity of K, $\mathrm{cone}(K) \subseteq \mathrm{hom}(K)$. Therefore,

$$\mathrm{clo}[\mathrm{hom}(C)] = \mathrm{clo}[\mathrm{cone}(K)] \supseteq \mathrm{clo}[\mathrm{cone}(C)] \supseteq \mathrm{clo}[\mathrm{hom}(C)],$$

and the claim follows from Lemma 2.2 iii). □

2.1. Dual Cone Calculus. Any subset $C \subseteq \mathbb{R}^n$ is associated with a dual cone $C^* = \{y \in \mathbb{R}^n : \langle x, y\rangle \geq 0,\ \forall\, x \in C\}$. When K_1, K_2 are closed polyhedral cones, then the dual cone formula

$$(K_1 \cap K_2)^* = K_1^* + K_2^* \tag{2.2}$$

applies. In particular, this formula holds true for all closed cones $K_1, K_2 \subseteq \mathbb{R}^2$, since all cones in $\mathbb{R}^2$ are polyhedral. The following property of dual cones is also well known,

$$C^* = (\mathrm{clo}[\mathrm{cone}(C)])^*, \tag{2.3}$$

$$\tag{2.4}$$

In this section we set out to generalizing the relation (2.2) to the case where K_1 is merely a homogenization-convex set and K_2 is a closed convex cone with nonempty interior.

LEMMA 2.6. *Let $C \subseteq \mathbb{R}^2$ be homogenization-convex and $K \subseteq \mathbb{R}^2$ a closed convex cone with nonempty interior. Then*

$$\mathrm{clo}[\mathrm{cone}(C \cap K)] = \mathrm{clo}[\mathrm{cone}(C) \cap K]. \tag{2.5}$$

Proof. We only need to prove the inclusion $\supseteq$, since the reverse relation is trivial. Let $x \in \mathrm{cone}(C) \cap K \setminus \{0\}$. Then there exist $x_1, x_2 \in C$ and $\lambda_1, \lambda_2 \geq 0$ such that $x = \lambda_1 x_1 + \lambda_2 x_2$. If either λ_1 or λ_2 is zero or if $x_1, x_2 \in K$, then it is trivially true that $x \in \mathrm{cone}(C \cap K)$. Furthermore, if x_1, x_2 are linearly dependent, then $x = \tau x_i$ for some $\tau > 0$ and $i \in \{1, 2\}$, and by homogeneity of K, $x_i \in C \cap K$ and $x \in \mathrm{cone}(C \cap K)$. We may therefore assume that x_1, x_2 are linearly independent, $\lambda_1, \lambda_2 > 0$, and that $x_1 \notin K$.

Like all closed convex cones in $\mathbb{R}^2$, K is of the form $K = \{x : \phi_1(x) \geq 0,\ \phi_2(x) \geq 0\}$ for some linear forms $\phi_i : \mathbb{R}^2 \to \mathbb{R}$, $(i = 1, 2)$. We may furthermore assume that both are nonzero, as the case $\phi_1 = 0 = \phi_2$ is trivial, and the case $\phi_1 \neq 0 = \phi_2$ follows from a simplification of the argument we are about to give. Without loss of generality, we may assume that $\phi_1(x_1) < 0$. Since $0 \leq \phi_1(x) = \lambda_1\phi_1(x_1) + \lambda_2\phi_1(x_2)$, we then have $\phi_1(x_2) > 0$.

