CENTER FOR CYBERNETICS STUDIES

The University of Texas
Austin, Texas 78712

37

Graduate School of Business

Research Report
CS #37

THE GENERALIZED LATTICE POINT PROBLEM

by

Fred Glover*
Darwin Klingman

August 1970

* King Resources Professor at the University of Colorado

This research was partly supported by ONR Contract NR-047-021, Task N00014-67-A-0126-0009 with the Center for Cybernetic Studies, The University of Texas. Reproduction in whole or in part is permitted for any purpose of the United States Government.

CENTER FOR CYBERNETIC STUDIES

A. Charnes, Director

Abstract

The generalized lattice point problem, posed by Charnes and studied by M. J. L. Kirby, H. Love and others, is a linear program whose solutions are constrained to be extreme points of a specified polytope. We show how to exploit this and more general problems by convexity (or intersection) cut strategies without resorting to standard problem augmenting techniques such as introducing 0-1 variables. In addition, we show how to circumvent "degeneracy" difficulties inherent in this problem without relying on perturbation (which provides uselessly shallow cuts) by identifying nondegenerate subregions relative to which cuts may be effectively defined. Finally, we give results that make it possible to obtain strengthened cuts for problems with special structures.

1. Introduction

The generalized lattice point (GLP) problem (first posed by A. Charnes [6])[1] subsumes a wide variety of combinatorial programming problems, including the mixed 0-1 problem of integer linear programming.

The GLP problem, while remarkably easy to state, is not amenable to solution by standard optimization methods because of its generality and the non-convexity of its feasible region. For instance, solution approaches based on certain "purification" and "decomposition" principles have been studied by several researchers,[2] but may fail because it is not possible to generate the appropriate tableau information or because insufficient criteria are manufactured to distinguish local from global optima.

The solution method presented in this paper for the GLP problem is based on the "convexity cut" ideas developed by Young [16] and Balas [1] (in the context of mixed integer programming) and by Tui [15] (in the context of concave programming),[3] utilizing the extended conceptual framework due to Glover [8].

The key obstacle to applying convexity cuts to the GLP problem is the occurance of certain "degeneracy" difficulties which are not present in the integer programming applications and which are not resolvable by perturbation. In fact, we show that perturbation is useless in the creation of convexity cuts, which demonstrates that reliance on perturbation to overcome degeneracy in the use of these cuts, as proposed in the concave programming context, is unfortunately unworkable.

The work of this paper enlarges the applicability of convexity cuts in two ways. First, we show how to resolve the fundamental degeneracy difficulties associated with the GLP problem without relying on perturbation. Second, we give results for strengthening these cuts for GLP problems with special structures.

2. Problem Statement

We define the GLP problem to be that of finding a vector $x \in R^n$ to

$$\begin{aligned} &\text{Minimize } cx \\ &\text{Subject to: } Ax \le b \\ &\qquad\qquad\quad\ x \ge 0 \\ &\text{and} \quad x \text{ lies on an } n-q \text{ dimensional face of the polytope } Q = \{x\colon Dx \le d\},^4 \end{aligned}$$

where $n \ge q \ge 0$ and all matrices and vectors are, of course, assumed to be dimensioned conformably. By the statement that x lies on an $n - q$ dimensional face of Q we mean we mean $x \in Q$ and $D^i x - d_i = 0$ for at least q linearly independent rows D^i of D (where $q \le \text{rank}(D)$).

For simplicity in the following we shall assume that the inequality $x \ge 0$ is absorbed into the matrix inequality $Ax \le b$ (except for any components of this inequality that are already contained in $Dx \le d$) and that none of the row inequalities of $Ax \le b$ duplicate those of $Dx \le d$. (Duplicate inequalities may, of course, be deleted from $Ax \le b$ without altering the problem.)

Introducing slack variables $u = d - Dx$ and $v = b - Ax$, we define the polytopes

$$P = \{x\colon v \ge 0\}$$

$$Q = \{x\colon u \ge 0\}$$

Unclassified
Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Center for Cybernetic Studies The University of Texas at Austin	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP

3. REPORT TITLE

The Generalized Lattice Point Problem

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)
Center for Cybernetic Studies Report - August 1970

5. AUTHOR(S) (First name, middle initial, last name)

Fred Glover
Darwin Klingman

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
August 1970	34	17

8a. CONTRACT OR GRANT NO. NR-047-021 b. PROJECT NO. N00014-67-A-0126-0009 c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) Center for Cybernetic Studies Report #37 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

10. DISTRIBUTION STATEMENT

This document has been approved for public release and sale; its distribution is unlimited.

11. SUPPLEMENTARY NOTES Not applicable.	12. SPONSORING MILITARY ACTIVITY Code 434 Office of Naval Research Washington, D.C. 20360

13. ABSTRACT

The generalized lattice point problem, posed by Charnes and studied by M.J.L. Kirby, H. Love and others, is a linear program whose solutions are constrained to be extreme points of a specified polytope. We show how to exploit this and more general problems by convexity (or intersection) cut strategies without resorting to standard problem augmenting techniques such as introducing 0 - 1 variables. In addition, we show how to circumvent "degeneracy" difficulties inherent in this problem without relying on perturbation (which provides uselessly shallow cuts) by identifying nondegenerate subregions relative to which cuts may be effectively defined. Finally, we give results that make it possible to obtain strengthened cuts for problems with special structures.

DD FORM 1473 1 NOV 65 (PAGE 1)
S/N 0101-807-6801

Unclassified
Security Classification

Unclassified
Security Classification

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Mathematical programming						
Linear programming						
Integer programming						
Non-convex programming						
0 - 1 integer programming						

DD FORM 1 NOV 65 1473 (BACK)
(PAGE 2)

Unclassified
Security Classification

and also define

$$Q^* = \{x : x \text{ lies on an n-q dimensional face of } Q\} .$$

The GLP problem may then be written

$$\underset{x \in P \cap Q^*}{\text{Minimize}}\ cx .$$

Corresponding to the GLP problem, we shall define a relaxed problem which arises by replacing Q^* with Q to yield the ordinary linear program

$$\underset{x \in P \cap Q}{\text{Minimize}}\ cx .$$

Clearly, if an optimal solution to the relaxed problem is feasible for the GLP problem, then it must be optimal for the GLP problem.

