Math. Program., Ser. B 96: 341–375 (2003)

Digital Object Identifier (DOI) 10.1007/s10107-003-0389-3

#74

Ralph E. Gomory · Ellis L. Johnson

T-space and cutting planes

Received: May 24, 2001 / Accepted: January 2003
Published online: March 28, 2003 –

Abstract. In this paper we show how knowledge about T-space translates directly into cutting planes for general integer programming problems. After providing background on Corner Polyhedra and on T-space, this paper examines T-space in some detail. It gives a variety of constructions for T-space facets, all of which translate into cutting planes, and introduces continuous families of facets. In view of the great variety of possible facets, no one of which can be dominated either by any other or by any combination of the others, a measure of merit is introduced to provide guidance on their usefulness. T-spaces based on higher dimensional groups are discussed briefly as is the idea of going beyond cutting planes to iterated approximations of Corner Polyhedra.

Background, literature, and the contributions of this paper

Corner Polyhedra based on finite groups were introduced by Gomory [1]. That paper introduced finite T-space, Master Polyhedra and many of the fundamental properties of Corner Polyhedra, and made clear the possibility of generating cutting planes for small groups. It also indicated some methods whereby knowledge of facets for small groups could be lifted up to form facets for groups of any size. Gomory and Johnson [2] introduced continuous functions and the use of the unit interval (Mod 1), in place of the finite groups of [1]. They introduced infinite dimensional T-space and the fundamental concepts of valid inequality, subadditivity, minimality, and facets, which we will use here. Although this work was motivated by the desire to escape the limitations of finite groups, and especially of small finite groups, and to provide systematic practical methods for cutting planes, it was not until 1972 that Gomory and Johnson [3] spelled out the direct connection of the T-space theory with the practical issue of generating cutting planes. A useful rigorous treatment of infinite dimensional T-space can be found in [6]. Further work on the mixed integer case appears in [5]. There is related material in [7] and information on practical experience with the earliest cutting planes in [8].

In Section 1 of this paper we review the needed results from the papers [1], [2] and [3]. We show the direct connection of T-space with cutting planes and with the hierarchy of valid inequalities. We obtain the needed results from [2] and [3] by arguments that set the stage for Section 2.

In Section 2 we introduce the Facet Theorem, the cylindrical space S, and the Interval Lemma. These are the main tools that underlie the more detailed exploration of T-space that follows.

Ralph E. Gomory: Sloan Foundation, New York, NY, USA

Ellis L. Johnson: Georgia Institute of Technology, School of Industrial and System Engineering, 765 Ferst Street, Atlanta, GA, USA 30332-0205

In Section 3 we start the detailed exploration of T-space. We give a first construction (Construction 1) for a simple family of facets. Facets in this family are inequalities that vary continuously with one parameter. There is a facet for each parameter value, no one facet dominates another, and each one generates inequalities for integer programming.

We next introduce a more complex construction (Construction 2) for a two parameter family of facets. In proving that Construction 2 does yield facets, we introduce lemmas relating to subadditivity that are more generally useful. Then, using this material, we introduce a third construction (Construction 3) that gives us a two parameter family of three-slope facets.

Next we show how homomorphisms of the unit interval can be used to generate facets that are multiple replicas of some starting facet.

If we were dealing with a finite polyhedron, any facet would be linked to any other by a sequence of intermediate facets that are adjacent to each other. Since our T-space polyhedron is infinite dimensional, the passage from one facet to another through intermediate facets is continuous. So there should be a continuous path of intermediate facets linking any two of our facets. We give an example of such a path of intermediate facets linking an important three-slope facet through a variety of different looking intermediate facets to the basic mixed integer facet.

Having shown the great variety and extent of the possible facets in Section 3, we need some measures of their goodness. In Section 4 we introduce two new measures for facets, a Merit Index and an Intersection Index. The Merit Index is related to the number of paths on a facet and therefore in some rough sense to its size. The Intersection Index reflects the number of paths two facets have in common, and therefore relates to how much they intersect or how close they are. We define these indices and compute them for some of the facets and families of facets that have been produced.

In Section 5 we briefly discuss higher dimensions, and in Section 6 we give some perspectives that are suggested by this work and indicate some possible areas for further activity. We also state an important and challenging conjecture.

1. The cutting plane process and its justification

1.1. A practical construction process for cutting planes

We will start by showing a practical construction process for both pure integer and mixed integer problems. Then, later in this paper we show how to construct the class of special functions, referred to here as the π-functions, that are used in the construction process and that appear in Figure 1.

Assume that we have a basic feasible solution of the linear programming problem so that we have basic and non-basic variables. We will work with the non-basic variables to produce a new valid inequality which is not satisfied at the present basic feasible solution in which, of course, all the non-basic variables are zero.

Choose a row of the transformed matrix corresponding to a basic integer variable whose value in the basic solution is fractional. Then: (1) For a non-basic *integer* variable t, find the fractional part f of its coefficient. Use the π-function value $\pi(f)$ as the coefficient of t in a new inequality. (2) For a non-integer non-basic variable with positive

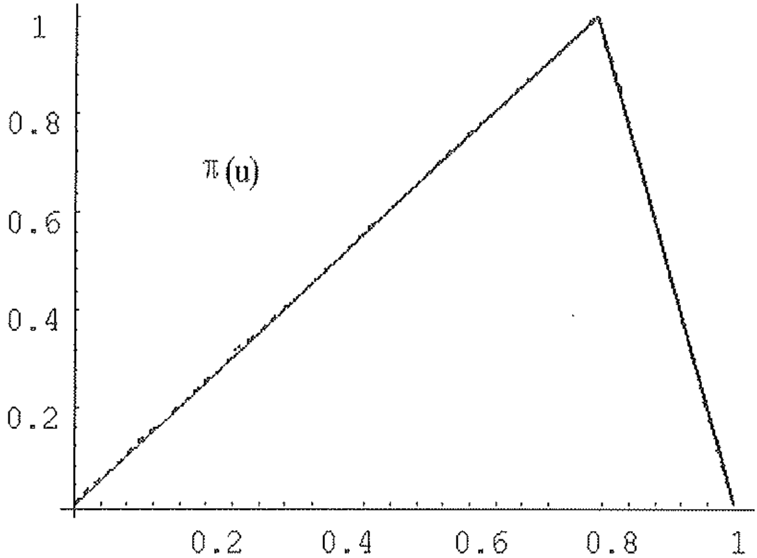


Fig. 1.

coefficient multiply the coefficient by the right hand slope of π at 0, and enter the result as the coefficient of that variable in the new inequality. (3) For a non-integer non-basic variable with negative coefficient multiply that coefficient by the (negative) slope of π at 1. Then complete the new inequality by choosing 1 as its right-hand side. The basic variable does not appear in the inequality.

Example 1. Suppose a row of the transformed matrix is $x_1 + 4.72t_1 - 2.93t_2 + 0.51t_3 + 0.14t_4 + 1.1t^+ - 1.4t^- = 2.79$ where x_1 is a basic, integer variable, $t_1, \ldots, t_4$ are non-basic integer variables, t^+ and t^- are non-basic continuous variables, and 2.79 is the value of x_1 in the basic solution. The fractional parts of the coefficients of $t_1, \ldots, t_4$ are $(0.72, -0.07, -0.51, -0.14)$.

New Inequality: Using $\pi(u)$ values and slopes 1/0.79 and $-(1/0.21)$ gives the inequality.

$$0.9114t_1 + 0.0886t_2 + 0.6456t_3 + 0.1772t_4 + 1.3924t^+ + 6.6667t^- \geq 1$$

This process yields a new inequality if π is a *valid inequality* for the Group Problem Mod 1. This will be explained in more detail in section 1.3.

