Journal of Convex Analysis
Volume 28 (2021), No. 4, 1015–1032

Separating Hyperplanes of Convex Sets

Valeriu Soltan
Dept. of Mathematical Sciences, George Mason University, Fairfax, VA 22030, U.S.A.
vsoltan@gmu.edu

Received: July 13, 2020
Accepted: August 31, 2020

Combining existing approaches, we provide a uniform way to describe all hyperplanes which separate (properly, or strongly) a given pair of nonempty convex sets K_1 and K_2 in the n-dimensional Euclidean space. The method is based on considering $(n-1)$-dimensional subspaces which bound the set $K_1 - K_2$ and then using properties of the polar cone $(K_1 - K_2)^\circ$. First, we characterize all separating hyperplanes with given normal vectors, and then those containing a given point. We also describe the union of all hyperplanes separating (properly separating) a given pair of convex sets.

Keywords: Separation, hyperplane, convex, cone.

2010 Mathematics Subject Classification: 52A20, 90C25.

1. Introduction

Support and separation of convex sets are among the most important topics in convexity. Studied by Minkowski [9, 10] on the turn of 20th century, they became useful tools in many mathematical disciplines, especially in convex geometry, linear analysis, convex analysis, and optimization. In what follows, we consider separation of convex sets in the Euclidean space $\mathbb{R}^n$.

We recall that nonempty convex sets K_1 and K_2 in $\mathbb{R}^n$ are *separated* by a hyperplane $H \subset \mathbb{R}^n$ provided K_1 and K_2 lie in the opposite closed halfspaces determined by H. On a historical note, Minkowski did not use the term *separating hyperplane*; this expression (*trennende Ebene*) can be traced in the paper of Brunn [1].

The concept of separation was gradually refined in the literature. The established terminology in this regard is mainly due to Klee [7] and Rockafellar [13]. Of main interest to us are the following two types: convex sets K_1 and K_2 are called *properly* separated by a hyperplane H if they are separated by H such that $K_1 \cup K_2 \not\subset H$, and they are called *strongly* separated by H if there are suitable open neighborhoods $U_\rho(K_1)$ and $U_\rho(K_2)$ of these sets which are separated by H. These two types of separation are the most popular in the literature due to the following simple criteria (see [13, Section 11]).

Proposition 1.1. *If K_1 and K_2 are convex sets in $\mathbb{R}^n$, then the following assertions hold.*

(1) *There is a hyperplane properly separating K_1 and K_2 if and only if*
$$\operatorname{rint} K_1 \cap \operatorname{rint} K_2 = \varnothing.$$

(2) *There is a hyperplane strongly separating K_1 and K_2 if and only if the inequality $\delta(K_1, K_2) > 0$ holds.* □

ISSN 0944-6532 / $ 2.50

Clearly, Proposition 1.1 is existential, and its proof provides ways to construct a single or a specific family of separating hyperplanes of the necessary type. On the other hand, several articles, [3]–[6] and [12], contain various descriptions of the family of all hyperplanes separating a pair of disjoint polytopes or polyhedra. These descriptions use finite algorithms and, generally, are not suitable for extensions to the case of arbitrary convex sets.

The purpose of this article is to describe all hyperplanes separating (properly, or strongly) a pair of convex sets K_1 and K_2 in $\mathbb{R}^n$. Reducing the argument to a description of all $(n-1)$-dimensional subspaces bounding a given convex set (Section 3), we characterize all separating hyperplanes with a given direction of their normal vectors (Section 4). Based on these results, we further characterize all separating hyperplanes containing a given point; finally, we describe the domain of all points which admit separating hyperplanes through them (Section 5).

2. Preliminaries

This section contains necessary definitions, notation, and results on convex sets in the n-dimensional Euclidean space $\mathbb{R}^n$ (see, e. g., [13] and [16] for details). The elements of $\mathbb{R}^n$ are called vectors (or points). In what follows, o stands for the zero vector of $\mathbb{R}^n$. We denote by $[u,v]$ and (u,v) the closed and open segments with endpoints $u,v \in \mathbb{R}^n$. Also, $u\cdot v$ will mean the dot product of u and v. The open ρ-neighborhood of X, denoted $U_\rho(X)$, is the union of all open balls $U_\rho(x)$ of radius $\rho > 0$ centered at $x \in X$. Nonempty sets X_1 and X_2 in $\mathbb{R}^n$ are called *strongly disjoint* provided $U_\rho(X_1) \cap U_\rho(X_2) = \varnothing$ for a suitable $\rho > 0$; the latter occurs if and only if the *inf*-distance $\delta(X_1, X_2)$, defined by

$$\delta(X_1, X_2) = \inf\{\|x_1 - x_2\| : x_1 \in X_1, x_2 \in X_2\},$$

is positive. For a nonempty set $X \subset \mathbb{R}^n$, the notations $\operatorname{span} X$, $X^\perp$, $\operatorname{cl} X$, and $\operatorname{int} X$, stand, respectively, for the *span*, *orthogonal complement*, *closure*, and *interior* of X. The *affine span* of X, denoted $\operatorname{aff} X$, is the intersection of all planes containing X, and $\dim X$ is defined as the *dimension* of the plane $\operatorname{aff} X$. Also, the *direction space* of X is defined by $\operatorname{dir} X = \operatorname{aff} X - \operatorname{aff} X$.

By an *r-dimensional plane* L in $\mathbb{R}^n$, where $0 \le r \le n$, we mean a translate of a suitable r-dimensional subspace S of $\mathbb{R}^n$: $L = c + S$, where $c \in \mathbb{R}^n$. It is known that a nonempty set $M \subset \mathbb{R}^n$ is a plane if and only if $(1-\lambda)x + \lambda y \in M$ whenever $x, y \in M$ and $\lambda \in \mathbb{R}$.

A *hyperplane* is a plane of dimension $n-1$; it can be described as

$$H = \{x \in \mathbb{R}^n : x\cdot e = \gamma\}, \quad e \ne o, \quad \gamma \in \mathbb{R}. \tag{1}$$

Consequently, a hyperplane of the form (1) is a translate of the hypersubspace

$$S = \{x \in \mathbb{R}^n : x\cdot e = 0\}, \quad e \ne o. \tag{2}$$

Nonzero multiples te of the vector e in (1) and (2) are called normal vectors of H and S.

Every hyperplane of the form (1) determines a pair of opposite closed halfspaces

$$V_1 = \{x \in \mathbb{R}^n : x \cdot e \le \gamma\} \quad \text{and} \quad V_2 = \{x \in \mathbb{R}^n : x \cdot e \ge \gamma\}. \tag{3}$$

In what follows, K means a *nonempty convex set* in $\mathbb{R}^n$. The relative interior and relative boundary of K are denoted $\operatorname{rint} K$ and $\operatorname{rbd} K$, respectively. It is known that K has nonempty relative interior, and that $\operatorname{rbd} K \ne \varnothing$ if and only if K is not a plane. Furthermore, both sets $\operatorname{cl} K$ and $\operatorname{rint} K$ are convex and

$$\begin{aligned} \operatorname{cl} K &= \operatorname{cl}(\operatorname{rint} K), \ \operatorname{rint} K = \operatorname{rint}(\operatorname{cl} K), \\ \operatorname{aff} K &= \operatorname{aff}(\operatorname{cl} K) = \operatorname{aff}(\operatorname{rint} K). \end{aligned} \tag{4}$$

We recall that a nonempty set C in $\mathbb{R}^n$ is a *cone* with apex $a \in \mathbb{R}^n$ if $a + \lambda(x - a) \in C$ whenever $\lambda \ge 0$ and $x \in C$. (Obviously, this definition implies that $a \in C$, although a stronger condition $\lambda > 0$ can be beneficial; see, e. g., [8].) The cone C is called *convex* if it is a convex set.

Given a convex cone $C \subset \mathbb{R}^n$ with apex o, the set $\operatorname{lin} C = C \cap (-C)$ is called the *lineality space* of C; it is the largest subspace contained in C. Consequently, $C \ne \operatorname{lin} C$ if and only if C is not a subspace.

