Separation theorems for convex polytopes and finitely-generated cones derived from theorems of the alternative

David Bartl

Department of Mathematics, Faculty of Science, University of Ostrava, 30. dubna 22, 701 03 Ostrava, Czech Republic

A R T I C L E   I N F O

Article history:
Received 25 November 2011
Accepted 13 December 2011
Available online 5 January 2012
Submitted by P. Šemrl

AMS classification:
52A05
15A39
15A06
12J15

Keywords:
Gordan’s Theorem
Motzkin’s Theorem
Stiemke’s Theorem
Tucker’s Theorem
Residual theorem for linear equations

A B S T R A C T

We derive from Motzkin’s Theorem that a point can be strongly separated by a hyperplane from a convex polytope and a finitely-generated convex cone. We state a similar result for Tucker’s Theorem of the alternative. A generalisation of the residual existence theorem for linear equations which has recently been proved by Rohn [8] is a corollary. We state all the results in the setting of a general vector space over a linearly ordered (possibly skew) field.

© 2011 Elsevier Inc. All rights reserved.

1. Introduction

Rohn [8, Theorem 2] has recently proved the residual existence theorem for linear equations:

Theorem 1 (Residual existence theorem for linear equations). Let a matrix $A \in \mathbb{R}^{r \times s}$, a point $b \in \mathbb{R}^r$, and a finite subset $X = \{x_1, \ldots, x_m\} \subseteq \mathbb{R}^s$ be given. Then the inequality

$$\max_{x_i \in X} p^T (Ax_i - b) \geq 0 \quad (1)$$

* Tel.: +420 59 709 2133.
E-mail address: bartl@osu.cz

0024-3795/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
holds for all \( p \in \mathbb{R}^T \) if and only if the system
\[
Ax = b
\] has a solution in the convex hull \( \text{conv} X \) of the set \( X \).

Rohn [8] proves that result using Gordan’s Theorem of the alternative [4]. Nonetheless, we know a generalisation of Gordan’s Theorem into the setting of a general vector space over a linearly ordered (even possibly skew) field [1, Theorem 5.1]. The question whether Theorem 1 can be generalised accordingly is therefore apparent.

It turns out however, that Rohn’s Theorem 1 is a separation theorem actually: either the point \( b \) lies in the polytope \( A(\text{conv} X) \), the image of the polytope \( \text{conv} X \) via \( A \), the linear mapping \( A : \mathbb{R}^3 \to \mathbb{R}^T \) with \( A : x \mapsto Ax \) induced by the matrix \( A \), cf. (2), or there exists a hyperplane that strongly separates the polytope \( A(\text{conv} X) \) and the point \( b \). To see the latter, note that the negation of (1) is equivalent to the fact that there exists a linear form \( \eta : \mathbb{R}^T \to \mathbb{R} \) with \( \eta : y \mapsto p^T y \) induced by the row vector \( p^T \), and a constant \( \varepsilon \in \mathbb{R} \) such that \( \eta(y) < \varepsilon < \eta(b) \) for all \( y \in A(\text{conv} X) \).

Gordan’s Theorem of the alternative, which is used to prove the result, is generalised by Motzkin’s Theorem [6, 7, Theorem D6, p. 60], cf. [10, Corollary 2A (ii)]. We can obtain yet a more general separation result thus. Stiemke’s Theorem [9] and Tucker’s Theorem of the alternative [10, Corollary 2A (i)] are, in a sense, dual to Gordan’s Theorem and Motzkin’s Theorem, respectively. Applying that theorem instead, we obtain another separation result. Finally, we can generalise Rohn’s Theorem 1 and obtain a new related result as a corollary.

2. Concepts and notation

Let \( F \) be a linearly ordered, possibly skew, field: the field either may or may not be commutative. The fields \( \mathbb{R} \) and \( \mathbb{Q} \) of the real and rational numbers, respectively, are examples of a commutative linearly ordered field. A scalar \( \lambda \in F \) is non-negative or positive iff \( \lambda \geq 0 \) or \( \lambda > 0 \), respectively.