1.2. Origins of the inequalities: T-space and the master polyhedron

We start with an additive Abelian group G which can be finite or infinite. We form sums of elements $u \in G$ by taking for each $u \in G$ a nonnegative integer $t(u)$ and then adding up the $t(u)u$ for all group elements u. We only consider sums having a finite number of non-zero $t(u)$. We refer to the vector $\{t(u)\}$ as a path that leads to the sum element $u_0 = \sum t(u)u$.

If G is a finite group, as it was in [1], we can form a vector space having one dimension for each non-zero group element. There is a one-to-one correspondence between the non-zero group elements and the coordinates of vectors in the vector space. The various paths $\{t(u)\}$ are then represented by the integer lattice points in the first quadrant of that finite dimensional vector space. We refer to this space as T-space (Figure 2). If G is infinite, as it was in [2] and [3], T-space formed in the same way is infinite dimensional, and the paths form a regular integer lattice in that infinite dimensional space. In this paper G will always be the real numbers (Mod 1), so T-space will be infinite dimensional. We can single out those paths that add up to any one particular group element u_0, that we will call the *right-hand side element* or *rhs*. These paths form a sub-lattice in the integer lattice of paths in T-space. In Figure 2, the dots represent all the paths in a two dimensional T-space. Those paths that add up to the rhs have circles around them. The convex hull of these first quadrant elements forms a polyhedron that is uniquely determined by the group G and by the choice of u_0. We call this polyhedron the Master Polyhedron $P(G, u_0)$.

Figure 2 illustrates a corner polyhedron with two variables. If we could visualize T-space over the interval [0,1), it would be an infinite dimensional version of this polyhedron.

1.3. Valid inequalities and subadditivity

For a given function $\pi(u)$ defined on G with zero element $\underline{0}$, we will say π is a *valid function* if it leaves the polyhedron $P(G, u_0)$ on one side, as illustrated by the heavy line in Figure 2.

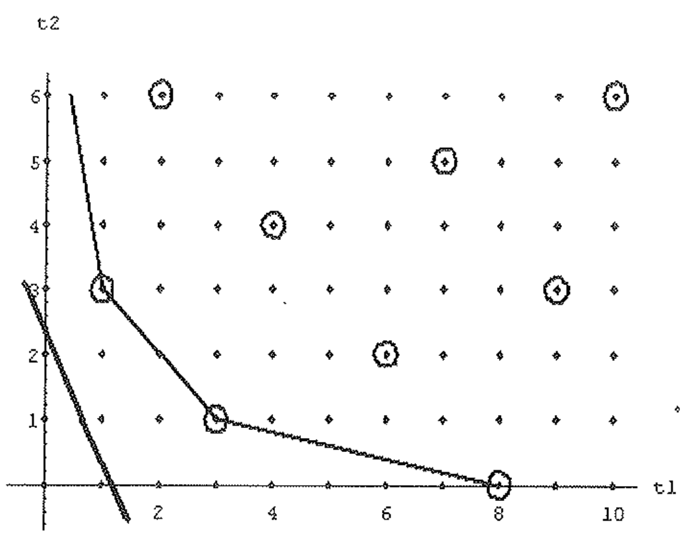


Fig. 2.

More formally, we will say that π is a valid function with rhs element u_0 if and only if:

$$\pi(u) \text{ is continuous and nonnegative, } \pi(\underline{0}) = 0, \pi(u_0) = 1, \text{ and} \tag{1}$$
$$\sum t(u)u = u_0 \text{ implies } \sum \pi(u)t(u) \geq 1.$$

The valid functions $(\pi, \pi(u_0))$ that we study are always *subadditive*. Subadditivity simply means that $\pi(u_1) + \pi(u_2) \geq \pi(u_1 + u_2)$. Validity and subadditivity are naturally and tightly connected. If π is subadditive and has $\pi(\underline{0}) = 0$, then, for any path $\sum t(u)u = u_0$, we have $\sum \pi(u)t(u) \geq \pi(u_0)$. If $\pi(u_0) > 0$ we can divide both sides of the inequality by $\pi(u_0)$ to make the right-hand side 1 and obtain a valid inequality. So subadditivity implies validity. While it is possible to construct valid inequalities that are not subadditive, they can always be strengthened into subadditive ones, so we will work exclusively with functions $\pi(u)$ that are subadditive. Our functions $\pi(u)$ will always have bounded slopes and are typically made up of a finite number of line segments. [1]Throughout the paper, we will often refer to a valid function π as a *valid inequality* $\pi(u)$, to demonstrate the analogy between these valid functions for infinite polyhedra and traditional valid inequalities for discrete polyhedra.

1.4. Cutting planes

We next discuss the connection between subadditive functions $\pi(u)$ and cutting planes for integer programming.

Consider the m row, $m+n$ column integer programming problem whose constraints are $Bx + Nt = b$. Here x and t are the basic and non-basic variables and b is the vector of right-hand sides. All variables are required to be integers. Then the transformed problem is $Ix + (B^{-1}N)t = B^{-1}b$ where I is the identity matrix. If the x are to be integer, the non-basic variables t must satisfy $(B^{-1}N)t = B^{-1}b$ (Mod 1).

If we denote the columns of $B^{-1}N$ by c_i and $B^{-1}b$ by c_0 we have

$$\sum_i t_i c_i = c_0 \text{ (Mod 1)}\,. \tag{2}$$

This looks very much like the path condition of (1) with the c_i playing the role of the u's, the t_i playing the role of the $t(u)$, and the c_0 the role of the rhs element u_0. The only difference is that the c_i and c_0 are m-vectors being added (Mod 1) while the u and u_0 are real numbers being added (Mod 1).

However there are simple mappings that send vectors of m-space into elements of G while preserving addition. For example, if we choose from any m-vector the k^{th} component, and then map that element (Mod 1) into G, this mapping χ of m-vectors into group elements will be addition preserving. It also sends the 0-vector into the $\underline{0}$ of G. Mappings, such as χ or $n\chi$, that are addition preserving, satisfy $\chi(g_1) + \chi(g_2) = \chi(g_1 + g_2)$, and

[1] That we need only consider functions with bounded slope can be proven theoretically [6]. It is easily shown that a subadditive function π has bounded slopes if its slopes at the origin and at 1 are bounded. What a bounded slope at the origin means is that the quotient $\pi(u)/u$, for $u > 0$, does not become unbounded as u becomes small.

map some group into G, are called *group characters*. Taking any such mapping χ we have from (2)

$$\sum_i \chi(c_i t_i) = \chi(c_0). \tag{3}$$

Since each $\chi(c_i t_i)$ is a group element of G, the reals (Mod 1), the t_i of the integer solution have given us in (3) a path in G to the rhs element $\chi(c_0)$. For any valid $\pi(u)$ then we can use subadditivity to obtain:

$$\sum_i \pi(\chi(c_i t_i)) \geq 1. \tag{4}$$

For the variables t_i that are integers, since we are adding t_i copies of the same group element,

$$\pi(\chi(c_i))t_i \geq \pi(\chi(c_i t_i)) \text{ and} \tag{5}$$
$$\sum_i \pi(\chi(c_i))t_i \geq 1.$$

Since the $\pi(\chi(c_i))$ in (5) are non-negative constants, this is a new cutting plane involving the non-basic variables t_i. So we have proved:

Theorem 1. *All Integer Cutting Plane Theorem: If $\pi(u)$ is subadditive, and $\chi(c_i)$ is a character sending the columns of $B^{-1}N$ into G, then the inequality (5) on the non-basic variables t is a valid cutting plane for the all integer programming problem whose constraints are $Bx + Nt = b$.*

If we take as the character $\chi(c_j)$, $\chi(c_j) = F(c_{j,k})$ the fractional part of the element $c_{j,k}$, then this is, for the integer variables, the cutting plane generation process described at the beginning of Section 1.