The *recession cone* of a convex set $K \subset \mathbb{R}^n$ is defined by

$$\operatorname{rec} K = \{e \in \mathbb{R}^n : x + \lambda e \in K \text{ whenever } x \in K \text{ and } \lambda \ge 0\}.$$

If K is closed, then $\operatorname{rec} K$ is a closed convex cone with apex o; it is distinct from $\{o\}$ if and only if K is unbounded. The *lineality space* of K is the subspace defined by $\operatorname{lin} K = \operatorname{lin}(\operatorname{rec} K)$. In a standard way, the (negative) *polar cone* of a nonempty set $X \subset \mathbb{R}^n$ is defined by

$$X^\circ = \{e \in \mathbb{R}^n : x \cdot e \le 0 \ \text{ for all } \ x \in X\}.$$

We will need the following assertions (see [14] and [15] for additional results).

(P1) If $K \subset \mathbb{R}^n$ is a closed convex set and T is the orthogonal complement of $\operatorname{lin} K$, then $K = \operatorname{lin} K + K \cap T$.

(P2) K° is a closed convex cone with apex o, and $\operatorname{lin} K^\circ = K^\perp$.

(P3) $K^\circ = \{o\}$ if and only if $o \in \operatorname{int} K$, and K° is a subspace if and only if $o \in \operatorname{rint} K$.

(P4) If $C \subset \mathbb{R}^n$ is a convex cone with apex o, then a vector $e \in \mathbb{R}^n$ belongs to $\operatorname{rint} C^\circ$ if and only if $x \cdot e < 0$ for all vectors $x \in \operatorname{cl} C \setminus \operatorname{lin}(\operatorname{cl} C)$.

(P5) If C_1 and C_2 are convex cones in $\mathbb{R}^n$, with common apex o, then

$$(C_1 + C_2)^\circ = (C_1 \cup C_2)^\circ = C_1^\circ \cap C_2^\circ.$$

For a convex set $K \subset \mathbb{R}^n$ and a point $a \in \mathbb{R}^n$, the *generated cone* $C_a(K)$ with apex a is defined by

$$C_a(K) = \{a + \lambda(x - a) : x \in K, \ \lambda \ge 0\}.$$

Both sets $C_a(K)$ and $\operatorname{cl} C_a(K)$ are convex cones with apex a. By $C_o(K)$ we will denote the generated cone with the zero apex o.

(P6) $C_a(K)$ is a plane if and only if $a \in \operatorname{rint} K$ (in particular, $C_a(K) = \mathbb{R}^n$ if and only if $a \in \operatorname{int} K$).

(P7) $C_a(\operatorname{cl} K) \subset \operatorname{cl} C_a(K)$ and $C_a(\operatorname{rint} K) = \{a\} \cup \operatorname{rint} C_a(K)$.

(P8) If K is bounded and $a \notin \operatorname{cl} K$, then $C_a(\operatorname{cl} K) = \operatorname{cl} C_a(K)$.

(P9) $K^\circ = (C_o(K))^\circ = (\operatorname{cl} C_o(K))^\circ$ and $\operatorname{cl} C_o(K) = (K^\circ)^\circ$.

3. Bounding hypersubspaces

We will say that a hypersubspace $S \subset \mathbb{R}^n$, that is, an $(n-1)$-dimensional subspace of the form (2), *bounds* a convex set $K \subset \mathbb{R}^n$ provided K is contained in one of the homogeneous closed halfspaces

$$E_1 = \{x \in \mathbb{R}^n : x \cdot e \le 0\} \quad \text{and} \quad E_2 = \{x \in \mathbb{R}^n : x \cdot (-e) \le 0\}. \tag{5}$$

Furthermore, S *properly* bounds K if S bounds K such that $K \not\subset S$, and S *strongly* bounds K if S bounds a suitable ρ-neighborhood $U_\rho(K)$ of K.

Our motivation for considering bounding hypersubspaces is due to the following lemma (see, e. g., [16, Theorem 10.7]).

Lemma 3.1. *For convex sets K_1 and K_2 in $\mathbb{R}^n$ and a hypersubspace $S \subset \mathbb{R}^n$, the assertions below are equivalent.*

(1) *There is a translate of S which separates* (*respectively, properly or strongly*) *K_1 and K_2.*

(2) *S bounds* (*respectively, properly or strongly*) *the set $K_1 - K_2$.* □

Theorem 3.2. *For a convex set $K \subset \mathbb{R}^n$, the assertions below hold.*

(1) *There is a hypersubspace bounding K if and only if any of the following two conditions is satisfied:*
   (a) $o \notin \operatorname{int} K$,
   (b) $K^\circ \ne \{o\}$.

(2) *A hypersubspace of the form* (2) *bounds K if and only if one of the vectors e and $-e$ belongs to $K^\circ \setminus \{o\}$.*

Proof. (1) Assume first that a hypersubspace $S \subset \mathbb{R}^n$ bounds K. Expressing S in the form (2), we conclude that one of the halfspaces E_1 and E_2 from (5) contains K. If $K \subset E_1$, then $e \in E_1^\circ \subset K^\circ$. Similarly, $-e \in E_2^\circ \subset K^\circ$ if $K \subset E_2$. In either case, $K^\circ \ne \{o\}$. The equivalence of conditions (a) and (b) follows from (P3).

Conversely, suppose that $K^\circ \ne \{o\}$. Choosing a vector $e \in K^\circ \setminus \{o\}$, we obtain that $K \subset \{e\}^\circ = E_1$. The latter inclusion shows that the hypersubspace (2) bounds K.

(2) This assertion follows from the above argument □

Theorem 3.3. *For a convex set $K \subset \mathbb{R}^n$, the assertions below hold.*

(1) *There is a hypersubspace properly bounding K if and only if any of the following three conditions is satisfied:*
   (a) $o \notin \operatorname{rint} K$,
   (b) *K° is not a subspace,*
   (c) $K^\circ \setminus \operatorname{lin} K^\circ \ne \varnothing$.

(2) *A hypersubspace of the form* (2) *properly bounds K if and only if one of the vectors e and $-e$ belongs to $K^\circ \setminus \operatorname{lin} K^\circ$.*

Proof. (1) Assume first that a hypersubspace $S \subset \mathbb{R}^n$ properly bounds K. Expressing S in the form (2), we conclude that one of the halfspaces E_1 and E_2 from (5), say E_1, properly contains K. Then

$$\operatorname{rint} K \subset \operatorname{int} E_1 = \{x \in \mathbb{R}^n : x \cdot e < 0\}$$

(see [16, Corollary 9.18]). Since $o \in \operatorname{bd} E_1$, one has $o \notin \operatorname{rint} K$. So, condition (a) is satisfied. The equivalence of conditions (a)–(c) follows from the property (P3).

Conversely, suppose that one of conditions (a)–(c), say (c), is satisfied. Choose any vector $e \in K^\circ \setminus \operatorname{lin} K^\circ$. Then $e \neq o$ and, by Theorem 3.2, the hypersubspace S of the form (2) bounds K. Then one of the halfspaces E_1 and E_2 from (5) contains K. Assume that $K \subset E_1$ (the case $K \subset E_2$ is similar). Since $e \notin \operatorname{lin} K^\circ$, (P2) shows that $e \notin K^\perp$. Hence $K \not\subset \{e\}^\perp = S$. Summing up, S properly bounds K.

(2) This assertion is a summary the arguments from the proof of part (1). □

The following lemma is necessary for the proof of Theorem 3.5.

Lemma 3.4. *Let $K \subset \mathbb{R}^n$ be a convex set such that $o \notin \operatorname{cl} K$, and let $D = \operatorname{cl} C_o(K)$. If z is the metric projection of o on $\operatorname{cl} K$, then*

$$\delta(\operatorname{lin} D, \operatorname{cl} K) = \|z\|.$$

Proof. Since $o \in \operatorname{lin} D$, one has $\delta(\operatorname{lin} D, \operatorname{cl} K) \le \delta(o, \operatorname{cl} K) = \|z\|$. Assume, for contradiction, that $\delta(\operatorname{lin} D, \operatorname{cl} K) < \|z\|$. Then there are points $u \in \operatorname{lin} D$ and $v \in \operatorname{cl} K$ for which $\|u - v\| < \|z\|$. Let $w = v - u$. Because $\operatorname{lin} D = \operatorname{lin}(\operatorname{cl} K)$ (see [16, Theorem 7.9]), the property (P1) gives

$$w = v - u \in v + \operatorname{lin} D = v + \operatorname{lin}(\operatorname{cl} K) \subset \operatorname{cl} K.$$

So,
$$\delta(o, \operatorname{cl} K) \le \|o - w\| = \|u - v\| < \|z\|,$$

contrary to the choice of z. Thus $\delta(\operatorname{lin} D, \operatorname{cl} K) = \|z\|$. □

Theorem 3.5. *For a convex set $K \subset \mathbb{R}^n$, the assertions below hold.*

(1) *There is a hypersubspace strongly bounding K if and only if $o \notin \operatorname{cl} K$.*

(2) *If $o \notin \operatorname{cl} K$ and e is a vector in $\mathbb{R}^n$ such that one of the vectors e and $-e$ belongs to $\operatorname{rint} K^\circ$, then the hypersubspace (2) strongly bounds K.*

(3) *If a hypersubspace of the form (2) strongly bounds K, then one of the vectors e and $-e$ belongs to $K^\circ \setminus \operatorname{lin} K^\circ$. If, additionally, K is a bounded set, then one of these vectors belongs to $\operatorname{rint} K^\circ$.*

Proof. (1) Assume first that a hypersubspace $S \subset \mathbb{R}^n$ strongly bounds K. Then S bounds a suitable open ρ-neighborhood $U_\rho(K)$ of K. Denote by V the closed halfspace determined by S and containing $U_\rho(K)$. Because the set $U_\rho(K)$ is open, one has $U_\rho(K) \subset \operatorname{int} V$. Therefore, $o \notin U_\rho(K)$, due to the inclusion $o \in S$. Thus $o \notin \operatorname{cl} K$.

