ELSEVIER

Discrete Applied Mathematics 89 (1998) 3–44

DISCRETE
APPLIED
MATHEMATICS

Disjunctive programming: Properties of the convex hull of feasible points ☆

Egon Balas*

Graduate School of Industrial Administration, Carnegie Mellon University, Pittsburgh, PA 15213, USA

Abstract

In this paper we characterize the convex hull of feasible points for a disjunctive program, a class of problems which subsumes pure and mixed integer programs and many other nonconvex programming problems. Two representations are given for the convex hull of feasible points, each of which provides linear programming equivalents of the disjunctive program. The first one involves a number of new variables proportional to the number of terms in the disjunctive normal form of the logical constraints; the second one involves only the original variables and the facets of the convex hull. Among other results, we give necessary and sufficient conditions for an inequality to define a facet of the convex hull of feasible points. For the class of disjunctive programs that we call facial, we establish a property which makes it possible to obtain the convex hull of points satisfying n disjunctions, in a sequence of n steps, where each step generates the convex hull of points satisfying one disjunction only. 1998 Published by Elsevier Science B.V. All rights reserved.

Keywords: Lift-and-project; Sequential convexification; Facial disjunctive programs; Reverse polars

1. Introduction: Disjunctive programming

By disjunctive programming we mean linear programming with disjunctive constraints. Integer programs (pure or mixed), and a host of other nonconvex programming problems (general quadratic programs, separable nonlinear programs, etc.) can be stated as linear programs with logical conditions. By logical conditions we mean, in the present context, statements about linear inequalities involving the operations "and" (conjunction), "or" (disjunction), "complement of" (negation). The operation "if...then" (implication) is known to be equivalent to a disjunction. The operations of conjunction and negation applied to linear inequalities give rise to (convex)

☆ Research supported under Grant #GP 37510X of the National Science Foundation and Contract N00014-67-A-0314-0007 NR 047-048 with the U.S. Office of Naval Research.

* E-mail: eb17+@andrew.cmu.edu.

0166-218X/98/$19.00 1998 Published by Elsevier Science B.V. All rights reserved.
PII: S0166-218X(98)00136-X

polyhedral sets and hence leave the problem of optimizing a linear form subject to such constraints within the realm of linear programming. That is why we view the disjunctions as the crucial element in a logical condition, and call this whole area of mathematical programming, *disjunctive programming*.

Several special cases of disjunctive programming have been studied in the past. Work on such problems includes the papers by Glover and Klingman [8, 9], Owen [13], Zwart [16] and others. The more recent work of Glover [7] is also highly relevant to our topic. Finally, Jeroslow's recent contribution to general cutting plane theory [12] provides many insights that are useful in our context too.

In our own earlier work on the subject [3–5], we addressed the general problem of obtaining valid cutting planes from arbitrary logical conditions brought to disjunctive normal form. The family of cutting planes that we obtained includes improved versions of many earlier cuts of the literature and also new cuts with some attractive features (low computational cost, coefficients of different signs, etc.). The disjunctive programming formulation seems to be particularly helpful in taking advantage of problem structure where such structure originates in the "logical" nature of the physical conditions that the problem constraints are meant to translate (like in the case of multiple choice constraints, set partitioning, etc.).

The discovery by Jeroslow [11] of the fact that the family of cutting planes introduced in our paper [4] is exhaustive, i.e., comprises all valid cutting planes for the given problem, provided at least part of the motivation for the present paper, which studies the properties of the (closed) convex hull H of the feasible points of a disjunctive program (DP). Our initial goal was simply to characterize the facets of the convex hull, i.e., find necessary and sufficient conditions for a member of the family of inequalities introduced in [4] to define a facet of H. Once this goal was achieved, however, other interesting developments were obtained as a by-product of our investigations; so that the paper in its present form goes beyond the mere characterization of facets.

In Section 2 we give a first representation of H. This leads to the formulation of a linear programming equivalent of (DP), which has a block diagonal structure with as many blocks as there are terms in the disjunction to which the constraints of (DP) are reduced when expressed in disjunctive normal form. Some properties of this linear programming equivalent of (DP) are established and its connections with a branch and bound procedure for (DP) are discussed.

Section 3 starts out with a characterization of the family of all valid inequalities for a given disjunctive program. This family is closely connected to the "reverse" polar of the feasible set of (DP), and the remainder of Section 3 is used to investigate the properties of the latter. The results of this are then used in Section 4 to characterize the facets of H. The characterization is a constructive one, i.e., it provides the tools for calculating the facets by solving a large linear program. Since the size of this linear program is proportional to the number of terms in the disjunctive normal form of (DP), the cost of calculating a facet may be prohibitive when the number of terms is large, but is quite acceptable when there are only a few terms in the disjunction.

This situation has led us to the question as to when H can be obtained via generating a sequence of "partial" convex hulls; i.e., expressing the constraints of (DP) in conjunctive normal form, when is it possible to generate H by first generating the convex hull H_1 of the feasible points of the linear program and one disjunction only; then generating the convex hull H_2 of feasible points of H_1 (which is a polyhedral set) and one disjunction only (another one), etc. It turns out, and this is the subject matter of Section 5, that such a procedure is valid for the class of disjunctive programs that we call facial, and which subsumes the most important cases of disjunctive programming. In terms of a mixed integer program $(\mathrm{IP})_n$ with n 0–1 variables, defined on a linear constraint set F_0, this means that if H_k is the (closed) convex hull of the set of points satisfying F_0 and the constraints $x_j = 0$ or 1 for the first k variables, then $(\mathrm{IP})_n$ is equivalent to the mixed integer program $(\overline{\mathrm{IP}})_{n-k}$ in which the constraints of F_0 are replaced by the facets of H_k, and only the last $n-k$ variables are integer constrained. This result establishes new connections between branch and bound and cutting planes and opens up promising possibilities for new hybrid algorithms.

The mathematical tools used in our paper are those of convex analysis, mainly concepts and results related to polarity. In this sense the present paper can be viewed as a continuation of our earlier work on convex analysis as applied to integer and nonconvex programming [1, 2].

For an arbitrary set $S \subseteq R^n$, we will denote by $\mathrm{cl}\, S$, $\mathrm{conv}\, S$, $\mathrm{aff}\, S$, $\mathrm{lh}\, S$, $\mathrm{cone}\, S$, $\dim S$, and $\mathrm{lin}\, S$, the closure, the convex hull, the affine hull, the linear hull, the conical hull, the dimension and the lineality of S. For a polyhedral set $S \subseteq R^n$, we will denote by $\mathrm{vert}\, S$ and $\mathrm{dir}\, S$ the set of vertices (extreme points) and the set of extreme direction vectors of S, respectively. For definitions and background material on these and related concepts the reader is referred to [15] or [14] (see also [10]), but we have tried to make the paper reasonably self-contained.

The disjunctive programming problem (DP) can be stated as the problem of minimizing a linear function cx, $c \in R^n$, $x \in R^n$, subject to

$$\begin{aligned} &Ax \geqslant a_0, \\ &x \geqslant 0, \\ &\bigvee_{h \in Q} (D^h x \geqslant d_0^h), \end{aligned} \tag{1'}$$

where A is $m \times n$, $a_0 \in R^m$, D^h is $m_h \times n$, $d_0^h \in R^{m_h}$, $h \in Q$, Q is a (not necessarily finite) index set, and the last condition requires that x satisfies at least one of the systems $D^h x \geqslant d_0^h$, $h \in Q$ (see [4] for illustrations of how various integer and other nonconvex programs can be brought to this form).

The linear program (LP) associated with (DP), i.e., the problem

$$\min\{cx \mid Ax \geqslant a_0,\ x \geqslant 0\} \tag{LP}$$

can be thought of as being expressed in the nonbasic variables x_j, $j \in J = \{1, \ldots, n\}$, associated with a given basic solution (in terms of the nonbasic variables, the solution is $x = 0$). This solution is feasible for (LP) if and only if $a_0 \leqslant 0$, and it is optimal for (LP), if and only if $c \geqslant 0$.

While the condition

$$\bigvee_{h \in Q} (D^h x \geqslant d_0^h)$$

is the disjunctive normal form of the logical constraints of (DP) (which may initially have been stated in a completely different form), the *disjunctive normal form* of the constraint set (1′) in its entirety is

$$\bigvee_{h \in Q} \begin{pmatrix} A^h x \geqslant a_0^h \\ x \geqslant 0 \end{pmatrix}, \tag{1}$$

where

$$A^h = \begin{pmatrix} A \\ D^h \end{pmatrix} \quad \text{and} \quad a_0^h = \begin{pmatrix} a_0 \\ d_0^h \end{pmatrix}, \quad h \in Q.$$

We will also need the *conjunctive normal form* of the constraint set (1′), namely,

$$\begin{pmatrix} Ax \geqslant a_0 \\ x \geqslant 0 \end{pmatrix} \wedge \left\{ \bigwedge_{j \in S} \left[\bigvee_{i \in Q_j} (d^i x \geqslant d_{i0}) \right] \right\}, \tag{2}$$

where $d^i \in R^n$ and d_{i0} is a scalar, $i \in Q_j$, $j \in S$. The connection between (1) and (2) is that each system $D^h x \geqslant d_0^h$, $h \in Q$, of (1) has $|S|$ inequalities, exactly one from each disjunction $\bigvee_{i \in Q_j} (d^i x \geqslant d_{i0})$ of (2) and that all distinct systems $D^h x \geqslant d_0^h$ with this property are present in (1); so that $Q = \prod_{j \in S} Q_j$, where $\prod$ stands for cartesian product. Since the operations $\wedge$ ("and", conjunction) and $\vee$ ("or", disjunction) are distributive with respect to each other [i.e., if A, B, C are inequalities, $A \wedge (B \vee C) = (A \wedge B) \vee (A \wedge C)$ and $A \vee (B \wedge C) = (A \vee B) \wedge (A \vee C)$], any logical condition involving these operations can be brought to any of the above forms, and each of the two forms can be obtained from the other one.

We illustrate the meaning of these two forms on the case when (DP) is a 0–1 program in n variables. Then the disjunctive normal form is

$$Ax \geqslant a_0,$$

$$x \geqslant 0,$$

$$x = x^1 \vee \cdots \vee x = x^q,$$

where $q=2^n$ and $x^1,\ldots,x^q$ are all the 0–1 points; whereas the conjunctive normal form is the customary

$$Ax \geqslant a_0,$$
$$x \geqslant 0,$$
$$\bigwedge_{j\in N}[(x_j=0)\vee(x_j=1)],$$

where $N=\{1,\ldots,n\}$.

2. The linear programming equivalent of (DP)

Let F be the feasible set of (DP), i.e., the set of points satisfying the constraints of (DP). Expressing the latter in the disjunctive normal form (1),

$$F=\left\{x\in R^n \;\middle|\; \bigvee_{h\in Q}(A^h x \geqslant a_0^h, x\geqslant 0)\right\}.$$

Denoting

$$F_h=\{x\in R^n \mid A^h x\geqslant a_0^h,\ x\geqslant 0\},$$

we then have

$$F=\bigcup_{h\in Q} F_h,$$

where each F_h is a polyhedral set.

If Q is finite and each F_h is bounded, then $\operatorname{conv} F$, the convex hull of F, is itself a polytope. However, the sets F_h, or some of them, may be unbounded; in which case $\operatorname{conv} F$ may not be closed, i.e., may not be polyhedral. The closure of $\operatorname{conv} F$, which we will briefly call the closed convex hull of F and denote $\operatorname{clconv} F$, is a polyhedral set whenever Q is finite. From now on, we assume that Q is finite. Let

$$Q^*=\{h\in Q \mid F_h\neq\emptyset\}.$$

Theorem 2.1. *If $F\neq\emptyset$, then*

$$\operatorname{clconv} F=\left\{x\in R^n \;\middle|\; \begin{array}{ll} x=\sum_{h\in Q^*}\xi^h, & \\ A^h\xi^h-a_0^h\xi_0^h\geqslant 0, & h\in Q^*, \\ \sum_{h\in Q^*}\xi_0^h=1, & (\xi^h,\xi_0^h)\geqslant 0,\ h\in Q^* \end{array}\right\}.$$

Proof. Let S denote the set on the right-hand side of the equality, so that the statement in the theorem is $\operatorname{clconv} F=S$.

If Q is finite and $F \neq \emptyset$, then Q^* is nonempty and finite, and

$$\operatorname{clconv} F = \operatorname{clconv}\left(\bigcup_{h \in Q^*} F_h\right).$$

(i) We first show that $\operatorname{conv} F \subseteq S$. If $x \in \operatorname{conv} F$, then x is a convex combination of at most $|Q^*|$ points, each belonging to a different F_h; i.e.,

$$x = \sum_{h \in Q^*} \lambda^h u^h, \quad \lambda^h \geqslant 0, \ h \in Q^*$$

with

$$\sum_{h \in Q^*} \lambda^h = 1,$$

where for each $h \in Q^*$, $A^h u^h \geqslant a_0^h$, $u^h \geqslant 0$. But if x, λ^h, u^h, $h \in Q^*$, satisfy the above constraints, then $x, \xi_0^h = \lambda^h$, $\xi^h = u^h \lambda^h$, $h \in Q^*$, clearly satisfy the constraints of S. Hence, $x \in \operatorname{conv} F \Rightarrow x \in S$.

(ii) Next, we show that $S \subseteq \operatorname{clconv} F$. Let $\bar{x} \in S$, with associated vectors $(\bar{\xi}^h, \bar{\xi}_0^h)$, $h \in Q^*$. Let

$$Q_1^* = \{h \in Q^* \mid \bar{\xi}_0^h > 0\}, \qquad Q_2^* = \{h \in Q^* \mid \bar{\xi}_0^h = 0\}.$$

For $h \in Q_1^*$, $\bar{\xi}^h / \bar{\xi}_0^h$ is a solution to $A^h x \geqslant a_0^h$, $x \geqslant 0$, i.e. $(\bar{\xi}^h / \bar{\xi}_0^h) \in F_h$, hence

$$(\bar{\xi}^h / \bar{\xi}_0^h) = \sum_{i \in U_h} \mu^{hi} u^{hi} + \sum_{k \in V_h} v^{hk} v^{hk}$$

for some $u^{hi} \in \operatorname{vert} F_h$, $i \in U_h$, and $v^{hk} \in \operatorname{dir} F_h$, $k \in V_h$, where U_h and V_h are finite index sets, $\mu^{hi} \geqslant 0$, $i \in U_h, v^{hk} \geqslant 0$, $k \in V_h$, and $\sum_{i \in U_h} \mu^{hi} = 1$; or, setting $\mu^{hi} \bar{\xi}_0^h = \theta^{hi}$, $v^{hk} \bar{\xi}_0^h = \sigma^{hk}$,

$$\bar{\xi}^h = \sum_{i \in U_h} \theta^{hi} u^{hi} + \sum_{k \in V_h} \sigma^{hk} v^{hk}$$

with $\theta^{hi} \geqslant 0$, $i \in U_h$, $\sigma^{hk} \geqslant 0$, $k \in V_h$, and $\sum_{i \in U_h} \theta^{hi} = \bar{\xi}_0^h$.

For $h \in Q_2^*$, either $\bar{\xi}^h = 0$, or $\bar{\xi}^h$ is a nontrivial solution to the homogeneous system $Ax \geqslant 0$, $x \geqslant 0$, hence

$$\bar{\xi}^h = \sum_{k \in V_h} \sigma^{hk} v^{hk}, \quad \sigma^{hk} \geqslant 0, \ k \in V_h$$

for some $v^{hk} \in \operatorname{dir} F_h$, $k \in V_h$. Let $\bar{Q}_2^* = \{h \in Q_2^* \mid \bar{\xi}^h \neq 0\}$. Then

$$\begin{aligned}\bar{x} &= \sum_{h \in Q^*} \bar{\xi}^h \\ &= \sum_{h \in Q_1^*} \left(\sum_{i \in U_h} \theta^{hi} u^{hi} + \sum_{k \in V_h} \sigma^{hk} v^{hk}\right) + \sum_{h \in \bar{Q}_2^*} \left(\sum_{k \in V_h} \sigma^{hk} v^{hk}\right)\end{aligned}$$

with

$$\sum_{h\in Q_1^*}\sum_{i\in U_h}\theta^{hi}=\sum_{h\in Q_1^*}\bar{\xi}_0^h=1,$$

i.e., $\bar{x}$ is the convex combination of finitely many points and directions of F. Hence $\bar{x}\in \operatorname{clconv} F$. This proves that $S\subseteq \operatorname{clconv} F$.