Let \( W \) be a vector space over the linearly ordered field \( F \). As we shall work here with both left and right vector spaces, we state explicitly that we mean that \( W \) is a left vector space over \( F \). That is, vectors \( x \in W \) are multiplied by scalars \( \lambda \in F \) from the left. Likewise, when \( Z \) is a right vector space over the linearly ordered field \( F \), vectors \( \xi \in Z \) are multiplied by scalars \( \lambda \in F \) from the right.

Given a (left) vector space \( W \) over the field \( F \), let \( W^\ast \) denote the algebraic dual of \( W \), which is the space of all linear forms on \( W \). We note that \( W^\ast \) is a right vector space over the field \( F \). Let \( W^\circ \) be any subspace of \( W^\ast \) such that, for any non-zero vector \( x \in W \), there exists a linear form \( \alpha \in W^\ast \) with \( \alpha(x) \neq 0 \). (For example, if \( W \) is a real Banach space, then \( W^\ast \) can be its topological dual, the space of all continuous linear functionals on \( W \).)

Then \( (W, W^\circ) \) is a dual pair of spaces. Let \( \langle \cdot, \cdot \rangle \) denote the canonical pairing, which means that we put \( \langle x, \alpha \rangle = \alpha(x) \) for any \( x \in W \) and \( \alpha \in W^\circ \).

Let \( Z \) be a right vector space over \( F \) and let \( A : Z \to F \) be a (right) linear form on \( Z \). Note that a scalar \( \lambda \in F \) induces a right linear form \( \hat{\lambda} : F \to F \) with \( \hat{\lambda} : \xi \mapsto \lambda \xi \) for any \( \xi \in F \). Therefore, the right linear form \( A : Z \to F \) and that \( \hat{\lambda} : F \to F \) can be composed. The resulting right linear form will be denoted by \( \lambda A \). We have \( \langle \lambda A(\xi) \rangle = \lambda \langle A(\xi) \rangle \) for all \( \xi \in Z \). Let \( O \) denote the zero linear form \( O : Z \to F \) with \( O : \xi \mapsto 0 \) for all \( \xi \in Z \).

In the sequel, we shall need Motzkin’s Theorem [6,7, Theorem D6, p. 60], cf. [10, Corollary 2A (ii)] and Tucker’s Theorem [10, Corollary 2A (i)]. The next result states these theorems as part (i) and part (ii), respectively. We obtain Gordan’s Theorem [4] and Stiemke’s Theorem [9] as a special case of part (i), Motzkin’s Theorem, and part (ii), Tucker’s Theorem, respectively, by putting \( n = 0 \), i.e., vanishing the system \( B_1(\xi), \ldots, B_n(\xi) \leq 0 \), in the result.

**Theorem 2.** Let \( Z \) be a right vector space over a linearly ordered (possibly skew) field \( F \). Let \( A_1, \ldots, A_m : Z \to F \) and \( B_1, \ldots, B_n : Z \to F \) be (right) linear forms. Then:

(i) Motzkin’s Theorem. There does not exist any \( \xi \in Z \) such that \( A_1(\xi), \ldots, A_m(\xi) < 0 \) and \( B_1(\xi), \ldots, B_n(\xi) \leq 0 \) if and only if \( \lambda_1 A_1 + \cdots + \lambda_m A_m + \mu_1 B_1 + \cdots + \mu_n B_n = 0 \) for
3. Separation based on Motzkin’s Theorem

We say that a set \( P \subseteq W \) is a polytope iff it is a convex hull, \( P = \text{conv} \, X \), of a finite set of points \( X = \{x_1, \ldots, x_m\} \subseteq W \).

Let a polytope \( P \subseteq W \) and a point \( x \in W \) be given. Assuming that \( x \) is not in \( P \), we ask whether the point \( x \) and the polytope \( P \) can be strongly separated by a hyperplane. That is, we seek for a linear form \( \alpha \in W^* \) and a constant \( \varepsilon \in F \) such that \( \alpha(p) < \varepsilon < \alpha(x) \) for all \( p \in P \). It turns out that we can answer yet a more general question.

We say that a set \( C \subseteq W \) is a finitely-generated cone iff it is a (convex) conical hull, \( C = \text{cone} \, Y \), of a finite set of points \( Y = \{y_1, \ldots, y_n\} \subseteq W \). The Minkowski sum of the polytope \( P \) and cone \( C \) is the set \( P + C = \{p + c : p \in P, c \in C\} \).