Conversely, suppose that $o \notin \operatorname{cl} K$ and denote by z the metric projection of o on $\operatorname{cl} K$. A well-known assertion of convex geometry (proved by Carathéodory [2, p. 198] for the case of convex bodies) states that the hypersubspace $S = \{x \in \mathbb{R}^n : x \cdot z = 0\}$ strongly bounds K.

(2) Assuming that $o \notin \mathrm{cl}\, K$, choose a vector $e \in \mathbb{R}^n$ such that $e \in \mathrm{rint}\, K^\circ$ (the case $-e \in \mathrm{rint}\, K^\circ$ is similar). By (P6), the generated cone $C_o(K)$ is not a subspace. Consequently, the closed convex cone $D = \mathrm{cl}\, C_o(K)$ also is not a subspace. By (P9), $K^\circ = D^\circ$, whence $e \in \mathrm{rint}\, D^\circ$. Therefore, (P4), applied to the case $C = D$, shows that the hypersubspace S of the form (2) satisfies the condition $D \cap S = \mathrm{lin}\, D$. If z is the metric projection of o on $\mathrm{cl}\, K$, then, with $\rho = \|z\|$, Lemma 3.4 implies that the both-way infinite open cylinder

$$U_\rho(\mathrm{lin}\, D) = \mathrm{lin}\, D + U_\rho(o) = \{x \in \mathbb{R}^n : \delta(x, \mathrm{lin}\, D) < \rho\}$$

is disjoint from $\mathrm{cl}\, K$. Therefore, the closed set $F = D \setminus U_\rho(\mathrm{lin}\, D)$ contains $\mathrm{cl}\, K$. The condition $D \cap S = \mathrm{lin}\, D$ implies that

$$F \cap S \subset (D \setminus \mathrm{lin}\, D) \cap S = (D \cap S) \setminus (\mathrm{lin}\, D \cap S) = \varnothing.$$

Hence $F \cap S = \varnothing$, and F lies in the open halfspace

$$W = \{x \in \mathbb{R}^n : x \cdot e < 0\}.$$

We assert that F and S are strongly disjoint. For this, consider the subspace $T = (\mathrm{lin}\, D)^\perp$. Since $\mathrm{lin}\, D = \mathrm{lin}\,(\mathrm{cl}\, K)$ (see [16, Theorem 7.9]), the property (P1) gives

$$D = \mathrm{lin}\, D + D \cap T, \quad \mathrm{cl}\, K = \mathrm{lin}\, D + \mathrm{cl}\, K \cap T. \tag{6}$$

It is easy to see that

$$U_\rho(\mathrm{lin}\, D) = \mathrm{lin}\, D + U_\rho(\mathrm{lin}\, D) \cap T.$$

This argument and (6) give

$$\begin{aligned} F = D \setminus U_\rho(\mathrm{lin}\, D) &= (\mathrm{lin}\, D + D \cap T) \setminus (\mathrm{lin}\, D + U_\rho(\mathrm{lin}\, D) \cap T) \\ &= \mathrm{lin}\, D + (D \setminus U_\rho(\mathrm{lin}\, D)) \cap T = \mathrm{lin}\, D + F \cap T. \end{aligned} \tag{7}$$

Let $G = (F \cap T) \cap S_\rho(o)$, where $S_\rho(o) \subset \mathbb{R}^n$ is the sphere of radius ρ centered at o. The set G is compact as the intersection of the closed set $F \cap T$ and the compact set $S_\rho(o)$. Obviously,

$$F \cap T = \{\lambda x : \lambda \ge 1,\ x \in G\}. \tag{8}$$

Because G, as a subset of F, is disjoint from S, a compactness argument implies the existence of a positive scalar r such that $\delta(G, S) \ge r$. If $y \in F \cap T$, then, by (8), $y = \lambda x$ for a suitable scalar $\lambda \ge 1$ and a point $x \in G$. Therefore,

$$\delta(y, S) = \lambda \delta(x, S) \ge \lambda r \ge r \quad \forall\, y \in F \cap T. \tag{9}$$

Finally, choose any point $u \in F$. The equalities (7) show that u can be written as $u = w + y$, where $w \in \mathrm{lin}\, D$ and $y \in F \cap T$. Let f denote the orthogonal projection of $\mathbb{R}^n$ onto S. Since $w \in S$, the linearity of f gives

$$f(y) = f(u - w) = f(u) - f(w) = f(u) - w.$$

This argument and (9) imply

$$\begin{aligned}\delta(u,S) &= \|u-f(u)\| = \|(w+y)-f(u)\| \\ &= \|y-(f(u)-w)\| = \|y-f(y)\| = \delta(y,S) \geq r.\end{aligned}$$

Thus $$\delta(F,S) = \inf\{\delta(u,S) : y \in F\} \geq r.$$

Due to $\mathrm{cl}\, K \subset F$, we have $\delta(\mathrm{cl}\, K, S) \geq \delta(F,S) \geq r$. These inequalities and the inclusion $F \subset W$ show that the open r-neighborhood $U_r(\mathrm{cl}\, K)$ is contained in W. Summing up, S strongly bounds $\mathrm{cl}\, K$.

(3) If a hypersubspace S of the form (2) strongly bounds K, then Theorem 3.3 shows that one of the vectors e and $-e$ is in $K^\circ \setminus \mathrm{lin}\, K^\circ$.

Assume, additionally, that K is a bounded set, and let $e \in K^\circ \setminus \mathrm{lin}\, K^\circ$. Suppose, for contradiction, that $e \notin \mathrm{rint}\, K^\circ$. Then $e \in \mathrm{rbd}\, K^\circ$. By (P4), there is a nonzero vector $x \in \mathrm{cl}\, C_o(K)$ satisfying the condition $x \cdot e = 0$ (equivalently, $x \in S$). The property (P8) shows that $x \in C_o(\mathrm{cl}\, K)$. So, there is a vector $z \in \mathrm{cl}\, K$ with the property $x = \lambda z$ for a suitable scalar $\lambda > 0$. Furthermore, $z = \lambda^{-1} x \in S$ (due to $x \in S$), contrary to the assumption that $\mathrm{cl}\, K$ and S are strongly disjoint. Hence $e \in \mathrm{rint}\, K^\circ$, as desired. □

The following example shows that, if the set K in assertion (3) of Theorem 3.5 is unbounded, then none of the inclusions $e \in K^\circ \setminus \mathrm{lin}\, K^\circ$ or $-e \in K^\circ \setminus \mathrm{lin}\, K^\circ$ can be replaced with $e \in \mathrm{rint}\, K^\circ$ or $-e \in \mathrm{rint}\, K^\circ$, respectively.

Example 3.6. Let a closed convex set $K \subset \mathbb{R}^3$ be given by

$$K = \{(x,y,z) : x \geq 1, y \geq 1\}.$$

Then $$K^\circ = \{(x,y,0) : x \leq 0, y \leq 0\} \quad \text{and} \quad \mathrm{lin}\, K^\circ = \{o\}.$$

The yz-coordinate plane of $\mathbb{R}^3$, given by the equation $x = 0$, strongly bounds K, while its normal vector $e = (1,0,0)$ belongs to $\mathrm{rbd}\, K^\circ$.

4. Hyperplanes with given normals

The definitions of hyperplane separation can be formulated in analytic terms, as shown in the following obvious lemma.