Since $\operatorname{conv} F\subseteq S\subseteq \operatorname{clconv} F$ and S is a closed set while $\operatorname{clconv} F$ is the smallest closed set containing $\operatorname{conv} F$, clearly $S=\operatorname{clconv} F$. □

Corollary 2.1.1. *If $\{x\in R^n \mid Ax\geqslant a_0,\ x\geqslant 0\}$ is bounded, then Theorem* 2.1 *remains true when Q^* is replaced by Q.*

Proof. If the hypothesis holds, then the homogeneous system $A^h\xi^h-a_0\xi_0^h\geqslant 0, (\xi^h,\xi_0^h)\geqslant 0$, has no nontrivial solution for any $h\in Q-Q^*$. For if $(\xi^h,\xi_0^h)\neq(0,0)$ satisfies this homogeneous system, then either $\xi_0^h>0$ and $A^h(\xi^h/\xi_0^h)\geqslant a_0$, $(\xi^h/\xi_0^h)\geqslant 0$, contrary to the definition of $Q-Q^*$; or $\xi_0^h=0$, $\xi^h\neq 0$ and $A^h\xi^h\geqslant 0$, $\xi^h\geqslant 0$, contrary to the boundedness hypothesis. Thus, for every vector (x,ξ) satisfying the constraints of S, there corresponds a vector $(x,\xi,0,\dots,0)$ satisfying the constraints of the set $S(Q)$, obtained from S by replacing Q^* with Q, where the zeroes represent the components $h\in Q-Q^*$; and conversely, every vector satisfying the constraints of $S(Q)$ has zero components for all $h\in Q-Q^*$, i.e., is of the form $(x,\xi,0,\dots,0)$, where (x,ξ) satisfies the constraints of S. □

Corollary 2.1.2. *If the linear program* (LP) *has a finite optimum, then the disjunctive program* (DP) *is equivalent to the linear program*

$$\begin{array}{lll}
\min & & \displaystyle\sum_{h\in Q} c\xi^h \\
(\mathscr{LP}) & \text{s.t.} & A^h\xi^h-a_0^h\xi_0^h\geqslant 0,\quad h\in Q,\\
& & \displaystyle\sum_{h\in Q}\xi_0^h=1,\\
& & (\xi^h,\xi_0^h)\geqslant 0,\quad \forall h\in Q
\end{array}$$

in the sense that

(i) *if x is a vertex of $\operatorname{clconv} F$, then there exists $k\in Q$ such that ξ defined by $(\xi^k,\xi_0^k)=(x,1)$ and $(\xi^h,\xi_0^h)=(0,0)$, $\forall h\in Q-\{k\}$, is a vertex of P, the feasible set of $(\mathscr{LP})$;*

(ii) *if ξ is a vertex of P, then there exists $k\in Q$ such that $\xi_0^k=1$, $(\xi^h,\xi_0^h)=(0,0)$ for $h\in Q-\{k\}$, and $x=\xi^k$ is a vertex of F_k.*

(iii) *x is an optimal solution to* (DP) *if and only if ξ as defined in* (i) *is an optimal solution $(\mathscr{LP})$.*

Proof. (i) If $\bar{x}\in \operatorname{vert}\operatorname{clconv} F$, then $\bar{x}\in F_k$, hence $A^k\bar{x}\geqslant a_0^k$, $\bar{x}\geqslant 0$, for some $k\in Q$. Therefore $\bar{\xi}\in P$, where $\bar{\xi}$ is defined by $(\bar{\xi}^k,\bar{\xi}_0^k)=(x,1)$, $(\bar{\xi}^h,\bar{\xi}_0^h)=(0,0)$, $h\in Q-\{k\}$. Also, $\bar{\xi}$ is an extreme point (hence a vertex) of P. To show this, we assume it to be false; then

$$\bar{\xi}=\sum_{i=1}^{p}\lambda_i\eta_i,\quad \lambda_i\geqslant 0,\ i=1,\dots,p,\quad \sum_{i=1}^{p}\lambda_i=1$$

and $\lambda_i>0$ for at least one $i\in\{1,\dots,p\}$, where $\eta_i\neq\bar{\xi}$, $i=1,\dots,p$, are points of P. If (η_i^h,η_{i0}^h) denotes the component of η_i corresponding to (ξ^h,ξ_0^h), then from $(\bar{\xi}^h,\bar{\xi}_0^h)=0$, $\forall h\in Q-\{k\}$, and the condition $(\xi^h,\xi_0^h)\geqslant 0$, $\forall h\in Q$, for all $\xi\in P$, it follows that $\lambda_i>0\Rightarrow(\eta_i^h,\eta_{i0}^h)=0$, $\forall h\in Q-\{k\}$. Hence, the only component of the vectors η_i (i such that $\lambda_i>0$) that can differ from the corresponding component of $\bar{\xi}$, is (η_i^k,η_{i0}^k). But then $\bar{\xi}^k=\bar{x}$ itself is a convex combination of vectors η_i^k such that $A^k\eta_i^k\geqslant a_0^k$, $\eta_i^k\geqslant 0$, i.e. of points of F_k; which contradicts the fact that $\bar{x}\in\operatorname{vert}F_k$.

(ii) Conversely, if $\bar{\xi}$ is a vertex of P, then $\bar{\xi}_0^k=1$ for some $k\in Q$ and $\bar{\xi}_0^h=0$, $\forall h\in Q-\{k\}$; for otherwise $\bar{\xi}$ is the convex combination of as many points of P as there are positive $\bar{\xi}_0^h$. From $\bar{\xi}_0^h=0$, $\forall h\in Q-\{k\}$, it follows that $\bar{\xi}^h=0$, $\forall h\in Q-\{k\}$; for, if $\bar{\xi}^h>0$ and $\bar{\xi}_0^h=0$, then $\bar{\xi}=\frac{1}{2}(\bar{\xi}'+\bar{\xi}'')$, for $\bar{\xi}'\in P$, $\bar{\xi}''\in P$, where $\bar{\xi}'$ and $\bar{\xi}''$ are obtained from $\bar{\xi}$ by replacing $\bar{\xi}^h$ with 0 and $2\bar{\xi}^h$, respectively; which contradicts the assumption that $\bar{\xi}$ is an extreme point of P. Further, $\bar{\xi}^k\in F_k$ and $\bar{\xi}^k$ is a vertex of F_k, for otherwise it is the convex combination of some vertices and extreme direction vectors of F_k, which implies that $\bar{\xi}$ is the corresponding convex combination of points and directions of P.

(iii) Finally, if the linear program (LP) has a finite optimum, then the homogeneous system $Ax\geqslant 0$, $x\geqslant 0$ does not have a nontrivial solution $\bar{x}$ such that $c\bar{x}<0$; hence none of the homogeneous systems $A^hx\geqslant 0$, $x\geqslant 0$ has such a solution and therefore $\sum_{h\in Q}c\xi^h$ is bounded from below on P. Furthermore, from (i) and (ii), $\bar{x}$ minimizes cx over $\operatorname{clconv}F$, if and only if $\bar{\xi}$, as defined in (i), minimizes $\sum_{h\in Q}c\xi^h$ over P. □

In Section 5 we give a stronger version of Corollary 2.2 for the class of disjunctive programs called facial.

The linear program ($\mathscr{LP}$) has $q\times(n+1)$ structural variables and $q\times(m+n+1)+\sum_{h\in Q}m_h+1$ constraints (where $q=|Q|$), all but the last one of which are homogeneous inequalities, while the last one is a convexity equation. Though this is a large, unwieldy problem, it has a block-angular structures with q blocks (subproblems) and a single coupling constraint. Its coefficient matrix, after introducing slack vectors, is of the form

$$\bar{A}=\begin{pmatrix} A^1 & -a_0^1 & -I^1 & & & & \\ & & & \ddots & & & \\ & & & & A^q & -a_0^q & -I^q \\ & 1 & & & & 1 & \end{pmatrix},$$

where I^h is the identity matrix of order $m+m_h$, and the blanks are zero matrices of suitable dimensions.

Since each block of $\bar{A}$ contains a copy of the coefficient matrix of (LP), and differs from the other blocks only in its lower part corresponding to $(D^h,-d_0^h)$, if one wants to think about solving $(\mathscr{LP})$, the most natural approach seems to be some decomposition-oriented version of the dual simplex method, which would start with a dual feasible solution obtained from an optimal solution to (LP). In order to assess the potential merits of such an approach, we give a necessary and sufficient condition for a basis for $(\mathscr{LP})$ to be dual feasible. The condition relates a basic dual feasible solution of $(\mathscr{LP})$, to solutions of the problems

$$(\mathrm{LP}_h) \qquad \min\{cx \mid A^h x \geqslant a_0^h,\ x \geqslant 0\}$$

$h \in Q$, and their duals.

Let us denote by $(\overline{\mathrm{LP}}_h)$ the subproblem of $(\mathscr{LP})$ indexed by h, i.e., the homogenized version of (LP_h):

$$(\overline{\mathrm{LP}}_h) \qquad \min\{c\xi^h \mid A^h\xi^h - a_0^h\xi_0^h \geqslant 0,\ (\xi^h, \xi_0^h) \geqslant 0\}.$$

We assume, as before, that $F \neq \emptyset$, and (LP) has a finite minimum.

Theorem 2.2. *Every basis for $(\mathscr{LP})$ is of the form (modulo row and column permutations)*

$$B = \begin{pmatrix} B^1 & & & & \\ & \ddots & & & \\ & & B^{q-1} & & \\ & & & B^q & -a_0^q \\ \delta^1 & \cdots & \delta^{q-1} & 0 & 1 \end{pmatrix},$$

where B^q is a basis for $(\overline{\mathrm{LP}}_q)$, and for each $h \in Q - \{q\}$, B^h is a basis for $(\overline{\mathrm{LP}}_h)$, while δ^h is the unit vector with 1 in the position corresponding to the column $-a_0^h$ if B^h contains $-a_0^h$, and $\delta^h = 0$ if B^h does not contain $-a_0^h$. The blanks are zeroes.

The basis B is dual feasible if and only if

(i) B^q is dual feasible for (LP_q).

(ii) If $\eta = (\eta^1, \ldots, \eta^q; \eta_0)$ is the solution associated with B to the dual of $(\mathscr{LP})$, then for each $h \in Q - \{q\}$, η^h is a feasible solution to the dual of (LP_h), with

$$\eta^h a_0^h \geqslant cx^q, \quad \forall h \in Q - \{q\},$$

where x^q is the solution to (LP_q) associated with B^q.

Proof. The coefficient matrix $\bar{A}$ of $(\mathscr{LP})$ has $r = q \times m + \sum_{h \in Q} m_h + 1$ rows. $\bar{A}$ is of full row rank, since it contains a $r \times r$ triangular submatrix of the form

$$\left(\begin{array}{ccc|c} -I^1 & & & \\ & \ddots & & \\ & & -I^q & -a_0^q \\ \hline & & & 1 \end{array}\right),$$

where all the blanks are zero matrices. Further, any basis matrix B for $(\mathscr{LP})$ must contain, for each $h \in Q$, a submatrix B^h of order $(m + m_h)$, which is a basis for $(\overline{\mathrm{LP}}_h)$; or else the rows of B corresponding to $(\overline{\mathrm{LP}}_h)$ are linearly dependent. This amounts to $q \times m + \sum_{h \in Q} m_h = r - 1$ columns; hence B has an additional column. Let $\bar{A}^k = (A^k, -a_0^k, -I^k)$ be the submatrix of $\bar{A}$ containing the nonzero entries of this additional column. Then the $(m + m_k) \times (m + m_k + 1)$ submatrix of $\bar{A}^k$ contained in B is (after suitable column permutations) of the form $\bar{B}^k = (B^k, -a_0^k)$, where B^k is a basis for (LP_k); for if $\bar{B}^k$ does not contain $-a_0^k$, or does not contain a basis for (LP_k), then the $(m + m_k + 1)$ columns of B corresponding to the columns of $\bar{B}^k$ are linearly dependent (in the first case, since all of them have nonzero entries in only $m + m_k$ rows; in the second, since those $m + m_k$ of them which have nonzero entries only in $\bar{B}^k$, are linearly dependent.

Finally, the vectors δ^h in the last row of B, corresponding to the matrices B^h, are of the form defined in the theorem, since each B^h, $h \neq k$, may or may not contain the column $-a_0^h$. Hence, B is a matrix of the form required by the theorem, where it is assumed, without loss of generality, that $k = q$.

Now, B is dual feasible if and only if the inequalities

$$\eta^h A^h \leqslant c,$$

$$\eta_0 - \eta^h a_h^0 \leqslant 0,$$

$$\eta^h \geqslant 0$$

hold for all $h \in Q$ for the scalar η_0 and the vectors η^h, $h \in Q$, defined by

$$\eta_0 \delta^h + \eta^h B^h = c(B^h), \quad \forall h \in Q - \{q\},$$

$$\eta^q B^q = c(B^q),$$

$$\eta_0 - \eta^q a_0^q = 0,$$

where $c(B^h)$ is the vector whose components are those c_j associated with the columns of B^h (for the slack vectors and the columns $-a_0^h$, $c_j = 0$). Solving the last two equations, we find that

$$\eta^q = c(B^q)\ (B^q)^{-1}$$

and

$$\begin{aligned}\eta_0 &= c(B^q)\ (B^q)^{-1} a_0^q \\ &= cx^q,\end{aligned}$$

where x^q is the solution to (LP_q) associated with B^q.

Substituting for η_0 in the above inequalities then yields the result that B is a dual feasible basis for $(\mathscr{LP})$ if and only if

$$\eta^h A^h \leqslant c, \quad \eta^h \geqslant 0$$

and

$$\eta^h a_0^h \geqslant cx^q$$

for all $h \in Q$, i.e., if and only if B^q is dual feasible for (LP_q), and for each $h \in Q$, η^h is a feasible solution to the dual of (LP_h), with $\eta^h a_0^h \geqslant cx^q$. □

From Theorem 2.2, an optimal solution $\bar{x}$ to (LP), with associated (dual feasible) basis B^0, can be used to generate the solution $\bar{\xi}$ to $(\mathscr{LP})$, defined by $(\bar{\xi}^k, \bar{\xi}_0^k) = (\bar{x}, 1)$ for some $k \in Q$, and $(\bar{\xi}^h, \bar{\xi}_0^h) = (0, 0)$ for all $h \in Q - \{k\}$, with the associated dual feasible basis B as defined in Theorem 2.2, where

$$B^h = \left(\begin{array}{c|c} B^0 & 0 \\ \hline D^h & -I \end{array} \right), \quad h \in Q$$

and where the column containing $-a_0^q$ is to be replaced by the column of $(\mathscr{LP})$ containing $-a_0^k$.

Starting with this basis and performing dual simplex pivots one obtains an optimal solution to $(\mathscr{LP})$ as soon as one of the vectors ξ^h, say for $h = s$, becomes primal feasible for the corresponding problem (LP_s). Then the solution $(\xi^s, \xi_0^s) = (\xi^s, 1)$, $(\xi^h, \xi_0^h) = (0, 0)$, $\forall h \in Q - \{s\}$, is primal (and dual) feasible for $(\mathscr{LP})$, hence optimal.

Such a procedure would be analogous to applying the dual simplex method "in parallel" to the q subproblems (LP_h), in the sense that one would always pivot in the subproblems with smallest objective function value, until one of them becomes feasible; then its solution is optimal for (DP).

When q is large, solving $(\mathscr{LP})$ is costly. On the other hand, one might be tempted to believe that if one happens to guess which term of the disjunction (1) yields an optimal solution $\bar{x}$ to (DP), and solves the corresponding linear program (LP_k), then $(\mathscr{LP})$ can be used to price out the other subproblems without actually solving them, so as to prove optimality. Theorem 2.2 shows that such hopes are likely to be unfounded: while $(\mathscr{LP})$ can indeed be used to price out the subproblems, due to the high degree of degeneracy of $(\mathscr{LP})$, a very large number of bases can be associated with the same optimal solution; and of all these bases, only those which correspond to the requirements of Theorem 2.2 are dual feasible, i.e., will prove the optimality of the solution. Furthermore, from (ii) it seems that finding a dual feasible basis for a *given* optimal solution requires as many dual simplex pivots as are needed to raise the objective function value $\eta^h a_0^h$ of the dual of each (LP_h) (hence also of the primal) to the level of $c\bar{x}$.

While these are important drawbacks when q is large, we will presently show that the disjunction (1) can be imposed gradually rather than all at once, i.e., the problem $(\mathscr{LP})$ can be built up step by step. This can be done by replacing the disjunctive normal form (1) with the conjunction of several disjunctions, each of which has fewer terms than q. There is a variety of forms in which a logical constraint can be expressed, and if the disjunctive normal form (1) is at one end of the spectrum, with $|Q|$ rather large, at the other end of the spectrum one has the *conjunctive* normal form, (2), where each $|Q_j|$ is as small as possible, but the corresponding disjunction represents only one of $|S|$ disjunctions which have to hold. Since, as mentioned in Section 1, $Q = \prod_{j\in S} Q_j$, the smaller the sets Q_j (for given Q), the larger the set S, and vice versa.

The following algorithm solves the disjunctive program (DP) in finitely many steps, by building up the linear program ($\mathscr{LP}$) step by step. We state it for the disjunctive condition expressed in the conjunctive normal form (2), but the algorithm can be adapted in an obvious way to any intermediate form.

0. Solve (LP) and append to the optimal simplex tableau one of the disjunctions $j \in S$ of (2). Set up the problem ($\mathscr{LP}$) corresponding to this disjunctive program, using copies of the optimal basis of (LP) to construct a dual feasible starting basis for ($\mathscr{LP}$). Go to 1.
1. Perform dual simplex pivots (or their equivalent in some decomposition framework) on ($\mathscr{LP}$), until a primal feasible solution $\bar{\xi}$ is obtained. Let ($\overline{\mathrm{LP}}_k$) be the subproblem corresponding to $(\bar{\xi}^k, \bar{\xi}_0^k) \neq 0$, and B^k the (dual feasible) basis for (LP_k) associated with $\bar{\xi}^k$. Go to 2.
2. Append to ($\overline{\mathrm{LP}}_k$) a disjunction $j \in S$ of (2), not yet represented among the constraints of ($\overline{\mathrm{LP}}_k$), i.e., replace ($\overline{\mathrm{LP}}_k$) by $|Q_j|$ new subproblems, each of which consists of ($\overline{\mathrm{LP}}_k$), plus the homogenized version of one of the terms of disjunction j. Use the last (dual feasible) basis B, and copies of B^k, to construct a dual feasible basis for the expanded ($\mathscr{LP}$). Go to 1.

The procedure stops when step 2 cannot be carried out, since all the disjunctions of (2) are represented among the constraints of (LP_k). Then $\bar{\xi}^k$ is an optimal solution to the disjunctive program (DP).

This algorithm is analogous to a "parallel" version of branch and bound. A more thorough exploration of its potential merits and drawbacks would exceed the framework of the present paper. Therefore, here we will not pursue this further, but rather turn to the problem of exploring the set of all valid inequalities for (DP), i.e., the problem of identifying the convex hull of F.