Given yet a point \( x \in W \) which is not in \( P + C \), we ask if the point \( x \) and the set \( P + C \) can be strongly separated by a hyperplane, i.e., whether there exists a linear form \( \alpha \in W^* \) and a constant \( \varepsilon \in F \) such that \( \alpha(p + c) < \varepsilon < \alpha(x) \) for all \( p \in P \) and \( c \in C \).

Lemma 1. Let \( P = \text{conv} \{x_1, \ldots, x_m\} \) and \( C = \text{cone} \{y_1, \ldots, y_n\} \) be a polytope and a finitely-generated cone, respectively, in the vector space \( W \). Let \( x \in W \) be a point. Then there exists an \( \alpha \in W^* \) such that \( \alpha(x_1), \ldots, \alpha(x_m) < \alpha(x) \) and \( \alpha(y_1), \ldots, \alpha(y_n) \leq 0 \) if and only if the point \( x \) and the set \( P + C \) can be strongly separated by a hyperplane.

Proof. There exists an \( \alpha \in W^* \) such that \( \alpha(x_1), \ldots, \alpha(x_m) < \alpha(x) \) and \( \alpha(y_1), \ldots, \alpha(y_n) \leq 0 \) if and only if there exists a linear form \( \alpha \in W^* \) and a constant \( \varepsilon \in F \) such that \( \alpha(x_1), \ldots, \alpha(x_m) \prec \varepsilon < \alpha(x) \) and \( \alpha(y_1), \ldots, \alpha(y_n) \leq 0 \).

So \( \alpha(x_1) - \varepsilon, \ldots, \alpha(x_m) - \varepsilon < 0 \) and \( \alpha(y_1), \ldots, \alpha(y_n) \leq 0 \). By Proposition 1 (i), we equivalently have that \( \lambda_1 (\alpha(x_1) - \varepsilon) + \cdots + \lambda_m (\alpha(x_m) - \varepsilon) + \mu_1 \alpha(y_1) + \cdots + \mu_n \alpha(y_n) < 0 \) for all non-negative \( \lambda_1, \ldots, \lambda_m \in F \) with \( \lambda_1 + \cdots + \lambda_m = 1 \) and for all non-negative \( \mu_1, \ldots, \mu_n \in F \).

Hence, equivalently, we have that \( \alpha(x_1) + \cdots + \lambda_m x_m + \mu_1 y_1 + \cdots + \mu_n y_n < \varepsilon \) for all non-negative \( \lambda_1, \ldots, \lambda_m \in F \) with \( \lambda_1 + \cdots + \lambda_m = 1 \) and for all non-negative \( \mu_1, \ldots, \mu_n \in F \). Recalling that \( \varepsilon < \alpha(x) \), we equivalently have that \( \alpha(p + c) < \varepsilon < \alpha(x) \) for all \( p \in P \) and \( c \in C \), which equivalently means that the point \( x \) and the set \( P + C \) can be strongly separated by a hyperplane. □
Theorem 3. Let \( P = \text{conv}\{x_1, \ldots, x_m\} \) and \( C = \text{cone}\{y_1, \ldots, y_n\} \) be a polytope and a finitely-generated cone, respectively, in the vector space \( W \). Let \( x \in W \) be a point. Then \( x \notin P + C \) if and only if the point \( x \) and the set \( P + C \) can be strongly separated by a hyperplane.

Proof. We have \( x \notin P + C \) if and only if \( x \neq \lambda_1 x_1 + \cdots + \mu_m x_m + \lambda_1 y_1 + \cdots + \mu_n y_n \) or

\[
\lambda_1 (x_1 - x) + \cdots + \lambda_m (x_m - x) + \mu_1 y_1 + \cdots + \mu_n y_n \neq 0
\]

for all non-negative \( \lambda_1, \ldots, \lambda_m \in F \) with \( \lambda_1 + \cdots + \lambda_m = 1 \) and for all non-negative \( \mu_1, \ldots, \mu_n \in F \). Equivalently, by Motzkin’s Theorem 2 (i), there exists an \( \alpha \in W^* \), which is variable on the right vector space \( W^* \), such that

\[
\langle x_1 - x, \alpha \rangle, \ldots, \langle x_m - x, \alpha \rangle < 0 \quad \text{and} \quad \langle y_1, \alpha \rangle, \ldots, \langle y_n, \alpha \rangle \leq 0
\]

or \( \alpha(x_1), \ldots, \alpha(x_m) < \alpha(x) \) and \( \alpha(y_1), \ldots, \alpha(y_n) \leq 0 \). Lemma 1 finishes the proof. \( \square \)