3. Tableau Representation of the GLP Problem

We assume (without loss of generality) that the problem constraints and variables are indexed so that the equations $u = d - Dx$ and $v = b - Ax$ have the form

$$\begin{aligned} u_I &= d_I - D_I x \\ x_I &= 0 + x_I \end{aligned}$$

and

$$\begin{aligned} x_{II} &= 0 + x_{II} \\ v_I &= b_I - A_I x \end{aligned}$$

where $u = \binom{u_I}{x_I}$, $v = \binom{x_{II}}{v_I}$ and $x = \binom{x_I}{x_{II}}$. Thus, defining $y = \binom{u}{v}$, the conditions $u \geq 0$ and $v \geq 0$ may alternately be written

$$y = \begin{pmatrix} u_I \\ x \\ v_I \end{pmatrix} \geq 0 .$$

These notational assumptions insure that the linear programming "column tableau" representation for the relaxed problem has a convenient structure that simplifies the identification of the "current x vector." Specifically, the initial column tableau representation is given by

$$\begin{aligned} \text{maximize } x_o &= 0 + c(-x) \\ \text{subject to } \quad u &= d + D(-x) \\ v &= b + A(-x) \\ u \geq 0&, \; v \geq 0 \end{aligned}$$

or, in more compressed form, by

$$\begin{aligned} \text{Maximize } x_o &= 0 + c(-x) \\ \text{subject to } \quad y &= B_o + B(-x) \\ y &\geq 0 \end{aligned}$$

where $B_o = \binom{d}{b}$ and $B = \binom{D}{A}$. (Maximizing $x_o = -cx$, of course, corresponds to minimizing cx.) It may be noted that the "middle equations" of this tableau representation are simply the identity equations

$$x_j = -(-x_j), \quad j = 1, \ldots, n .$$

The components of x are the nonbasic variables for this initial tableau, and setting $x = 0$ uniquely determines the linear programming extreme point (or "basic") solution: $x_o = 0$, $y = B_o$. This solution is defined to be primal feasible if $B_o \geq 0$, in which case it follows that $x \in P \cap Q$. The solution is defined to be dual feasible if $c \geq 0$. By standard linear programming theory, a basic solution that is both primal and dual feasible is optimal for the relaxed problem.

We write the more general current tableau representation of the problem in the analogous form

$$\begin{aligned} \text{Maximize} \quad & x_0 = c_0 + c(-t) \\ \text{subject to} \quad & y = B_0 + B(-t) \\ & y \geq 0 \end{aligned}$$

where the y vector is the same as before, but where t denotes any set of "current nonbasic variables" (hence its components constitute a subset of the components of y). The scalar c_0, the vectors c and B_0, and the matrix B will depend upon the composition of t. This composition may be identified by n of the equations of $y = B_0 - Bt$, which have the form

$$y_i = -(-t_j)$$

where of course j depends on i. These again are identity equations, corresponding to the previously indicated identity equations of the initial tableau.

Example of the GLP Problem in Tableau Form

To illustrate the foregoing tableau representation of the GLP problem consider the problem

$$\begin{aligned} \text{Minimize} \quad & -3x_1 + 4x_2 - 6x_3 \\ \text{subject to} \quad & 2x_1 - 1x_2 + 2x_3 \leq 7 \\ & x_1, x_2, x_3 \geq 0 \end{aligned}$$

where x is required to lie on a one dimensional face of the polytope defined by

$$\begin{aligned} x_1 + x_2 + x_3 &\le 4 \\ -x_1 - x_2 - x_3 &\le -3 \\ x_1 \qquad\qquad &\ge 0 . \end{aligned}$$

Then (eliminating $x_1 \ge 0$ since it is accommodated in Q to follow),

$$P = \{x: 2x_1 - x_2 + 2x_3 \le 7,\ x_2 \ge 0,\ x_3 \ge 0\}$$

$$Q = \{x: x_1 + x_2 + x_3 \le 4,\ -x_1 - x_2 - x_3 \le -3,\ x_1 \ge 0\}$$

and the tableau representation is:

Maximize $x_o = 0 - 3(-x_1) + 4(-x_2) - 6(-x_3)$

subject to

$$\begin{aligned}
y_1 &= 4 + 1(-x_1) + 1(-x_2) + 1(-x_3) \\
y_2 &= -3 - 1(-x_1) - 1(-x_2) - 1(-x_3) \\
y_3 &= 0 - 1(-x_1) \\
y_4 &= 0 \qquad\qquad - 1(-x_2) \\
y_5 &= 0 \qquad\qquad\qquad\qquad - 1(-x_3) \\
y_6 &= 7 + 2(-x_1) - 1(-x_2) + 2(-x_3) \\
y &\ge 0
\end{aligned}$$

where $y_1 = u_1,\ y_2 = u_2,\ y_3 = u_3 = x_1,\ y_4 = v_1 = x_2,\ y_5 = v_2 = x_3,\ y_6 = v_3$.

In the detached coefficient form

$$\begin{array}{ccc} & & -t \\ x_o = & \boxed{c_o} & \boxed{c} \\ y = & \boxed{B_o} & \boxed{B} \end{array}$$

the foregoing becomes

				$-x_1$	$-x_2$	$-x_3$
		x_o =	0	-3	4	-6
y =	u =	u_1 =	4	1	1	1
		x_2 =	-3	-1	-1	-1
		x_1 =	0	-1	0	0
	v =	x_2 =	0	0	-1	0
		x_3 =	0	0	0	-1
		v_3 =	7	2	-1	2*

where the double line separates the "u" and "v" sections of the tableau.

<u>Optimality Criteria for the GLP Problem in the Tableau Representation</u>

To illustrate how to identify the optimality criteria for the GLP problem in terms of the tableau format, we first pivot on the coefficient 2 indicated by the asterisk in the preceding tableau to obtain

		$-x_1$	$-x_2$	$-v_3$
x_o =	21	3	1	3
u_1 =	1/2	0	3/2	-1/2
u_2 =	1/2	0	-3/2	1/2
x_1 =	0	-1	0	0
x_2 =	0	0	-1	0
x_3 =	7/2	1	-1/2	1/2
v_3 =	0	0	0	-1

This tableau gives an optimal linear programming solution to the relaxed problem, since it is both primal and dual feasible ($B_0 \geq 0$ and $c \geq 0$, respectively).

To determine whether this solution to the relaxed problem also solves the GLP problem we must translate the feasibility requirement of the GLP problem into an equivalent requirement applicable to the linear programming tableau. The statement that a vector $x = x'$ lies on an $n - q$ dimensional face of Q (i.e., $x' \in Q^*$) may be expressed in the terminology of the column tableau for which (i) the basic solution ($y = B_0$) yields $x = x'$, and (ii) at least q of the current non-basic variables (components of t) are also components of u.