3. The family of valid inequalities for (DP)

A constraint B is said to be *a consequence of, or implied by* a constraint A, if every x that satisfies A also satisfies B. We are interested in the family of inequalities

$$\alpha x \geqslant \alpha_0$$

implied by the constraint set of a general disjunctive program (DP). The family of all such inequalities includes of course all valid cutting planes for (DP). On the other hand, the set of points satisfying all members of the family is precisely $\operatorname{clconv} F$, the closed convex hull of the set of feasible solutions to (DP). A characterization of this family is given in the next theorem, which is an easy but important generalization of a classical result. The "if" part is Theorem 1 of our paper [4], the "only if" part is due to R. Jeroslow [11].

Let Q^* be the set of those $h \in Q$ such that $\{x \in R^n \mid A^h x \geqslant a_0^h,\ x \geqslant 0\} \neq \emptyset$. We will assume that $Q^* \neq \emptyset$; i.e., (DP) has a solution.

Theorem 3.1. *The inequality $\alpha x \geqslant \alpha_0$ is a consequence of the constraints*

$$\begin{aligned} &Ax \geqslant a_0,\\ &x \geqslant 0,\\ &\bigvee_{h \in Q} (D^h x \geqslant d_0^h) \end{aligned} \tag{3'}$$

if and only if there exists a set of vectors $\theta^h \in R^m$, $\sigma^h \in R^{m_h}$, $\theta^h \geqslant 0$, $\sigma^h \geqslant 0$, $h \in Q^*$, such that

$$\alpha \geqslant \theta^h A + \sigma^h D^h, \quad \forall h \in Q^*$$

and

$$\alpha_0 \leqslant \theta^h a_0 + \sigma^h d_0^h, \quad \forall h \in Q^*.$$

Proof. Rewriting (1′) in the disjunctive normal form

$$\bigvee_{h \in Q} \begin{pmatrix} Ax \geqslant a_0 \\ x \geqslant 0 \\ D^h x \geqslant d_0^h \end{pmatrix}, \tag{1}$$

we note that $\alpha x \geqslant \alpha_0$ is a consequence of (1) if and only if it is a consequence of each (conjunctive) system $h \in Q^*$ of (1). But according to a well-known result on linear inequalities (see for instance Theorem 1.4.4 of [15], or Theorem 22.3 of [14]), this is the case if and only if the conditions stated in the theorem hold. □

Remark 3.1. *If the ith inequality of a system $h \in Q^*$ of (1) is replaced by an equation, the ith component of θ^h is to be made unconstrained. If the variable x_j in (1) is let to be unconstrained, the jth inequality of each system $\alpha \geqslant \theta_A^h + \sigma^h D^h$, $h \in Q^*$, is to be replaced by the corresponding equation.*

With these changes, the theorem remains true.

An inequality $\alpha x \geqslant \alpha_0$ satisfying the conditions of Theorem 3.1, i.e., a "valid" inequality $\alpha x \geqslant \alpha_0$, may or may not be a cutting plane, i.e., may or may not cut off a nonempty subset of F_0, the feasible set of the linear program (LP). If $\alpha_0 > 0$, however, then the inequality is not only a valid cutting plane, but one which cuts off the current solution defined by $x = 0$.

For given α_0, the family of inequalities $\alpha x \geqslant \alpha_0$ implied by (1) is isomorphic to the family of vectors $\alpha \in F^{\#}(\alpha_0)$, where

$$F^{\#}(\alpha_0) = \{y \in R^n \mid yx \geqslant \alpha_0,\ \forall x \in F\},$$

since $\alpha x \geqslant \alpha_0$ is implied by (1) if and only if $\alpha \in F^{\#}(\alpha_0)$. From Theorem 3.1,

$$F^{\#}(\alpha_0) = \left\{ y \in R^n \;\middle|\; \begin{array}{l} y \geqslant \theta^h A + \sigma^h D^h, \;\; h \in Q^* \\ \text{for some } \theta^h \geqslant 0, \;\; \sigma^h \geqslant 0, \;\; h \in Q^*, \\ \text{such that } \theta^h a_0 + \sigma^h d_0^h \geqslant \alpha_0 \end{array} \right\}.$$

In view of its relationship to ordinary polar sets, we will call $F^{\#}(\alpha_0)$ the *reverse polar* of F (scaled with α_0). Indeed, the ordinary polar set of F is

$$F^0 = \{y \in R^n \mid yx \leqslant 1, \;\; \forall x \in F\}$$

and if we denote by $F^0(\alpha_0)$ the scaled polar of F, i.e., the set obtained from F^0 by replacing 1 with α_0, then the relationship between $F^{\#}$ and F^0 is given by $F^{\#}(\alpha_0) = -F^0(-\alpha_0)$.

The size, as opposed to the sign, of α_0, is of no interest to us in the present context. Therefore, we will distinguish only between the three cases $\alpha_0 > 0$ (or $\alpha_0 = 1$), $\alpha_0 = 0$ and $\alpha_0 < 0$ (or $\alpha_0 = -1$), and whenever the sign of α_0 makes no difference, we will simply write $F^{\#}$ for $F^{\#}(\alpha_0)$.

Next, we derive some basic properties of reverse polars, which we need in order to characterize the facets of $\operatorname{conv} F$. Most of these properties are parallel to those of ordinary polars, but some are different. Though here we state them for $F^{\#}$, they are valid for the reverse polars of arbitrary sets whose closed convex hull is polyhedral. Moreover, those properties which do not specifically refer to polyhedra, carry over to arbitrary sets, modulo a closure operation.

Some properties follow immediately from the definitions. Thus, for arbitrary sets $S \subseteq R^n$, $T \subseteq R^n$, one has

(a) $(\lambda S)^{\#} = (1/\lambda) S^{\#}, \quad -\infty < \lambda < \infty$,

(b) $S \subseteq T \Rightarrow S^{\#} \supseteq T^{\#}$,

(c) $(S \cup T)^{\#} = S^{\#} \cap T^{\#}$.

Before stating the next theorem, we note that for an arbitrary set S and closed halfspace H^+, $S \subseteq H^+ \Rightarrow \operatorname{clconv} S \subseteq H^+$.

Theorem 3.2. (i) *If* $\alpha_0 > 0$, *then*

$$0 \in \operatorname{clconv} F \Leftrightarrow F^{\#} = \emptyset \Leftrightarrow F^{\#} \text{ is bounded.}$$

(ii) *If* $\alpha_0 \leqslant 0$, *then* $F^{\#} \neq \emptyset$, *and*

$$0 \in \operatorname{int} \operatorname{clconv} F \Leftrightarrow F^{\#} \text{ is bounded.}$$

Proof. (i) Let $\alpha_0 > 0$. If $F^{\#} \neq \emptyset$, then there exists $y \in R^n$ such that $xy \geqslant \alpha_0$, $\forall x \in F$; hence $xy \geqslant \alpha_0$, $\forall x \in \operatorname{clconv} F$. Thus, the hyperplane $yx = \alpha_0$ separates 0 from $\operatorname{clconv} F$, i.e., $0 \notin \operatorname{clconv} F$. Therefore, if $0 \in \operatorname{clconv} F$, then $F^{\#} = \emptyset$; and of course, $F^{\#}$ is bounded when it is void. Conversely, if $0 \notin \operatorname{clconv} F$, then there exists a hyperplane $ax = \alpha_0$ separating 0 from $\operatorname{clconv} F$, i.e., such that $ax \geqslant \alpha_0$, $\forall x \in \operatorname{clconv} F$, which implies $a \in F^{\#}$, i.e. $F^{\#} \neq \emptyset$. Also, it implies that $\lambda a \in F^{\#}$, $\forall \lambda > 1$, i.e., $F^{\#}$ is unbounded.

(ii) Follows from the corresponding property of ordinary polar sets (and cones), and the fact that $F^{\#}(\alpha_0) = -F^0(-\alpha_0)$. $\square$

Since F_0 is contained in the nonnegative orthant, so is F and $\operatorname{clconv} F$. Thus, $\operatorname{clconv} F$ has at least one vertex, and if unbounded, it has at least one extreme direction.

Now, let the vertices (extreme points) and extreme direction vectors of $\operatorname{clconv} F$ be denoted by

$$\operatorname{vert}\operatorname{clconv} F = \{u_1, \ldots, u_p\}$$

and

$$\operatorname{dir}\operatorname{clconv} F = \{v_1, \ldots, v_q\},$$

respectively.

Theorem 3.3. *$F^{\#}$ is the convex polyhedral set*

$$F^{\#} = \left\{ y \in R^n \,\middle|\, \begin{array}{ll} u_i y \geqslant \alpha_0, & i = 1, \ldots, p \\ v_i y \geqslant 0, & i = 1, \ldots, q \end{array} \right\}.$$

Proof.

$$\begin{aligned} F^{\#} &= \{ y \in R^n \mid xy \geqslant \alpha_0,\ \forall x \in F \} \\ &= \{ y \in R^n \mid xy \geqslant \alpha_0,\ \forall x \in \operatorname{clconv} F \} \\ &= \left\{ y \in R^n \,\middle|\, \begin{array}{ll} uy \geqslant \alpha_0, & u \in \operatorname{vert}\operatorname{clconv} F \\ vy \geqslant 0, & v \in \operatorname{dir}\operatorname{clconv} F \end{array} \right\}. \qquad \square \end{aligned}$$

For arbitrary sets S and T, we denote by $S + T$ the Minkowski sum of S and T, i.e.

$$S + T = \{ x \mid x = s + t,\ s \in S,\ t \in T \}.$$

Theorem 3.4. *Assume $F^{\#} \neq \emptyset$. Then*

$$F^{\#\#} = \begin{cases} \operatorname{cl}(\operatorname{conv} F + \operatorname{cone} F) & \text{if } \alpha_0 > 0, \\ \operatorname{clcone} F & \text{if } \alpha_0 = 0, \\ \operatorname{clconv}(F \cup \{0\}) & \text{if } \alpha_0 < 0. \end{cases}$$

Proof.

$$\begin{aligned} F^{\#\#} &= \{ x \in R^n \mid xy \geqslant \alpha_0,\ \forall y \in F^{\#} \} \\ &= \{ x \in R^n \mid y \in F^{\#} \Rightarrow xy \geqslant \alpha_0 \} \\ &= \left\{ x \in R^n \,\middle|\, \left\{ \begin{array}{ll} u_i y \geqslant \alpha_0, & i = 1, \ldots, p \\ v_i y \geqslant 0, & i = 1, \ldots, q \end{array} \right\} \Rightarrow xy \geqslant \alpha_0 \right\} \quad \text{(From Theorem 3.3)}, \end{aligned}$$

where u_i and v_i stand for the vertices and extreme direction vectors of $\operatorname{clconv} F$, respectively. But from the basic result in linear inequalities mentioned in the proof of Theorem 3.1, $xy \geqslant \alpha_0$ is a consequence of the constraints in the inner brackets (whose system is solvable by the assumption) if and only if there exists a set of multipliers θ_i, $i=1,\dots,p$, σ_i, $i=1,\dots,q$, such that

$$x = \sum_{i=1}^{p} \theta_i u_i + \sum_{i=1}^{q} \sigma_i v_i$$

and $\theta_i \geqslant 0$, $i=1,\dots,p$, $\sigma_i \geqslant 0$, $i=1,\dots,q$, with

$$\sum_{i=1}^{p} \theta_i \alpha_0 \geqslant \alpha_0$$

if $\alpha_0 \neq 0$. Dividing through with α_0 when $\alpha_0 \neq 0$, we conclude that $F^{\#\#}$ is the set of points $x \in R^n$ of the form

$$x = \sum_{i=1}^{p} \theta_i u_i + \sum_{i=1}^{q} \sigma_i v_i, \quad \theta_i \geqslant 0, \ \sigma_i \geqslant 0, \ \forall i$$

with

$$\sum_{i=1}^{p} \theta_i \begin{cases} \geqslant 1 & \text{if } \alpha_0 > 0, \\ \geqslant 0 & \text{if } \alpha_0 = 0, \\ \leqslant 1 & \text{if } \alpha_0 < 0. \end{cases}$$

But these are precisely the expressions for the three sets claimed in the theorem to be equal to $F^{\#\#}$ in the respective cases. □

Remark 3.4.1. If $F^{\#} = \emptyset$ (*which, from Theorem* 3.2, *is only possible when* $\alpha_0 > 0$), then $F^{\#\#} = R^n$.

Since F is contained in the nonnegative orthant, so is $F^{\#\#}$ unless $F^{\#} = \emptyset$; and thus $F^{\#\#}$ has at least one vertex, and if unbounded, it has at least one extreme direction.

Theorem 3.5. $F^{\#\#\#} = F^{\#}$.

Proof. If $\alpha_0 \leqslant 0$, this follows from the corresponding property of ordinary polars and the fact that $F^{\#}(\alpha_0) = -F^0(-\alpha_0)$.

If $\alpha_0 > 0$ and $0 \in \operatorname{clconv} F$, then $F^{\#} = \emptyset$, $F^{\#\#} = R^n$, and $F^{\#\#\#} = \emptyset = F^{\#}$.

If $\alpha_0 > 0$ and $0 \notin \operatorname{clconv} F$, then

$$\begin{aligned} F^{\#\#\#} &= \operatorname{cl}(\operatorname{conv} F + \operatorname{cone} F)^{\#} \quad \text{(from Theorem 3.4)} \\ &= \{y \in R^n \mid xy \geqslant \alpha_0, \quad \forall x \in \operatorname{cl}(\operatorname{conv} F + \operatorname{cone} F)\} \\ &= \{y \in R^n \mid xy \geqslant \alpha_0, \quad \forall x \in F\} \\ &= F^{\#}, \end{aligned}$$

since

$$\mathrm{cl}(\mathrm{conv}\,F + \mathrm{cone}\,F) = \left\{ x \in R^n \;\middle|\; \begin{array}{l} x = \sum_{i=1}^{p} \theta_i u_i + \sum_{i=1}^{q} \sigma_i v_i \\ \sum_{i=1}^{p} \theta_i \geqslant 1, \;\; \theta_i \geqslant 0, \;\; \sigma_i \geqslant 0, \forall i \end{array} \right\}. \qquad \square$$

For an arbitrary subset S of R^n, the linear hull of S, denoted $\mathrm{lh}\,S$, is the subspace of R^n generated by S, i.e., the set of all linear combinations of points of S. Clearly, $\mathrm{lh}\,S$ is the smallest subspace of R^n containing S. The affine hull of S, denoted $\mathrm{aff}\,S$, is the affine manifold (linear variety) generated by S, i.e., the set of all such linear combinations of points of S, where the sum of coefficients equals 1.

Theorem 3.6. *Assume $F^{\#} \neq \emptyset$. Then* $\mathrm{aff}\,F^{\#\#} = \mathrm{lh}\,F^{\#\#} = \mathrm{lh}\,F$.

Proof. From Theorem 3.4, if $x \in F$ then $\lambda x \in F^{\#\#}$ for all λ such that $\lambda \geqslant 1$ if $\alpha_0 > 0$, $\lambda \geqslant 0$ if $\alpha_0 = 0$, and $0 \leqslant \lambda \leqslant 1$ if $\alpha_0 < 0$. Now $\mathrm{aff}\,F^{\#\#}$ contains all the lines through pairs of distinct points of $F^{\#\#}$; and for every $x \in F \subseteq F^{\#\#}$, $x \neq 0$, there exists a point $\lambda x \in F^{\#\#}$, such that $\lambda > 1$ (if $\alpha_0 \geqslant 0$) or $\lambda < 1$ (if $\alpha_0 \leqslant 0$). But then for every $x \in F$, $x \neq 0$, $\mathrm{aff}\;F^{\#\#}$ contains the entire line λx, i.e., the line through x and 0. But the set characterized by this property is precisely the linear hull of F, i.e., $\mathrm{aff}\,F^{\#\#} \supseteq \mathrm{lh}\,F$. Since this shows $\mathrm{aff}\,F^{\#\#}$ to be a linear space, it follows that $\mathrm{aff}\,F^{\#\#} = \mathrm{lh}\,F^{\#\#}$. Further, since $\mathrm{lh}\,F \supseteq \mathrm{cone}\,F$ and $\mathrm{lh}\,F \supseteq \mathrm{conv}\,F$ and since $\mathrm{lh}\,F$ is closed, it follows that $\mathrm{lh}\,F \supseteq F^{\#\#}$ for all three cases of Theorem 3.4, and hence $\mathrm{lh}\,F \supseteq \mathrm{lh}\,F^{\#\#} = \mathrm{aff}\,F^{\#\#}$. $\square$

Remark 3.6.1. *If $F^{\#} = \emptyset$, then* $\mathrm{aff}\,F^{\#\#} = \mathrm{lh}\,F^{\#\#} \supseteq \mathrm{lh}\,F$.

Now, let L be the lineality space of $F^{\#}$, defined to be the largest linear subspace of R^n contained in $C(F^{\#})$, the (closed) cone of all the directions of $F^{\#}$, also called the recession cone (see [14], Ch. 8) or characteristic cone (see [15], Ch. 3) of $F^{\#}$. The dimension of L is the lineality of $F^{\#}$, denoted $\mathrm{lin}\,F^{\#}$.

For any linear subspace S of R^n, let $S^{\perp}$ denote the orthogonal complement of S, i.e.,

$$S^{\perp} = \{y \in R^n \mid xy = 0, \;\; \forall x \in S\}.$$

The next theorem states that L, the largest subspace contained in $C(F^{\#})$, is the orthogonal complement of $\mathrm{lh}\,F$, the smallest subspace of R^n containing F.