Note that in the reasoning that led to the Cutting Plane Theorem we did not use the non-negativity of the basic variables, only their integrality. The geometric object obtained by relaxing the non-negativity constraint on the basic variables is called the Corner Polyhedron associated with this basic solution, see [1], [2], [3] and [4]. It is the faces of these Corner Polyhedra that we will be pursuing in our quest for cutting planes. In Appendix A we extend the analysis to cover the treatment of non-integer variables.

1.5. Hierarchy of inequalities – minimality

If every valid inequality π gives an inequality or cutting plane, or perhaps several depending on the choice of row k, we want to be able to generate the functions π. Since we will soon see that they are in fact rather easy to generate, we will also need to decide which are better than others so we will describe a hierarchy of inequalities.

There are three levels of inequalities, valid, minimal and facet (or extreme). Geometrically they are quite easy to understand. Valid is the weakest level of inequality, it merely means that the inequality leaves the polyhedron to one side.

A *minimal* inequality is stronger, it means that there is no other valid inequality that is uniformly better. More formally π is minimal if there is no valid π_1 with $\pi_1 \leq \pi$, and

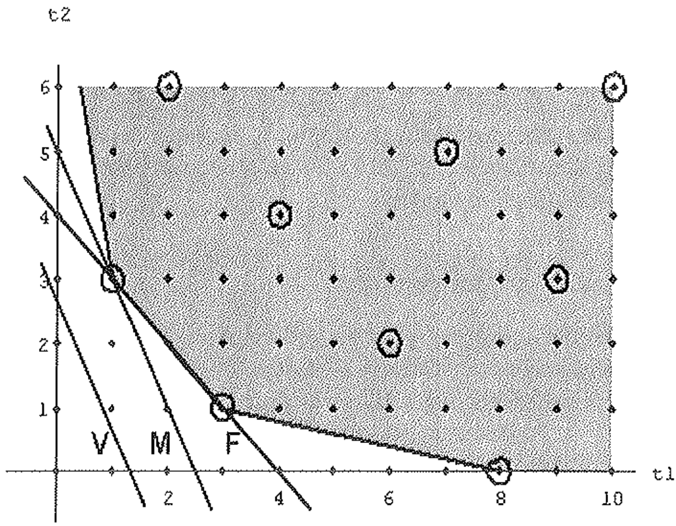


Fig. 3.

$\pi_1(u) < \pi(u)$ for some u. In practice the extremely useful criterion for minimality is the one derived by Gomory and Johnson [2, 3]. In Figure 3, the three types of inequalities are illustrated.

Theorem 2. *Minimality Theorem: The necessary and sufficient condition for a valid $\pi(u)$ to be minimal is that $\pi(u)$ is subadditive and that the symmetry condition, $\pi(u) + \pi(u_0 - u) = \pi(u_0) = 1$, holds for all $u \in G$, where u_0 is the rhs element in (1).*

Sketch of proof for the finite group case: The concept behind this theorem is clear enough when we are dealing with *finite* groups G. Suppose π has symmetry. If π is not minimal there is a $\pi_1 \leq \pi$ with $\pi_1(u) < \pi(u)$ for some u. Then $\pi(u) + \pi(u_0 - u) = \pi(u_0) = 1$ turns into $\pi_1(u) + \pi_1(u_0 - u) < 1$, which means that π_1 is not valid. So symmetry implies minimality. Now suppose for an element u_1 in a finite group with n elements we did not have symmetry. Then $\pi(u_1) + \pi(u_0 - u_1) = 1 + \delta$. If we reduce the value of π on u_1 by δ/n to form π_1, we have for any path to u_0 that uses u_1 r times,

$$\pi(u_1) + [(r-1)\pi(u_1) + \sum_{(u \neq u_1)} t(u)\pi(u)] \geq \pi(u_1) + \pi(u_0 - u_1) = 1 + \delta.$$

Since changing π to π_1 decreases the left side by $\delta r/n$ and we can assume $r < n$, we have

$$\pi_1(u_1) + (r-1)\pi_1(u_1) + \sum_{(u \neq u_1)} t(u)\pi_1(u) \geq 1 + (1 - r/n)(\delta) \geq 1,$$

which shows that the new smaller π_1 is a valid inequality. Therefore π was not minimal. Therefore minimality implies symmetry and ends the proof. The proof for continuous groups involves the same ideas but is much more complex.

The symmetry condition is quite strong. First of all, it implies that any minimal $\pi(u)$ passes through the two points $\frac{1}{2}(u_0, 1)$ and $\frac{1}{2}(1 + u_0, 1)$ that are halfway up the two straight lines connecting (0,0) and $(u_0, 1)$ in Figure 4. In Figure 4, $\pi(u)$ is the dark line, and P1 and P2 mark the two halfway points. Second, it implies that if half the curve $\pi(u)$ is known, so is the other half through $\pi(u) = 1 - \pi(u_0 - u)$. We will refer to u and $(u_0 - u)$ as *complementary* points. All minimal $\pi(u)$ must exhibit the odd sort of complementary symmetry seen in Figure 4.

1.6. Hierarchy of inequalities – facet definition

Facets are stronger still and they are what we will emphasize. The intuitive idea of a facet is that it is a valid inequality π whose contact with $P(G, u_0)$ is maximal. Facets are always minimal. Since facets are valid inequalities we always have for any path $\{t(u)\}$, $\sum \pi(u)t(u) \geq 1$. However facets usually have lots of paths $\{t(u)\}$ for which equality holds, i.e., $\sum \pi(u)t(u) = 1$. We will use the phrase *a path* $\{t(u)\}$ *lies on an inequality* π to mean that $\sum \pi(u)t(u) = 1$. We will sometimes denote this by $t \in \pi$.

In this paper[2] we use as our concept of facet the idea that π is a facet when there is no other inequality π^* which has all the paths on it that π has, plus an additional path or paths. More formally, let $P(\pi)$ denote the set of paths lying on an inequality, then:

Facet Definition: *π is a facet if and only if*

$$P(\pi^*) \supset P(\pi) \text{ implies } \pi^* = \pi.$$

While this turns out to be a useful definition of a facet, we can not list all possible paths to find out if a $\pi(u)$ is or is not a facet. Without looking at the paths themselves, we need to know if a given inequality $\pi(u)$ is or is not a facet. Some simple observations will lead us closer to that goal.

1.7. Path properties

Paths lying on a face have some special properties. Consider elements u_1 and u_2 that make up part of a path p lying on a valid subadditive inequality $\pi(u)$. We will denote the fact that u_1 and u_2 are part of the path p by $(u_1, u_2) \in p$, so we have $(u_1, u_2) \in p \in \pi$. Suppose that u_1 and u_2 on p had a strict inequality $\pi(u_1) + \pi(u_2) > \pi(u_1 + u_2)$. Then the substitution of the element $(u_1 + u_2)$ for the two separate elements u_1 and u_2 would produce a new path to u_0 that would have a smaller sum $\sum \pi(u)t(u)$, a sum less than 1. But since $\pi(u)$ is valid, $\sum \pi(u)t(u) \geq 1$ for *any* path $\{t(u)\}$. This is a contradiction. We conclude:

Lemma 1. *Pairs Lemma: Paths lying on subadditive valid inequalities contain only pairs u_1 and u_2 that satisfy $\pi(u_1) + \pi(u_2) = \pi(u_1 + u_2)$.*

With this background established we are in a position to create the tools needed for investigating T-space and for facet creation.

[2] While this definition is different from, although eventually equivalent to, the one used in [3] it is the one that leads much more easily and intuitively to the Facet Theorem which we need for facet construction.