The next corollary follows from Theorem 3 in a standard way. If \( n' = n'' = 0 \) in the corollary, then Gordan’s Theorem of the alternative [4] is sufficient to prove it and the result yields a separation theorem for two convex polytopes. If \( m'' = 1 \) and \( n'' = 0 \), then the corollary coincides with last Theorem 3.

Corollary 1. Let \( P' = \text{conv}\{x'_1, \ldots, x'_m\} \) with \( P'' = \text{conv}\{y'_1, \ldots, y'_n\} \) and \( C' = \text{cone}\{y'_1, \ldots, y'_n\} \) with \( C'' = \text{cone}\{y''_1, \ldots, y''_m\} \) be two polytopes and two finitely-generated cones, respectively, in the vector space \( W \). Then \( (P' + C') \cap (P'' + C'') = \emptyset \) if and only if the sets \( P' + C' \) and \( P'' + C'' \) can be strongly separated by a hyperplane, i.e., if and only if there exists a linear form \( \alpha \in W^* \) and constants \( \epsilon', \epsilon'' \in F \) such that \( \alpha(x' + c') \leq \epsilon' < \epsilon'' \leq \alpha(x'' + c'') \) for all \( x' \in P' \) and \( c' \in C' \) with \( x'' \in P'' \) and \( c'' \in C'' \).

Proof. We obviously have \((P' + C') \cap (P'' + C'') = \emptyset \) if and only if 0 \( \notin (P' + C') - (P'' + C'') = (P' - P'') + (C' - C'') = P + C \) where \( P = \{p' - p'' : p' \in P', p'' \in P''\} \) and \( C = \{c' - c'' : c' \in C', c'' \in C''\} \).

It is an exercise to show that \( P = \text{conv}\{x'_i - x''_i : i' \in \{1, \ldots, m'\}, i'' \in \{1, \ldots, m''\}\} \) is a convex polytope and that \( C = \text{cone}\{y'_i, \ldots, y''_i\} \) is a centrally-generated convex cone.

Equivalently, by Theorem 3, there exists a linear form \( \alpha \in W^* \) and a constant \( \epsilon \in F \) so that \( \alpha(x' + c' - y'' - c'') < \epsilon < 0 \) for all \( x' \in P' \) and \( c' \in C' \) with \( p'' \in P'' \) and \( c'' \in C'' \). Necessarily, we have \( \alpha(c') \leq 0 \) and \( \alpha(c'') \geq 0 \) for all \( c' \in C' \) and \( c'' \in C'' \). Hence, considering \( c' = c'' = 0 \) and putting \( \epsilon' = \max_{i=1}^{m'} \alpha(x'_i) \) and \( \epsilon'' = \min_{i=1}^{m''} \alpha(y''_i) \), we obviously have \( \epsilon' - \epsilon'' < \epsilon < 0 \). It follows that \( \alpha(x' + c') \leq \epsilon' < \epsilon'' \leq \alpha(x'' + c'') \) for all \( x' \in P' \) and \( c' \in C' \) with \( x'' \in P'' \) and \( c'' \in C'' \), which means we are done. \( \square \)

4. Separation based on Tucker’s Theorem

We say that a subset \( M \) of a vector space \( W \) over a linearly ordered (possibly skew) field \( F \) is relatively absorbing at a point \( x \in M \), or that the point \( x \) is in the relative algebraic interior of \( M \), writing \( x \in \text{rel alg int } M \), if and only if, for any \( y \) from the affine hull of \( M \), there exists a positive \( \varepsilon \in F \) such that \((1 - \lambda) x + \lambda y \in M \) for all \( \lambda \in F \) between \( -\varepsilon \) and \( \varepsilon \), i.e., such that \( -\varepsilon < \lambda < \varepsilon \). See [3, Definitions II.1.5 and III.1.6, p. 45 and 109], [5, Definition 1.10, p. 34]. The next lemma is an exercise, so we omit its proof here.