Theorem 3.7. *Assume $F^{\#} \neq \emptyset$. Then $\mathrm{lh}\,F = L^{\perp}$.*

Proof. If $x \in \mathrm{lh}\,F$, then $x = \sum_{i=1}^{r} \lambda^i u^i$ for some $u^i \in F$, $i = 1, \dots, r$. Since $L \subseteq C(F^{\#})$, $u^i y \geqslant 0$, $i = 1, \dots, r$, $\forall y \in L$. But L is a linear space, i.e., $y \in L \Rightarrow -y \in L$; thus we also have $u^i(-y) \geqslant 0$, $i = 1, \dots, r$. Hence, $u^i y = 0$, $i = 1, \dots, r$, and therefore $xy = 0$, $\forall y \in L$. This proves that $\mathrm{lh}\,F \subseteq L^{\perp}$, which implies $(\mathrm{lh}\,F)^{\perp} \supseteq L$.

On the other hand,

$$
\begin{aligned}
y \in (\mathrm{lh}\, F)^{\perp} &\Rightarrow xy = 0, \quad \forall x \in F \quad (\text{since } F \subseteq \mathrm{lh}\, F)\\
&\Rightarrow x(\lambda y) = 0, \quad \forall x \in F,\ \forall \lambda\\
&\Rightarrow \lambda y \in C(F^{\#}), \quad \forall \lambda\\
&\Rightarrow y \in L \quad (\text{since } L \text{ is the largest subspace contained in } C(F^{\#}))
\end{aligned}
$$

i.e., $(\mathrm{lh}\, F)^{\perp} \subseteq L$. □

Corollary 3.7.1. $\dim F^{\#\#} + \mathrm{lin}\, F^{\#} = n$.

Proof. $\dim F^{\#\#} = \dim \mathrm{aff}\, F^{\#\#} = \dim \mathrm{lh}\, F$, and $\mathrm{lin}\, F^{\#} = \dim L$. From Theorem 3.7, $(\mathrm{lh}\, F) + L = R^n$. □

Corollary 3.7.2.

$$
\dim F^{\#\#} = \begin{cases} \dim F & \text{if } 0 \in \mathrm{aff}\ F,\\ (\dim F) + 1 & \text{if } 0 \notin \mathrm{aff}\ F. \end{cases}
$$

Proof. If $0 \in \mathrm{aff}\, F$, then $\mathrm{aff}\, F = \mathrm{lh}\, F = \mathrm{aff}\, F^{\#\#}$, and thus $\dim F = \dim F^{\#\#}$. This is of course always the case when $\dim F = n$.

If $0 \notin \mathrm{aff}\, F$, then $\dim(\mathrm{lh}\, F) = \dim(\mathrm{aff}\, F) + 1$, and from $\mathrm{lh}\, F = \mathrm{aff}\, F^{\#\#}$ it follows that $\dim F^{\#\#} = (\dim F) + 1$. □

Corollary 3.7.3. $F^{\#} = (F^{\#} \cap \mathrm{lh}\, F) + L$.

Proof. Follows from $\mathrm{lh}\, F = L^{\perp}$. □

Corollary 3.7.4. *The lowest-dimensional faces of $F^{\#}$ are those of dimension n-$\dim F^{\#\#}$.*

Proof. n-$\dim F^{\#\#} = \dim L$ is a lower bound on the dimension of a face of $F^{\#}$, since the lineality space of $F^{\#}$ is also the lineality space of each face of $F^{\#}$. But $F^{\#} = (F^{\#} \cap L^{\perp}) + L$, $\mathrm{lin}(F^{\#} \cap L^{\perp}) = 0$, and if x is a vertex (0-dimensional face) of $F^{\#} \cap L^{\perp}$, then $x + L$ is a face of dimension n-$\dim F^{\#\#}(= \dim L)$ of $F^{\#}$. □

4. The facets of the convex hull of feasible points

An inequality $\alpha x \geqslant \alpha_0$, with $\alpha \in R^n$, $\alpha \neq 0$, defines a facet $[(d-1)$-dimensional face$]$ of a d-dimensional polyhedral set $S \subseteq R^n$, if $\alpha x \geqslant \alpha_0$, $\forall x \in S$, and $\{x \in S \mid \alpha x = \alpha_0\}$ is a facet of S, i.e., $\alpha x = \alpha_0$ for exactly d affinely independent points x of S. We say "exactly" in order to exclude the case when $\alpha x = \alpha_0$ is a singular supporting hyperplane for S, i.e., contains all of S. For the sake of brevity, and in keeping with the terminology used in integer programming, we will call the inequality $\alpha x \geqslant \alpha_0$ itself a facet when it defines a facet.

Theorem 4.1.

$$\left\{\begin{array}{l} \alpha x \geqslant \alpha_0 \ \textit{is a facet} \\ \textit{of } F^{\#\#}, \textit{ and } \alpha \in \operatorname{lh} F \end{array}\right\} \Leftrightarrow \left\{\begin{array}{l} \alpha_0 \neq 0, \ \alpha \neq 0, \ \alpha \in \operatorname{vert}(F^{\#} \cap L^{\perp}) \\ \alpha_0 = 0, \ \alpha \neq 0, \ \alpha \in \operatorname{dir}(F^{\#} \cap L^{\perp}) \end{array}\right\}.$$

Proof. Denote $\dim F^{\#\#} = d$.

(i) Let $\alpha_0 \neq 0$. Then from Theorems 3.5 and 3.3,

$$\begin{aligned} F^{\#} &= \{ y \in R^n \mid xy \geqslant \alpha_0, \ \forall x \in F^{\#\#} \} \\ &= \left\{ y \in R^n \,\middle|\, \begin{array}{ll} uy \geqslant \alpha_0, & \forall u \in \operatorname{vert} F^{\#\#} \\ vy \geqslant 0, & \forall v \in \operatorname{dir} F^{\#\#} \end{array} \right\} \end{aligned}$$

and hence

$$F^{\#} \cap L^{\perp} = \left\{ y \in R^n \,\middle|\, \begin{array}{ll} uy \geqslant \alpha_0, & \forall u \in \operatorname{vert} F^{\#\#} \\ vy \geqslant 0, & \forall v \in \operatorname{dir} F^{\#\#} \\ wy = 0, & \forall w \in L \end{array} \right\}.$$

Then $\alpha \in \operatorname{vert}(F^{\#} \cap L^{\perp})$ if and only if $\alpha \in F^{\#} \cap L^{\perp}$, and α satisfies with equality a subset of the inequalities of the constraint set defining $F^{\#} \cap L^{\perp}$, the rank of the coefficient matrix of this subset being d. Further, $\alpha \neq 0$ if and only if this subset of inequalities is not homogeneous, i.e., has at least one nonzero right-hand side coefficient. On the other hand, $\alpha x \geqslant \alpha_0$, where $\alpha_0 \neq 0$ and $\alpha \in \operatorname{lh} F$, is a facet of $F^{\#\#}$ if and only if (i) $\alpha x \geqslant \alpha_0$, $\forall x \in F^{\#\#}$, i.e., $\alpha \in F^{\#}$, and, since $\alpha \in \operatorname{lh} F = L^{\perp}$, $\alpha \in F^{\#} \cap L^{\perp}$; (ii) $\alpha x = \alpha_0$ for exactly d affinely independent points of $F^{\#\#}$. But condition (ii) holds if and only if $\alpha u = \alpha_0$ for r vertices u of $F^{\#\#}$, and $\alpha v = 0$ for s extreme direction vectors v of $F^{\#\#}$, with $r \geqslant 1$ (since $\alpha \neq 0$) and $r + s \geqslant d$, such that the $n \times (r+s)$ matrix whose columns are the vectors u and v, has rank d. Since the two conditions are identical, the statement is true for $\alpha_0 \neq 0$.

(ii) Now, let $\alpha_0 = 0$. Then from Theorems 3.5 and 3.3

$$\begin{aligned} F^{\#} &= \{ y \in R^n \mid xy \geqslant 0, \quad \forall x \in F^{\#\#} \} \\ &= \{ y \in R^n \mid vy \geqslant 0, \quad \forall v \in \operatorname{dir} F^{\#\#} \} \end{aligned}$$

and thus

$$F^{\#} \cap L^{\perp} = \left\{ y \in R^n \,\middle|\, \begin{array}{ll} vy \geqslant 0, & \forall v \in \operatorname{dir} F^{\#\#} \\ wy = 0, & \forall w \in L \end{array} \right\}.$$

Then $\alpha \in R^n$, $\alpha \neq 0$, is an extreme direction vector of $F^{\#} \cap L^{\perp}$, if and only if $\alpha \in F^{\#} \cap L^{\perp}$ and α satisfies with equality a subset of the inequalities of the constraint set defining $F^{\#} \cap L^{\perp}$, the rank of the coefficient matrix of this subset being $d - 1$. On the other hand, if $\alpha \neq 0$, $\alpha \in \mathrm{lh}\, F$, $\alpha x \geqslant 0$ is a facet of $F^{\#\#}$ (which in this case is a polyhedral cone) if and only if (i) $\alpha x \geqslant 0$, $\forall x \in F^{\#\#}$, i.e. $\alpha \in F^{\#}$ and, since $\alpha \in \mathrm{lh}\, F = L^{\perp}$, $\alpha \in F^{\#} \cap L^{\perp}$; (ii) $\alpha x = 0$ for exactly $d-1$ affinely independent points $x \neq 0$ of $F^{\#\#}$ (the dth point being 0). It is easy to check that again, the two conditions are identical; hence the statement is true also for the case $\alpha_0 = 0$. □

The condition $\alpha \in \mathrm{lh}\, F$ is necessary since, whenever $F^{\#\#}$ is less than full dimensional, each facet of $F^{\#\#}$ defined by a hyperplane H can also be defined by any other member of the family of hyperplanes H' such that $(\mathrm{lh}\, F) \cap H' = (\mathrm{lh}\, F) \cap H$. Thus, we represent this family of hyperplanes by its (unique) member whose normal lies in $\mathrm{lh}\, F$.

The following result will also be useful for the actual calculation of facets.

Corollary 4.1.1. *Let $g \in R^n$ and let $\alpha x \geqslant \alpha_0$ be a facet of $F^{\#\#}$, with $\alpha \in \mathrm{lh}\, F$. Then $g \in \{x \in F^{\#\#} \mid \alpha x = \alpha_0\}$ if and only if $gx = \alpha_0$ is a supporting hyperplane for $F^{\#}$ which contains the vertex (if $\alpha_0 \neq 0$), or extreme direction vector (if $\alpha_0 = 0$), α of $F^{\#} \cap L^{\perp}$.*

Proof. From Theorem 4.1, $\alpha \in \mathrm{vert}\, F^{\#} \cap L^{\perp}$ if $\alpha_0 \neq 0$, and $\alpha \in \mathrm{dir}\, F^{\#} \cap L^{\perp}$, if $\alpha_0 = 0$. In both cases $\alpha \in F^{\#}$.

(i) Let $g \in F^{\#\#}$, $\alpha g = \alpha_0$.

$$g \in F^{\#\#} \Rightarrow F^{\#\#\#} \subseteq g^{\#} \quad \text{(from property (6) of polars)}$$

$$\Rightarrow F^{\#} \subseteq \{x \in R^n \mid gx \geqslant \alpha_0\} \quad \text{(from Theorem 3.3)}$$

and since $g\alpha = \alpha_0$, $\alpha \in F^{\#}$, it follows that $gx = \alpha_0$ is a supporting hyperplane for $F^{\#}$, which contains α.

(ii) Suppose $gx = \alpha_0$ is a supporting hyperplane for $F^{\#}$ which contains α. Then

$$gx \geqslant \alpha_0, \quad \forall x \in F^{\#},$$

i.e.,

$$g \in F^{\#\#}, \quad \text{and} \quad \alpha g = \alpha_0, \text{ i.e., } g \in \{x \in F^{\#\#} \mid \alpha x = \alpha_0\}. \qquad \square$$

Theorem 4.1 characterizes the facets of $F^{\#\#}$; our main interest, however, lies in the facets of $\mathrm{clconv}\, F$ rather than $F^{\#\#}$. Next, we characterize the facets of $\mathrm{clconv}\, F$ in terms of the reverse polar $F^{\#}$. We do this separately for the two cases when $\alpha_0 \neq 0$ and $\alpha_0 = 0$.

Theorem 4.2. *Assume $0 \in \mathrm{aff}\, F$. If $\alpha_0 \neq 0$, then $\alpha x \geqslant \alpha_0$ is a facet of* clconv F *if and only if it is a facet of $F^{\#\#}$.*

Proof. (i) If $\alpha_0 > 0$, then

$$\alpha x \geqslant \alpha_0, \quad \forall x \in \operatorname{clconv} F \Leftrightarrow \alpha x \geqslant \alpha_0, \quad \forall x \in \operatorname{cl}(\operatorname{conv} F + \operatorname{cone} F) = F^{\#\#}.$$

If $\alpha_0 < 0$, then

$$\alpha x \geqslant \alpha_0, \quad \forall x \in \operatorname{clconv} F \Leftrightarrow \alpha x \geqslant \alpha_0, \quad \forall x \in \operatorname{clconv}(F \cup \{0\}) = F^{\#\#}.$$

In both cases, $\alpha x \geqslant \alpha_0$ is a supporting halfspace for $\operatorname{clconv} F$ if and only if it is a supporting halfspace for $F^{\#\#}$.

(ii) Next, we show that if $\alpha_0 \neq 0$ and $\alpha x \geqslant \alpha_0$ is a supporting halfspace for $F^{\#\#}$, then

$$\{x \in \operatorname{clconv} F \mid \alpha x = \alpha_0\} = \{x \in F^{\#\#} \mid \alpha x = \alpha_0\}.$$

The relation $\subseteq$ is obvious, since $\operatorname{clconv} F \subseteq F^{\#\#}$. To show the converse, we assume it to be false, and let $x \in F^{\#\#} - \operatorname{clconv} F$ satisfy $\alpha x = \alpha_0$. From the definition of $F^{\#\#}$, $x = \lambda u$ for some $u \in \operatorname{clconv} F$, and $\lambda > 1$ if $\alpha_0 > 0$, $0 < \lambda < 1$ if $\alpha_0 < 0$. In each case, $\alpha x = \alpha_0$ implies $\alpha u = (1/\lambda)\alpha x < \alpha_0$ for some $u \in \operatorname{clconv} F \subseteq F^{\#\#}$, which contradicts the assumption that $\alpha x \geqslant \alpha_0$, $\forall x \in F^{\#\#}$.

Thus, if $\alpha_0 \neq 0$ and $\alpha x \geqslant \alpha_0$ is a supporting halfspace for $F^{\#\#}$, then the hyperplane $\alpha x = \alpha_0$ contains d affinely independent point of clconv F if and only if it contains d affinely independent points of $F^{\#\#}$. Since $\dim F^{\#\#} = d$ and if $0 \in \operatorname{aff} F$, $\dim F = d$ too (Corollary 3.7.2), the theorem follows. □

If $0 \notin \operatorname{aff} F$, then $\dim F = d - 1$, and each facet $\alpha x = \alpha_0$ ($\alpha_0 \neq 0$) of $F^{\#\#}$ is a singular supporting hyperplane, rather than a facet, of $\operatorname{clconv} F$. Each facet of $\operatorname{clconv} F$ is in this case the intersection of $\alpha x = \alpha_0$ with some other facet of $F^{\#\#}$, hence a $(d-2)$-dimensional face of $F^{\#\#}$, corresponding to an edge (one-dimensional face) of $F^{\#} \cap L^{\perp}$.

If $F^{\#\#}$ is full-dimensional, then $L = \emptyset$ and $L^{\perp} = R^n$, hence vert $(F^{\#} \cap L^{\perp}) = \operatorname{vert} F^{\#}$. If $d = \dim F^{\#\#} < n$, then $\operatorname{lin} F^{\#} = n - d > 0$, and $F^{\#}$ has no vertices. However, there is a one-to-one correspondence between vertices α of $F^{\#} \cap L^{\perp}$, and $(n-d)$-dimensional faces of the form $\alpha + L$, of $F^{\#}$ (these are the lowest-dimensional faces of $F^{\#}$). Hence, from Theorem 4.2 we have the following.

Corollary 4.2.1. *Assume $0 \in \operatorname{aff} F$. The inequality $\alpha x \geqslant \alpha_0$, where $\alpha_0 \neq 0$ and $\alpha \in \operatorname{lh} F$, is a facet of $\operatorname{clconv} F$, if and only if $\alpha \neq 0$ and $\alpha \in \operatorname{vert}(F^{\#} \cap L^{\perp})$, i.e., $\alpha + L$ is a $(n-d)$-dimensional face of $F^{\#}$. In particular, if $d = n$, then $\alpha x \geqslant \alpha_0$ is a facet of $\operatorname{clconv} F$ if and only if $\alpha \neq 0$ is a vertex of $F^{\#}$.*

Next, we turn to the case where $\alpha_0 = 0$.

Theorem 4.3. *If $\alpha x \geqslant 0$ is a facet of $\operatorname{clconv} F$, and $\alpha \in \operatorname{lh} F$, then $\alpha \neq 0$ and $\alpha \in \operatorname{dir}(F^{\#} \cap L^{\perp})$.*

Conversely, if $\alpha \neq 0$ and $\alpha \in \operatorname{dir}(F^{\#} \cap L^{\perp})$, then $\alpha x \geqslant 0$ is either a facet or a $(d-2)$-dimensional face of $\operatorname{clconv} F$. In the latter case, the $(d-2)$-dimensional face is the *intersection of two adjacent facets of the form* $\alpha^1 x \geqslant \alpha_0^1$ *and* $\alpha^2 x \geqslant \alpha_0^2$, *with* $\alpha_0^1 > 0$, $\alpha_0^2 < 0$, *and* $\alpha = \alpha^1/\alpha_0^1 - \alpha^2/\alpha_0^2$.

Proof. When $\alpha_0 = 0$, $F^{\#\#} = \mathrm{clcone}\, F$. Hence, if $\alpha x \geqslant 0$, $\forall x \in \mathrm{clconv}\, F$, then $\alpha x \geqslant 0$, $\forall x \in F^{\#\#}$; i.e., if $\alpha x \geqslant 0$ is a supporting halfspace for $\mathrm{clconv}\, F$, then it is a supporting halfspace for $F^{\#\#}$ too. Further, since $\mathrm{clconv}\, F \subseteq F^{\#\#}$, if $\alpha x \geqslant 0$ is a facet for $\mathrm{clconv}\, F$, then it is a facet for $F^{\#\#}$; hence, from Theorem 4.1, $\alpha \neq 0$, $\alpha \in \mathrm{dir}(F^{\#} \cap L^{\perp})$.