2. Tools for facet creation

2.1. Facet theorem

Our main tools for facet creation will be the Facet Theorem and the Interval Lemma. To prove the Facet Theorem we need a lemma connecting equalities which relate two elements of the same path, with all the equalities of the form $\pi(u_1)+\pi(u_2) = \pi(u_1+u_2)$ satisfied by a function $\pi(u)$.

Denote by $E(p)$ the set of all equalities $\pi(u_1) + \pi(u_2) = \pi(u_1 + u_2)$ where both u_1 and u_2 are used in the path p. Denote by $E(\pi)$ the set of all possible inequalities $\pi(u_1) + \pi(u_2) \geq \pi(u_1 + u_2)$, that are satisfied as *equalities* by π. Here u_1 and u_2 are any elements of G. We can now state the useful lemma.

Lemma 2. *Equations Lemma: If π is subadditive and minimal, the set of equalities $E(p)$ obtained from all paths p lying on π is the same as the set of equalities $E(\pi)$. Symbolically*

$$\bigcup_{p\in\pi} E(p) = E(\pi).$$

Proof. Clearly any equality $\pi(u_1) + \pi(u_2) = \pi(u_1 + u_2)$ obtained by the Pairs Lemma from a path $p \in \pi$ is part of $E(\pi)$. Now consider *any* equality $\pi(u_1) + \pi(u_2) = \pi(u_1 + u_2)$ from $E(\pi)$. Since π is minimal, $\pi(u_1 + u_2) + \pi(u_0 - (u_1 + u_2)) = 1$. Therefore $\pi(u_1) + \pi(u_2) + \pi(u_0 - (u_1 + u_2)) = 1$. Therefore u_1 and u_2 lie on this path of length three, and their equality is included among the equalities obtained from all paths. This ends the proof.

We have now laid the groundwork for the Facet Theorem which we will use repeatedly below to identify facets.

Theorem 3. *Facet Theorem: If π is subadditive and minimal, and if the set $E(\pi)$ of all equalities has no solution other than π itself, then π is a facet.*

Proof. Suppose π is the only solution to the equations $E(\pi)$, but it is not a facet. Then there is a facet π^* that contains all the paths on π and at least one additional path p^* that does not lie on π. However, by the Equations Lemma, all the paths on π already generate $E(\pi)$. Therefore π^* satisfies the equations of $E(\pi)$. But $E(\pi)$ has only the one solution π. So $\pi = \pi^*$ and π is a facet.

We will use the Facet Theorem extensively to show that the valid inequalities we construct in the next section are in fact facets.

2.2. The cylindrical space S and a diagram

In constructing facets we will make extensive use of the space S of all points (u, h) where $u \in G$ and h is any real number. Figure 4 shows the diagram we use to represent S and inequalities $(u, \pi(u))$ in S. The h values are plotted vertically, and the group elements, represented by their real values between 0 and 1 are plotted horizontally. In

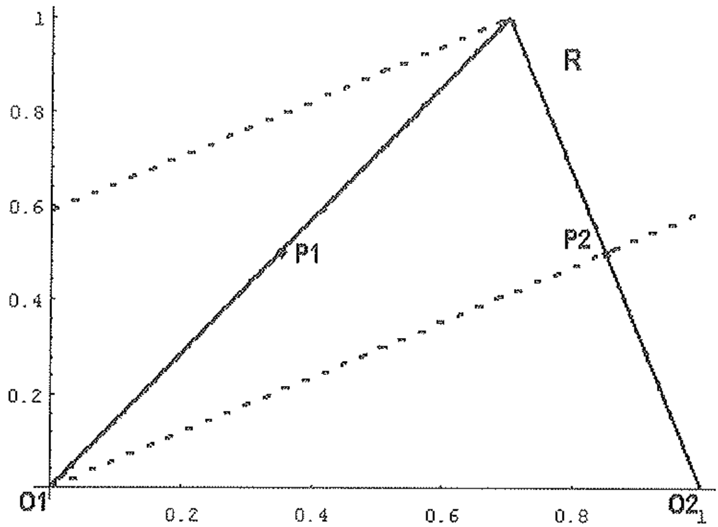


Fig. 4.

the diagrams, points lying on the vertical line through the origin in S, $u = \underline{0}$, always appears twice. The origin $O = (\underline{0}, 0)$ in S, appears once in the diagram at (0,0), which we will name $O1$, and again at (1,0) which we refer to as $O2$. If u is a group element then the corresponding number in the interval $0 \le x < 1$ we denote by $\eta(u)$.

Any point in S, such as R in Figure 4, has its corresponding group element $u(R)$ and its height $h(R)$. In S we can add any two points A and B, taking care that the horizontal components are added (Mod 1). Since the vertical components are added in the usual way, S resembles an infinite cylinder. In S, in the immediate neighborhood of the origin or of some other point P, the behavior of vectors and line segments is exactly the same as in the plane. For example the concepts of slope and of length in S are the same as in the plane and can easily be made rigorous. However for larger objects, the topology of the cylinder makes a difference.

In S, the origin is connected to a point A not by one straight line, but by many. Some lines head off clockwise, some counterclockwise, and lines can wind around the cylinder different numbers of times. In fact, for any point $A = (u, h)$, instead of there being one straight line from the origin to A, with one uniquely determined slope, there is a straight line from the origin to A having slope $h/(\eta(u) + n)$, where n is *any* integer. In Figure 4 we see three such lines from (0,0) to (0.7,1). For $n = 0$, we get the line O, $P1$, R with slope 1/(0.7), for $n = -1$ we get the line $O2$, $P2$, R with slope $1/(-0.3)$, and for $n = 1$ we get the dashed line ($O1$, $P2$, (0, .6), R) with slope 1/(1.7).

We have to keep these possibilities in mind as we work in S. However they all depend on the cylindrical topology. Whenever we explicitly confine ourselves to a region of S that can be mapped one-one onto a region of the plane, line segments and vectors have the usual planar properties. For example, within a vertical strip of the cylinder with width strictly less than 1, two points have only one connecting straight line segment lying entirely within that region. When such a region is specified, the line segment or vector lying entirely within it will be referred to as the *direct* vector or line segment for that region.

2.3. *S*-Vectors

In creating facets we will often need to multiply a vector by a non-integer scalar. Multiplying a vector by an integer is merely repeated vector addition, and therefore gives a unique result. However, in S, there is no unique meaning to multiplying a vector v by a non-integer scalar λ because of the multiple lines. In Figure 4, R is (0.7, 1). Both $P1 = (0.35, 0.5)$ and $P2 = (0.85, 0.5)$ are vectors $\frac{1}{2}R$. Multiplying either by 2 produces R. If we had chosen λ as $\frac{1}{10}$ instead of $\frac{1}{2}$ we would have had 10 choices for $(\frac{1}{10})R$, and if λ were irrational there would be a countable infinity of choices.

To deal with this we attach a slope s to the vector v, and then interpret λv as changing the distance to the origin, measured down the line of slope s, by a factor of λ. Of course we can only choose among slopes s that do connect the point R to the origin. More precisely from a vector $v = (u, h)$ we define an *s-vector*, $\{v, s\}$, as a vector $v = (u, h)$ and one of the slopes, $s = h/(\eta(u) + n)$ for some integer n. We define $\lambda\{v, s\}$ by $(\lambda h/s(\text{ Mod } 1), \lambda h) = (\lambda(\eta(u) + n)(\text{Mod } 1), \lambda h)$. For example, multiplying $R = (0.7, 1)$ for the slope $s = 1/(-0.3)$ by $1/2$ yields $(0.5(-0.3) \text{ Mod } 1, 0.5) = (0.85, 0.5)$, which is $P2$. Doing the same for the pair (0.7,1) and $s = 1/0.7$ yields $P1$.