We first treat the case $\phi_2(x_1) \geq 0$. The linear independence of x_1 and x_2 implies that $y_1 := \xi x_1 + (1-\xi)x_2 \neq 0$, where $\xi = \phi_1(x_2)/(\phi_1(x_2) - \phi_1(x_1)) \in (0,1)$. By construction, $\phi_1(y_1) = 0$. The homogenization-convexity of C furhter implies $\mathfrak{q}(y_1) \in [\mathfrak{q}(x_1), \mathfrak{q}(x_2)] \subseteq \mathrm{clo}[\mathfrak{q}(C)]$. This shows the existence of a sequence $(y_1^n)_{n\in\mathbb{N}} \subset C$ such that $\phi_1(y_1^n) > 0$ and $\mathfrak{q}(y_1^n) \to \mathfrak{q}(y_1)$. Defining $\rho := \lambda_1/(\lambda_1+\lambda_2)$ and $z := \rho x_1 + (1-\rho)x_2$, we have $x = (\lambda_1+\lambda_2)z$ and $\phi_1(z) = (\lambda_1+\lambda_2)^{-1}\phi_1(x) \geq 0$. Since $\phi_1(y_1) = 0$, it must be the case that $\rho \leq \xi$, so that $\eta := \rho/\xi \in (0,1]$, and furthermore, $z = \eta y_1 + (1-\eta)x_2$. Since $\phi_2(x_1), \phi_2(z) \geq 0$ and $y_1 \in [x_1, z]$, we also have $\phi_2(y_1) \geq 0$, so that $y_1 \in K$. Since K has nonempty interior and $y_1 \neq 0$, we have $y_1^n \in C \cap K$ for all $n \gg 1$, and without loss of generality, we may assume that this holds for all $n \in \mathbb{N}$. Next, if $\phi_2(x_2) \geq 0$, set $y_2 = x_2$ and $y_2^n = x_2$ for all $n \in \mathbb{N}$. Otherwise, interchanging the roles of x_1 and x_2 and of ϕ_1 and ϕ_2, a repeat of the above construction yields the existence of a point $y_2 \in K \setminus \{0\}$ and of a sequence $(y_2^n)_{n\in\mathbb{N}} \subset C \cap K$ such that $\mathfrak{q}(y_2^n) \to \mathfrak{q}(y_2)$ and $z \in [y_1, y_2]$. This shows

$$\begin{aligned} x = (\lambda_1+\lambda_2)z &\in \mathrm{cone}(\{y_1, y_2\}) \\ &\subseteq \mathrm{clo}[\mathrm{cone}(\{y_i^n : n \in \mathbb{N}, i = 1,2\})] \\ &\subseteq \mathrm{clo}[\mathrm{cone}(C \cap K)]. \end{aligned}$$

It remains to treat the case $\phi_2(x_1) < 0$. In this case, $x \in K$ implies $x_2 \in K$. The above construction can then be repeated using the point x_1 for both ϕ_1 and ϕ_2, revealing the existence of points $y_i \neq 0$ such that $\phi_i(y_i) = 0$ and $z \in [y_i, x_2]$, $(i = 1,2)$. Without loss of generality, we may assume that $y_2 \in [y_1, x_2]$, whence $y_2 \in K$ and there exists a sequence $(y_2^n)_{n\in\mathbb{N}} \subseteq C \cap K$ such that $\mathfrak{q}(y_2^n) \to \mathfrak{q}(y_2)$. We therefore have

$$\begin{aligned} x = (\lambda_1+\lambda_2)z &\in \mathrm{cone}(\{y_2, x_2\}) \\ &\subseteq \mathrm{clo}[\mathrm{cone}(\{y_2^n : n \in \mathbb{N}\} \cup \{x_2\})] \\ &\subseteq \mathrm{clo}[\mathrm{cone}(C \cap K)]. \end{aligned}$$

In summary, we have established that $\mathrm{clo}[\mathrm{cone}(C\cap K)] \supseteq \mathrm{cone}(C) \cap K \setminus \{0\}$. Our claim now follows by taking closures on both sides of this inclusion. ☐

We are now ready to state and prove the main result of this paper, for the purpose of which we are going to make the following regularity assumption,

$$\mathrm{clo}\big[\mathrm{cone}(C) \cap K\big] = \mathrm{clo}\big[\mathrm{cone}(C)\big] \cap K. \tag{2.6}$$

THEOREM 2.7. *Let $C \subseteq \mathbb{R}^2$ be homogenization-convex and $K \subseteq \mathbb{R}^2$ a closed convex cone such that the regularity assumption* (2.6) *holds. Then*

$$(C \cap K)^* = C^* + K^*.$$

Proof. Using Lemma 2.6 and the classical dual cone calculus formulas, we find

$$\begin{aligned}(C \cap K)^* &\overset{(2.3)}{=} (\mathrm{clo}\,[\mathrm{cone}\,(C \cap K)])^* \\ &\overset{(2.5)}{=} (\mathrm{clo}\,[\mathrm{cone}(C) \cap K])^* \\ &\overset{(2.6)}{=} (\mathrm{clo}\,[\mathrm{cone}(C)] \cap K)^* \\ &\overset{(2.2)}{=} (\mathrm{clo}\,[\mathrm{cone}(C)])^* + K^* \\ &\overset{(2.3)}{=} C^* + K^*.\end{aligned}$$

□

Next, let us give a sufficient criterion that is easier to check than Condition (2.6).