Conversely, if $\alpha \neq 0$, $\alpha \in \mathrm{dir}(F^{\#} \cap L^{\perp})$, then from Theorem 4.1, $\alpha x \geqslant 0$ defines a facet f_* for $F^{\#\#}$, with $\alpha \in \mathrm{lh}\, F$. Since $\mathrm{clconv}\ F \subseteq F^{\#\#}$, $\alpha x \geqslant 0$ is a supporting halfspace for $\mathrm{clconv}\, F$. The facet f_* of $F^{\#\#} = \mathrm{clcone}\, F$ contains either a facet of $\mathrm{clconv}\, F$, or a $(d-2)$-dimensional face; since f_* itself is a polyhedral cone which can only be generated by a set of the same dimension as that of f_*, or one less. A $(d-2)$-dimensional face of a polyhedron is known to be the intersection of two adjacent facets, and the necessary and sufficient condition for $\alpha x = 0$ to be satisfied by all x satisfying $\alpha^1 x = \alpha_0^1$ and $\alpha^2 x = \alpha_0^2 (\alpha^1 \neq \alpha^2)$, is that $\alpha = \alpha^1 \lambda_1 + \alpha^2 \lambda_2$ for some λ_1, λ_2 such that $\alpha_0^1 \lambda_1 + \alpha_0^2 \lambda_2 = 0$. Setting $\lambda_1 = 1/\alpha_0^1$ yields $\lambda_2 = -1/\alpha_0^2$ and $\alpha = \alpha^1/\alpha_0^1 - \alpha^2/\alpha_0^2$. □

We now turn to the problem of actually calculating facets of $\mathrm{conv}\, F$. To this end, we will use the expression given for $F^{\#}$ in Section 3, namely,

$$F^{\#}(\alpha_0) = \left\{ y \in R^n \;\middle|\; \begin{array}{l} y \geqslant \theta^h A + \sigma^h D^h, \quad \forall h \in Q^*, \\ \text{for some } \theta^h \geqslant 0, \quad \sigma^h \geqslant 0, \;\; h \in Q^*, \\ \text{such that } \theta^h a_0^h + \sigma^h d_0^h \geqslant \alpha_0, \quad h \in Q^* \end{array} \right\}.$$

From Theorem 4.3, all facets of $F^{\#\#}$ of the form $\alpha x \geqslant \alpha_0$ can be obtained by finding the vertices (if $\alpha_0 \neq 0$) or extreme direction vectors (if $\alpha_0 = 0$) of $F^{\#} \cap L^{\perp}$. In each of the two cases, the vertex (extreme direction vector) α of $F^{\#} \cap L^{\perp}$ corresponds to a lowest-dimensional face $\alpha + L$ of $F^{\#}$, where L is the lineality space of $F^{\#}$, and there exists a supporting hyperplane to $F^{\#}$ whose intersection with $F^{\#}$ is precisely $\alpha + L$. If such a hyperplane is known, then the face $\alpha + L$ of $F^{\#}$ can be found by maximizing or minimizing the associated linear function over $F^{\#}$, i.e., by solving a linear program. Moreover, minimizing any linear function which is bounded from below on $F^{\#}$ yields a lowest-dimensional face of $F^{\#}$.

If $0 \in \mathrm{clconv}\, F$, then there are no facets of the form $\alpha x \geqslant \alpha_0$, with $\alpha_0 > 0$, at all. Since we can always choose an expression for (DP) such that $0 \notin \mathrm{clconv}\, F$, (see Section 1), we will assume this to be the case.

The problem of defining the class of linear functions which are bounded from below on $F^{\#}$ is best approached by setting up the linear program with a hypothetical minimand, say gy, and then looking at its dual. Using the more compact notation

$$A^h = \begin{pmatrix} A \\ D^h \end{pmatrix}, \qquad a_0^h = \begin{pmatrix} a_0 \\ d_0^h \end{pmatrix}, \qquad u^h = (\theta^h, \sigma^h)$$

the two problems can be stated as

$$P_1^*(g,\alpha_0)\qquad \begin{array}{ll} \min & yg \\ \text{s.t.} & y - u^h A^h \geqslant 0, \\ & \quad u^h a_0^h \geqslant \alpha_0, \quad h \in Q^*, \\ & \quad u^h \geqslant 0 \end{array}$$

and

$$P_2^*(g,\alpha_0)\qquad \begin{array}{ll} \max & \sum_{h\in Q^*} \alpha_0 \xi_0^h \\ \text{s.t.} & a_0^h \xi_0^h - A^h \xi^h \leqslant 0, \quad h \in Q^*, \\ & \quad \sum_{h\in Q^*} \xi^h = g, \\ & \quad \xi_0^h \geqslant 0, \xi^h \geqslant 0, \quad h \in Q^*. \end{array}$$

The latter problem is, as one would expect it to be, closely related to the representation of clconv$F$ introduced in Section 2. We are interested in characterizing the class of vectors $g \in R^n$ for which $P_1^*(g,\alpha_0)$ has a finite minimum. This of course is the same as the class of $g \in R^n$ for which $P_2^*(g,\alpha_0)$ is feasible and has a finite maximum. Since $F^\# \neq \emptyset$ by assumption, $P_1^*(g,\alpha_0)$ has an optimal solution if and only if $P_2^*(g,\alpha_0)$ has a feasible solution. We will denote the objective function value of $P_2^*(g,\alpha_0)$ by ζ.

Theorem 4.4. (i) *If $g \in \text{clcone}\,F$, $g \neq 0$, then for every $\lambda > 0$ such that $\lambda g \in \text{clconv}\,F$ (and such λ always exists), $P_2^*(g,\alpha_0)$ has a feasible solution $\bar{\xi}$, with $\sum_{h\in Q^*} \bar{\xi}_0^h = \lambda^{-1}$. Conversely, if $\bar{\xi}$ is a feasible solution to $P_2^*(g,\alpha_0)$ with $\sum_{h\in Q^*} \bar{\xi}^h = \lambda^{-1}$, then $g \in$ clcone$F$, $g \neq 0$, and $\lambda g \in \text{clconv}\,F$.*

(ii) *If $\alpha_0 \neq 0$, a feasible solution $\bar{\xi}$ to $P_2^*(g,\alpha_0)$ is optimal if and only if it has value $\bar{\zeta} = \alpha_0 \bar{\lambda}^{-1}$, where*

$$\bar{\lambda} = \begin{cases} \min\{\lambda \mid \lambda g \in \text{clconv}\,F\} & \text{if } \alpha_0 > 0, \\ \max\{\lambda \mid \lambda g \in \text{clconv}\,F\} & \text{if } \alpha_0 < 0. \end{cases}$$

(iii) *If $\alpha_0 = 0$, any feasible solution $\bar{\xi}$ to $P_2^*(g,0)$ is optimal, with value $\bar{\zeta} = 0$; but $P_1^*(g,0)$ has an optimal solution $(\bar{y},\bar{u})$, with $\bar{y}g = 0$ and $\bar{y} \neq 0$, if and only if $g \in bd$ clcone$\,F$.*

Proof. (i) Let $g \in \text{clcone}\,F$, $g \neq 0$. Then $\lambda g \in \text{clconv}\,F$ for some scalar $\lambda > 0$, and for any such λ we have

$$\lambda g = \sum_{h\in Q^*} \left(\sum_{i\in U_h} \theta^{hi} u^{hi} + \sum_{k\in V_h} \sigma^{hk} v^{hk} \right)$$

with

$$\sum_{h\in Q^*}\sum_{i\in U_h}\theta^{hi}=1,\quad \theta^{hi}\geqslant 0,\ i\in U_h,\ \ \sigma^{hk}\geqslant 0,\ k\in V_h,$$

where $u^{hi}\in \operatorname{vert} F_h$, $i\in U_h$, and $v^{hk}\in \operatorname{dir} F_h$, $k\in V_h$, $h\in Q^*$.

Now, let us define $\xi(\lambda)$ by

$$\xi^h(\lambda)=1/\lambda\left(\sum_{i\in U_h}\theta^{hi}u^{hi}+\sum_{k\in V_h}\sigma^{hk}v^{hk}\right),\qquad \xi^h_0(\lambda)=1/\lambda\sum_{i\in U_h}\theta^{hi}.$$

Then $\xi^h(\lambda)\geqslant 0$, $\xi^h_0(\lambda)\geqslant 0$, $\forall h\in Q^*$, and

$$g=\sum_{h\in Q^*}\xi^h(\lambda),$$

$$\begin{aligned}A^h\xi^h(\lambda) &= 1/\lambda\sum_{i\in U_h}\theta^{hi}A^hu^{hi}+1/\lambda\sum_{k\in V_h}\sigma^{hk}A^hv^{hk}\\ &\geqslant 1/\lambda\sum_{i\in U_h}\theta^{hi}a^h_0\quad \text{(since } A^hu^{hi}\geqslant a^h_0, i\in U_h, \text{ and } A^hv^{hk}\geqslant 0, k\in V_h)\\ &= a^h_0\xi^h_0(\lambda),\end{aligned}$$

hence $\xi(\lambda)$ is a feasible solution to $P_2^*(g,\alpha_0)$, with value

$$\zeta(\lambda)=\sum_{h\in Q^*}\alpha_0\xi^h_0(\lambda)=\alpha_0/\lambda.$$

Also, $\sum_{h\in Q^*}\xi^h_0(\lambda)=\lambda^{-1}$.

Conversely, let $\hat{\xi}$ be a feasible solution to $P_2^*(g,\alpha_0)$, with $\hat{\xi}^h_0\neq 0$ for at least one $h\in Q^*$, and with value $\hat{\zeta}$. Let

$$Q_1^*=\{h\in Q\mid \hat{\xi}^h_0>0\},\qquad Q_2^*=\{h\in Q\mid \hat{\xi}^h_0=0\},$$

where $Q_1^*\neq\emptyset$ by assumption. Then

$$\begin{aligned}g&=\sum_{h\in Q^*}\hat{\xi}^h\\ &=\sum_{h\in Q_1}\hat{\xi}^h_0(\hat{\xi}^h/\hat{\xi}^h_0)+\sum_{h\in Q_2^*}\hat{\xi}^h,\end{aligned}$$

where $(\hat{\xi}^h/\hat{\xi}^h_0)\in F_h\subseteq \operatorname{clconv} F$, $h\in Q_1^*$, whereas for $h\in Q_2^*$, either $\hat{\xi}^h=0$, or $\hat{\xi}$ is a direction vector of the recession cone (characteristic cone) of F_h, hence of $\operatorname{clconv} F$. Thus $g\in \operatorname{clcone} F$. Further, setting

$$\lambda=\left(\sum_{h\in Q^*}\hat{\xi}^h_0\right)^{-1},$$

we have

$$\lambda g = \sum_{h \in Q_1^*} \lambda \hat{\xi}_0^h (\hat{\xi}^h / \hat{\xi}_0^h) + \sum_{h \in Q_2^*} \lambda \hat{\xi}^h$$

with

$$\sum_{h \in Q_1^*} \lambda \hat{\xi}_0^h = 1,$$

since $\hat{\xi}_0^h = 0$, $\forall h \in Q_2^*$; hence $\lambda g \in \operatorname{clconv} F$, and $g \neq 0$.

This proves (i).

(ii) Let $\alpha_0 \neq 0$. From (i), any feasible solution $\bar{\xi}$ satisfies $\sum_{h \in Q^*} \bar{\xi}_0^h = \lambda^{-1}$ for some λ such that $\lambda g \in \operatorname{clconv} F$. Since $\bar{\zeta} = \alpha_0 \lambda^{-1}$, $\bar{\xi}$ is clearly optimal if and only if $\bar{\zeta} = \alpha_0 \bar{\lambda}^{-1}$, where $\bar{\lambda}$ is as defined in the theorem. This proves (ii).

(iii) Let $\alpha_0 = 0$. If $g \in bd \operatorname{clcone} F$, then $\alpha g = 0$ for some facet $\alpha x \geqslant 0$ of $F^{\#\#}$, where $\alpha \neq 0$, $\alpha \in \operatorname{lh} F$. From Corollary 4.1.1, α is then an extreme direction vector of $F^{\#} \cap L^{\perp}$, contained in the supporting hyperplane $gy = 0$ to $F^{\#}$. Therefore, $\alpha = \bar{y}$ for some optimal solution $(\bar{y}, \bar{u})$ of $P_1^*(g, 0)$, such that $\bar{y} \neq 0$ and $\bar{y}g = 0$.

Conversely, if $(\bar{y}, \bar{u})$ is an optimal solution to $P_1^*(g, 0)$, with $\bar{y} \neq 0$ and $\bar{y}g = 0$, then $gy \geqslant 0$, $\forall y \in F^{\#}$; hence $g \in F^{\#\#}(0) = \operatorname{clcone} F$; $g\bar{y} = 0$ for $\bar{y} \in F^{\#}$, i.e., $g \in bd \operatorname{clcone} F$. □

Next, we examine the connection between $F^{\#}, F^{\#} \cap L^{\perp}$, which are subsets of $\mathbb{R}^n$, and the feasible set of $P_1^*(g, \alpha_0)$ i.e., the set

$$U = \left\{ (y,u) \;\middle|\; \begin{array}{ll} y - u^h A^h \geqslant 0, & \\ u^h a_0^h \geqslant \alpha_0, & h \in Q^* \\ u^h \geqslant 0, & \end{array} \right\},$$

which belongs to a higher dimensional space.

First, we recall that whenever $L \neq \emptyset$, i.e., $\operatorname{lin} F^{\#} > 0$, $F^{\#}$ has no vertices, while its smallest dimensional faces are those of dimension equal to $\operatorname{lin} F^{\#}$, which are in a one-to-one correspondence with the vertices of $F^{\#} \cap L^{\perp}$. On the other hand, the (polyhedral) set U is always pointed, i.e., always has vertices (its recession come $C(U)$ contains no lines). To see this, notice that if there exists $(\hat{y}, \hat{u})$ such that $(\lambda\hat{y}, \lambda\hat{u}) \in C(U)$ for all λ, then from $u^h \geqslant 0, h \in Q^*$ one has $\hat{u}^h = 0$, $\forall h \in Q^*$, which in turn implies $\hat{y} = 0$; i.e., there exists no line contained in $C(U)$.

The question arises then, what corresponds in U to a vector $y \in F^{\#}$, $y \notin L^{\perp}$; or, more specifically, what corresponds in $C(U)$ to a line in $L \subset C(F^{\#})$; and the answer is, a pair of directions. Let $\bar{y} \in F^{\#}$, $\bar{y} \notin L^{\perp}$; then $\bar{y}$ has a unique expression of the form $\bar{y} = \hat{y} + \tilde{y}$, where $\hat{y} \in F^{\#} \cap L^{\perp}$ and $\tilde{y} \in L$. In other words, this expression is such that $\hat{y} + \lambda\tilde{y} \in F^{\#}$ for any λ. In terms of U, this implies that (i) there exists $\hat{u}$ such that $(\hat{y}, \hat{u}) \in U$; and (ii) there exists a *pair* of vectors $\tilde{u}, \tilde{\tilde{u}}$, such that

$$(\hat{y}, \hat{u}) + (\lambda\tilde{y}, \lambda\tilde{u}) \in U \quad \text{and} \quad (\hat{y}, \hat{u}) + (-\lambda\tilde{y}, \lambda\tilde{\tilde{u}}) \in U$$

for all $\lambda \geqslant 0$; i.e., a pair of vectors $\tilde{u}, \tilde{\tilde{u}}$ satisfying

$$\begin{array}{lcl} \tilde{y} - \tilde{u}^h A^h \geqslant 0 & & -\tilde{y} - \tilde{\tilde{u}}^h A^h \geqslant 0, \\ \tilde{u}^h a_0^h \geqslant 0 & \text{and} & \tilde{\tilde{u}}^h a_0^h \geqslant 0, \\ \tilde{u}^h \geqslant 0 & & \tilde{\tilde{u}}^h \geqslant 0, \\ \tilde{y}_i < 0 \Rightarrow \tilde{u}^h a_i^h < 0 & & -\tilde{y}_i < 0 \Rightarrow \tilde{\tilde{u}}^h a_i^h < 0 \end{array}$$

for all $h \in Q^*$ (here a_i^h is the ith column of A^h).

Another way of looking at this is as follows. Let $\dim L = \operatorname{lin} F^\# = l$, and A_L be a $n \times l$ matrix whose columns generate the subspace L; i.e.,

$$L = \{x \in R^n \mid x = A_L \lambda\},$$

where λ is an arbitrary l-vector. Then

$$\begin{aligned} L^\perp &= \{y \in R^n \mid yx = 0, \ \forall x \in L\} \\ &= \{y \in R^n \mid yA_L = 0\}, \end{aligned}$$

i.e., $L^\perp$ is the null-space of A_L^{T}, the transpose of A_L. Then

$$F^\# \cap L^\perp = \left\{ y \in R^n \,\middle|\, \begin{array}{ll} yA_L = 0, & \\ y - u^h A^h \geqslant 0, & \\ u^h a_0^h \geqslant \alpha_0, & h \in Q^* \\ u^h \geqslant 0 & \end{array} \right\}$$

and

$$U \cap L^\perp = \left\{ (y, u) \,\middle|\, \begin{array}{ll} yA_L = 0, & \\ y - u^h A^h \geqslant 0, & \\ u^h a_0^h \geqslant \alpha_0, & h \in Q^* \\ u^h \geqslant 0, & \end{array} \right\}.$$

Now, let $\alpha_0 \neq 0$. A point $\bar{y} \in F^\#$ is a vertex of $F^\# \cap L^\perp$ if and only if (a) $\bar{y} \in L^\perp$, and (b) there exists $p \in L^\perp$ such that $\bar{y}$ is the unique point which minimizes py on $F^\#$ and satisfies (a). Accordingly, if $(\bar{y}, \bar{u})$ is an optimal solution to $P_1^*(g, \alpha_0)$ for some $g \in \operatorname{clcone} F$, then $\bar{y} \in \operatorname{vert}(F^\# \cap L^\perp)$ if and only if

(i) $\bar{y} \in L^\perp$; and

(ii) there exists $\gamma \in L^\perp$ such that $y = \bar{y}$ for every optimal solution (y, u) to $P_1^*(g + \gamma, \alpha_0)$ for which $y \in L^\perp$.