Lemma 4.1. *For convex sets K_1 and K_2 in $\mathbb{R}^n$ and a hyperplane H of the form* (1)*, the assertions below hold.*

(1) *H separates K_1 and K_2 if and only if e and γ can be chosen such that*

$$\sup\{x \cdot e : x \in K_1\} \leq \gamma \leq \inf\{x \cdot e : x \in K_2\}. \tag{10}$$

(2) *H properly separates K_1 and K_2 if and only if e and γ can be chosen to satisfy both* (10) *and*

$$\inf\{x \cdot e : x \in K_1\} < \sup\{x \cdot e : x \in K_2\}. \tag{11}$$

(3) *H strongly separates K_1 and K_2 if and only if e and γ can be chosen such that*

$$\sup\{x \cdot e : x \in K_1\} < \gamma < \inf\{x \cdot e : x \in K_2\}. \tag{12}$$

According to Lemma 4.1, the existence of a hyperplane H which separates (properly, or strongly) convex sets K_1 and K_2 is equivalent to the existence of a suitable vector $e \neq o$ and a respective scalar γ in the description (1) of H. In this section, we divide the procedure of choice of e and γ into consecutive steps: first, we describe the domain of such vectors e, and then find limits for the values of γ. In doing so, it is convenient to view H as a translate of the hypersubspace (2).

Theorem 4.2. *For convex set K_1 and K_2 in $\mathbb{R}^n$, the assertions below hold.*

(1) *There is a hyperplane separating K_1 and K_2 if and only if any of the following two conditions is satisfied:*

(a) $o \notin \operatorname{int}(K_1 - K_2)$,

(b) $(K_1 - K_2)^\circ \neq \{o\}$.

(2) *There is a translate of the hypersubspace* (2) *separating K_1 and K_2 if and only if one of the vectors e and $-e$ belongs to $(K_1 - K_2)^\circ \setminus \{o\}$.*

(3) *For any vector $e \in (K_1 - K_2)^\circ \setminus \{o\}$, one has*

$$\sup\{x \cdot e : x \in K_1\} \leq \inf\{x \cdot e : x \in K_2\}. \tag{13}$$

Consequently, any scalar γ satisfying the inequalities (10) *can be used in the description* (1) *of a hyperplane which separates K_1 and K_2.*

Proof. (1) Assume first that a hyperplane $H \subset \mathbb{R}^n$ separates K_1 and K_2. Expressing H in the form (1), we may consider H as a translate of the hypersubspace S of the form (2). According to Lemma 3.1, S bounds the convex set $K_1 - K_2$. Consecutively, Theorem 3.2 shows that both conditions (a) and (b) are satisfied.

Conversely, if any of conditions (a) and (b) is satisfied, then Theorem 3.2 implies the existence of a hypersubspace bounding $K_1 - K_2$. By Lemma 3.1, there is a hyperplane separating K_1 and K_2.

(2) The proof of this assertion is similar to part 1).

(3) This assertion follows from Lemma 4.1 and the above parts 1) and 2). □

Theorem 4.3. *For convex sets K_1 and K_2 in $\mathbb{R}^n$, the assertions below hold.*

(1) *There is a hyperplane properly separating K_1 and K_2 if and only if any of the following three conditions is satisfied:*

(a) $o \notin \operatorname{rint}(K_1 - K_2)$,

(b) $(K_1 - K_2)^\circ$ *is not a subspace,*

(c) $(K_1 - K_2)^\circ \setminus \operatorname{lin}(K_1 - K_2)^\circ \neq \varnothing$.

(2) *There is a translate of the hypersubspace* (2) *properly separating K_1 and K_2 if and only if one of the vectors e and $-e$ belongs to $(K_1 - K_2)^\circ \setminus \operatorname{lin}(K_1 - K_2)^\circ$.*

(3) *For any vector $e \in (K_1 - K_2)^\circ \setminus \operatorname{lin}(K_1 - K_2)^\circ$, both inequalities* (11) *and* (13) *hold. Consequently, any scalar γ satisfying the inequalities* (10) *can be used in the description* (1) *of a hyperplane which properly separates K_1 and K_2.*

Proof. (1) According to Proposition 1.1, K_1 and K_2 are properly separated by a hyperplane if and only if $\operatorname{rint} K_1 \cap \operatorname{rint} K_2 = \varnothing$, which happens if and only if

$$o \notin \operatorname{rint} K_1 - \operatorname{rint} K_2 = \operatorname{rint}(K_1 - K_2)$$

(see [13], Corollary 6.6.2). The equivalence of conditions (a)–(c) follows from Theorem 3.3.

(2) Assume first that a hyperplane $H \subset \mathbb{R}^n$ properly separates K_1 and K_2. Expressing H in the form (1), we may consider H as a translate of the hypersubspace S of the form (2). According to Lemma 3.1, S properly bounds the convex set $K_1 - K_2$. Consecutively, Theorem 3.2 shows that one of the vectors e and $-e$ is in $(K_1 - K_2)^\circ \setminus \operatorname{lin}(K_1 - K_2)^\circ$.

Conversely, if one of the vectors e and $-e$ is in $(K_1 - K_2)^\circ \setminus \operatorname{lin}(K_1 - K_2)^\circ$, then Theorem 3.2 implies that the hypersubspace (2) properly bounds $K_1 - K_2$. By Lemma 3.1, there is a translate of S properly separating K_1 and K_2.

3) This assertion follows from Lemma 4.1 and the above parts 1) and 2). □

Remark 4.4. Each of the forbidden inclusions $e, -e \in \operatorname{lin}(K_1 - K_2)^\circ$ in assertion 2) of Theorem 4.3 is equivalent to the fact that $K_1 - K_2$ is contained in the hypersubspace S of the form (2). The latter happens if and only if a hyperplane of the form (1) separating K_1 and K_2 contains both these sets. Theorem 4.5 below gives a more explicit connection between these arguments. □

We will need the following properties (S1)–(S4) of affine spans and their direction spaces (see, e. g., [16, Chapter 1] for details): given nonempty set X and Y in $\mathbb{R}^n$, one has

(S1) If $o \in X$, then $\operatorname{aff} X = \operatorname{span} X = \operatorname{dir} X$.

(S2) $\operatorname{aff}(\lambda X + \mu Y) = \lambda \operatorname{aff} X + \mu \operatorname{aff} Y$ for any choice of scalars λ, μ.

(S3) $\operatorname{aff} Y \subset \operatorname{aff} X$ whenever $Y \subset X$.

(S4) $\operatorname{dir} X = \operatorname{span}(X - X) = \operatorname{aff} X - Y$ provided $Y \subset \operatorname{aff} X$.

Theorem 4.5. *For nonempty sets X_1 and X_2 in $\mathbb{R}^n$, the plane* $\operatorname{aff}(X_1 \cup X_2)$ *is a translate of the subspace* $\operatorname{span}(X_1 - X_2)$. *Equivalently,*

$$\operatorname{dir}(X_1 \cup X_2) = \operatorname{span}(X_1 - X_2). \tag{14}$$

Proof. If $X_1 = X_2$, then, by (S4),

$$\operatorname{dir}(X_1 \cup X_2) = \operatorname{dir} X_1 = \operatorname{span}(X_1 - X_1) = \operatorname{span}(X_1 - X_2).$$

So, we may assume that $X_1 \neq X_2$. Then $\operatorname{card}(X_1 \cup X_2) \geq 2$, and one can choose distinct points $u_1 \in X_1$ and $v_1 \in X_2$. Because $o = v_1 - v_1 \in X_1 \cup X_2 - X_2$, the property (S1) gives

$$\operatorname{aff}(X_1 \cup X_2 - X_2) = \operatorname{aff}(\{o\} \cup (X_1 \cup X_2 - X_2)) = \operatorname{span}(X_1 \cup X_2 - X_2).$$

Using this argument and properties (S2)–(S4), we obtain

$$\begin{aligned} \operatorname{span}(X_1 - X_2) \subset \operatorname{span}(X_1 \cup X_2 - X_2) &= \operatorname{aff}(X_1 \cup X_2 - X_2) \\ &= \operatorname{aff}(X_1 \cup X_2) - \operatorname{aff} X_2 = \operatorname{dir}(X_1 \cup X_2) \\ &= \operatorname{aff}(X_1 \cup X_2) - v_1 = \operatorname{aff}(X_1 \cup X_2 - v_1) \\ &= \operatorname{span}(X_1 \cup X_2 - v_1). \end{aligned}$$