In the foregoing example $n - q = 1$ and so $q = 2$. However, only 1 of the components of u (i.e., x_1) is among the current nonbasic variables in the final tableau. Moreover, there is no other tableau yielding the same solution in which 2 components of u are nonbasic, due to the absence of primal degeneracy. (We show how to take care of degeneracy in later sections.) The existence of such a tableau is both necessary and sufficient for the current basic solution to lie on a one-dimensional face of Q, and hence for the solution to satisfy the desired optimality criteria. Thus, the basic solution of the current tableau does not solve the GLP problem.

The main results of this paper provide a method for "cutting away" such a basic solution to the relaxed problem when it is thus determined to be infeasible for the GLP problem. The background for these results is given in the next section.

4. Convexity Cuts: Preliminaries

The convexity cuts, in general, may be applied to a problem in which it is desired to find a vector y contained in the intersection of the cone

$$C = \{y: y = B_0 - Bt, \quad t \geq 0\}$$

and some arbitrary set S. The basic characterization of the convexity cuts is as follows:

Convexity Cut Lemma ([8])[6] Assume there is a convex set R whose interior contains the vertex B_0 of C, but does not contain any points of S. Then, given numbers $t_j^* > 0$ such that

$$B_0 - B_j t_j^* \in R \text{ for all } j = 1, \ldots, n$$

the cut

$$\sum_j \left(\frac{1}{t_j^*}\right) t_j \geq 1 \qquad (1)$$

is satisfied by all $y = B_0 - Bt$ such that $y \in C \cap S$.

One way to visualize the assertion of the lemma is to think of extending the j^{th} edge (half line), $y = B_0 - B_j t_j$, $t_j \geq 0$, of the cone C a "distance" $t_j^* > 0$ such that the endpoint of the extended edge remains in the convex set R. If this is done for each edge, and a hyperplane is passed through the endpoints of these edges (disregarding B_0), then the convexity cut lemma says that all points y of $C \cap S$ must lie in the associated half space that does not contain B_0. Whenever B has rank n, as it does for the linear programming tableau, the solution $y = B_0$ is uniquely determined for $t = 0$, and cannot be obtained for any $y \in C$ satisfying (1), which implies $t \geq 0$ and $t \neq 0$.

The basic solution strategy implied by the convexity cut lemma is to obtain a "cone representation" of the feasible region, such as that automatically given by the linear programming tableau, and to determine a convex set R and numbers $t_j^* > 0$ as indicated in the lemma. The problem constraints are then augmented by requiring (1) to hold, i.e., by adjoining the cut (1) to the linear programming tableau and pivoting to a new basic solution (thus obtaining a new cone representation). The process repeats until $B_o \in S$, whereupon $y = B_o$ satisfies $y \in C \cap S$.

<u>Convexity Cut Strategy Exemplified for the GLP Problem</u>

We illustrate the foregoing (without yet specifying the convex region R on which the numbers t_j^* are based) for the example GLP problem of the preceding section. The final linear programming tableau previously obtained, which gives the desired representation, is

		$-x_1$	$-x_2$	$-V_3$
$x_o =$	21	3	1	3
$u_1 =$	1/2	0	3/2	-1/2
$u_2 =$	1/2	0	-3/2	1/2
$u_3 = x_1 =$	0	-1	0	0
$v_1 = x_2 =$	0	0	-1	0
$v_2 = x_3 =$	7/2	1	-1/2	1/2
$v_3 =$	0	0	0	-1
$v_4 =$	-1	0	-3*	-1

An appropriate convexity cut for this problem has been adjoined as a new constraint row which appears at the bottom of the tableau. The cut is derived (using procedures to be specified later) from the values $t_1^* = \infty$, $t_2^* = 1/3$, and $t_3^* = 1$ (where $t_1 = x_1$, $t_2 = x_2$, $t_3 = v_3$). We have denoted the slack variable that transforms the convexity cut inequality into an equality by v_4, thereby indicating that the cut is being used to augment the set of constraints defining the polytope P.

Applying the dual simplex method to this tableau yields the pivot element indicated by the asterisk, and thus gives rise to the new tableau

		$-x_1$	$-v_4$	$-v_3$
$x_o =$	20 2/3	3	1/3	8/3
$u_1 =$	0	0	1/2	-1
$u_2 =$	1	0	-1/2	1
$u_3 = x_1 =$	0	-1	0	0
$v_1 = x_2 =$	1/3	0	-1/3	1/3
$v_2 = x_3 =$	22/6	1	-1/6	4/6
$v_3 =$	0	0	0	-1
$v_4 =$	0	0	-1	0

At this point the updated cut row (i.e., the v_4 row) can optionally be dropped. The current tableau is primal and dual feasible and hence is once again "locally optimal". Using the ideas presented in Section 3, we investigate whether the current basic solution solves the GLP problem. Only one of the "u variables" is nonbasic (i.e., x_1), which is one less than the number required. However, the current tableau is degenerate in Q (i.e., in the u variables) since the basic variable u_1 receives a value of 0. Moreover, both v_4 and v_3 have nonzero coefficients in the u_1 row, and hence it is possible to pivot on either of these coefficients to obtain a new tableau in which both u_1 and x_1 are nonbasic, leaving the solution $y = B_0$ unchanged. Thus the current tableau in fact provides an optimal solution to the example GLP problem.

We now develop the fundamental notions and procedure for implementing the convexity cuts and for accommodating the considerations illustrated in this example.

5. Determining a Convexity Cut for the GLP Problem

Corresponding to the sets P, Q, and Q*, we introduce counterpart sets $\bar{P}$, $\bar{Q}$, and $\bar{Q}^*$, defined by:

$$\bar{P} = \{y:\ x \in P\}$$

$$\bar{Q} = \{y:\ x \in Q\}$$

$$\bar{Q}^* = \{y:\ x \in Q^*\}$$

where

$$y = \begin{pmatrix} u_I \\ x \\ v_I \end{pmatrix}$$

We note that $\bar{P}$ and $\bar{Q}$ are also given by $\bar{P} = \{y: v \geq 0\}$ and $\bar{Q} = \{y: u \geq 0\}$ (where $y = \binom{u}{v}$). Thus, for example, $y \in \bar{P} \cap \bar{Q}$ may equivalently be expressed as $y \geq 0$.