LEMMA 2.8. *Let $C \subseteq \mathbb{R}^2$ and $K \subseteq \mathbb{R}^2$ a convex cone. If $C \cap \mathrm{int}[K] \neq \emptyset$, then Condition* (2.6) *holds.*

Proof. The proof works in arbitrary normed vector spaces V. We only need to prove that the inclusion $\supseteq$ holds in (2.6), the reverse relation being trivial. Let $x_0 \in C \cap \mathrm{int}[K]$, and let $(x_n)_{\mathbb{N}} \subset \mathrm{cone}(C)$ be a sequence such that $x_n \to x \in K$. Then for every $\varepsilon > 0$ we have $x_n + \varepsilon x_0 \in \mathrm{cone}(C) \cap K$ for all n large enough. Therefore, $x + \varepsilon x_0 \in \mathrm{clo}[\mathrm{cone}(C) \cap K]$. This being true for all $\varepsilon > 0$, we have $x \in \mathrm{clo}[\mathrm{cone}(C) \cap K]$, as claimed. □

COROLLARY 2.9. *$C \subseteq \mathbb{R}^2$ be homogenization-convex, and let $\psi, \phi : \mathbb{R}^2 \to \mathbb{R}$ be linear forms, with ϕ chosen such that there exists $x_0 \in C$ where $\phi(x_0) > 0$. Then the following conditions are equivalent,*

i) $\psi(x) \geq 0$ *for all* $x \in C$ *such that* $\phi(x) \geq 0$,
ii) *there exists* $\xi \geq 0$ *such that* $\psi(x) - \xi\phi(x) \geq 0$ *for all* $x \in C$.

Proof. This is a special case of Theorem 2.7 with $K = \{x : \phi(x) \geq 0\}$ and where the sufficient criterion of Lemma 2.8 applies. □

Next, we lift Corollary 2.9 into arbitrary real Hilbert spaces, resulting in the following result.

THEOREM 2.10. *Let $(V, \langle\cdot,\cdot\rangle)$ be a real Hilbert space, $\psi, \phi : V \to \mathbb{R}$ continuous linear forms, $W := (\ker(\phi) \cap \ker(\psi))^{\perp}$ and π_W the orthogonal projection of V onto W along $\ker(\phi) \cap \ker(\psi)$. Let C be a subset of V such that $\phi(x_0) > 0$ for some $x_0 \in C$ and such that $\pi_W C$ is homogenization-convex in W. Then the following conditions are equivalent:*

i) $\psi(x) \geq 0$ *for all* $x \in C$ *such that* $\phi(x) \geq 0$,
ii) *there exists* $\xi \geq 0$ *such that* $\psi(x) - \xi\phi(x) \geq 0$ *for all* $x \in C$.

Proof. Applying Corollary 2.9 to $\phi|_W, \psi|_W$ and $\pi_W C$ on the two-dimensional subspace W, we find that i) $\Leftrightarrow$ $\psi|_W(x) \geq 0$ for all $x \in \pi_W C$ such that $\phi|_W(x) \geq 0 \Leftrightarrow \exists\, \xi \geq 0$ such that $\psi|_W(x) - \xi\phi|_W(x) \geq 0$ for all $x \in \pi_W C \Leftrightarrow$ ii). □

It is important to understand that Theorem 2.10 is more than just a generalization of Corollary 2.9 to arbitrary real Hilbert spaces, for rather than assuming that C be homogenization-convex in V (if the definition is appropriately extended to arbitrary Hilbert spaces), the theorem merely gets away with the weaker assumption that the projected set $\pi_W C$ be homogenization-convex. This distinction is crucial, as in our proof of the S-Lemma, C is not homogenization-convex, while $\pi_W C$ *is* homogenization-convex due to the two dimensional nature of W. In fact, $\pi_W C$ is in general not homogenization-convex when $\dim(W) \geq 3$, and this is the main reason why the S-Lemma does not hold for quadratic functions copositive with more than one quadratic form.