2.4. *Interval lemma*

The inequalities $\pi(u)$ that we deal with always consist of successive straight line segments, many having the same slopes. To prove that $\pi(u)$ is a facet we will show that the inequalities that $\pi(u)$ satisfies determine its values uniquely. Then we will apply the Facet Theorem. In showing the uniqueness of $\pi(u)$ we will make extensive use of the Interval Lemma.

Lemma 3. *Interval Lemma: Let $U = [u_1, u_2]$, $V = [v_1, v_2]$, and $U + V = [u_1 + v_1, u_2 + v_2]$ be three closed intervals on G. If, whenever $u \in U$ and $v \in V$, we have $\pi(u) + \pi(v) = \pi(u + v)$, then $\pi(u)$ must be a straight line with some constant slope s, for all u in U, V, and $U + V$.*

Interval Lemma would be obvious if we knew that $\pi(u)$ had a derivative instead of being merely continuous. In that case, differentiating the equality $\pi(u) + \pi(v) = \pi(u + v)$ with respect to u would yield $\pi'(u) = \pi'(u + v)$. This means that for a fixed u and varying v the slope at $u + v$ remains the same and produces a straight line segment. So the only difficulty lies in proving the lemma without differentiability. Since this Lemma is simply a restatement on G of a well known result about linear functions, we will not give the proof here.

3. Families of facets

3.1. A family of facets

We now construct our first family of facets. Let $R = (u_0, 1)$. Let v_1 be the s-vector from the origin O to R with slope $1/(\eta(u_0))$; this is $(O1, R)$ in Figure 5. Let v_2 be the s-vector from O to R with slope $1/(\eta(u_0) - 1)$; this is $(O2, R)$ in Figure 5.

Construction 1: First construct the point $P1 = (1/2)v_1$. Then for some λ, $0 \leq \lambda \leq 1$ construct the points $A = P1 + (1/2)\lambda v_2$ and $AA = P1 - (1/2)\lambda v_2$. In order to produce a single valued non-negative $\pi(u)$, λ is limited to the values that keep A and AA in the diagram, specifically $\eta(A) > 0$ and $h(AA) \geq 0$. This requires that $\lambda \leq \text{Min}\{1, \eta(u_0)/(1 - \eta(u_0))\}$

Theorem 4. Construction 1 Theorem: *The $\pi(u)$ formed by the direct segments connecting the successive pairs of points in the sequence O, A, AA, R, O is a facet.*

Remarks. (1) $\pi(u)$ appears in the diagram as $O1$, A, AA, R, $O2$ because of the two appearances of O. (2) In this theorem, and in others to follow, each direct segment is defined relative to a specific vertical strip. For each segment it is the vertical strip containing all points (u, h) with $\eta(u)$ on or between the $\eta(u)$ of the endpoints. (3) Define w_1 to the the direct vector from O to A (this is $O1$ to A in Figure 5) and w_2 to be the direct vector from O to $P2$ (this is $O2$ to $P2$ in Figure 5). Thus, $w_2 = (1/2)v_2$. (4) When we have two points p_1 and p_2 on a line segment, we will use $u[p_1, p_2]$ to refer to the group interval $[u(p_1), u(p_2)]$ directly below that segment.

Proof: (A) Minimality and Subadditivity.

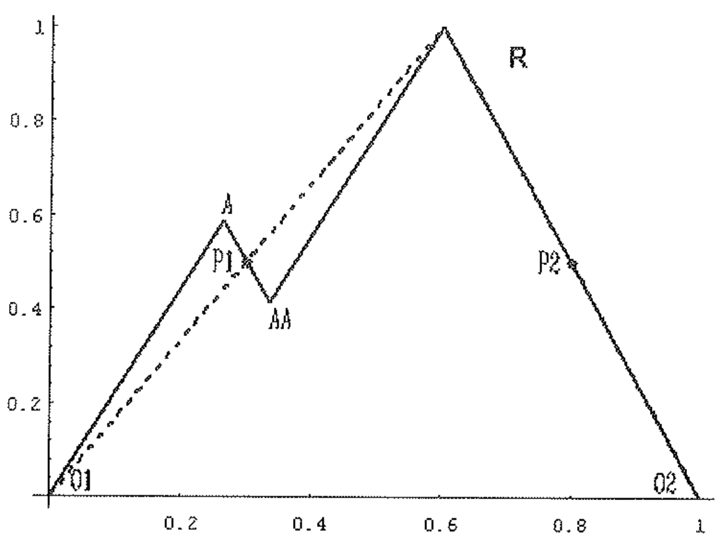


Fig. 5.

(1) Minimality: Since $A + AA = 2P1 = R$, A and AA are complementary points. The points on $[O, A]$ are complementary to those on $[AA, R]$ and the intervals $[A, P1]$ and $[P1, AA]$ are complementary as are $[R, P2]$ and $[P2, O]$. Since π is symmetric, by the Minimality Theorem, it is minimal. (2) Subadditivity: To check subadditivity we first observe that any point on π is either of the form $O1 + \rho w_1$, or the form $O2 + \rho w_2$, or of the form $AA + \rho w_1$, or $AA + \rho w_2$ for $\rho \geq 0$. Therefore any sum $p_1 + p_2$ is of the form (a) $O + \rho w_1 + \tau w_2$, (b) $AA + \rho w_1 + \tau w_2$, or (c) $2AA + \rho w_1 + \tau w_2$ with $\rho, \tau \geq 0$. We can readily see from Figure 5 that any sum of form (a) or (b) will lie on or above π. By construction, the point $AA = P1 - \lambda w_2$. Therefore, $2AA = 2P1 - 2\lambda w_2 = R - \lambda v_2 = (1 - \lambda)v_2$, since $v_2 = 2w_2 = R$. Hence, $2AA$ lies on $[R, O]$, (R, $O2$ in Figure 5), for $\lambda \leq 1$. So the vectors of the form (c) are really of the form $\rho w_1 + (\tau + 2(1 - \lambda))w_2$, which always lies on or above π. This proves subadditivity.

Since π is minimal and subadditive, the next step is to show that $\pi(u)$ is the only possible solution to all the equalities $E(\pi)$ and then apply the Facet Theorem.

Proof (B): Uniqueness of the Solution. Consider any $\pi^*(u)$ that satisfies all the equations satisfied by $\pi(u)$. Consider the direct segment $[O, (1/2)w_1]$. Take the corresponding interval on G, $u[O, (1/2)w_1]$ as both the interval U and as the interval V in the Interval Lemma. The interval $U + V$ then is $u[O, A]$. Since for any points $(u, \pi(u))$ and $(v, \pi(v))$ on the segment $[O, A]$ we have $\pi(u) + \pi(v) = \pi(u + v)$, we must also have $\pi^*(u) + \pi^*(v) = \pi^*(u + v)$ for those points. Then by the Interval Lemma, π^* must be a straight line segment on $U + V = u[O, w_1]$ with some slope s_1.

We next consider the direct segment $[AA, R]$. We again take $u[O, (1/2)w_1]$ as U, but, we now take $[AA, AA + (1/2)w_1]$ as V. The resulting $U + V$ is $u[AA, AA + w_1] = u[AA, R]$. Again, since we have $\pi(u) + \pi(v) = \pi(u + v)$ for $u \in U$ and $v \in V$, we must also have $\pi^*(u) + \pi^*(v) = \pi^*(u + v)$ for all the points on those intervals. Applying the Interval Lemma, we find that π^* must be a straight line of fixed slope on $u[AA, R]$. Furthermore, from the Interval Lemma, the slope of π^* must be the same as it is in $U = u[O, (1/2)w_1]$, so it is s_1.

Continuing to use the Interval Lemma we can show that on the intervals $u[A, AA]$ and $u[R, O2]$, π^* must be made up of straight line segments with slope s_2.

The slope s_2 is uniquely determined by the condition that the segment $[R, O2]$ descends from 1 to 0 over the interval $u[R, O2]$. Since π also meets that same slope condition, π^* is not only a straight line but also has the same slope as π on $[R, O]$.