Our first goal is to specify the set S of the convexity cut lemma in a form relevant to the GLP problem. This may be done rather easily, since it is sufficient to define S so that the statement $y \in C \cap S$ is equivalent to the statement $x \in P \cap Q^*$. Thus, we let $S = \{y: y \geq 0 \text{ and } y \in \bar{Q}^*\}$, which implies $S = C \cap S$, since $y = B_0 - Bt$. (noting that $y \geq 0$ implies $t \geq 0$ in the linear programming tableau).

Our second goal is to specify the convex set R relative to which the convexity cut is determined. It is instructive in pursuing this goal to identify the "natural" alternative for the set R. By the convexity cut lemma, the interior of R must contain B_0 but no points of S. Thus, a plausible first candidate for R is $\bar{Q}$. This set is clearly convex and, moreover, every y that satisfies $y \in \bar{Q}^*$ (and hence $y \in S$) lies on the "boundary" of $\bar{Q}$, but not in its interior. However, there is no assurance that $\bar{Q}$ will contain B_0 in its interior. In fact, this will almost never occur, for $y = B_0$ can be in the interior of $\bar{Q}$ $(=\{y: u \geq 0\})$ only if $u > 0$. This latter implies not only the absence of "degeneracy in Q" (as discussed in the preceding section), but also that all of the components of u are basic. This condition can never be satisfied in tableaus that are feasible (or even "partially feasible") for the GLP problem, since one condition of feasibility is that $u_i = 0$ for at least q of the components of u.

The set $\bar{P}$ can be quickly dismissed as a viable alternative for R, since, althoug... , it may quite possibly contain points of S in its interior. The set $\bar{P} \cap \bar{Q}$ overcomes this latter limitation, but also fails, since B_o is always on its boundary (and in any case this choice would not be as good as $\bar{Q}$, if $\bar{Q}$ were acceptable, since $\bar{P} \cap \bar{Q} \subset \bar{Q}$).

A more plausible approach is that of "perturbing" B_o (when necessary) from the boundary to the interior of $\bar{Q}$, whereupon $\bar{Q}$ becomes an acceptable choice for R. However, such perturbation, which very satisfactorily combats degeneracy in ordinary linear programming, is totally useless in the convexity cut context. To see this, note that perturbation gives rise to one of the following two possibilities. First, there may be an edge from the perturbed B_o that stays inside R for no more than an "ϵ distance." Consequently, the value of t^*_j for this edge is minutely small, and the convexity cut eliminates a correspondingly small "ϵ neighborhood" of B_o. Disregarding the numerical difficulties of dealing with an enormously large cut coefficient $1/t^*_j$ (which might be accommodated, for example, by using a non-Archimedean ordering), the new B_o obtained after re-optimizing with the dual simplex method must be essentially the same as the perturbed B_o, and the cut method gets nowhere.

The second possible consequence of perturbation is the opposite one, that an edge from the perturbed B_o now stays for some distance in the interior of R when it would have continued along a boundary of R in the absence of perturbation. (The case in which all edges lead from a boundary into the interior of R is automatically legitimate by footnote 6 and does not require perturbation; hence we exclude it.)[7] But then points of S on the boundary of R may be by-passed by the cut. Thus, once again, perturbation is of no use.

Our fundamental result, which specifies a superset of $\bar{Q}$ that is acceptable in the role of R, and which does not rely on perturbation, is given to follow.

Theorem:

Let $y' = \binom{u'}{v'} = B_o \geq 0$, and define the index set $M = \{i: u_i' \neq 0\}$. Then, if y' is not feasible for the GLP problem ($B_o \notin S$), the set $R = \{y: u_i \geq 0, i \in M\}$ satisfies the assumptions of the convexity cut lemma, i.e., R contains B_o but no points of S in its interior.

The set R specified in the foregoing theorem is the superset of $\bar{Q} = \{y: u \geq 0\}$ that results by disregarding all of the inequalities $u_i = d_i - D^i x \geq 0$ defining $\bar{Q}$ except those that are inactive (i.e., for which $D^i x < d_i$) at the solution $y' = B_o$. Thus, the theorem rather surprisingly (and pleasantly) states that one can ignore degeneracy (and indeed all active constraints) in $\bar{Q}$ when determining R.

The key observation we shall use to prove this theorem is the following:

Lemma:

Let y^* be a feasible solution for the GLP problem (i.e., $y^* \in S$). Then if $y' = B_o \geq 0$ and $B_o \notin S$, there is an index k such that $u^*_k = 0$ and $u'_k \neq 0$.

Proof:

We shall say that a tableau verifies $y^* \in S$ (or $y' \in S$) if y^* (or y') is obtained as the current basic solution and at least q of the components of

u are nonbasic in this tableau. Relative to a given tableau that verifies $y^* \in S$, define $M^* = \{i: u_i \text{ is nonbasic}\}$. Then, for the sake of contradiction, suppose the lemma is false, whereupon $u^*_i = 0$ implies $u'_i = 0$ for all $i \in M^*$. By the basis exchange theorem[8] there must exist a sequence of pivots from the current tableau (in which $y' = B_0$) to a new tableau in which u_i is nonbasic for all $i \in M^*$. Moreover, each pivot is executed by selecting an $i \in M^*$ such that u_i is basic, and then performing a basis exchange that leaves u_i nonbasic (replacing u_i in the basis by some variable that is basic in the tableau that verifies $y^* \in S$, but which is nonbasic in the current tableau before the pivot). Each of these successive pivots leaves B_0 unchanged, and hence the end result is a tableau that verifies $y' \in S$. But this violates the assumption that y' is not feasible for the GLP problem, thereby completing the proof.

It is interesting to consider a "separating hyperplane" interpretation of the lemma by visualizing its implications geometrically in the space of the y vectors. In this context the lemma can be restated in the form: if y^* is feasible for the GLP problem but y' is not (where $y' = B_0 \geq 0$), then there is at least one constraint hyperplane defining $\bar{Q}$ (i.e., a hyperplane of the form $\{y: u_k = 0 \text{ for some } k\}$) that passes through y^* but that does not pass through y'. This interpretation prompts the following speculation. The half space $\{y: u_k \geq 0\}$ corresponding to a constraint hyperplane that passes through y^* but not $y' = B_0$ must contain B_0 but not y^* in its interior. Moreover, if y^* is optimal for the GLP problem, then all other feasible solutions can be thrown away and S redefined to consist of the single point y^*. It immediately follows that the half space $\{y: u_k \geq 0\}$ provides an acceptable R for determining a convexity cut.