Note further that if the set C is actually convex (rather than just homogenization-convex), Theorem 2.10 becomes a special case of Farkas' Theorem, see [11].

2.2. Proof of the S-Lemma. Next, we shall see that, despite its Farkas flavour, Theorem 2.10 is in fact a generalisation of the S-Lemma, and (2.6) is a weaking of the standard regularity assumption: denoting the set of real symmetric $n \times n$ matrices by $\mathscr{S}_n$, and combining the tools developed above, we obtain a proof of Theorem 1.1:

Proof. Let g be given by $g(x) = x^{\mathrm{T}}Qx + 2\ell^{\mathrm{T}}x + c$, where $Q \in \mathscr{S}_n$, $\ell \in \mathbb{R}^n$ and $c \in \mathbb{R}$. Then $g(x) = \langle A, [\begin{smallmatrix} x & 1 \end{smallmatrix}]^{\mathrm{T}}[\begin{smallmatrix} x & 1 \end{smallmatrix}]\rangle$, where $\langle A, X\rangle = \mathrm{tr}(A^{\mathrm{T}}X)$ is the trace inner product defined on the space $\mathscr{S}_{n+1}$ of symmetric $(n+1)\times(n+1)$ matrices, and where

$$A = \begin{bmatrix} Q & \ell \\ \ell^{\mathrm{T}} & c \end{bmatrix}.$$

Likewise, there exists $B \in \mathscr{S}_{n+1}$ such that $h(x) = \langle B, [\begin{smallmatrix} x & 1 \end{smallmatrix}]^{\mathrm{T}}[\begin{smallmatrix} x & 1 \end{smallmatrix}]\rangle$. Let $C \subset \mathscr{S}_{n+1}$ be defined by $C = \{zz^{\mathrm{T}} : z = [\begin{smallmatrix} x & 1 \end{smallmatrix}]^{\mathrm{T}}, x \in \mathbb{R}^n\}$. Using the notation just introduced, the claim of the theorem is that the following two conditions are equivalent,

i) $\langle A, X\rangle \geq 0$ for all $X \in C$ such that $\langle B, X\rangle \geq 0$,
ii) there exists $\xi \geq 0$ such that $\langle A - \xi B, X\rangle \geq 0$ for all $X \in C$.

We note that $\psi : X \mapsto \langle A, X\rangle$ and $\phi : X \mapsto \langle B, X\rangle$ are linear forms on $\mathscr{S}_{n+1}$. Furthermore, if $X_0 = [\begin{smallmatrix} x_0 & 1 \end{smallmatrix}]^{\mathrm{T}}[\begin{smallmatrix} x_0 & 1 \end{smallmatrix}]$, then $X_0 \in C$ and $\phi(X_0) = h(x_0) > 0$. Thus, the equivalence of i) and ii) follows from Theorem 2.10 if it can be established that $\pi_W C$ is homogenization-convex, where π_W is the orthogonal projection of $(\mathscr{S}_{n+1}, \langle\cdot,\cdot\rangle)$ onto $W := (\ker(\phi) \cap \ker(\psi))^{\perp} = \mathrm{span}\{A, B\}$. Let $X_1, X_2 \in C$. Then $X_i = [\begin{smallmatrix} x_i & 1 \end{smallmatrix}]^{\mathrm{T}}[\begin{smallmatrix} x_i & 1 \end{smallmatrix}]$ for some $x_i \in \mathbb{R}^n$, $(i = 1, 2)$. For $t \in \mathbb{R}$, define $x(t) := x_2 - t(x_2 - x_1)$ and $X(t) := [\begin{smallmatrix} x(t) & 1 \end{smallmatrix}]^{\mathrm{T}}[\begin{smallmatrix} x(t) & 1 \end{smallmatrix}] = G_0 + tG_1 + t^2G_2$, where