Once s_2 is determined, the slope s_1 of the other segments of π^* is also uniquely determined by the condition that $O1$, A, AA, R, rise up from $O1$ to R. Since this condition applies to both π and π^* they both have the same slope on $u[O1, A]$ and $u[AA, R]$ as well. So $\pi^* = \pi$. Therefore, by the Facet Theorem, π is a facet.

Since λ, was arbitrarily chosen, this construction generates a one parameter family of facets. Representative members of the family shown in Figure 6 are $(O1, A, AA, R, O2)$, $(O1, B, BB, R, O2)$, $(O1, C, CC, R, O2)$, and of course the original mixed integer facet $(O1, P1, R, O2)$.

What is the geometric meaning of a one parameter family? Geometrically the $\pi(u)$ is the normal to a face of the polyhedron. In finite T-space if we move from one face to one of the adjacent faces, the normal takes a discontinuous jump. Here it changes

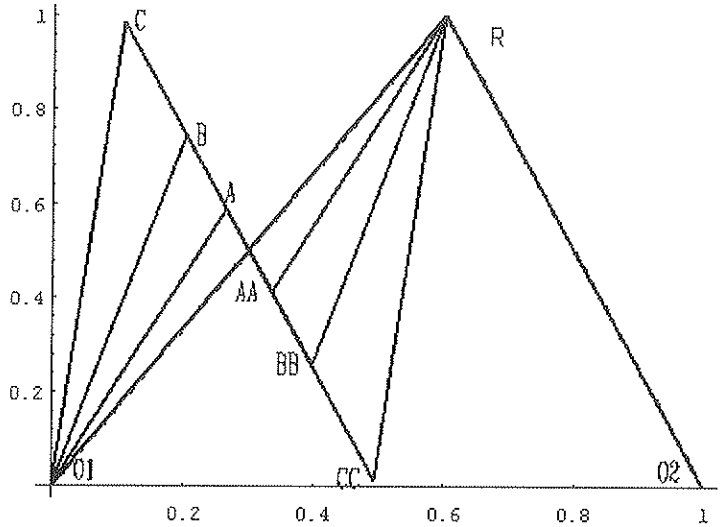


Fig. 6.

continuously. So as we change the one parameter in a one parameter family of facets we are moving steadily and continuously over the surface of the polyhedron. We are moving continuously through a sequence of normals $\pi(u)$ of adjacent facets.

3.2. More complicated two-slope facet families

The family introduced by Construction 1 is a two-slope family. The Gomory-Johnson Two Slope Theorem, see [2], simplifies the construction of much more complicated two slope families.

Theorem 5. *Gomory-Johnson Two Slope Theorem: If $\pi(u)$ is subadditive, minimal, and has only two slopes, then it is a facet.*

The use of the Interval Lemma and the Facet Theorem now enables us to give a fairly short proof of this theorem which is in Appendix B. This theorem opens up the possibility of a wide variety of facets.

We now proceed to the construction of a more complicated family of facets (Figure 7). Again the rhs element is u_0 and $R = (u_0, 1)$. The vector v_1 connects $O1$ to R with slope $1/u_0$ and the vector v_2 connects $O2$ to R with slope $1/((\eta(u_0) - 1)$.

Construction 2: Choose positive integers m and n and then locate the points $p_1 = (1/m)v_1$ and $q_1 = (1/n)v_2$. Through O ($O1$ in Figure 7) construct a line L^+ of positive slope s^+ greater than the slope of v_1. Through O ($O2$ in Figure 7) also construct a line L^- of negative slope s^- more negative than the slope of v_2.

Next through p_1 construct a line of slope s^-. Within the vertical strip containing $O, p1$, and R this line uniquely intersects L^+ at a point A_1 which determines the s^+

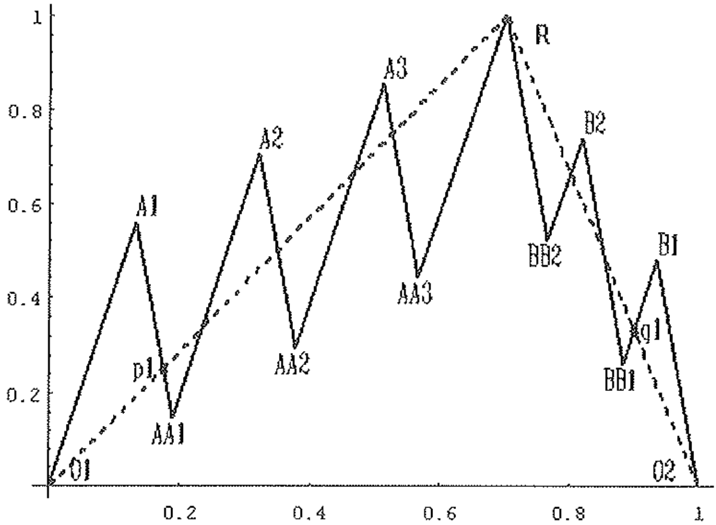


Fig. 7.

vector $(O1, A_1) = w_1$. Repeat this construction in the strip $R, q_1, O2$ by putting a line of slope s^- through $(O2)$ and a line of slope s^+ through q_1. These lines intersect at a point B_1 which determines the s^- vector $(O, B_1) = w_2$ $(O2, B1$ in Figure 7).

Since w_1 and p_1 lie on the same line of slope s^-, the vector $w_1 - p_1 = \lambda w_2$ for some λ. Choose $\lambda_2 = m/(m-1)\lambda$, and form the point AA_1 by adding $-\lambda_2 w_2$ to w_1. Then form A_2 by adding w_1 to AA_1, and then form AA_2 by adding $-\lambda_2 w_2$ to A_2. Continue adding w_1 and $-\lambda w_2$ by turns until $A_m = R$ is reached. (We will show below that this construction actually produces $A_m = R$).

Apply the same construction to the R, $O2$ side. There we define λ_1 by $\lambda_1 w_1 = (n/(n-1))(vw_2 - q_1)$, then add $\lambda_1 w_1$ to B_1 to form BB_1, then add w_2, to form B_2, and continue until $B_n = R$ is reached.

Theorem 6. Construction 2 Theorem: *For any integers m and n, and for s^+ and s^- as described in the construction, the π formed by the direct line segments of the line $O, A_1, AA_1, A_2, \ldots, ..A_{m-1}, AA_{m-1}R, BB_{n-1}, B_{n-1}, \ldots\ldots B_2$, BB_1, B_1, O, is a facet if and only if $0 \le \lambda_1 \le 1$ and $0 \le \lambda_2 \le 1$.*

Figure 7 is an example with $m = 4$ and $n = 3$. Construction 1 is the special case $n = 2, m = 1$.

Remark. π has only the two slopes s^+ and s^-, so the Gomory-Johnson Two-Slope Theorem can be applied. In proving that theorem, the Facet Theorem and the Interval Lemma have already done their work. So we need only prove minimality and subadditivity to establish that $\pi(u)$ is a facet. Doing that will be helped by some general remarks on minimality and subadditivity.

3.3. *Minimality*

Because of the Minimality Theorem, minimality follows from symmetry so we need only check whether the symmetry relation, $\pi(u) + \pi(u_0 - u) = 1$, holds for every u. However, for piecewise linear $\pi(u)$ the symmetry relation will clearly hold for every u if and only if it holds for the endpoints of each linear segment. Checking minimality for piecewise linear functions, therefore, is always reduced to the relatively easy task of checking symmetry for the u-values of the endpoints. In our diagrams checking symmetry for endpoints is simply showing that each endpoint can be paired with another endpoint such that the two add up to R, the rhs point.

3.4. *Subadditivity*

Once minimality is established, checking of subadditivity can also be reduced to checking endpoints. In fact only some endpoints need to be considered.