Lemma 2. Let \( P = \text{conv}\{x_1, \ldots, x_m\} \) be a polytope in the vector space \( W \). Then \( x \in \text{rel alg int } P \) if and only if \( x = \lambda_1 x_1 + \cdots + \lambda_m x_m \) for some positive \( \lambda_1, \ldots, \lambda_m \in F \) with \( \lambda_1 + \cdots + \lambda_m = 1 \).

Now, let \( P \subseteq W \) and \( C \subseteq W \) be a polytope and a centrally-generated cone, respectively. Consider the relative algebraic interior of \( P \) and add (in the Minkowski sense) the whole cone \( C \) to it. Given a point \( x \in W \) which is not in the sum, we ask whether the point \( x \) and the set \( \text{rel alg int } P + C \) can be semi-
strictly separated by a hyperplane. That is, we seek for a linear form \( \alpha \in W^* \) such that \( \alpha(p + c) < \alpha(x) \) for all \( p \in \text{rel alg int } P \) and \( c \in C \).

**Corollary 1.** Let \( P = \text{conv}\{x_1, \ldots, x_m\} \) and \( C = \text{cone}\{y_1, \ldots, y_n\} \) be a polytope and a finitely-generated cone, respectively, in the vector space \( W \). Let \( x \in W \) be a point. Then there exists an \( \alpha \in W^* \) such that \( \alpha(x_1), \ldots, \alpha(x_m) \leq \alpha(x) \), not all \( \alpha(x_1), \ldots, \alpha(x_m) = \alpha(x) \), and \( \alpha(y_1), \ldots, \alpha(y_n) \leq 0 \) if and only if the point \( x \) and the set \( \text{rel alg int } P + C \) can be semi-strictly separated by a hyperplane.

**Lemma 3.** Let \( P = \text{conv}\{x_1, \ldots, x_m\} \) and \( C = \text{cone}\{y_1, \ldots, y_n\} \) be a polytope and a finitely-generated cone, respectively, in the vector space \( W \). Let \( x \in W \) be a point. Then there exists an \( \alpha \in W^* \) such that \( \alpha(x_1), \ldots, \alpha(x_m) \leq \alpha(x) \), not all \( \alpha(x_1), \ldots, \alpha(x_m) = \alpha(x) \), and \( \alpha(y_1), \ldots, \alpha(y_n) \leq 0 \) if and only if the point \( x \) and the set \( \text{rel alg int } P + C \) can be semi-strictly separated by a hyperplane.

**Theorem 4.** Let \( P = \text{conv}\{x_1, \ldots, x_m\} \) and \( C = \text{cone}\{y_1, \ldots, y_n\} \) be a polytope and a finitely-generated cone, respectively, in the vector space \( W \). Let \( x \in W \) be a point. Then \( x \notin \text{rel alg int } P + C \) if and only if the point \( x \) and the set \( \text{rel alg int } P + C \) can be semi-strictly separated by a hyperplane.

**Corollary 2.** Let \( P' = \text{conv}\{x'_1, \ldots, x'_{m'}\} \) with \( P'' = \text{conv}\{x''_1, \ldots, x''_{m''}\} \) and \( C' = \text{cone}\{y'_1, \ldots, y'_{n'}\} \) with \( C'' = \text{cone}\{y''_1, \ldots, y''_{n''}\} \) be two polytopes and two finitely-generated cones, respectively, in the vector space \( W \). Then \( (\text{rel alg int } P' + C') \cap (\text{rel alg int } P'' + C'') = \emptyset \) if and only if the sets \( \text{rel alg int } P' + C' \) and \( \text{rel alg int } P'' + C'' \) can be strictly separated by a hyperplane, i.e., if and only if there exists a linear form \( \alpha \in W^* \) and a constant \( \varepsilon \in \mathbb{F} \) such that \( \alpha(p' + c') < \varepsilon < \alpha(p'' + c'') \) for all \( p' \in \text{rel alg int } P' \) and \( c' \in C' \), \( p'' \in \text{rel alg int } P'' \) and \( c'' \in C'' \).