Of course, such a half space would undoubtedly be extremely difficult to identify. In fact, its identification would render a cutting approach superfluous, since after q steps of specifying an appropriate index k for which one could require u_k to be nonbasic, the GLP problem would be solved. Nevertheless, the fact that these half spaces exist makes it possible to specify an R that can be readily identified, as we now demonstrate in the following proof.

Proof of the theorem:

First, it is immediate that R as specified in the theorem contains B_o in its interior. Thus, we must show that this interior contains none of the points of S. Let $M' = \{k: u'_k \neq 0$ and $u^*_k = 0$ for some $y^* \in S\}$, and let R^* be the intersection of the half spaces $\{y: u_k \geq 0\}$, $k \in M'$. By the lemma, every $y^* \in S$ must lie on a hyperplane $\{y: u_k = 0\}$ for some $k \in M'$. Thus $y^* \in S$ implies y^* cannot be contained in the interior of R^*. But $M' \subset M$, and hence the set R of the theorem (which is the intersection of the half spaces $\{y: u_k \geq 0\}$, $k \in M$) must be contained in R^*. This completes the proof.

6. Implementation of the Convexity Cuts for the GLP Problem

Having specified an appropriate R , two tasks remain in order to implement the convexity cuts. The tasks are:

1) To provide explicit criteria for determining when $B_o \in S$.

2) To specify the calculation of the numbers t_j^*.

We consider these tasks in sequence.

Determining whether $B_o \in S$

The first task, applied to a primal feasible tableau, involves determining whether there exists a series of degenerate pivots that will yield at least q components of u nonbasic (if this is not already the case). Sometimes simple inspection of the tableau will provide the answer, and sometimes it may actually be necessary to carry out several degenerate pivots. The key is this: as long as there is a degenerate basic u_i whose tableau row contains a nonzero coefficient in the column of a nonbasic v_i, then a pivot may be made on this nonzero coefficient yielding a new tableau in which one more component of u is nonbasic (and B_o remains unchanged). If there are no such u_i, then the current tableau contains the maximum number of nonbasic u_i for any tableau with the same B_o.

In particular, if the current tableau has r of the u_i nonbasic, and if fewer than q - r of the nonbasic v_i have a nonzero coefficient in at least one of the degenerate u_i rows, then it is clear that $B_o \in S$ (without having to pivot). On the other hand, if it is known from the structure of the constraint inequality $Dx \leq d$ that all 0-valued u_i can be made nonbasic, then checking for $B_o \in S$ becomes merely a matter of counting (see Section 7). Other structures can, of course, provide other shortcuts for checking whether $B_o \in S_o$. However, it seems intuitively reasonable that the convexity cut will tend to be stronger if a tableau is used that has a maximum number of u_i nonbasic.

Calculating t^*_j

The second task, calculating the t^*_j values, corresponds precisely to the task of determining the amount a nonbasic variable can be increased without violating primal feasibility in the application of the primal simplex method. However, primal feasibility in the present context is restricted to the tableau rows corresponding to "basic non-degenerate u_i" (i.e., $i \in M$). Thus, specifically, it follows that

$$t^*_j = \underset{\substack{i \in M \\ b_{ij} > 0}}{\text{Min.}} \{b_{io}/b_{ij}\}$$

where b_{io} is the i^{th} component of B_o and b_{ij} is the ij^{th} component of B. (By convention $t^*_j = \infty$ if $b_{ij} \leq 0$ for all $i \in M$.) This value is, of course, always positive since $b_{io} > 0$ for all $i \in M$.

7. Special Structures and Strengthened Cuts

The issue of special structures was briefly touched on in the preceding section, where we remarked that for some structures the problem of checking whether $y \in Q^*$ reduces to checking the number of 0-valued u_i. Such structures are of interest for more than their ability to facilitate checking $y \in \bar{Q}^*$, however, as we now indicate.

We shall way that a GLP problem has the counting property provided that the condition "$D^i x = d_i$ (hence $u_i = 0$) for q or more of the rows D^i of D" implies that at least q of these rows D^i are linearly independent. By definition, the counting property implies $y \in \bar{Q}^*$ if and only if $y \in \bar{Q}$ and $u_i = 0$ for at least q components of u. The 0 - 1 mixed integer programming problem obviously has the counting property since in this case $Dx \leq d$ summarizes the constraints $x_j \leq 1$ and $-x_j \leq 0$, $j = 1, \ldots, q$. Similarly, it is clear that any problem has the counting property for which $Dx \leq d$ summarizes sets of constraint pairs $H^i x \leq h_i$ and $-H^i x \leq k_i$, where all of the H^i are linearly independent and $k_i \neq -h_i$. Rather than attempt to characterize more general structures that give rise to the counting property, however, we note that these problems are subsumed in a larger class of "quasi-GLP" problems in which $\bar{Q}^*$ is redefined to be $\{y: y \in \bar{Q}$ and $u_i = 0$ for at least q components of $u\}$, ignoring whether the associated rows D^i of D are linearly independent. The fundamental results of Section 5 apply as readily to the quasi-GLP problem as to the GLP problem itself, and hence the previously indicated convexity cuts can be the same for both problems. (The altered definition of $\bar{Q}^*$, of course, alters the definition of S correspondingly.)

A more important observation concerns the ability to obtain stronger cuts for the GLP problem. Suppose that the matrix D (upon re-indexing) is partitioned into two submatrices

$$D = \begin{pmatrix} D' \\ D'' \end{pmatrix}$$

where the rank of D' is $q' < q$. Then, if x lies on an $n - q$ dimensional face of D, it must also lie on an $n - p$ dimensional face of D", where $p = q - q'$. Thus, whenever fewer than p of the components of u associated with D" can be made nonbasic in the current basic solution, then a convexity cut can be determined relative to $D''x \le d''$ instead of $Dx \le d$. This must yield a cut at least as strong as the one for $Dx \le d$, since the half spaces whose intersection determines R are reduced in number whenever $D'x \ne d'$, thereby enlarging R.