$$\begin{aligned} G_0 &= [\begin{smallmatrix} x_2 \\ 1 \end{smallmatrix}][\begin{smallmatrix} x_2 \\ 1 \end{smallmatrix}]^{\mathrm{T}}, \\ G_1 &= -[\begin{smallmatrix} x_2 \\ 1 \end{smallmatrix}][\begin{smallmatrix} x_2 - x_1 \\ 0 \end{smallmatrix}]^{\mathrm{T}} - [\begin{smallmatrix} x_2 - x_1 \\ 0 \end{smallmatrix}][\begin{smallmatrix} x_2 \\ 1 \end{smallmatrix}]^{\mathrm{T}}, \\ G_2 &= [\begin{smallmatrix} x_2 - x_1 \\ 0 \end{smallmatrix}][\begin{smallmatrix} x_2 - x_1 \\ 0 \end{smallmatrix}]^{\mathrm{T}}. \end{aligned}$$

Let $E_1, E_2 \in \mathscr{S}_{n+1}$ be an orthonormal basis of W. Then $\pi_W X(t) = a(t)E_1 + b(t)E_2$, where $a(t) = \langle G_0, E_1\rangle + t\langle G_1, E_1\rangle + t^2\langle G_2, E_1\rangle$ and $b(t) = \langle G_0, E_2\rangle + t\langle G_1, E_2\rangle + t^2\langle G_2, E_2\rangle$. Defining $T := \{[\begin{smallmatrix} a(t) & b(t) \end{smallmatrix}]^{\mathrm{T}} : t \in \mathbb{R}\}$, Lemma 2.3 shows that T is homogenization-convex in $\mathbb{R}^2$. By virtue of Lemma 2.2 iv), this implies that $\pi_W C$ is homogenization-convex, as claimed. □

2.3. Generalization to Hilbert Space Kernels. The proof given above generalizes to infinite-dimensional spaces:

THEOREM 2.11. *Let $(V, \langle\cdot,\cdot\rangle)$ be a real Hilbert space, and let $g, h : V \to \mathbb{R}$ be continuous quadratic functions defined on V by*

$$\begin{aligned} g &: x \mapsto c_g + 2\langle v_g, x\rangle + \langle x, M_g x\rangle, \\ h &: x \mapsto c_h + 2\langle v_h, x\rangle + \langle x, M_h x\rangle, \end{aligned}$$

where $M_g, M_h : V \to V$ are self-adjoint operators, $v_g, v_h \in V$ and $c_g, c_h \in \mathbb{R}$. Let h us further assume that there exists $x_0 \in V$ where $h(x_0) > 0$. Then g is copositive with h if and only if there exists $\xi \geq 0$ such that $g(x) - \xi h(x) \geq 0$ for all $x \in V$.

Proof. The proof is identical to that of Theorem 1.1 bar the following construction: let $H := V \oplus \mathbb{R}$, where $\oplus$ denotes the direct sum of Hilbert spaces, and let us write $\langle \cdot, \cdot \rangle_H$ for the inner product on H. Let $\mathscr{S}$ be the space of self-adjoint operators on H. By the Hellinger-Toeplitz Theorem, such operators are automatically continuous, and it is easy to see that $A, B \in \mathscr{S}$, where

$$\begin{aligned} A &: (x, \tau) \mapsto (M_g x + \tau v_g, \langle v_g, x \rangle + \tau c_g) \\ B &: (x, \tau) \mapsto (M_h x + \tau v_h, \langle v_h, x \rangle + \tau c_h). \end{aligned}$$