So, to prove (14), it suffices to show that the subspaces

$$S = \operatorname{span}(X_1 - X_2) \quad \text{and} \quad T = \operatorname{span}(X_1 \cup X_2 - v_1)$$

coincide. Since $S \subset T$, the equality $S = T$ occurs if and only if $\dim S = \dim T$. Due to the obvious inequality $\dim S \le \dim T$, it remains to show that $\dim S \ge \dim T$. For this, let $k = \dim T$. Clearly, $k \ge 1$ because $\operatorname{card}(X_1 \cup X_2) \ge 2$. Let $z_1 = u_1 - v_1$ and choose in $X_1 \cup X_2 - v_1$ a basis $Z = \{z_1, z_1, \dots, z_k\}$ for T. We can write

$$z_i = u_i - v_1, \quad \text{where} \;\; u_i \in X_1 \cup X_2, \quad i = 2, \dots, k.$$

Renumbering the vectors $z_2, \dots, z_k$, we assume that the set $U_1 = \{u_1, \dots, u_p\}$ is in X_1 and the set $U_2 = \{u_{p+1}, \dots, u_k\}$ is in X_2 ($U_2 = \varnothing$ if $p = k$). We state that the set of vectors

$$u_1 - v_1,\; u_2 - v_1,\; \dots,\; u_p - v_1,\; u_1 - u_{p+1},\; \dots,\; u_1 - u_k \tag{15}$$

is linearly independent. Since the case $p = k$ is trivial (in this case the set (15) is Z), we will assume that $p < k$. So, let

$$\lambda_1(u_1 - v_1) + \lambda_2(u_2 - v_1) + \cdots + \lambda_p(u_p - v_1)$$
$$+ \lambda_{p+1}(u_1 - u_{p+1}) + \cdots + \lambda_k(u_1 - u_k) = o$$

for unknown scalars $\lambda_1, \dots, \lambda_k$. Rewriting this equation as

$$(\lambda_1 + \lambda_{p+1} + \cdots + \lambda_k)(u_1 - v_1) + \lambda_2(u_2 - v_1) + \cdots + \lambda_p(u_p - v_1)$$
$$-\lambda_{p+1}(u_{p+1} - v_1) - \cdots - \lambda_k(u_k - v_1) = o,$$

or, as

$$(\lambda_1 + \lambda_{p+1} + \cdots + \lambda_k)z_1 + \lambda_2 z_2 + \cdots + \lambda_p z_p$$
$$-\lambda_{p+1} z_{p+1} - \cdots - \lambda_k z_k = o,$$

and based on the linear independence of Z, we obtain that $\lambda_i = 0$ for all $i = 1, \dots, k$. Hence the set (15) is linearly independent. Since this set is contained in $X_1 - X_2$, it follows that $\dim S \ge k$, as desired. □

Theorem 4.6. *For convex set K_1 and K_2 in $\mathbb{R}^n$, the assertions below hold.*

(1) *There is a hyperplane strongly separating K_1 and K_2 if and only if we have $o \notin \operatorname{cl}(K_1 - K_2)$.*

(2) *If $o \notin \operatorname{cl}(K_1 - K_2)$ and e is a vector in $\mathbb{R}^n$ such that one of the vectors e and $-e$ belongs to $\operatorname{rint}(K_1 - K_2)^\circ$, then a suitable translate of the hypersubspace* (2) *strongly separates K_1 and K_2.*

(3) *If there is a translate of the hypersubspace* (2) *strongly separating K_1 and K_2, then one of the vectors e and $-e$ belongs to $(K_1 - K_2)^\circ \setminus \operatorname{lin}(K_1 - K_2)^\circ$. If, additionally, both K_1 and K_2 are bounded sets, then one of these vectors belongs to $\operatorname{rint}(K_1 - K_2)^\circ$.*

(4) *If $o \notin \operatorname{cl}(K_1 - K_2)$, then for any vector $e \in \operatorname{rint}(K_1 - K_2)^\circ$, the inequality*

$$\sup\{x \cdot e : x \in K_1\} < \inf\{x \cdot e : x \in K_2\} \tag{16}$$

holds. Consequently, any scalar γ satisfying the inequalities (12) *can be used in the description* (1) *of a hyperplane which strongly separates K_1 and K_2.*

Proof. (1) According to Proposition 1.1, K_1 and K_2 are strongly separated by a hyperplane if and only if $\delta(K_1, K_2) > 0$. The latter happens if and only if there is scalar $\rho > 0$ such that the open ρ-neighborhoods $U_\rho(K_1)$ and $U_\rho(K_2)$ are disjoint.

Equivalently, $$o \notin U_\rho(K_1) - U_\rho(K_2) = U_{2\rho}(K_1 - K_2),$$

which happens if and only if $o \notin \mathrm{cl}\,(K_1 - K_2)$.

(2) Let $o \notin \mathrm{cl}\,(K_1 - K_2)$ and e be a vector in $\mathbb{R}^n$ such that one of the vectors e and $-e$ belongs to $\mathrm{rint}\,(K_1 - K_2)^\circ$. According to Theorem 3.5, the hypersubspace (2) strongly bounds $K_1 - K_2$. Now, Lemma 3.1 implies that a suitable translate of this hypersubspace strongly separates K_1 and K_2.

(3) If a translate of the hypersubspace (2) strongly separates K_1 and K_2, then this translate properly separates K_1 and K_2. By Theorem 4.3, one of the vectors e and $-e$ belongs to $(K_1 - K_2)^\circ \setminus \mathrm{lin}\,(K_1 - K_2)^\circ$.

Suppose, additionally, that both K_1 and K_2 are bounded sets. Then the set $K_1 - K_2$ is also bounded. Because the hypersubspace (2) strongly bounds $K_1 - K_2$, Theorem 3.5 shows that one of the vectors e and $-e$ belongs to $\mathrm{rint}\,(K_1 - K_2)^\circ$.

(4) This assertion follows from Lemma 4.1 and the above parts (2) and (3). $\square$

5. Hyperplanes through given points

In this section, we describe all hyperplanes containing a given point $z \in \mathbb{R}^n$ and separating (properly separating) a pair of convex sets K_1 and K_2 in $\mathbb{R}^n$. For this, we reduce our argument to the separation of convex cones with apex z generated by these sets.

Lemma 5.1. *Convex sets K_1 and K_2 in $\mathbb{R}^n$ are separated (properly separated) by a hyperplane through a given point $z \in \mathbb{R}^n$ if and only if the generated cones $C_z(K_1)$ and $C_z(K_2)$ are separated (properly separated) by a hyperplane through z.*

Proof. (i) Assume first that K_1 and K_2 in $\mathbb{R}^n$ are separated by a hyperplane $H \subset \mathbb{R}^n$ through z. Denote by V_1 and V_2 the closed halfspaces determined by H and containing K_1 and K_2, respectively. If H has the form (1), then $z \cdot e = \gamma$. Without loss of generality, we may assume that V_1 and V_2 are given by (3). We assert that $C_z(K_i) \subset V_i$, $i = 1, 2$. Indeed, if $u \in C_z(K_1)$, then $u = z + \lambda(x - z)$ for a suitable scalar $\lambda \geq 0$ and a point $x \in K_1$. The inclusion $x \in K_1 \subset V_1$ gives $x \cdot e \leq \gamma$.

Therefore, $$u \cdot e = (z + \lambda(x - z)) \cdot e = \gamma + \lambda(x \cdot e - \gamma) \leq \gamma.$$

Hence $u \in V_1$, implying the inclusion $C_z(K_1) \subset V_1$. Similarly, $C_z(K_2) \subset V_2$. Thus H separates $C_z(K_1)$ and $C_z(K_2)$.

Conversely, if a hyperplane through z separates $C_z(K_1)$ and $C_z(K_2)$, then the inclusions $K_i \subset C_z(K_i)$, $i = 1, 2$, imply that H separates K_1 and K_2.

(ii) Assume now that K_1 and K_2 are properly separated by a hyperplane H. By the above argument, the cones $C_z(K_1)$ and $C_z(K_2)$ are separated by H. If, for instance, $K_1 \not\subset H$, then $C_z(K_1) \not\subset H$, implying that H properly separates $C_z(K_1)$ and $C_z(K_2)$.