Theorem 7. *Subadditivity Checking Theorem: If $\pi(u_1) + \pi(u_2) \geq \pi(u_1 + u_2)$ whenever u_1 and u_2 are convex endpoints of π, then if $\pi(u)$ is piecewise linear and minimal, it is also subadditive.*

By a convex endpoint of $\pi(u)$ we mean an endpoint where both the line segments that end there, would lie below $\pi(u)$ if they were extended. A concave endpoint is one where the two line segments, if extended, would lie above $\pi(u)$. The points 0 and 1 are convex endpoints.

Proof. The proof consists of showing that if subadditivity is violated for u_1 and u_2, then there are convex endpoints u_1' and u_2' that also violate subadditivity and by at least as much as u_1 and u_2. Consider the cases below where $u_3 = u_1 + u_2$, $p_1 = p(u_1)$, $p_2 = p(u_2)$, and $p_3 = p(u_3)$.

Case 1. Both p_1 and p_2 are interior points or concave endpoints. If both are interior points, then we can slide p_1 a small amount δ to the right and p_2 the same amount to the left, or visa-versa. Then one of $\pi(u_1 + \delta) + \pi(u_2 - \delta)$ or $\pi(u_1 - \delta) + \pi(u_2 + \delta)$ will be smaller than or equal to the original sum $\pi(u_1) + \pi(u_2)$. If the slope of $\pi(u)$ at u_1 is smaller than the slope at u_2, then the first sum, $\pi(u_1 + \delta) + \pi(u_2 - \delta)$, will be smaller than $\pi(u_1) + \pi(u_2)$. The value of δ can be increased until one of $u_1 + \delta$, $u_2 - \delta$ is an endpoint. If the endpoint is a concave endpoint, then we can continue to increase δ. The reason is that moving to the right through a concave endpoint causes the slope to become smaller and moving to the left causes the slope to become larger.

If either or both of p_1, p_2 were initially concave endpoints, the same argument works because we need only have one or both conditions: the right-slope at u_1 is smaller than the left-slope at u_2, or the right-slope at u_2 is smaller than the left-slope at u_1.

Case 2. One of p_1, p_2 is a convex endpoint, the other is an interior point, and p_3 is an interior point or a convex endpoint: The same sliding argument works except that now

p_1 and p_3 both have the same δ added to them. Then one of $\pi(u_1+\delta)-\pi(u_3+\delta)$ or $\pi(u_1-\delta)-\pi(u_3-\delta)$ will be smaller than or equal to the original difference $\pi(u_1)-\pi(u_3)$. Increasing δ will either move $u_1+\delta$ to a convex endpoint or $u_3+\delta$ to a concave endpoint.

Case 3. One of p_1, p_2 is a convex endpoint, the other is an interior point, and p_3 is a concave endpoint: The minimality condition $\pi(u)+\pi(u_0-u)=1$ implies that complementary points u and u_0-u are either both interior points or one is a concave endpoint and the other is a convex endpoint. Thus $p_3 = p(u_0-u_3)$ is a convex endpoint. Suppose that p_1 is interior and p_2 is a convex endpoint. Using $\pi(u_1) = 1-\pi(u_0-u_1)$ and $\pi(u_3)=1-\pi(u_0-u_3)$, we have $\pi(u_1)+\pi(u_2)<\pi(u_3)$ implies $1-\pi(u_0-u_1)+\pi(u_2)<1-\pi(u_0-u_3)$, or $\pi(u_2)+\pi(u_0-u_3)<\pi(u_0-u_1)$. Both p_2 and $p_3 = p(u_0-u_3)$ are convex endpoints, completing the proof.

Now let us turn to the proof of the Construction 2 Theorem.

Proof of Minimality. Minimality will follow easily if we can show that the construction of the successive segments does lead to R. From the construction we have $A_m = mv_1-(m-1)\lambda_2 v_2$. Also by construction we have $p_1 = v_1-((m-1)/m)\lambda_2 w_2$. Multiplying this by m yields $R = mp_1 = mw_1-(m-1)\lambda_2 w_2 = A_m$.

Therefore $w_1-\lambda w_2+w_1-\lambda w_2\cdots+w_1-\lambda w_2+w_1 = R$ is a path of vectors summing to R. If we break this sum after a w_1, the terms on the left will add up to one of the A_i and the remaining terms, as can be seen by reversing their order, will add up to one of the AA_i. Since together they add up to R, we have proved their complementarity. Similarly arguments can be made if we break the sum after any of the terms $-\lambda w_2$, or if we deal with the corresponding sums from the other side, $[O_2, R]$. This establishes the complementarity of all the segment end points and hence, by the Minimality Theorem, establishes minimality.

Proof of Subadditivity. In proving subadditivity in Construction 2 we will often reach a stage, as we did in Construction 1, where we can show that the sum q of two points p_1 and p_2 on π is at a position that can be reached from a different point p on π by adding positive multiples of the upward pointing vectors w_1 and w_2 to p. An upward pointing vector simply means that the vector (u, h) has $h>0$. We would like to conclude that q can not be under π, so that the sum of p_1 and p_2 satisfies subadditivity.

We will be able to do that using the Separation Lemma.

Lemma 4. *Separation Lemma: Let π be piecewise linear with the slopes s of all segments satisfying $s^+\geq s\geq s^-$. Let w^+ be an upward pointing s-vector with slope s^+, and w^- be an upward pointing s-vector with slope s^-. Then if the point p lies on π, the point $q = p+w^++w^-$ can not lie below π.*

Proof. Choose any very small ε. Then, because of the steep slopes of w^+ and w^-, $p+\varepsilon w_1$ and $p+\varepsilon w_2$ both lie on or above π for any p on π. Let $w = \varepsilon w^+ + \varepsilon w^-$. Since the vectors $\varepsilon w_1, \varepsilon w_2$, are upward pointing, w is an upward pointing vector. Since all three vectors lie in a small planar region E around the origin O, ordinary planar geometry applies. We find the slope τ of the direct vector from O to w in E, and from the s-vector

w with slope τ. τ is either $\geq s+$ or $\leq s-$. Because of its slope, $p + w$ also lies on or above π for any p on π.

Now $1/\varepsilon = n + r$, where n is an integer and r is a real, $0 \leq r \leq 1$. Within E, vectors combine as in the plane, so for r, $r\varepsilon w^+ + r\varepsilon w^- = rw$. For any integer n we also have $n\varepsilon w^+ + n\varepsilon w^- = nw$. Adding the two equations we get $w^+ + w^- = (n+r)w$ and then, adding p to both sides, $q = p + (n + r)w$. This means we have a straight line segment $[p, q]$ with slope τ from p to q. If we were to traverse $[p, q]$ from p toward a q that was strictly below π, there would be a last point q^* on $[p, q]$ and on or above π. Of course q^* would be on π. Then we would have, for some small δ, q^* on π and $q^* + \delta w$ strictly under π. Since τ is the slope of $[p, q]$ this can not happen. This contradiction ends the proof.

We can now continue the proof of subadditivity. By the Subadditivity Checking Theorem, we need only consider pairs of points p_1 that are either (1) an AA_i and a BB_j, (2) an AA_i and an AA_j, or (3) a BB_i and a BB_j.

Case (1). A pair AA_i, BB_j. Assume that $i = j + r$ with $r \geq 0$. Then $AA_i + BB_j = r(w_1 - \lambda_2 w_2) + j(w_1 - \lambda_2 w_2) + j(w_2 - \lambda_1 w_1) = AA_r + j(1 - \lambda_1)w_1 + j(1 - \lambda_2)w_2 = AA_r + \tau_1 w_1 + \tau_2 w_2$. Since AA_r lies on $\pi(u)$, the Separation Lemma, with $w^+ = \tau_1 w_1$ and $w^- = \tau_2 w_2$ tells us that $AA_i + BB_j$ does not lie below $\pi(u)$.