Lemma 3, Theorem 4, and Corollary 2 can be proved analogously as Lemma 1, Theorem 3, and Corollary 1, respectively, taking Lemma 2 into account. That is why we omit the proofs here.

5. A generalisation of the residual existence theorem for linear equations

Let \( A : W \to H \) be a linear operator (i.e., mapping) where both \( W \) and \( H \) are (left) vector spaces over a linearly ordered (possibly skew) field \( F \).

Let \( H^* \) be any subspace of the algebraic dual \( H^# \) of the space \( H \) such that, for any non-zero vector \( h \in H \), there exists a linear form \( \eta \in H^* \) with \( \eta(h) \neq 0 \). It follows \( (H, H^*) \) is a dual pair of spaces.

The following result is a generalisation of Rohn’s residual existence theorem for linear equations [8, Theorem 2], see Theorem 1 above.

**Theorem 5.** Let a linear mapping \( A : W \to H \), a point \( b \in H \), and finite subsets \( X = \{x_1, \ldots, x_m\} \subseteq W \) and \( Y = \{y_1, \ldots, y_n\} \subseteq W \) be given. Then the linear equation

\[
Ax = b
\]

has a solution in the set \( \text{conv } X + \text{cone } Y \) if and only if

\[
\eta(Ay_1), \ldots, \eta(Ay_n) \leq 0 \quad \text{implies} \quad \max_{x_i \in X} \eta(Ax_i - b) \geq 0
\]

for all \( \eta \in H^* \), which holds if and only if there does not exist any hyperplane strongly separating the set \( A(\text{conv } X + \text{cone } Y) \) and the point \( b \).

**Proof.** Note first that the linear image \( A(\text{conv } X) \) of the polytope \( \text{conv } X \) is a polytope in the space \( H \). More generally, note that \( A(\text{conv } X + \text{cone } Y) = \text{conv } A(X) + \text{cone } A(Y) \), so that the linear image of the Minkowski sum of a polytope and finitely-generated cone yields again a set which is the Minkowski sum of a polytope and finitely-generated cone in the space \( H \).

The linear equation \( Ax = b \) has a solution in the set \( \text{conv } X + \text{cone } Y \) if and only if the point \( b \) is in the set \( A(\text{conv } X + \text{cone } Y) \). By Theorem 3, equivalently, the set \( A(\text{conv } X + \text{cone } Y) \) and the point \( b \) cannot be strongly separated by a hyperplane. By Lemma 1, equivalently, there does not exist any \( \eta \in H^* \) such that \( \eta(Ax_1), \ldots, \eta(Ax_m) < \eta(b) \), or \( \eta(Ax_1 - b), \ldots, \eta(Ax_m - b) < 0 \), and \( \eta(Ay_1), \ldots, \eta(Ay_n) \leq 0 \), which equivalently means that \( \eta(Ay_1), \ldots, \eta(Ay_n) \leq 0 \) implies \( \max_{x_i \in X} \eta(Ax_i - b) \geq 0 \) for all \( \eta \in Y^* \). □
Taking Lemma 2 into account, the proof of next Theorem 6 is analogous to that of Theorem 5. We omit it therefore.

**Theorem 6.** Let a linear mapping \( A : W \to H \), a point \( b \in H \), and finite subsets \( X = \{x_1, \ldots, x_m\} \subseteq W \) and \( Y = \{y_1, \ldots, y_n\} \subseteq W \) be given. Then the linear equation

\[
Ax = b
\]

has a solution in the set \( \text{rel alg int conv } X + \text{cone } Y \) if and only if

\[
\eta(Ay_1), \ldots, \eta(Ay_n) \leq 0 \quad \text{and} \quad \max_{x_i \in X} \eta(Ax_i - b) \leq 0
\]

implies

\[
\eta(Ax_1 - b), \ldots, \eta(Ax_m - b) = 0
\]

for all \( \eta \in H^* \), which holds if and only if there does not exist any hyperplane semi-strictly separating the set \( A(\text{rel alg int conv } X + \text{cone } Y) \) and the point \( b \).

**Acknowledgement**

The author wishes to thank the anonymous referee for useful suggestions, which helped to improve the manuscript.

**References**