Conversely, suppose that the cones $C_z(K_1)$ and $C_z(K_2)$ are properly separated by a hyperplane H through z. Then K_1 and K_2 are separated by H. Let, for instance, $C_z(K_1) \not\subset H$. We state that $K_1 \not\subset H$. Indeed, assume, for contradiction, that $K_1 \subset H$. Since any point $u \in C_z(K_1)$ can be written as an affine combination $u = z + \lambda(x - z)$ of points z and $x \in K_1 \subset H$, we should obtain the inclusion $u \in H$, and thus $C_z(K_1) \subset H$, contrary to the assumption. Hence $K_1 \not\subset H$, implying that H properly separates K_1 and K_2. □

Theorem 5.2. *Given convex sets K_1 and K_2 in $\mathbb{R}^n$ and a point $z \in \mathbb{R}^n$, assume $D_1 = C_z(K_1) - z$ and $D_2 = C_z(K_2) - z$. The following assertions hold.*

(1) *There is a hyperplane through z separating K_1 and K_2 if and only if the cones D_1 and D_2 satisfy any of the following three conditions:*
   (a) $o \notin \operatorname{int}(D_1 - D_2)$,
   (b) $(D_1 - D_2)^\circ \neq \{o\}$,
   (c) $D_1^\circ \cap (-D_2^\circ) \neq \{o\}$.

(2) *There is a hyperplane through z properly separating K_1 and K_2 if and only if the cones D_1 and D_2 satisfy any of the following three conditions:*
   (d) $o \notin \operatorname{rint}(D_1 - D_2)$,
   (e) $(D_1 - D_2)^\circ$ *is not a subspace,*
   (f) $D_1^\circ \cap (-D_2^\circ)$ *is not a subspace.*

Proof. (1) Assume that a hyperplane $H \subset \mathbb{R}^n$ through z separates K_1 and K_2. According to Lemma 5.1, H separates the cones $C_z(K_1)$ and $C_z(K_2)$. By Theorem 4.2, both conditions (a) and (b) are satisfied. The equivalence of conditions (b) and (c) follows from (P5):

$$(D_1 - D_2)^\circ = (D_1 \cup (-D_2))^\circ = D_1^\circ \cap (-D_2)^\circ = D_1^\circ \cap (-D_2^\circ).$$

Conversely, if any of the conditions (a)–(c) is satisfied, then Theorem 4.2 shows that the cones $C_z(K_1)$ and $C_z(K_2)$ are separated by a hyperplane. Finally, Lemma 5.1 implies that K_1 and K_2 are separated by a hyperplane.

(2) This part is similar to the above one and uses Theorem 4.3. □

Theorem 5.2 gives an important corollary (see below) on separation of convex sets whose closures have nonempty intersection. We recall that, given a convex set $K \subset \mathbb{R}^n$ and a point $z \in \operatorname{cl} K$, the normal cone of K at z is defined by any of the equalities below:

$$N_z(K) = (K - z)^\circ = C_o(K - z)^\circ = (C_z(K) - z)^\circ. \tag{17}$$

In view of Lemma 5.1, we can formulate the following corollary. We note that condition (c) in this corollary can be found in [11, Proposition 2.13].

Corollary 5.3. *Assume that the convex sets K_1 and K_2 in $\mathbb{R}^n$ satisfy the condition $\operatorname{cl} K_1 \cap \operatorname{cl} K_2 \neq \varnothing$, and let z be a point in $\operatorname{cl} K_1 \cap \operatorname{cl} K_2$. With $N_1 = N_z(K_1)$ and $N_2 = N_z(K_2)$, the assertions below hold.*

(1) *There is a hyperplane separating K_1 and K_2 if and only if any of the following three conditions is satisfied:*

(a) $o \notin \operatorname{int}(N_1 - N_2)$,

(b) $(N_1 - N_2)^\circ \neq \{o\}$,

(c) $N_1^\circ \cap (-N_2^\circ) \neq \{o\}$.

(2) *There is a hyperplane properly separating K_1 and K_2 if and only if any of the following three conditions is satisfied:*

(d) $o \notin \operatorname{rint}(N_1 - N_2)$,

(e) $(N_1 - N_2)^\circ$ *is not a subspace,*

(f) $N_1^\circ \cap (-N_2^\circ)$ *is not a subspace.*

Proof. Both assertions of Corollary 5.3 follow from Theorem 5.2 and the fact that any hyperplane separating K_1 and K_2 should contain z. □

Theorem 5.2 does not provide any information on the domain of points $z \in \mathbb{R}^n$ which belong to hyperplanes separating (properly separating) a given pair of convex sets. The description of this domain is given in Theorem 5.5 below. For a pair of disjoint convex sets K_1 and K_2, we let

$$F(K_1, K_2) = \{(1-\mu)x_1 + \mu x_2 : \mu > 1, x_1 \in K_1, x_2 \in K_2\}. \tag{18}$$

(Replacing in (18) the inequality $\mu > 1$ with $\mu \geq 1$, we obtain the definition of *penumbra* of K_2 with respect to K_1, introduced by Rockafellar [13, p. 22]).

Theorem 5.4. *For a pair of disjoint convex sets K_1 and K_2 in $\mathbb{R}^n$, the assertions below hold.*

(1) *The set $F(K_1, K_2)$ is convex, and $K_2 \subset \operatorname{cl} F(K_1, K_2)$.*

(2) $\operatorname{aff} F(K_1, K_2) = \operatorname{aff} F(K_2, K_1) = \operatorname{aff}(K_1 \cup K_2)$.

(3) *The sets $F(K_1, K_2)$ and $F(K_2, K_1)$ are disjoint.*

(4) *A hyperplane $H \subset \mathbb{R}^n$ separates (properly separates) K_1 and K_2 if and only if H separates (properly separates) $F(K_1, K_2)$ and $F(K_2, K_1)$.*

Proof. (1) For the inclusion $K_2 \subset \operatorname{cl} F(K_1, K_2)$, choose any points $x_1 \in K_1$ and $x_2 \in K_2$ and let $u(\mu) = (1-\mu)x_1 + \mu x_2$, where $\mu > 1$. Then $u(\mu) \in F(K_1, K_2)$ and $u(\mu) \to x_2$ as $\mu \to 1$. Hence $x_2 \in \operatorname{cl} F(K_1, K_2)$. Summing up, $K_2 \subset \operatorname{cl} F(K_1, K_2)$.

For the convexity of $F(K_1, K_2)$, choose any points $u, v \in F(K_1, K_2)$ and a scalar $\lambda \in [0, 1]$. Then

$$u = (1-\eta)x_1 + \eta x_2 \quad \text{and} \quad v = (1-\theta)y_1 + \theta y_2$$

for suitable points $x_1, y_1 \in K_1$, $x_2, y_2 \in K_2$ and scalars $\eta, \theta > 1$. One has

$$\begin{aligned}(1-\lambda)u + \lambda v &= (1-\lambda)((1-\eta)x_1 + \eta x_2) + \lambda((1-\theta)y_1 + \theta y_2) \\ &\in (1-\lambda)((1-\eta)K_1 + \eta K_2) + \lambda((1-\theta)K_1 + \theta K_2) \\ &= (1-\lambda)(1-\eta)K_1 + \lambda(1-\theta)K_1 + (1-\lambda)\eta K_2 + \lambda\theta K_2.\end{aligned}$$

Put $\gamma = (1-\lambda)\eta + \lambda\theta$. Clearly, $\gamma > 1$. Since $\xi K + \gamma K = (\xi + \gamma)K$ whenever K is convex and $\xi, \gamma \geq 0$ (see, e. g., [13, Theorem 3.2]), we have

$$\begin{aligned}
&(1-\lambda)(1-\eta)K_1 + \lambda(1-\theta)K_1 \\
&\quad = -((1-\lambda)(\eta-1)K_1 + \lambda(\theta-1)K_1) \\
&\quad = -((1-\lambda)(\eta-1) + \lambda(\theta-1))K_1 \\
&\quad = (1-\gamma)K_1.
\end{aligned}$$

Similarly, $$(1-\lambda)\eta K_2 + \lambda\theta K_2 = ((1-\lambda)\eta + \lambda\theta)K_2 = \gamma K_2.$$

Therefore,

$$\begin{aligned}
(1-\lambda)u + \lambda v \in (1-\gamma)K_1 + \gamma K_2 & \\
&= \{(1-\gamma)x_1 + \gamma x_2 : x_1 \in K_1, x_2 \in K_2\} \\
&\subset \{(1-\mu)x_1 + \mu x_2 : \mu > 1, x_1 \in K_1, x_2 \in K_2\} \\
&= F(K_1, K_2),
\end{aligned}$$

which proves the convexity of $F(K_1, K_2)$.