The special structures previously indicated very conveniently submit to such partitioning, since it suffices to let $D'x \le d'$ summarize any $q - 1$ pairs of the constraints $H^ix \le h_i$, $-H^ix \le k_i$; and, in fact, any choice of these pairs will yield a convexity cut (when $y \notin S$) provided $D''x \ne d''$. The best choices are, of course, those that assign as many of the unsatisfied constraints of $Dx \le d$ to $D'x \le d'$ as possible (provided at least one is assigned to $D''x \le d''$). Such choices are trivial for the 0 - 1 mixed integer problem for then $D''x \le d''$ can be selected to be any single pair of constraints $x_j \le 1$ and $-x_j \le 0$ such that $x_j \ne 0$ or 1 in the current basic solution. Convexity cuts determined in this fashion for the 0 - 1 problem correspond precisely to the mixed integer cuts proposed by Gomory in [12].[9]

By introducing a number of additional variables and constraints, a variety of GLP problems (and all quasi-GLP problems) can be given a 0 - 1 mixed integer formulation. Interestingly enough, however, it can be shown that the mixed integer cuts of [12] applied to such a formulation will be generally weaker than the convexity cuts indicated here.[10] A similar statement applies to certain "conjunctive-disjunctive" and "disjunctive-conjunctive" generalizations of the GLP problem, which subsume the general mixed integer problem in the same way that the GLP problem subsumes the 0 - 1 mixed integer problem. (See [11].)

8. Examples

To illustrate the ideas of the foregoing sections, we provide the following geometric and numerical examples:

Consider the GLP problem

$$\text{Minimize } 1x_1 + 1x_2$$

subject to

$$\begin{aligned} -1/2\,x_1 - 1x_2 &\le -1 \\ 5/2\,x_1 + 3x_2 &\le 23 \\ -5/2\,x_1 + 1x_2 &\le 1 \\ x_1, x_2 &\ge 0 \end{aligned}$$

and (x_1, x_2) is an extreme point (i.e., lies on a 0 dimensional face) of the polytope defined by

$$\begin{aligned} 1/2\,x_1 - 1x_2 &\le 1 \\ -2x_1 - 1x_2 &\le -2 \\ 5/2\,x_1 - 1x_2 &\le 9 \end{aligned} .$$

Then

$$P = \{(x_1, x_2): -1/2x_1 - 1x_2 \le -1,\ 5/2x_1 + 3x_2 \le 23,\ -5/2x_1 + 1x_2 \le 1,$$
$$x_1 \ge 0,\ x_2 \ge 0\}$$

and

$$Q = \{(x_1, x_2): 1/2x_1 - 1x_2 \le 1,\ -2x_1 - 1x_2 \le -2,\ 5/2\,x_1 - 1x_2 \le 9\},$$

and the tableau representation is:

Maximize $x_o = 0 + 1(-x_1) + 1(-x_2)$

subject to

$$u_1 = 1 + 1/2(-x_1) - 1(-x_2)$$
$$u_2 = -2 - 2(-x_1) - 1(-x_2)$$
$$u_3 = 9 + 5/2(-x_1) - 1(-x_2)$$
$$v_1 = 0 - 1(-x_1) + 0(-x_2)$$
$$v_2 = 0 + 0(-x_1) - 1(-x_2)$$
$$v_3 = -1 - 1/2(-x_1) - 1(-x_2)$$
$$v_4 = 23 + 5/2(-x_1) + 3(-x_2)$$
$$v_5 = 1 - 5/2(-x_1) + 1(-x_2)$$

Table 1 illustrates the problem in the detached coefficient form

		$-x_1$	$-x_2$
x_0	0	1	1
u_1	1	1/2	-1
u_2	-2	-2	-1
u_3	9	5/2	-1
v_1	0	-1	0
v_2	0	0	-1
v_3	-1	-1/2	-1
v_4	23	5/2	3
v_5	1	-5/2	1

Table 1

The geometry of this problem is illustrated in Figure 1. (From this diagram it can be seen that the GLP problem has only one feasible solution, namely, the point at the intersection of the hyperplanes $u_1 = 0$ and $u_3 = 0$.)

The inequalities $u_i \geq 0$, $i = 1, 2, 3$, and $v_i \geq 0$, $i = 1, 2, \ldots, 5$ define the L. P. region. The inequalities $v_1 = x_1 \geq 0$ and $v_2 = x_2 \geq 0$ are redundant in this example.

As shown in Table 2, the optimal L. P. solution occurs at the intersection of the hyperplanes $u_2 = 0$ and $v_3 = 0$. At this point all other u_i and v_i are positive since the L. P. solution does not lie on their associated hyperplanes. Thus the intersection of the half spaces $u_1 \geq 0$ and $u_3 \geq 0$ contains the L. P. solution in its interior and serves as a suitable choice for R. The convexity cut