An optimal solution to $P_1^*(g, \alpha_0)$ which satisfies (i) and (ii), will be called *regular*.

If $\alpha_0 = 0$, then $F^\#$ is a cone, and $F^\# \cap L^\perp$ a pointed cone. A vector $\bar{y} \in F^\#$ defines an extreme direction (i.e., is an extreme direction vector) of $F^\# \cap L^\perp$ if and only if (a) $\bar{y} \neq 0$, $\bar{y} \in L^\perp$; and (b) there exists $p \in L^\perp$ such that, up to a positive multiplier, $\bar{y}$ is the unique point which minimizes py on $F^\#$ and satisfies (a). Hence, if $(\bar{y}, \bar{u})$ is an optimal solution to $P_1^*(g,0)$ for some $g \in bd\,\mathrm{clcone}\,F$, then $\bar{y}$ is an extreme direction vector of $F^{\#} \cap L^{\perp}$ if and only if

(i) $\bar{y} \neq 0$, $y \in L^{\perp}$; and

(ii) there exists $\gamma \in L^{\perp}$ such that $y = \lambda \bar{y}$, $\lambda > 0$, for every optimal solution (y,u) to $P_1^*(g+\gamma,0)$ for which $y \neq 0, y \in L^{\perp}$.

An optimal solution to $P_1^*(g,0)$ which satisfies (i) and (ii) will again be called *regular*.

Lemma 4.1. *If $P_1^*(g,\alpha_0)$ has an optimal solution, it has a regular optimal solution.*

Proof. Let $\alpha_0 \neq 0$. From Theorem 4.4, if $P_1^*(g,\alpha_0)$ has an optimal solution, then $g \in \mathrm{clcone}\,F$, $g \neq 0$; hence $g \in L^{\perp}$. Further, if $(\bar{y},\bar{u})$ is an optimal solution to $P_1^*(g,\alpha_0)$, then for every $y \in (\bar{y}+L)$ there exists some u such that (y,u) is an optimal solution to $P_1^*(g,\alpha_0)$. Hence, if $P_1^*(g,\alpha_0)$ has an optimal solution (y,u), then it has one with $y \in L^{\perp}$. Also, if $(\bar{y},\bar{u})$ is an optimal solution to $P_1^*(g,\alpha_0)$, with $\bar{y}g = \alpha_0 \bar{\lambda}^{-1}$, then $\bar{\lambda}gx = \alpha_0$ is a supporting hyperplane to $F^{\#} \cap L^{\perp}$ which contains $\bar{y}$. If $\bar{y}$ is the only point of $F^{\#} \cap L^{\perp}$ contained in $\bar{\lambda}gx = \alpha_0$, then $(\bar{y},\bar{u})$ is regular. Otherwise, $\bar{\lambda}gx = \alpha_0$ contains a face of $F^{\#} \cap L^{\perp}$ of dimension $d \geqslant 1$ and, since $F^{\#} \cap L^{\perp}$ is pointed, every such face contains at least one vertex of $F^{\#} \cap L^{\perp}$. Hence, the hyperplane $\bar{\lambda}gx = \alpha_0$ can be "tilted" so as to intersect $F^{\#} \cap L^{\perp}$ in only one point, i.e., the coefficients $\bar{\lambda}g$ can be replaced by $\bar{\lambda}g + \bar{\lambda}\gamma$, where $\gamma \in L^{\perp}$, so that this is achieved.

A perfectly analogous argument holds when $\alpha_0 = 0$. □

Theorem 4.5. *Let $g \in \mathrm{clcone}\,F$, $g \neq 0$, and*

$$\bar{\lambda} = \begin{cases} \min\{\lambda \mid \lambda g \in \mathrm{clconv}\,F\} & \text{if } \alpha_0 > 0, \\ \max\{\lambda \mid \lambda g \in \mathrm{clconv}\,F\} & \text{if } \alpha_0 < 0. \end{cases}$$

Then $\alpha x \geqslant \alpha_0$, where $\alpha_0 \neq 0$ and $\alpha \in l\mathrm{h}\,F$, is a facet of $F^{\#\#}$ containing the point $\bar{\lambda}g$, if and only if $\alpha = \bar{y}$ for some regular optimal solution $(\bar{y},\bar{u})$ to $P_1^*(g,\alpha_0)$.

Proof. From Theorem 4.1, $\alpha x \geqslant \alpha_0$, with $\alpha_0 \neq 0$ and $\alpha \in \mathrm{lh}\,F$, is a facet of $F^{\#\#}$ if and only if $\alpha \in \mathrm{vert}(F^{\#} \cap L^{\perp})$; and from Corollary 4.1.1, the facet $\alpha x \geqslant \alpha_0$ of $F^{\#\#}$ contains the point $\bar{\lambda}g$ if and only if $\bar{\lambda}gx = \alpha_0$ is a supporting hyperplane for $F^{\#}$, which contains the vertex α of $F^{\#} \cap L^{\perp}$. But this latter condition holds if and only if $\bar{\lambda}\alpha g = \min\{\bar{\lambda}yg \mid y \in F^{\#} \cap L^{\perp}\} = \alpha_0$, i.e., if and only if $\alpha = \bar{y}$ for some regular optimal solution $(\bar{y},\bar{u})$ to $P_1^*(g,\alpha_0)$, with objective function value $\bar{y}g = \alpha_0\bar{\lambda}^{-1}$. From Lemma 4.1, $P_1^*(g,\alpha_0)$ has a regular optimal solution if (and, obviously, only if) it has any optimal solution at all; and, finally, from Theorem 4.4, the optimal objective function value is $\bar{y}g = \alpha_0\bar{\lambda}^{-1}$. □

The analogous result for the case $\alpha_0 = 0$ follows.

Theorem 4.6. *Let $g \in bd\,\mathrm{clcone}\,F$, $g \neq 0$. Then $\alpha x \geqslant 0$, with $\alpha \in \mathrm{lh}\,F$, is a facet of $F^{\#\#}$ containing the point g, if and only if $\alpha = \lambda\bar{y}$ for some $\lambda > 0$ and some regular optimal solution $(\bar{y},\bar{u})$ to $P_1^*(g,0)$.*

Proof. From Theorems 4.1 and 4.3, $\alpha x \geqslant 0$, with $\alpha \in \mathrm{lh}\,F$, is a facet of $F^{\#\#}$ (hence, from Theorem 3.2, of $\mathrm{clcone}\,F$) if and only if $\alpha \in \mathrm{dir}(F^{\#} \cap L^{\perp})$; and from Corollary 4.1.1, $\alpha g = 0$ if and only if $gx = 0$ is a supporting hyperplane for $F^{\#}$ which contains the extreme direction vector α of $F^{\#} \cap L^{\perp}$. This latter condition holds, however, if and only if $\min\{yg \mid y \in F^{\#} \cap L^{\perp}\} = 0$, and this minimum is attained for $\bar{y} \neq 0$ such that $\alpha = \lambda \bar{y}$, with $\lambda > 0$; i.e., if and only if $\alpha = \lambda \bar{y}$ for some $\lambda > 0$ and some regular optimal solution $(\bar{y}, \bar{u})$ to $P_1^*(g, 0)$, with objective function value $\bar{y}g = 0$. From Lemma 4.1, $P_1^*(g, 0)$ has a regular optimal solution if (and only if) it has an optimal solution; and from Theorem 4.4, the optimal objective function value is $\bar{y}g = 0$. □

Using Theorems 4.5 and 4.6, all the facets of $F^{\#\#}$ can be obtained by solving the problem $P_1^*(g, \alpha_0)$, or its dual, for various vectors $g \in \mathrm{clcone}\,F$. From Theorems 4.2 and 4.3, each such facet is, or yields in conjuction with some other facet of $F^{\#\#}$, a facet of $\mathrm{clconv}\,F$. Note that, if $\alpha_0 \neq 0$ and $\bar{\lambda}$ is defined as in Theorem 4.5, and $\bar{\lambda}g$ is the convex combination of k vertices and extreme direction vectors of $\mathrm{clconv}\,F$, $(1 \leqslant k \leqslant n)$, then each of these vertices and extreme direction vectors are contained in each facet of $\mathrm{clconv}\,F$ that contains $\bar{\lambda}g$; and that each such facet can be obtained by solving $P_1^*(g, \alpha_0)$ or $P_2^*(g, \alpha_0)$. An analogous statement holds, of course, for the case when $\alpha_0 = 0$. Note also that if g is a vertex of $\mathrm{clconv}\,F$, then by solving $P_1^*(g, \alpha_0)$ for $\alpha_0 = 1, -1$ and 0 we obtain all the facets of $\mathrm{clconv}\,F$ containing the vertex g; furthermore, these facets correspond, for a given α_0, to alternative regular optimal solutions of the same linear program $P_1^*(g, \alpha_0)$; so that if one facet containing g, i.e., one regular optimal solution to $P_1^*(g, \alpha_0)$ is found, the other facets containing g are easy to obtain.

We will conclude this section with a few considerations on the practical solvability of $P_1^*(g, \alpha_0)$ or its dual. First, if $\alpha_0 = 0$, then $P_1^*(g, \alpha_0)$ has a homogeneous constraint set and thus has no nontrivial basic solution. On the other hand, if any nonzero vector (y, u) is an optimal solution, then so is $(\lambda y, \lambda u)$ for any $\lambda > 0$. Therefore $P_1^*(g, 0)$ can be normalized, for instance, by adding the constraint $eu^h \leqslant 1$, where $e = (1, \ldots, 1)$, to each subsystem $h \in Q^*$. The new problem will then have nontrivial basic solutions (if it has nontrivial solutions at all), with the same optimal objective function value as before.

Another problem that arises, whatever the value of α_0, is that the set $Q^* = \{h \in Q \mid F_h \neq \emptyset\}$ is usually not known. This difficulty can be circumvented by using the following result.

Let $P_1(g, \alpha_0)$, $P_2(g, \alpha_0)$, be the pair of dual linear programs obtained from $P_1^*(g, \alpha_0)$, $P_2^*(g, \alpha_0)$, by replacing Q^* with Q. Then we have the following correspondence between $P_1(g, \alpha_0)$, $P_2(g, \alpha_0)$ and their starred counterparts.

Theorem 4.7. (i) *If $P_1^*(g, \alpha_0)$ has a feasible solution, then $P_1(g, \alpha_0)$ has a feasible solution; and if $P_2^*(g, \alpha_0)$ has a feasible solution, then $P_2(g, \alpha_0)$ has a feasible solution.*

(ii) *If $P_2(g, \alpha_0)$ has an optimal solution $\bar{\xi}$ such that*

$$\bar{\xi}_0^h = 0, \quad \bar{\xi}^h \neq 0 \Rightarrow h \in Q^*,$$

then the vector obtained from $\bar{\xi}$ by removing the components $(\bar{\xi}^h, \bar{\xi}_0^h)$, $h \in Q - Q^*$, is an optimal solution to $P_2^*(g, \alpha_0)$; and all optimal solutions $(\bar{y}, \bar{u})$ to $P_1(g, \alpha_0)$ are optimal solutions to $P_1^*(g, \alpha_0)$.

Proof. (i) If $\alpha_0 \leqslant 0$, then $P_1(g, \alpha_0)$ is always feasible, just like $P_1^*(g, \alpha_0)$. If $\alpha_0 > 0$, then $P_1^*(g, \alpha_0)$ has a feasible solution, i.e., $F^{\#} \neq \emptyset$, if and only if $0 \notin \operatorname{clconv} F$ (see Theorem 3.2); which implies that the system $A^h x \geqslant a_0^h$, $x = 0$, has no solution for any $h \in Q^*$, and hence (from the definition of $Q - Q^*$) for any $h \in Q$. From Farkas' theorem it then follows that the system

$$u^h A^h \leqslant 0, \quad u^h a_0^h \geqslant \alpha_0, \quad u^h \geqslant 0$$

has a solution for each $h \in Q$. Let $\bar{u}^h$, $h \in Q$ be a set of solutions to these systems; then $y \in R^n$ defined by

$$\bar{y}_i = \max_{h \in Q} \bar{u}^h a_i^h, \quad i = 1, \ldots, n$$

(where a_i^h is the ith column of A^h), together with the vectors $\bar{u}^h$, $h \in Q$, is a feasible solution to $P_1(g, \alpha_0)$.

On the other hand, any feasible solution to $P_2^*(g, \alpha_0)$ trivially yields a feasible solution to $P_2(g, \alpha_0)$ by setting $(\xi^h, \xi_0^h) = (0, 0)$, $\forall h \in Q - Q^*$.

(ii) If $\bar{\xi}$ is an optimal solution to $P_2(g, \alpha_0)$, then $\bar{\xi}_0^h = 0$, $\forall h \in Q - Q^*$ (since if $\bar{\xi}_0^h > 0$, then $A^h \xi^h \geqslant a^h \bar{\xi}_0^h$, $\xi^h \geqslant 0$ has no solution). Therefore, if $\bar{\xi}$ satisfies the requirement of the theorem, then $\bar{\xi}^h = 0$, $\forall h \in Q - Q^*$, and the vector obtained from $\bar{\xi}$ by removing the components $(\bar{\xi}^h, \bar{\xi}_0^h)$, $h \in Q - Q^*$, is a feasible solution to $P_2^*(g, \alpha_0)$. It must also be optimal, since any better feasible solution to $P_2^*(g, \alpha_0)$ would yield a better feasible solution to $P_2(g, \alpha_0)$, according to (ii). Therefore, all optimal solutions $(\bar{y}, \bar{u})$ to $P_1(g, \alpha_0)$ have a value of $g\bar{y} = \sum_{h \in Q} \alpha_0 \bar{\xi}_0^h$. Since all feasible solutions to $P_1(g, \alpha_0)$ are trivially feasible for $P_1^*(g, \alpha_0)$, and the optimal value of a feasible solution to $P_1^*(g, \alpha_0)$ is also $\sum_{h \in Q} \alpha_0 \bar{\xi}_0^h$, it follows that all optimal solutions to $P_1(g, \alpha_0)$ are also optimal for $P_1^*(g, \alpha_0)$. $\square$

Thus, $P_2^*(g, \alpha_0)$ can be solved by solving $P_2(g, \alpha_0)$ (repeatedly, if necessary). If $\bar{\xi}$ is the optimal solution obtained and $\bar{\xi}$ has a component $(\bar{\xi}^h, \bar{\xi}_0^h)$ such that $\bar{\xi}_0^h = 0$, $\bar{\xi}^h \neq 0$, one has to check whether $A^h x \geqslant a_0^h$, $x \geqslant 0$ is feasible. If it is not, then Q has to be replaced by $Q - \{h\}$ and a new optimal solution has to be found. This procedure has to be repeated until a regular optimal solution is obtained which satisfies the condition of Theorem 4.7.

The larger the set Q, the more costly it becomes to solve the linear program $P_1(g, \alpha_0)$ or $P_2(g, \alpha_0)$. Thus, if the disjunctive program (DP) is a mixed integer 0–1 program with n 0–1 variables, $|Q| = 2^n$ and calculating facets of $\operatorname{clconv} F$ by solving $P_1(g, \alpha_0)$ or $P_2(g, \alpha_0)$ is obviously not a practical approach. On the other hand, when Q is small, $P_2(g, \alpha_0)$ is quite easy to solve and thus facets of $\operatorname{clconv} F$ can easily be calculated. Thus, it is easy to generate facets for a linear program with a single or just a few disjunctive constraints of the form $x_j = 0 \vee x_j = 1$, and it is also not very difficult to generate facets for a linear program with variables $0 \leqslant x_j \leqslant 1$, $j = 1, \dots, n$, and a disjunctive constraint of the form $x_1 = 1 \vee \cdots \vee x_k = 1$, where $1 \leqslant k \leqslant n$. The question is, however, whether facets obtained for such a problem could be used to solve the same linear program subject to additional disjunctive constraints.

The next section examines this question.

5. Facial disjunctive programs and the significance of "partial" convex hulls

Consider a disjunctive program (DP) with its logical constraints stated in conjunctive normal form, i.e.,

(DP) $\quad \min\{cx \mid x \in F^S\}$

where the superscript S refers to the set of disjunctions in (2), i.e.,

$$F^S = \left\{ x \in R^n \;\middle|\; \begin{array}{c} Ax \geqslant a_0 \\ x \geqslant 0 \\ \bigwedge_{j \in S} \left[\bigvee_{i \in Q_j} (d^i x \geqslant d_{i0}) \right] \end{array} \right\}.$$

(F^S is just a more specific notation for the set denoted earlier by F.)

Clearly, for any subset $T \subset S$, we have $F^T \supseteq F^S$. We wish to examine the connections between $\operatorname{clconv} F^S$ and $\operatorname{clconv} F^T$ for $T \subset S$. For lack of a better term, we will call $\operatorname{clconv} F^T$ a "partial" convex hull for (DP).