(2) First, we will prove the equality

$$\operatorname{aff} F(K_1, K_2) = \operatorname{aff}(K_1 \cup K_2).$$

Since any point $x \in F(K_1, K_2)$ can be written as an affine combination

$$x = (1-\mu)x_1 + \mu x_2, \quad \mu > 1,\ x_1 \in K_1,\ x_2 \in K_2,$$

the inclusion $F(K_1, K_2) \subset \operatorname{aff}(K_1 \cup K_2)$ is obvious. Consequently,

$$\operatorname{aff} F(K_1, K_2) \subset \operatorname{aff}(K_1 \cup K_2).$$

For the opposite inclusion, chose points $x_1 \in K_1$ and $x_2 \in K_2$. Put $z = -x_1 + 2x_2$. Clearly, $z = (1-2)x_1 + 2x_2 \in F(K_1, K_2)$. Because $K_2 \subset \operatorname{cl} F(K_1, K_2)$, one has, by (4),

$$\begin{aligned}
x_1 &= 2x_2 - (-x_1 + 2x_2) = 2x_2 + (1-2)z \\
&\in \operatorname{aff}(K_2 \cup F(K_1, K_2)) \subset \operatorname{aff}(K_2 \cup \operatorname{cl} F(K_1, K_2)) \\
&= \operatorname{aff}(\operatorname{cl} F(K_1, K_2)) = \operatorname{aff} F(K_1, K_2).
\end{aligned}$$

Summing up, $K_1 \subset \operatorname{aff} F(K_1, K_2)$. Hence $K_1 \cup K_2 \subset \operatorname{aff} F(K_1, K_2)$, which gives the desired inclusion

$$\operatorname{aff}(K_1 \cup K_2) \subset \operatorname{aff} F(K_1, K_2).$$

In a similar way,

$$\operatorname{aff} F(K_2, K_1) = \operatorname{aff}(K_2 \cup K_1) = \operatorname{aff}(K_1 \cup K_2).$$

(3) Assume, for contradiction, the existence of a point

$$u \in F(K_1, K_2) \cap F(K_2, K_1).$$

Then $$u = (1-\lambda)x_1 + \lambda x_2 = (1-\mu)y_2 + \mu y_1$$

for suitable scalars $\lambda, \mu > 1$ and points $x_1, y_1 \in K_1$ and $x_2, y_2 \in K_2$. Clearly,

$$x_2 = (1-\lambda^{-1})x_1 + \lambda^{-1}u, \quad y_1 = (1-\mu^{-1})y_2 + \mu^{-1}u. \tag{19}$$

Now, put $$v = \frac{\lambda-1}{\lambda+\mu-1}\, x_1 + \frac{\mu-1}{\lambda+\mu-1}\, y_2 + \frac{1}{\lambda+\mu-1}\, u,$$
$$\xi = \frac{\mu}{\lambda+\mu-1}, \quad \eta = \frac{\lambda}{\lambda+\mu-1}.$$

Obviously, $\xi, \eta \in [0,1]$. Based on (19) and the convexity of K_1, one has

$$\begin{aligned} v &= \Big(1 - \frac{\mu}{\lambda+\mu-1}\Big)x_1 + \frac{\mu}{\lambda+\mu-1}\Big(\frac{\mu-1}{\mu}\, y_2 + \frac{1}{\mu}u\Big) \\ &= (1-\xi)x_1 + \xi\big((1-\mu^{-1})y_2 + \mu^{-1}u\big) \\ &= (1-\xi)x_1 + \xi y_1 \in K_1. \end{aligned}$$

Similarly, $$\begin{aligned} v &= \Big(1 - \frac{\lambda}{\lambda+\mu-1}\Big)y_2 + \frac{\lambda}{\lambda+\mu-1}\Big(\frac{\lambda-1}{\lambda}\, x_1 + \frac{1}{\lambda}\, u\Big) \\ &= (1-\eta)y_2 + \eta\big((1-\lambda^{-1})x_1 + \lambda^{-1}u\big) \\ &= (1-\eta)y_2 + \eta x_2 \in K_2. \end{aligned}$$

Thus $v \in K_1 \cap K_2$, in contradiction with $K_1 \cap K_2 = \varnothing$.

(4) Assume first that a hyperplane $H \subset \mathbb{R}^n$ separates K_1 and K_2. Denote by V_1 and V_2 closed halfspaces determined by H and containing K_1 and K_2, respectively. Without loss of generality, we may suppose that V_1 and V_2 are given by (3) for a suitable nonzero vector $e \in \mathbb{R}^n$ and a scalar γ. We assert that $F(K_1, K_2) \subset V_2$ and $F(K_2, K_1) \subset V_1$.

Indeed, choose any point $u \in F(K_1, K_2)$. Then $u = (1-\mu)x_1 + \mu x_2$ for a suitable scalar $\mu > 1$ and points $x_1 \in K_1$ and $x_2 \in K_2$. Since $x_1 \in V_1$ and $x_2 \in V_2$, one has $x_1 \cdot e \le \gamma$ and $x_2 \cdot e \ge \gamma$. Therefore,

$$u\cdot e = (1-\mu)x_1\cdot e + \mu x_2\cdot e \ge (1-\mu)\gamma + \mu\gamma = \gamma.$$

Hence $u \in V_2$, and thus $F(K_1, K_2) \subset V_2$. Similarly, $F(K_2, K_1) \subset V_1$, which implies that H separates $F(K_1, K_2)$ and $F(K_2, K_1)$.

Conversely, assume that H separates the sets $F(K_1, K_2)$ and $F(K_2, K_1)$. Denote by V_1 and V_2 the closed halfspaces determined by H and containing $F(K_2, K_1)$ and $F(K_1, K_2)$, respectively. Since both sets V_1 and V_2 are closed, assertion (1) gives

$$K_1 \subset \operatorname{cl} F(K_2, K_1) \subset V_1 \quad \text{and} \quad K_2 \subset \operatorname{cl} F(K_1, K_2) \subset V_2.$$

Hence H separates K_1 and K_2.

Suppose, additionally, that K_1 and K_2 are properly separated by H. Then we obtain $K_1 \cup K_2 \not\subset H$, and assertion (1) implies that $\operatorname{cl} F(K_1, K_2) \cup \operatorname{cl} F(K_2, K_1) \not\subset H$. Since

H is a closed set, one has $F(K_1, K_2) \cup F(K_2, K_1) \not\subset H$. So, H properly separates $F(K_1, K_2)$ and $F(K_2, K_1)$.

Conversely, if H properly separates $F(K_1, K_2)$ and $F(K_2, K_1)$, then $K_1 \cup K_2 \not\subset H$. Indeed, otherwise, assertion (2) would give

$$F(K_1, K_2) \cup F(K_2, K_1) \subset \operatorname{aff} F(K_1, K_2) \cup \operatorname{aff} F(K_2, K_1) = \operatorname{aff}(K_1 \cup K_2) \subset H,$$

contrary to the assumption. □

Denote by $\mathcal{S}(K_1, K_2)$ (respectively, by $\mathcal{P}(K_1, K_2)$) the family of all hyperplanes separating (respectively, properly separating) convex sets K_1 and K_2. Put

$$S(K_1, K_2) = \cup (H : H \in \mathcal{S}(K_1, K_2)),$$
$$P(K_1, K_2) = \cup (H : H \in \mathcal{P}(K_1, K_2)).$$

Theorem 5.5. *Let K_1 and K_2 be convex sets in $\mathbb{R}^n$ such that* $\operatorname{rint} K_1 \cap \operatorname{rint} K_2 = \varnothing$.