Let $\{e_i : i \in \mathbb{N}\}$ be an orthonormal basis of H. The following operators are in $\mathscr{S}$,

$$E_{ij} : y \mapsto \frac{1}{1 + \delta_{ij}} \left(\langle e_i, y \rangle_H e_j + \langle e_j, y \rangle_H e_i \right),$$

where δ_{ij} is the Kronecker delta. Defining

$$\langle E_{ij}, E_{kl} \rangle_S := \begin{cases} 1 & \text{if } \{i, j\} = \{k, l\}, \\ 0 & \text{otherwise,} \end{cases}$$

the E_{ij} generate a Hilbert space $(S, \langle \cdot, \cdot \rangle_S)$ for which $\{E_{ij} : i, j \in \mathbb{N}\}$ is an orthonormal basis. In fact, S is the set of compact operators in $\mathscr{S}$, and the topology defined by the trace inner product $\langle \cdot, \cdot \rangle_S$ is the uniform topology, since $\langle E_{ij}, X \rangle_S = \langle e_i, X e_j \rangle_H$ for all $X \in S$. Every $x \in V$ defines an operator $R(x) \in \mathscr{S}$,

$$R(x) : z \mapsto \langle (x, 1), z \rangle_H (x, 1),$$

and if $(x, 1) = \sum_{i \in \mathbb{N}} \xi_i e_i$ then $R(x) = \sum_{ij} \xi_i \xi_j E_{ij}$ and $\sum_{ij} \xi_i^2 \xi_j^2 = (\sum_i \xi_i^2)(\sum_j \xi_j^2) < \infty$. This shows that $C := \{R(x) : x \in V\} \subset S$. Extending the map

$$\begin{aligned} \psi &: C \to \mathbb{R}, \\ R(x) &\mapsto \langle (x, 1), A(x, 1) \rangle_H \end{aligned}$$

by linearity and continuity, we obtain a bounded linear operator on the Hilbert space $(\mathrm{clo}[\mathrm{span}(C)], \langle \cdot, \cdot \rangle_S)$. Likewise, B defines a bounded linear operator ϕ on the same space. Replacing $\mathscr{S}_{n+1}$ by $\mathrm{clo}[\mathrm{span}(C)]$ in the proof of Section 2.2, a repetition of the arguments presented there proves the claim of Theorem 2.11. □

We remark that the condition

$$g(x) - \xi h(x) \geq 0, \quad \forall\, x \in V$$

is equivalent to requiring that

$$\begin{aligned} K &: V \times V \to \mathbb{R}, \\ &(x, y) \mapsto \langle x, (M_g - \xi M_h) y \rangle + \langle v_g - \xi v_h, x + y \rangle + c_g - \xi c_h \end{aligned}$$

be a positive definite kernel.

REFERENCES

[1] A. Ben-Tal and A. Nemirovski. Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications. *MPS-SIAM Series on Optimization*, SIAM, Philadelphia, PA, 2001.

[2] K. Derinkuyu and M.C. Pinar. On the S-Procedure and Some Variants. *Mathematical Methods of Operations Research*, Vol. 64, No. 1, pp.55-77, 2006.

[3] L.L. Dines. On the mapping of quadratic forms. *Bulletin of the AMS*, Vol.47, pp.494–498, 1941.

[4] P. Finsler. Über das Vorkommen definiter und semidefiniter Formen in Scharen quadratischer Formen. *Commentaria Mathematicae Helvetia*, Vol.9, pp.188–192, 1937.

[5] M.R. Hestenes and E.J. McShane. A theorem on quadratic forms and its application in the calculus of variations. *Transactions of the AMS*, Vol.47, pp.501–512, 1940.

[6] A. Megretsky and S. Treil. Power distribution in optimization and robustness of uncertain systems. *Journal of Mathematical Systems, Estimation and Control*, Vol.3, pp.301–319, 1993.

[7] I. Pólik and T. Terlaky. A survey of the S-Lemma. *SIAM Review*, Vol 49, pp.371–418, 2007.

[8] J.F. Sturm and S. Zhang. On cones of nonnegative quadratic functions. *Mathematics of Operations Research*, Vol.28, pp.246–267, 2003.

[9] V.A. Yakubovich. S-Procedure in nonlinear control theory. *Vestnik Leningrad University*, Vol.1, pp.62–77, 1971.

[10] Y.-X. Yuan. On a subproblem of trust region algorithms for constrained optimization. *Math. Prog.*, Vol 47, pp.53–63, 1990.

[11] R. T. Rockafellar. Convex Analysis. *Princeton University Press*, Princeton, N. J., 1970.