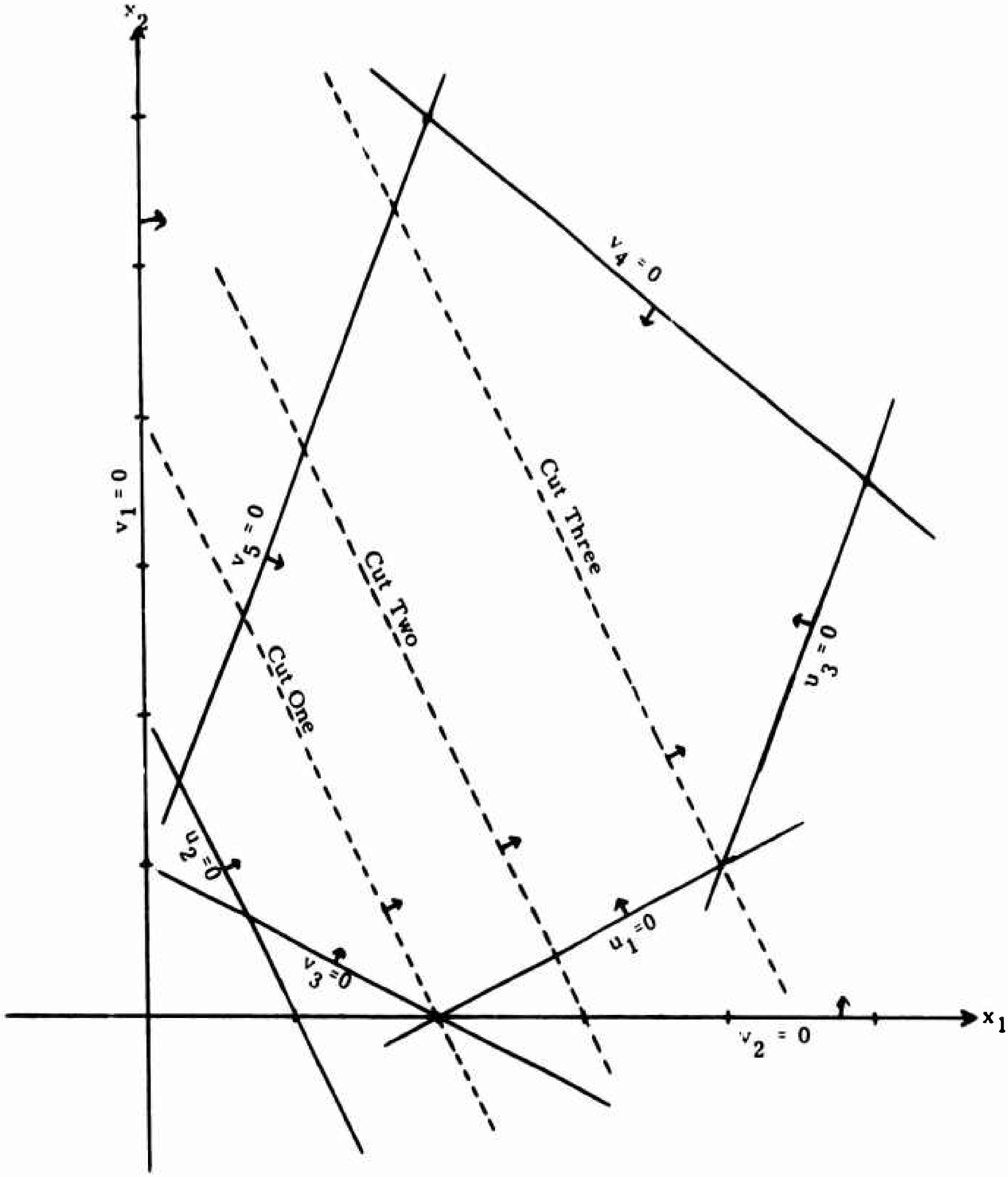

Figure 1

is thus obtained by extending the edges from the L. P. solution until they hit the first hyperplane $u_i = 0$ for $i = 1, 3$ (i.e., a boundary of R). As depicted in Figure 1, edge $u_2 = 0$ may be extended to infinity and edge $v_3 = 0$ may be extended to the hyperplane $u_1 = 0$. The cut is created by passing a hyperplane through the intersection points. Since one of the intersection points is along the displaced ray $u_2 = 0$ at infinity, the cut involves moving $u_2 = 0$ parallel to itself until it passes through the point of intersection of the hyperplanes $v_3 = 0$ and $u_1 = 0$. This is labeled Cut One in Figure 1.

The algebraic counterpart of extending an edge until it intersects R simply amounts to increasing the value of a current nonbasic variable and determining which one of the u_i in R becomes zero first.

		$-u_2$	$-v_3$
x_o	-4/3	1/3	2/3
u_1	4/3	2/3	-5/3
u_2	0	-1	0
u_3	8	2	-3
v_1	2/3	-2/3	2/3
v_2	2/3	1/3	-4/3
v_3	0	0	-1
v_4	58/3	1/12	7/3
v_5	2	-2	3

Table 2

To determine the cut, we identify the relevant equations for defining R from Table 2 to be:

$$u_1 = 4/3 - 2/3u_2 + 5/3v_3$$

$$u_3 = 8 - 2u_2 + 3v_3 .$$

Holding $v_3 = 0$ and allowing u_2 to vary (i.e., traveling along the hyperplane $v_3 = 0$), we find that

$$u_1 = 0 \quad \text{iff} \quad u_2 = 2$$

$$u_3 = 0 \quad \text{iff} \quad u_2 = 4 .$$

Thus, the hyperplane $u_1 = 0$ is intersected first (as may be seen visually in Figure 1). Furthermore, t^*_1 is equal to 2. Similar calculations involving v_3 yield:

$$u_1 = 0 \quad \text{if} \quad v_3 = -4/5$$

$$u_3 = 0 \quad \text{if} \quad v_3 = -8/3 .$$

Since all of the values for v_3 are negative, t^*_2 is taken to be $+\infty$ by convention.

The convexity cut is $(1/t^*_1)u_2 + (1/t^*_2)v_3 \geq 1$. Adding a slack v_6 to transform the convexity cut into an equality, we obtain

$$v_6 = -1 + 1/2u_2 + 0v_3 .$$

(Notice that the effect of the cut is to compel u_2 to be greater than or equal to 2. This likewise corresponds to the geometric portrayal in Figure 1.) Upon adjoining the convexity cut to Table 2 and iterating to the new L. P. optimum, Table 3 is obtained.

		$-v_6$	$-v_3$
x_o	-2	2/3	2/3
u_1	0	4/3	-5/3
u_2	2	-2	0
u_3	4	4	-3
v_1	2	-4/3	2/3
v_2	0	2/3	-4/3
v_3	0	0	-1
v_4	115/6	1/6	7/3
v_5	7	-4	3
v_6	0	-1	0

Table 3

The L. P. solution given by Table 3 is again infeasible for the GLP problem since none of the u_i are non-basic.

A new convexity cut is therefore calculated relative to the set $R = \{(x_1, x_2): u_2 \geq 0, u_3 \geq 0\}$, yielding $1 v_6 + 0 v_2 \geq 1$. (In Figure 1, this is called Cut Two.)

Adjoining this cut to Table 3 (with slack v_7) and finding the new optimal L. P. solution gives Table 4 to follow.

		$-v_7$	$-u_1$
x_o	-16/5	6/5	2/5
u_1	0	0	-1
u_2	4	-2	0
u_3	12/5	8/5	-9/5
v_1	14/5	-4/5	2/5
v_2	2/5	-2/5	-4/5
v_3	4/5	-4/5	-3/5
v_4	514/30	61/30	7/5
v_5	43/5	-8/5	9/5
v_6	1	-1	0
v_7	0	-1	0

Table 4

The solution provided by this table improves upon the previous solution since one of the u_i is now nonbasic. However, still another u_i must become nonbasic in order to provide a feasible solution for the GLP problem.

Consequently a new convexity cut is calculated relative to the set $R = \{(x_1, x_2): u_2 \geq 0,\ u_3 \geq 0\}$. (Note that this R is coincidentally the same one used to derive Cut Two. If u_1 had been made non-basic by replacing the non-basic variable v_3 in Table 3, Cut Two could have been avoided.) The convexity cut is $2/3 v_7 + 0 u_1 \geq 1$ (labeled Cut Three in Figure 1).

Adding this convexity cut to Table 4 and finding the optimal L. P. solution to the new problem, Table 5 is obtained.