Then $$P(K_1, K_2) = \mathbb{R}^n \setminus (F(\operatorname{rint} K_1, \operatorname{rint} K_2) \cup F(\operatorname{rint} K_2, \operatorname{rint} K_1)). \tag{20}$$

Furthermore, $$S(K_1, K_2) = \begin{cases} P(K_1, K_2) & \text{if} \quad \operatorname{aff}(K_1 \cup K_2) = \mathbb{R}^n, \\ \mathbb{R}^n & \text{if} \quad \operatorname{aff}(K_1 \cup K_2) \neq \mathbb{R}^n. \end{cases}$$

Proof. First, we assert that

$$F(\operatorname{rint} K_1, \operatorname{rint} K_2) \cap P(K_1, K_2) = \varnothing.$$

Indeed, let x be a point in $F(\operatorname{rint} K_1, \operatorname{rint} K_2)$, and H be a hyperplane properly separating K_1 and K_2. Assume, for contradiction, that $x \in H$. We express x as $x = (1-\gamma)x_1 + \gamma x_2$ for suitable $\gamma > 1$, $x_1 \in \operatorname{rint} K_1$, and $x_2 \in \operatorname{rint} K_2$. Since x_1 and x_2 belong to different closed halfspaces determined by H, the closed line segment $[x_1, x_2]$ meets H at some point y. Then $y = (1-\lambda)x_1 + \lambda x_2$ for a suitable scalar $\lambda \in [0, 1]$. Due to $\gamma \neq \lambda$, we have $x \neq y$. With both points x and y in H, the line l through x and y is contained in H. Obviously, $x_1, x_2 \in l \subset H$. The inclusion $x_i \in H \cap \operatorname{rint} K_i$ implies that $K_i \subset H$, $i = 1, 2$ (see [16, Theorem 9.3]). So, $K_1 \cup K_2 \subset H$, contrary to the assumption that H properly separates K_1 and K_2.

In a similar way,
$$F(\operatorname{rint} K_2, \operatorname{rint} K_1) \cap P(K_1, K_2) = \varnothing.$$

Therefore,

$$P(K_1, K_2) \subset \mathbb{R}^n \setminus (F(\operatorname{rint} K_1, \operatorname{rint} K_2) \cup F(\operatorname{rint} K_2, \operatorname{rint} K_1)).$$

For the opposite inclusion, choose any point

$$z \in \mathbb{R}^n \setminus (F(\operatorname{rint} K_1, \operatorname{rint} K_2) \cup F(\operatorname{rint} K_2, \operatorname{rint} K_1)) \tag{21}$$

and consider the generated cones $C_z(K_1)$ and $C_z(K_2)$. We assert that

$$\operatorname{rint} C_z(K_1) \cap \operatorname{rint} C_z(K_2) = \varnothing. \tag{22}$$

The property (P7) shows that (22) can be rewritten as

$$C_z(\operatorname{rint} K_1) \cap C_z(\operatorname{rint} K_2) = \{z\}. \tag{23}$$

To prove (23), assume, for contradiction, the existence of a point

$$u \in C_z(\operatorname{rint} K_1) \cap C_z(\operatorname{rint} K_2) \setminus \{z\}.$$

We can write $u = (1-\lambda_i)z + \lambda_i x_i$ for suitable scalars $\lambda_i > 0$ and points $x_i \in \operatorname{rint} K_i$, $i = 1, 2$. Then

$$x_i = (1-\lambda_i^{-1})z + \lambda_i^{-1}u \in (z, u\rangle, \quad i = 1, 2,$$

where $(z, u\rangle$ denotes the open halfline through u with endpoint z. Due to the relation $\operatorname{rint} K_1 \cap \operatorname{rint} K_2 = \varnothing$, the points x_1 and x_2 are distinct. So, one of them lies between z and the other one. Let, for instance, $x_2 \in (z, x_1)$. In this case, $x_2 = (1-\eta)x_1 + \eta z$, where $0 < \eta < 1$. Since $\eta^{-1} > 1$, we obtain

$$z = (1-\eta^{-1})x_1 + \eta^{-1}x_2 \in F(\operatorname{rint} K_1, \operatorname{rint} K_2),$$

contrary to the choice of z in (21). Summing up, the equality (22) holds.

By Proposition 1.1, there is a hyperplane $H \subset \mathbb{R}^n$ properly separating the cones $C_z(K_1)$ and $C_z(K_2)$. Since z belongs to both cones, it also belongs to H. Now, Lemma 5.1 shows that H properly separates K_1 and K_2. Thus $z \in H \subset P(K_1, K_2)$, which implies the desired inclusion

$$\mathbb{R}^n \setminus (F(\operatorname{rint} K_1, \operatorname{rint} K_2) \cup F(\operatorname{rint} K_2, \operatorname{rint} K_1)) \subset P(K_1, K_2).$$

For the description of $S(K_1, K_2)$, assume first that $\operatorname{aff}(K_1 \cup K_2) = \mathbb{R}^n$ and choose a hyperplane $H \subset \mathbb{R}^n$ separating K_1 and K_2. By Theorem 5.4, the convex sets $F(K_1, K_2)$ and $F(K_2, K_1)$ are separated by H and both have dimension n. Therefore, these sets are properly separated by H. The same theorem states that K_1 and K_2 are properly separated by H. This argument implies that $\mathcal{S}(K_1, K_2) = \mathcal{P}(K_1, K_2)$, and whence $S(K_1, K_2) = P(K_1, K_2)$.

Finally, let $\operatorname{aff}(K_1 \cup K_2) \neq \mathbb{R}^n$. Combining (4) and Theorem 5.4, we have

$$\begin{aligned} \operatorname{aff} F(\operatorname{rint} K_1, \operatorname{rint} K_2) &= \operatorname{aff} F(\operatorname{rint} K_2, \operatorname{rint} K_1) \\ &= \operatorname{aff}(\operatorname{rint} K_1 \cup \operatorname{rint} K_2) = \operatorname{aff}(K_1 \cup K_2). \end{aligned}$$

Therefore, (20) gives

$$\begin{aligned} &\mathbb{R}^n \setminus \operatorname{aff}(K_1 \cup K_2) \\ &= \mathbb{R}^n \setminus (\operatorname{aff} F(\operatorname{rint} K_1, \operatorname{rint} K_2) \cup \operatorname{aff} F(\operatorname{rint} K_2, \operatorname{rint} K_1)) \\ &\subset \mathbb{R}^n \setminus (F(\operatorname{rint} K_1, \operatorname{rint} K_2) \cup F(\operatorname{rint} K_2, \operatorname{rint} K_1)) \\ &= P(K_1, K_2) \subset S(K_1, K_2). \end{aligned} \tag{24}$$

If H is a hyperplane containing $\operatorname{aff}(K_1 \cup K_2)$, then H trivially separates K_1 and K_2. Hence $\operatorname{aff}(K_1 \cup K_2) \subset H \subset S(K_1, K_2)$. Combined with (24), the latter inclusion gives $S(K_1, K_2) = \mathbb{R}^n$. □

Acknowledgment. The author thanks the anonymous referee for helpful comments on the paper.

References

[1] H. Brunn: *Sätze über zwei getrennte Eikörper*, Math. Ann. 104 (1931) 300–324.

[2] C. Carathéodory: *Über den Variabilitätsbereich der Fourierschen Konstanten von positiven harmonischen Funktionen*, Rend. Circ. Mat. Palermo 32 (1911) 193–217.

[3] Z. R. Gabidullina: *A theorem on strict separability of convex polyhedra and its applications in optimization*, J. Optim. Theory Appl. 148 (2011) 550–570.

[4] A. I. Golikov, Yu. G. Evtushenko, S. Ketabchi: *On families of hyperplanes that separate polyhedra*, Comput. Math. Math. Phys. 45 (2005) 227–242.

[5] L. Grygarová: *A calculation of all separating hyperplanes of two convex polytopes*, Optimization 41 (1997) 57–69.

[6] L. Grygarová: *On a calculation of an arbitrary separating hyperplane of convex polyhedral sets*, Optimization 43 (1998) 93–112.

[7] V. L. Klee: *Separation and support properties of convex sets – a survey*, in: *Control Theory and the Calculus of Variations*, A. V. Balakrishnan (ed.), Academic Press, New York (1969) 235–303.

[8] J. Lawrence, V. Soltan: *On unions and intersections of nested families of cones*, Beitr. Algebra Geom. 57 (2016) 655–665.

[9] H. Minkowski: *Geometrie der Zahlen I*, Teubner, Leipzig (1896).

[10] H. Minkowski: *Gesammelte Abhandlungen, Band 2*, Teubner, Leipzig (1911).

[11] B. S. Mordukhovich, N. M. Nam: *An Easy Path to Convex Analysis and Applications*, Morgan & Claypool, Williston (2014).

[12] Y. Nievergelt: *Parameter spaces of separating hyperplanes*, Optimization 64 (2015) 1409–1438.

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

[14] V. Soltan: *Polarity and separation of cones*, Linear Algebra Appl. 538 (2018) 212–224.

[15] V. Soltan: *Moreau-type characterizations of polar cones*, Linear Algebra Appl. 567 (2019) 45–62.

[16] V. Soltan: *Lectures on Convex Sets*, 2nd edition, World Scientific, Hackensack (2020).