Our motivation is the following. In Section 4 we have shown that when T is small and each Q_j is small, a facet of $\operatorname{clconv} F^T$ is quite easy to generate. This is in sharp contrast with the facets of $\operatorname{clconv} F^S$, which are in general very expensive to generate. The question then arises, are such facets of a "partial" convex hull of any use? For instance, suppose one tries to solve a 0–1 program by branch and bound and after having solved the original linear program and the four subproblems corresponding to the constraints

$$\left\{ \begin{array}{c} x_1 = 0 \\ x_2 = 0 \end{array} \right\}, \quad \left\{ \begin{array}{c} x_1 = 0 \\ x_2 = 1 \end{array} \right\}, \quad \left\{ \begin{array}{c} x_1 = 1 \\ x_2 = 0 \end{array} \right\}, \quad \left\{ \begin{array}{c} x_1 = 1 \\ x_2 = 1 \end{array} \right\},$$

one generates all the facets of the "partial" convex hull H of the set of points satisfying the linear programming constraints and the disjunction

$$\left\{ \begin{array}{c} x_1 = 0 \\ x_2 = 0 \end{array} \right\} \vee \left\{ \begin{array}{c} x_1 = 0 \\ x_2 = 1 \end{array} \right\} \vee \left\{ \begin{array}{c} x_1 = 1 \\ x_2 = 0 \end{array} \right\} \vee \left\{ \begin{array}{c} x_1 = 1 \\ x_2 = 1 \end{array} \right\}.$$

If one now throws out the four subproblems, and starts again the branch and bound procedure, applying it this time to the 0–1 program over the "partial" convex hull H, can one safely assume that the variables x_1 and x_2 will only take on values 0 or 1, i.e., can the 0–1 constraints on x_1 and x_2 be removed from the new 0–1 program? The answer, contained in the developments to follow, is yes, for a large class of disjunctive programs (including 0–1 programs).

In this section, we focus our attention on a class of disjunctive programs which have an important special property.

Let

$$F_0 = \{x \in R^n \mid Ax \geqslant a_0,\ x \geqslant 0\}$$

and let the constraints of the disjunctive program (DP) be stated in the conjunctive normal form (2).

A disjunction

$$\bigvee_{i \in Q_j} (d^i x \geqslant d_{i0})$$

will be called *facial* or said to have the facial property with respect to F_0, if

$$F_i = F_0 \cap \{x \in R^n \mid d^i x \geqslant d_{i0}\}$$

is a face of F_0, for all $i \in Q_j$.

We will say that the disjunctive program (DP) is *facial*, if all the disjunctions in (2) are facial; i.e., if F_i is a face of F_0 for all $i \in Q_j$ and all $j \in S$. (A face of a polyhedral set P is the intersection of P with some of its boundary planes.)

The class of disjunctive programs which have the facial property includes the most important cases of disjunctive programming, like 0–1 programming, nonconvex quadratic and separable programming, and many others. In all of these cases the inequalities $d^i x \geqslant d_{i0}$ of each disjunction usually define facets, i.e., $(d-1)$-dimensional faces of F_0, where d is the dimension of F_0. The next theorem gives a necessary and sufficient condition for $d^i x \geqslant d_{i0}$ to have the more general property of defining a face of arbitrary dimension of F_0.

Theorem 5.1. *Let*

$$F_i = F_0 \cap \{x \in R^n \mid d^i x \geqslant d_{i0}\}.$$

If there exists $(\mu, \nu) \in R^m \times R^n$, *satisfying*

$$\begin{aligned} &\mu(-A) + \nu(-I) = d^i,\\ &\mu(-a_0) = d_{i0}, \\ &(\mu, \nu) \geqslant 0, \end{aligned} \tag{3}$$

then F_i is a face of F_0; namely

$$\begin{aligned} F_i &= \{x \in F_0 \mid d^i x = d_{i0}\}\\ &= \{x \in F_0 \mid a^h x = a_{h0},\ \forall h \in M^+;\ x_j = 0,\ \forall j \in N^+\}, \end{aligned}$$

where a^h is row h of A, and

$$M^+ = \{h \in M \mid \mu_h > 0\}, \qquad N^+ = \{j \in N \mid v_j > 0\}.$$

Conversely, if F_i is a face of F_0, and $F_0 \neq F_i \neq \emptyset$, then there exists $(\mu, v) \in R^m \times R^n$ satisfying (3).

Proof. 'From Farkas' theorem on linear inequalities, system (3) has a solution if and only if the homogeneous system

$$\begin{aligned} &-Ax - a_0 x_0 \leqslant 0, \\ &-x \leqslant 0, \\ &d^i x + d_{i0} x_0 > 0 \end{aligned} \tag{4}$$

has no solution.

(i) Assume there exists (μ, v) satisfying (3). Then (4) has no solution, hence neither does its nonhomogeneous correspondent, obtained by setting $x_0 = -1$. Therefore the inequalities $Ax \geqslant a_0$, $x \geqslant 0$, imply $d^i x \leqslant d_{i0}$, and thus

$$F_i = F_0 \cap \{x \in R^n \mid d^i x = d_{i0}\}.$$

Further, we claim that $x \in F_0$ satisfies $d^i x = d_{i0}$ if and only if it satisfies

$$\begin{aligned} &a^h x = a_{h0}, \quad \forall h \in M^+, \\ &x_j = 0, \quad \forall j \in N^+. \end{aligned}$$

The if part is obvious, since from (3) $d^i x = d_{i0}$ is the weighted sum of these equations, with weights $\mu_h > 0$, $h \in M^+$, and $v_j > 0$, $j \in N^+$. To see the only if part, assume it to be false. Then either $a^k \bar{x} \neq a_{k0}$ for some $\bar{x} \in F$ and $k \in M^+$ (case 1), or $\bar{x}_j \neq 0$ for some $\bar{x} \in F$ and $j \in N^+$ (case 2). We discuss case 1 only, since the same reasoning applies to case 2.

Since $\bar{x} \in F$, $a^k \bar{x} \neq a_{k0}$ implies $a^k \bar{x} > a_{k0}$. Multiplying this inequality by $-\mu_k$ yields

$$-\mu_k a^{k-} x < -\mu_k a_{k0},$$

whereas substituting for d^i and d_{i0} in $d^i \bar{x} = d_{i0}$ yields (after changing all signs)

$$\mu A \bar{x} + v \bar{x} = \mu a_0.$$

Adding the last two relations then results in

$$\mu' A \bar{x} + v \bar{x} < \mu' a_0,$$

with $\mu'_h = \mu_h$, $\forall h \neq k$, and $\mu'_k = 0$. But since $(\mu', v) \geqslant 0$, this last inequality contradicts $A\bar{x} + \bar{x} \geqslant a_0$, which follows from $\bar{x} \in F$.

Hence, if there exists (μ, v) as required in the theorem, then

$$F_i = F_0 \cap \{x \in R^n \mid a^h x = a_{h0}, \ \forall h \in M^+; \ x_j = 0, \ \forall j \in N^+\},$$

i.e., F_i is a face of F_0. (This face may be the empty set; it may also be F_0 itself, if F_0 is less than full dimensional.)

(ii) Assume now that F_i is a face of F_0, and $F_0 \neq F_i \neq \emptyset$. Then there exists at least one inequality $gx \geqslant g_0$ of F_0 (either of the form $a^k x \geqslant a_{k0}$, where a^k is row k of A, or of the form $x_j \geqslant 0$), which is not tight for every $x \in F_0$, but is tight for every $x \in F_0$ satisfying $d^i x \geqslant d_{i0}$. In other words, $(-g)x \geqslant -g_0$ is a consequence of the constraints $Ax \geqslant a_0$, $x \geqslant 0$ and $d^i x \geqslant d_{i0}$, but not of the constraints $Ax \geqslant a_0$ and $x \geqslant 0$. From that same well-known result on linear inequalities used in the proof of Theorem 3.1, this implies that the system

$$\begin{aligned} -g &= \theta A + \sigma I + \theta_0 d^i, \\ -g_0 &\leqslant \theta a_0 + \theta_0 d_{i0} \end{aligned} \tag{5}$$

has a solution $\theta \geqslant 0$, $\sigma \geqslant 0$, $\theta_0 \geqslant 0$, but the system obtained from (5) by deleting the last column has no solution $\theta \geqslant 0$, $\sigma \geqslant 0$. Hence (5) has a solution $\theta \geqslant 0$, $\sigma \geqslant 0$, $\theta_0 > 0$. Further (5) must have a solution in which the last inequality is tight, since otherwise there exists $\varepsilon > 0$ such that $(-g)x \geqslant -g_0 + \varepsilon$ is also a consequence of $Ax \geqslant a_0$, $x \geqslant 0$, and $d^i x \geqslant d_{i0}$, i.e., $F_i = \emptyset$, contrary to the assumption. Replacing the inequality in (5) with equality and dividing with $\theta_0 > 0$ yields the system

$$\theta'(-A) + \theta_0^{-1}(-g) + \sigma'(-I) = d^i,$$

$$\theta'(-a_0) + \theta_0^{-1}(-g_0) = d_{i0}$$

with $\theta' = \theta/\theta_0 \geqslant 0$, $\sigma' = \sigma/\theta_0 \geqslant 0$, which can be restated in the form (3), with $\lambda_0 = 0$, by setting

$$\mu_k = \begin{cases} \theta'_k + \theta_0^{-1} > 0 & \text{if } gx \geqslant g_0 \text{ is } a^k x \geqslant a_{k0}, \\ \theta'_k \geqslant 0 & \text{otherwise,} \end{cases}$$

$$\nu_j = \begin{cases} \sigma'_j + \theta_0^{-1} > 0 & \text{if } gx \geqslant g_0 \text{ is } x_j \geqslant 0, \\ \sigma'_j \geqslant 0 & \text{otherwise.} \end{cases}$$

Thus, if F_i is a face of F_0 and $F_0 \neq F_i \neq \emptyset$, then (3) has a solution, with $\lambda_0 = 0$. □

For the remainder of this section we assume that (DP) is facial. Further, we also assume that F_0 is bounded. This is an important restriction from the theoretical point of view, but inconsequential in practice, since boundedness can always be achieved by regularizing F_0.

One important consequence of the facial property of a disjunctive program (DP) is that every vertex of clconv F^S, the closed convex hull of feasible points of (DP), is a vertex of F_0, as shown in the next theorem.

Using again the disjunctive normal form, let

$$F^S = \left\{ x \in R^n \,\middle|\, \bigvee_{h \in Q} \begin{pmatrix} Ax \geqslant a_0 \\ x \geqslant 0 \\ D^h x \geqslant d_0^h \end{pmatrix} \right\},$$

where $Q = \prod_{j \in S} Q_j$. Denoting

$$F_h = \{x \in F_0 \mid D^h x \geqslant d_0^h\},$$

we have $F^S = \bigcup_{h \in Q} F_h$. Since F_0 is bounded, each set F_h is a polytope; hence $\operatorname{conv} F^S$ is closed, i.e., also a polytope.

Theorem 5.2. *vert* $\operatorname{conv} F^S = (\bigcup_{h \in Q} \operatorname{vert} F_h) \subseteq \operatorname{vert} F_0$.

Proof. From the facial property of (DP), each set F_h is the intersection of m_h faces of F_0 (where m_h is the number of rows of D^h), hence a face of F_0. But every face of a polytope is the closed convex hull of a subset of vertices of the polytope; hence $\operatorname{vert} F_h \subseteq \operatorname{vert} F_0$, and therefore

$$\operatorname{vert} \operatorname{conv} F^S \subseteq \left(\bigcup_{h \in Q} \operatorname{vert} F_h \right) \subseteq \operatorname{vert}\ F_0.$$

To show that vert $\operatorname{conv} F^S \supseteq (\bigcup_{h \in Q} \operatorname{vert} F_h)$, assume this to be false, and let $x \in \operatorname{vert} F_k$ for some $k \in Q$, but $x \notin \operatorname{vert} \operatorname{conv} F^S$. Then x is the convex combination of vertices of $\operatorname{conv} F^S$, hence of F_0; which contradicts the fact that x itself is a vertex of F_0. □

Our next result shows that Corollary 2.1.2 of Section 2 can be considerably strengthened for the case when (DP) has the facial property.

Corollary 5.2.1. *If (DP) is facial and bounded, then the vertices of clconv F^S are in a one-to-one correspondence with the vertices of the feasible set P of the linear program ($\mathscr{LP}$) of Corollary* 2.1.2.

Proof. If x is a vertex of $\operatorname{conv} F^S$, then from Corollary 2.1.2, ξ defined by $(\xi^k, \xi_0^k) = (x, 1)$ for some $k \in Q$, and $(\xi^h, \xi_0^h) = (0, 0)$, $\forall h \in Q - \{k\}$, is a vertex of P. Also, from Corollary 2.1.2, if ξ is a vertex of P, then $\xi^k = 1$ for some $k \in Q$, $(\xi^h, \xi_0^h) = (0, 0)$ for $h \in Q - \{k\}$, and $x = \xi^k$ is a vertex of $F_k = \{x \in F_0 \mid D^k x \geqslant d_0^k\}$. But if (DP) is facial and bounded, then from Theorem 5.2, $x \in \operatorname{vert} F_k \Rightarrow x \in \operatorname{vert} \operatorname{conv} F^S$. □

From a practical point of view, the most important consequence of the facial property is the fact, to be shown below, that the convex hull of F^S can be obtained via a step-by-step procedure which generates a sequence of "partial" convex hulls.

We first state an auxiliary result which we need in order to prove our next theorem.

Lemma 5.1. *Let $P_1,\ldots,P_r$ be a finite set of polytopes, $P=\bigcup_{h=1}^{r} P_h$, and let P be contained in the closed halfspace $H^+=\{x\in R^n \mid d^ix\leqslant d_{i0}\}$. Then*

$$H\cap \operatorname{conv} P=\operatorname{conv}(H\cap P),$$

where $H=\{x\in R^n \mid d^ix=d_{i0}\}$.

Proof. Let $H\cap \operatorname{conv} P\neq\emptyset$ (otherwise the lemma holds trivially). Clearly, $(H\cap P)\subseteq(H\cap \operatorname{conv} P)$, and therefore

$$\operatorname{conv}(H\cap P)\subseteq \operatorname{conv}(H\cap \operatorname{conv} P)=H\cap \operatorname{conv} P.$$

Next, we prove the inclusion $\supseteq$. Let $u^1,\ldots,u^p$ be the vertices of all the polytopes P_h, $h=1,\ldots,r$. Since r is finite and each P_h has a finite number of vertices, p is finite. Further, $\operatorname{conv} P$ is closed, and $\operatorname{vert}\operatorname{conv} P\subseteq(\bigcup_{h=1}^{r}\operatorname{vert} P_h)$.

Then

$$x\in \operatorname{conv} P\Rightarrow x=\sum_{k=1}^{p}\lambda^k u^k,\quad \sum_{k=1}^{p}\lambda^k=1,\quad \lambda^k\geqslant 0,\ k=1,\ldots,p$$

$$x\in H\Rightarrow d^ix=d_{i0},$$

$$P\subset H^+\Rightarrow d^iu^k\leqslant d_{i0},\quad k=1,\ldots,p.$$

We claim that in the above expression for x,

$$\lambda^k>0\ \Rightarrow\ d^iu^k=d_{i0}.$$

Indeed, if there exists $\lambda^k>0$ such that $d^iu^k<d_{i0}$, then

$$\begin{aligned} d^ix &= d^i\left(\sum_{k=1}^{p}\lambda^k u^k\right)\\ &< d_{i0}\left(\sum_{k=1}^{p}\lambda^k\right)\\ &= d_{i0},\end{aligned}$$

a contradiction. Hence, x is the convex combination of points $u^k\in H\cap P$, or $x\in \operatorname{conv}(H\cap P)$. $\square$

The above lemma is stated in terms of an arbitrary set P which is the union of a finite number of polytopes, and which is contained in a halfspace H^+. In terms of our problem, let F^S be expressed in the conjunctive normal form (2), and for any $T\subseteq S$, let

$$F^T=\left\{x\in F_0\ \middle|\ \bigwedge_{j\in T}\left[\bigvee_{i\in Q_j}(d^ix\geqslant d_{i0})\right]\right\}.$$

F^T is the union of a finite number of polytopes (F_0 is assumed to be bounded). While this is not obvious from the above expression, it becomes obvious when F^T is stated in disjunctive normal form.

Consider now any disjunction $\bigvee_{i \in Q_j}(d^i x \geqslant d_{i0})$ for an index $j \in S - T$, and denote

$$F_i = \{x \in F_0 \mid d^i x \geqslant d_{i0}\}.$$

From the facial property of (DP), we have

$$F_i = \{x \in F_0 \mid d^i x = d_{i0}\},$$

i.e.,

$$\operatorname{aff} F_i = \{x \in R^n \mid d^i x = d_{i0}\}$$

and $d^i x \leqslant d_{i0}$ for all $x \in F_0$, hence for all $x \in F^T$.
Applying the lemma yields

$$\operatorname{aff} F_i \cap \operatorname{conv} F^T = \operatorname{conv}[(\operatorname{aff} F_i) \cap F^T]$$

or, since

$$\operatorname{aff} F_i \cap \operatorname{conv} F^T = F_i \cap \operatorname{conv} F^T$$

and

$$\operatorname{aff} F_i \cap F^T = F_i \cap F^T,$$

we can restate the above result as follows.