		$-v_8$	u_1
x_0	-5	9/5	2/5
u_1	0	0	-1
u_2	7	-3	0
u_3	0	12/5	-9/5
v_1	4	-6/5	2/5
v_2	1	-3/5	-4/5
v_3	7/5	-3/5	-3/5
v_4	845/60	183/60	7/5
v_5	11	-12/5	9/5
v_6	5/2	-3/2	0
v_7	3/2	-3/2	0
v_8	0	-1	0

Table 5

The L. P. solution contained in Table 5 is feasible (and optimal) for the GLP problem since u_3 can obviously be made non-basic. Thus, the GLP problem is solved.[11]

FOOTNOTES

[1]The name "generalized lattice point problem" was coined for this problem by William M. Raike.

[2]Private communications, A. Charnes, Anthony Fiacco, Michael Kirby, and William W. Raike. More recently, Kirby, Love, and Swarup [13] have proposed solving this problem by procedures based on "extreme point ranking" and "parallel shifts of the objective hyperplane."

[3]The excellent work that pertains either directly or indirectly to the convexity cut ideas has been developed by Ragavachari, M. [14] and Burdet [5]. Recent extensions of interest are also to be found in [2, 3, 4, 9, 10, 13, 17].

[4]This statement is slightly more general than the original statement of Charnes, which requires x to be an extreme point of the polytope $Q' = \{x: D'x = d', x \geq 0\}$. It is interesting to note that the attempt to circumvent the non-negativity restriction on x in the definition of Q' encounters some difficulty. In particular, the standard device of replacing an unrestricted variable by the difference of two non-negative variables fails due to the introduction of extraneous extreme points. For example, the polytope $\{x: x_1 + x_2 = 1, x_2 + x_3 = 1, x_3 \geq 0\}$ has one extreme point: $x_1 = x_3 = 0$ and $x_2 = 1$. However, the corresponding polytope $\{y: y_1 - y_2 + y_3 - y_4 = 1, y_3 - y_4 + y_5 = 1, y \geq 0\}$, where $x_1 = y_1 - y_2$, $x_2 = y_3 - y_4$ and $x_3 = y_5$, has two extreme points: $y_1 = y_5 = 1$, $y_2 = y_3 = y_4 = 0$ and $y_3 = 1$, $y_1 = y_2 = y_4 = y_5 = 0$. Only the second of these corresponds to the extreme point of the original polytope.

[5]The rules for column tableau pivoting may be briefly stated: divide the pivot column by the negative of the pivot element; then add the appropriate multiple of the pivot column to each of the other columns so that the resulting "updated" pivot row will have 0's everywhere except in the pivot column.

[6]As observed in [8], the convexity cut lemma is also valid more generally if B_o is not in the interior of R, provided there is a deleted neighborhood of B_o such that all points of C in this neighborhood are in the interior of R.

[7]This case cannot arise for the GLP problem when $R = \bar{Q}$ if any components of u are nonbasic (unless, in fact, x contains only 1 component).

[8]The usual statement of the basis exchange theorem is (loosely) that, given any pair of bases, one can progress from one to the other through a sequence of bases in which any selected member of the second (not previously selected) replaces some member of the first (not previously replaced).

An easily proved consequence of this, which we rely on here, is the "reverse" theorem that states (again loosely) one can remove any selected member of the first basis (not previously selected) and find a member of the second (not previously found) to replace it.

[9]This connection is developed by Glover in [8].

[10]This assertion follows directly from considerations introduced in [9]. Note that this relationship to earlier cuts circumvents the finiteness issue for the present cuts.

[11]The fact that the successive convexity cuts are all parallel in this example is, of course, a fortuitous consequence of the problem structure and is not to be expected in general.

BIBLIOGRAPHY

[1] Balas, Egon. "The Intersection Cut - A Net Cutting Plane for Integer Programming," Management Sciences Research Report No. 187, Carnegie-Mellon University, October, 1969.

[2] ________. "Alternative Strategies for Using Intersection Cuts in Integer Programming," presented at the TIMS XVII International Conference, London, July, 1970.

[3] ________. "Intersection Cuts from Maximal Convex Extensions of the Ball and the Octahedron," Management Sciences Research Report No. 214, Carnegie-Mellon University, August, 1970.

[4] ________, Bowman, V. Joseph, Glover, Fred, and Sommer, David. "An Intersection Cut from the Dual of the Unit Hypercube," WP #90-69-7, Carnegie-Mellon University, April, 1970.

[5] Burdet, Claude A. "A Class of Cuts and Related Algorithms in Integer Programming," Working Paper, Carnegie-Mellon University, November, 1970.

[6] Charnes, A. Unpublished notes in mathematical programming course at Carnegie-Mellon University, Northwestern University and The University of Texas.

[7] ________ and Cooper, W. W. Management Models and Industrial Applications of Linear Programming, Volumes I and II. New York: John Wiley & Sons, 1961.

[8] Glover, Fred. "Convexity Cuts," Working paper, School of Business, The University of Texas, Austin, December, 1969.

[9] ________. "Convexity Cuts and Cut Search," MSRS 70-2, School of Business, University of Colorado at Boulder, July, 1970.

[10] ________ and Klingman, D. "Concave Programming Applied to a Special Class of 0 - 1 Integer Problems," Applied Mathematics for Management, No. AMM-11, WP 69-30, The University of Texas, Austin, January, 1969.

[11] ________, ________ and Stutz, Joel. "A Class of Disjunctive-Conjunctive Generalized Lattice Point Problems," Forthcoming, Center for Cybernetic Studies Research Report, The University of Texas.

[12] Gomory, R. "An Algorithm for the Mixed Integer Problem," Research Memorandum RM-2597, Rand Corporation, Santa Monica, 1960.

[13] Kirby, M.J.L., Love, H. R. and Swarup, Kanti. "Extreme Point Mathematical Programming," Research Report, Dalhousie University, February, 1970.

[14] Raghavachari, M. "On the Zero-One Integer Programming Problem," Operations Research Vol. 17 (1969).

[15] Tui, Hoang. "Concave Programming Under Linear Constraints," Soviet Mathematics (1964), pp. 1437-1440.

[16] Young, Richard. "New Cuts for a Special Class of 0 - 1 Integer Programs," Research Report in Applied Mathematics and System Theory, Rice University, Houston, Texas, November, 1968.

[17] _________. "Hypercylindrically Deduced Cuts in Zero-One Integer Programs," Rice University, Houston, Texas, June 1970.