Lemma 5.1′. *For any $T \subseteq S$, $i \in Q_j$, $j \in S - T$,*

$$F_i \cap \operatorname{conv} F^T = \operatorname{conv}(F_i \cap F^T).$$

Before we use the above result to prove the main theorem of this section, it will be useful to point out the fact that the statement in the lemma is not true for arbitrary disjunctive programs (i.e., those not having the facial property). Though F_i is always convex, it is *not* true in general that given $S_1 \subseteq R^n$, $S_2 \subseteq R^n$, S_1 and S_2 bounded, S_1 convex, then

$$S_1 \cap \operatorname{conv} S_2 = \operatorname{conv}(S_1 \cap S_2),$$

as illustrated by the following example:

Let

$$F_0 = \{x \in R^2 \mid 0 \leqslant x_1 \leqslant 1,\ 0 \leqslant x_2 \leqslant 1\},$$
$$F^T = \{x \in F_0 \mid (2x_1 + 2x_2 \leqslant 1) \vee (-2x_1 + 2x_2 \geqslant 1)\},$$
$$F_i = \{x \in F_0 \mid 2x_2 \geqslant 1\}.$$

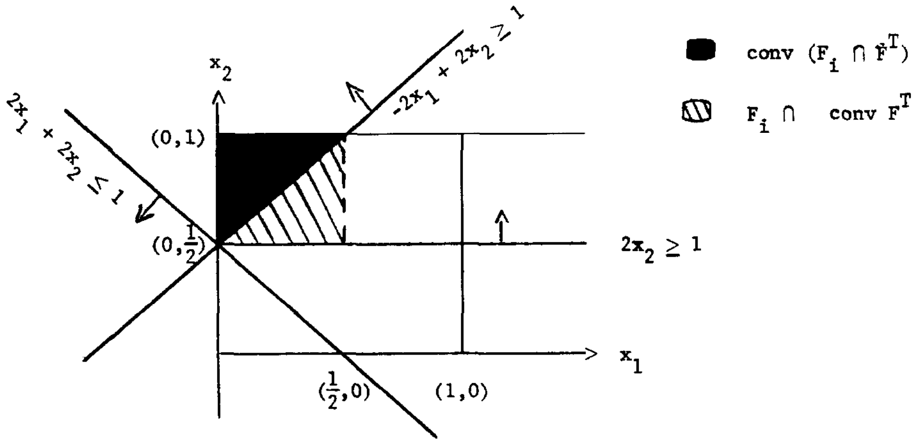

Fig. 1.

Then

$$F_i \cap \ \operatorname{conv} F^T = \{x \in R^2 \mid 0 \leqslant 2x_1 \leqslant 1,\ 1 \leqslant 2x_2 \leqslant 2\},$$

whereas

$$\begin{aligned}\operatorname{conv}(F_i \cap F^T) &= \{x \in R^n \mid -2x_1 + 2x_2 \geqslant 1,\ x_1 \geqslant 0,\ x_2 \leqslant 1\} \\ &\neq F_i \cap \operatorname{conv} F^T,\end{aligned}$$

as illustrated in Fig. 1.

Before stating our next theorem, we note the fact (to be used in the proof) that for arbitrary subsets S_1, S_2 of R^n,

$$\operatorname{conv}(S_1 \cup S_2) = \operatorname{conv}(\operatorname{conv} S_1 \cup \operatorname{conv} S_2).$$

Theorem 5.3. *For any* $T \subset S$,

$$\operatorname{conv}(F^{S-T} \cap \operatorname{conv} F^T) = \operatorname{conv} F^S.$$

Proof. For $R = T$ and $R = S - T$, we express F^R in disjunctive normal form, i.e.,

$$F^R = \left\{ x \in R^n \ \middle| \ \bigvee_{h \in Q_R} \begin{pmatrix} Ax \geqslant a_0 \\ x \geqslant 0 \\ D^h x \geqslant d_0^h \end{pmatrix} \right\}.$$

Denoting

$$F_h^R = \{x \in F_0 \mid D^h x \geqslant d_0^h\},$$

we have $F^R = \bigcup_{h\in Q_R} F_h^R$. Then

$$\begin{aligned} F^{S-T} \cap \operatorname{conv} F^T &= \left(\bigcup_{h\in Q_{S-T}} F_h^{S-T} \right) \cap \operatorname{conv} F^T \\ &= \bigcup_{h\in Q_{S-T}} (F_h^{S-T} \cap \operatorname{conv} F^T). \end{aligned}$$

Since

$$\begin{aligned} F_h^{S-T} &= \{x \in F_0 \mid D^h x \geqslant d_0^h\} \\ &= \bigcap_{i\in M_h} \{x \in F_0 \mid d_i^h x \geqslant d_{i0}^h\}, \end{aligned}$$

where M_h indexes the constraints of $D^h x \geqslant d_0^h$ and $d_i^h x \geqslant d_{i0}^h$ is the ith constraint of this set, by repeated application of Lemma 5.1′ we have

$$F_h^{S-T} \cap \operatorname{conv} F^T = \operatorname{conv}(F_h^{S-T} \cap F^T).$$

Hence,

$$\begin{aligned} \operatorname{conv}(F^{S-T} \cap \operatorname{conv} F^T) &= \operatorname{conv}\left\{ \bigcup_{h\in Q_{S-T}} \operatorname{conv}(F_h^{S-T} \cap F^T) \right\} \quad \text{(from Lemma 5.1′)} \\ &= \operatorname{conv} \bigcup_{h\in Q_{S-T}} (F_h^{S-T} \cap F^T) \\ &= \operatorname{conv}\left\{ F^T \cap \left(\bigcup_{h\in Q_{S-T}} F_h^{S-T} \right) \right\} \\ &= \operatorname{conv}(F^T \cap F^{S-T}) \\ &= \operatorname{conv} F^S. \qquad \square \end{aligned}$$

Next, we define an arbitrary ordering on S, i.e., we let

$$S = (j_1, \ldots, j_s),$$

where $s = |S|$.

Corollary 5.3.1.

$$\operatorname{conv}(F^{\{j_s\}} \cap \operatorname{conv}(\ldots \cap \operatorname{conv}(F^{\{j_2\}} \cap \operatorname{conv} F^{\{j_1\}}))\ldots)) = \operatorname{conv} F^S.$$

Proof. We use induction on s. For $s=1$ the statement is trivially true. Suppose the statement is true for $s=1,\ldots,q$; then for $s=q+1$ we have, denoting $\{j_1,\ldots,j_q\}=T$,

$$\begin{aligned}
&\operatorname{conv}(F^{\{j_{q+1}\}}\cap\operatorname{conv}(\ldots\cap\operatorname{conv}(F^{\{j_2\}}\cap\operatorname{conv}F^{\{j_1\}}))\ldots))\\
&\quad=\operatorname{conv}(F^{\{j_{q+1}\}}\cap\operatorname{conv}F^{T})\\
&\quad=\operatorname{conv}F^{T\cup\{j_{q+1}\}}\quad\text{(from Theorem 5.3.1)}\\
&\quad=\operatorname{conv}F^{S}.\qquad\square
\end{aligned}$$

Theorem 5.3 and Corollary 5.3.1 imply that for a bounded facial disjunctive program (DP), the convex hull of F^S, the feasible set of (DP), can be generated in $|S|$ stages, where $|S|$ is the number of disjunctions in the constraint set of (DP) when stated in the conjunctive normal form (2). At each stage, one has to generate a "partial" convex hull, namely one for a disjunctive program with only one elementary disjunction.

In terms of a 0–1 program, for instance, these results mean that the problem

$$\min\{cx\mid Ax\geqslant b,\ 0\leqslant x\leqslant e,\ x_j=0\text{ or }1,\ j=1,\ldots,n\},$$

where $e=(1,\ldots,1)$, is equivalent to (has the same feasible set as)

$$\min\{cx\mid Ax\geqslant b,\ \alpha^h x\geqslant\alpha_{h0},\ h\in H_1,\ 0\leqslant x\leqslant e,\ x_j=0\text{ or }1,\ j=2,\ldots,n\},$$

where H_1 is the index set for the facets of conv $F^{\{1\}}$, with

$$F^{\{1\}}=\{x\mid Ax\geqslant b,\ 0\leqslant x\leqslant e,\ x_1=0\text{ or }1\}.$$

A 0–1 program in n variables can thus be replaced by one in $n-1$ variables at the cost of introducing new linear inequalities. These inequalities (the facets of $\operatorname{conv}F^{\{1\}}$) are easy to generate, since the set Q in the pair of dual linear programs $P_1(g,\alpha_0)$, $P_2(g,\alpha_0)$ has only two elements.

If the optimal simplex tableau associated with the linear program

$$\text{(LP)}\qquad \min\{cx\mid Ax\geqslant b,\ 0\leqslant x\leqslant e\}$$

is of the form

$$x_i=a_{i0}+\sum_{j\in J}a_{ij}(-x_j),\quad i\in I\cup\{0\},$$

where I and J are the basic and nonbasic index sets, respectively, then the constraints of (LP), expressed in the nonbasic variables x_j, $j\in J$, become

$$\begin{aligned}
&-\sum_{j\in J}a_{ij}x_j\geqslant -a_{i0},\quad i\in I,\\
&\sum_{j\in J}a_{ij}x_j\geqslant a_{i0}-1,\quad i\in I\cap N,\\
&-x_j\geqslant -1,\quad j\in J\cap N,\\
&x_j\geqslant 0,\quad j\in J.
\end{aligned}\tag{6}$$

If the variable to be set to 0 or 1 is chosen to be

$$x_1 = a_{10} + \sum_{j\in J} a_{1j}(-x_j),$$

where $1 \in I \cap N$, then the pair of problems $P_1(g,\alpha_0)$, $P_2(g,\alpha_0)$ becomes

$$P_1(g,\alpha_0)\qquad \begin{array}{ll} \min & \displaystyle\sum_{j\in J} y_j g_j \\ \text{s.t.} & \displaystyle y_j + \sum_{i\in I} u_i^h a_{ij} - \sum_{i\in I\cap N} v_i^h a_{ij} + t_1^h a_{1j} + w_j^h \geqslant 0, \quad j \in J\cap N, \\ & \displaystyle y_j + \sum_{i\in I} u_i^h a_{ij} - \sum_{i\in I\cap N} v_i^h a_{ij} + t_1^h a_{1j} \geqslant 0, \quad j\in J\setminus N, \\ & \displaystyle -\sum_{i\in I} u_i^h a_{i0} + \sum_{i\in I\cap N} v_i^h (a_{i0}-1) + t_1^h(\delta^h - a_{10}) - \sum_{j\in J\cap N} w_j^h \geqslant \alpha_0, \\ & u_i^h \geqslant 0,\ i\in I;\ v_i^h \geqslant 0, i \in I\cap N;\ w_j^h \geqslant 0, j\in J\cap N, \quad h=1,2, \\ & y_j \text{ unconstrained, } j\in J;\ t_1^h \text{ unconstrained, } h=1,2 \end{array}$$

$$P_2(g,\alpha_0)\qquad \begin{array}{ll} \max & \alpha_0\xi_0^1 + \alpha_0\xi_0^2 \\ \text{s.t.} & \displaystyle -a_{i0}\xi_0^h + \sum_{j\in J} a_{ij}\xi_j^h \leqslant 0, \quad i\in I, \\ & \displaystyle (a_{i0}-1)\xi_0^h - \sum_{j\in J} a_{ij}\xi_j^h \leqslant 0, \quad i \in I\cap N, \\ & \displaystyle (\delta^h - a_{10})\xi_0^h + \sum_{j\in J} a_{1j}\xi_j^h = 0, \\ & -\xi_0^h + \xi_j^h \leqslant 0, \quad j\in J\cap N, \\ & \xi_j^1 + \xi_j^2 = g_j, \quad j \in J, \\ & \xi_0^h \geqslant 0, \xi_j^h \geqslant 0, \quad j\in J,\ h=1,2, \end{array}$$

where $\delta^h = 0$ if $h=1$ and $\delta^h = 1$ if $h=2$.

Were one to generate all the facets of $\operatorname{conv} F^{\{1\}}$, the same procedure could then be applied to replace the problem with $n-1$ 0–1 variables by one with $n-2$ 0–1 variables; and in n stages one would obtain the linear program over the convex hull of feasible 0–1 points. Note that at each stage one would have to generate facets of $\operatorname{conv} F^{\{j\}}$, where $F^{\{j\}}$ is the feasible set of a disjunctive program with only one disjunctive constraint $(x_j = 0 \vee x_j = 1)$. The only (but crucial) difficulty in the way of using this approach as a n-stage procedure to solve 0–1 programs in n variables, lies in the fact that the number of facets of each set $\operatorname{conv} F^{\{j\}}$ is very large. Nevertheless, by using some information as to which facets are likely to be binding in the region between the linear programming optimum and the integer optimum, one might be able to make the above approach efficient by generating at each step only a few facets of the current set $F^{\{j\}}$. Next, we outline a procedure based on this idea.

0. Solve (LP). Let I and J be the basic and nonbasic index sets associated with an optimal simplex tableau, and let F_0 be the feasible set of (LP) expressed in the nonbasic variables, i.e.,

$$F_0 = \{x \in R \mid x \text{ satisfies } (6)\},$$

where the components of x are indexed by J. Then, denoting by $\bar{c}$ the vector whose components are $\bar{c}_j = a_{0j}$, $j \in J$, (LP) can be stated as

$$\text{(LP)} \qquad \min\{\bar{c}x \mid x \in F_0\}.$$

For $i \in I \cap N$ and $\delta \in \{0,1\}$, let

$$F_{i,\delta} = \left\{ x \in F_0 \;\middle|\; \sum_{j \in J} a_{ij} x_j = a_{i0} - \delta \right\}$$

and

$$(\mathrm{LP}_{i,\delta}) \qquad \min\{\bar{c}x \mid x \in F_{i,\delta}\}.$$

Let i be the first index of $I \cap N$ (ordered arbitrarily) such that $0 < a_{i0} < 1$, and go to 1.

1. Solve ($\mathrm{LP}_{i,\delta}$) for $\delta = 0$ and 1. If there is no solution for some $\delta \in \{0,1\}$, set $x_i = 1 - \delta$ permanently, and solve the linear program (LP) in the remaining variables; then use the optimal simplex tableau to redefine I, J and F_0, let i be the first index of $I \cap N$ such that $0 < a_{i0} < 1$, and go to 1.

Otherwise, let x^δ be the component indexed by J of the optimal solution obtained for ($\mathrm{LP}_{i,\delta}$). (Given x^δ, the values of the remaining variables are uniquely determined by (6).) Without loss of generality (since the variable x_1 can be complemented if necessary), assume that $\bar{c}x^0 \leqslant \bar{c}x^1$.

If $x_j^0 = 0$ or 1, $\forall j \in J \cap N$, and $a_{i0} - \sum_{j \in J} a_{ij} x_j = 0$ or 1, $\forall i \in I \cap N$, stop: an optimal solution to the 0–1 program has been found. Otherwise go to 2.

2. Set $y = x^0$ and generate those facets of conv $(F_{i,0} \cup F_{i,1})$ containing x^0, by finding the regular optimal solutions of $P_2(x^0, \alpha_0)$ for $\alpha_0 = 1, -1$ and 0.

If $\alpha^k x \geqslant \alpha_0^k$, $k = 1, \ldots, p$ are the facets that were generated, let

$$F_0 \leftarrow \{x \in F_0 \mid \alpha^k x \geqslant \alpha_0^k, \; k = 1, \ldots, p\},$$

let i be the smallest index in $I \cap N$ such that $0 < x_i^0 < 1$, and go to 1.

There is no need to keep forever the facets generated at a given iteration. The easiest rule to follow seems to be the one that is customary in cutting plane methods, namely to keep only those facets whose associated slack variables leave the basis; and drop them as soon as the slack variable in question becomes basic again.

The procedure outlined above is based on the idea that the (translated) polyhedral cone defined by those facets of $\mathrm{conv}(F_{i,0} \cup F_{i,1})$ containing x^0 ($d-1$ of which also contain x^1) is a good approximation of $\mathrm{conv}(F_{i,0} \cup F_{i,1})$ in the vicinity of x^0. The question is, of course, how far does one have to get from x^0 before this approximation ceases to work. To put it in different terms, the facets of $\text{conv}(F_{i,0} \cup F_{i,1})$ generated at a given iteration will keep x_i integer for a number of subsequent iterations; the question is, for how many. Theoretical considerations do not seem to offer an answer to this question, which therefore can only be settled empirically.

Acknowledgements

I wish to thank Bob Jeroslow for his useful comments on an early draft of this paper.

References

[1] E. Balas, Integer programming and convex analysis, Math. Programming 2 (1972) 330–382.
[2] E. Balas, Nonconvex quadratic programming via generalized polars. MSRR No. 278, Carnegie–Mellon University, July 1972; Forthcoming in the SIAM J. Appl. Math.
[3] E. Balas, On the use of intersection cuts in branch and bound, Paper presented at the 8th Internat. Symp. on Mathematical Programming, Stanford, California, 27–31 August 1973.
[4] E. Balas, Intersection cuts from disjunctive constraints, MSRR No. 330, Carnegie-Mellon University. First draft August 1973, revised and expanded February 1974.
[5] E. Balas, Cutting planes from logical conditions, Paper presented at the SIGMAP – University of Wisconsin Nonlinear Programming Symp., Madison, Wisconsin, 15–17 April 1974.
[6] D. Gale, Theory of Linear Economic Models, McGraw-Hill, New York, 1960.
[7] F. Glover, Polyhedral annexation in mixed integer programming, MSRR No. 73-9, University of Colorado, August 1973.
[8] F. Glover, D. Klingman, The generalized lattice point problem, Oper. Res. 21 (1973) 141–156.
[9] F. Glover, D. Klingman, Improved Convexity Cuts for Lattice Point Problems, Research Report CS 133, University of Texas, April 1973.
[10] B. Grünbaum, Convex Polytopes, Interscience, New York, 1967.
[11] R.G. Jeroslow, Personal communication, March 1974.
[12] R.G. Jeroslow, Principles of Cutting Plane Theory: Part I, Carnegie-Mellon University, February 1974.
[13] G. Owen, Cutting planes for programs with disjunctive constraints, J. Optim. Theory Appl. 11 (1973) 49–55.
[14] R.T. Rockafellar, Convex Analysis, Princeton University Press, Princeton, 1970.
[15] J. Stoer, C. Witzgall, Convexity and Optimization in Finite Dimensions. I, Springer, Berlin, 1970.
[16] P.B. Zwart, Intersection cuts for separable programming, School of Engineering and Applied Science, Washington University, St. Louis, Missouri, January 1972.