Case (2a). A pair AA_i, AA_j with $i + j < m$. By the construction $AA_i + AA_j$ is AA_{i+j} which is on $\pi(u)$.

Case (2b). A pair AA_i, AA_j with $i + j \geq m$. We can continue the construction that ended at R by adding $-\lambda_2 w_2$ to R to obtain AA_m, and then add w_1 to that to obtain A_{m+1} and so on. Consider $i + j = m$. Since R is also $B_n = BB_{n-1} + w_2$, $AA_m = R - \lambda_2 w_2 = BB_{n-1} + w_2 - \lambda_2 w_2 = BB_n + (1 - \lambda_2)w_2$. Since BB_n lies on π, and $(1 - \lambda_2) \geq 0$ is one of the Construction Theorem conditions, the Separation Lemma tells us that AA_{m+1} can not lie below π.

It is important to note also that if $(1 - \lambda_2) < 0$, then $AA_m = BB_{n-1} + (1 - \lambda_2)w_2$ *does* lie under π. Therefore the condition $(1 - \lambda_2)$ is both necessary and sufficient for AA_m to avoid violating subadditivity.

Let us now advance to AA_{m+1} and compare that with BB_{n-2}. Since we start with $AA_m = BB_{n-1} + (1 - \lambda_2)w_2$, we can add $(w_1 - \lambda_2 w_2) + (w_2 - \lambda_1 w_1)$ to both sides to obtain $AA_{m+1} = BB_{n-2} + 2(1 - \lambda_2)w_2 + (1 - \lambda_1)w_1$. Since BB_{n-2} lies on π, we again use the Separation Lemma to assert that AA_{m+1} can not lie under π.

We can go on in this way until we reach the last BB which is $O2$. At that point we have $AA_{m+n+1} = O2 + \tau_1 w_1 + \tau_2 w_2$. At the next move we compare AA_{m+p+1} with AA_1. If we add AA_1 to both sides of this we get $AA_{m+p+2} = AA_1 + \tau_1 w_1 + \tau_2 w_2$ which establishes the subadditivity of elements adding to AA_{m+p+2}. We can continue around the perimeter of π in this way indefinitely.

Case (3). A pair BB_i, BB_j. We can see that this is exactly the same as Case (2) with the roles of the AA_i and BB_j reversed.

These cases exhaust the possibilities. By establishing subadditivity, we have proved the Two-Slope Construction Theorem.

Our empirical studies on finite dimensional T-space indicate that two-slope facet dominate among the largest faces, those that are hit most often in the shooting experiments which we describe in the section below on the Merit Index. However the experiments also turn up many facets with more slopes, so we will next examine three-slope facets.

3.5. Three slope facets

We now give the construction for a family of three-slope facets analogous to the two-slope families described above. Again we start with $O1$, $P1$, R, $P2$, $O2$, with right-hand side element $R = (u_0, 1)$ and halfway points $P1$ and $P2$ (Figure 8). We will refer to the direct vector $(O, P1)$ as v_1 and the direct vector $(O, P2)$ as v_2. For $\frac{1}{2} \geq \lambda_1 \geq 0$ we choose a point $\lambda_1 v_1$ on $[O, P1]$ which we denote by A. We denote its complement $R - \frac{1}{2}\lambda_1 v_1$ by B. Now we give a construction for a family of three slope facets.

Construction 3: Define the point A_1 by $A_1 = A + \lambda_2 v_2$, and the point B_1 by $B - \lambda_2 v_2$. Then π is defined by the direct line segments $O1$, A_1, A, B, B_1, R, $O2$ (Figure 8).

This construction gives us a π with slope s^+ on $[O, A_1]$ and slope s^- on $[A_1, A]$ and $[R, O]$. The segment $[A, B]$ has the third slope s_3. As in Construction 1, we limit the λ so that $\eta(A_1) > 0$ and $h(B_1) \geq 0$.

Theorem 8. *Three Slope Family Theorem: If π is constructed according to Construction 3, and if $\lambda_1 \leq 1/2$ and $0 \leq \lambda_2 \leq 1$, then π is a facet.*

Proof. We will follow our standard process of using the Interval Lemma to show that any π^* that satisfies the equalities satisfied by π must in fact be π itself. We start with the segment $[A, B]$.

Because $\lambda_1 \leq \frac{1}{2}$, there is a segment with A as its left end point and $P1 - A$ as the right endpoint. Applying the Interval Lemma to the segments $U = [A, P1 - A]$, $V = [P1 - A, P1]$ and $U + V = [P1, 2P1 - A] = [P1, B]$, we conclude that, since the segments cover $[A, B]$ and π^* is linear with the same slope in each one, and is required to be continuous, π^* must be linear with a single slope over $[A, B]$.

It remains to show that the slope of π^* is the same as the slope of π on $[A, B]$. However $[A, B]$ was constructed on the segment $[O, P1]$ passing through the origin O. Therefore π satisfies both $\pi(2A) = 2\pi(A)$ and $2\pi(P1) = \pi(R)$. These are two relations that π^* must also satisfy. However $\pi^*(2A) = 2\pi^*(A)$ implies that on $u[A, B]$ the linear π^* is part of a line that passes through the origin. In addition $\pi^*(P1) = (\frac{1}{2})\pi^*(R) = \frac{1}{2}$, so π^* passes through $P1$. However, in the vertical strip between O and R and containing $P1$, there is only one line passing through O and $P1$. So π and π^* must have the same slope.

We have now dealt with the segment having the third slope, what remains are the usual segments with slope s^+ and s^-. These are easily dealt with using the Interval Lemma as we did on the discussion of Construction 1.

We now know that the set $E(\pi)$ of all equalities has no solution other than π itself. If we can show that π is subadditive and minimal, we can apply the Facet Theorem.

Minimality: Construction 3 is easily seen to have produced a symmetric π, so π is minimal.

Subadditivity: Referring to the Subadditivity Checking Theorem, the only convex endpoints in π are the local minima at A and B_1. So the Subadditivity Checking Theorem applied here tells us that we need only check subadditivity for the three following cases.

Case 1: p_1 is A and p_2 is A. Since A is $\lambda_1 v_1$, $p_1 + p_2 = 2\lambda_1 v_1$. By the Separation Lemma this cannot be below π.

Case 2: p_1 is B_1 and p_2 is B_1. By the construction, $B_1 + B_1 = 2B - 2\lambda_2 v_2 = (B - A) + (B + A) - 2\lambda_2 v_2 = (B - A) + R - 2\lambda_2 v_2) = (B - A) + (2v_2 - 2\lambda_2 v_2) = (B - A) + 2(1 - \lambda_2)v_2$. However, $B - A$ lies on π and $2(1 - \lambda_2)v_2$ is a nonnegative multiple of v_2 so the Separation Lemma applies. Note that if $\lambda_2 > 1$, $2(1 - \lambda_2)v_2$ points inward from $B - A$ and $B_1 + B_2$ lies below π violating subadditivity. Therefore the condition $\lambda \leq 1$ is necessary.

Case 3: p_1 is B_1 and p_2 is A. Since by construction $B_1 = B - \lambda_2 v_2 = (R - A) - \lambda_2 v_2 = 2v_2 - A - \lambda_2 v_2 = -A + (2 - \lambda_2)v_2$, adding B_1 to A yields $(2 - \lambda_2)v_2$. By the Separation Lemma this can not be below π.

The three cases together prove subadditivity and end the proof of the Three Slope Family Theorem.

Another three-slope facet, $\pi 3(u)$, is shown as the first figure in Figure 10. $\pi 3(u)$ first emerged by being the most often hit three-slope facet in the shooting experiments.

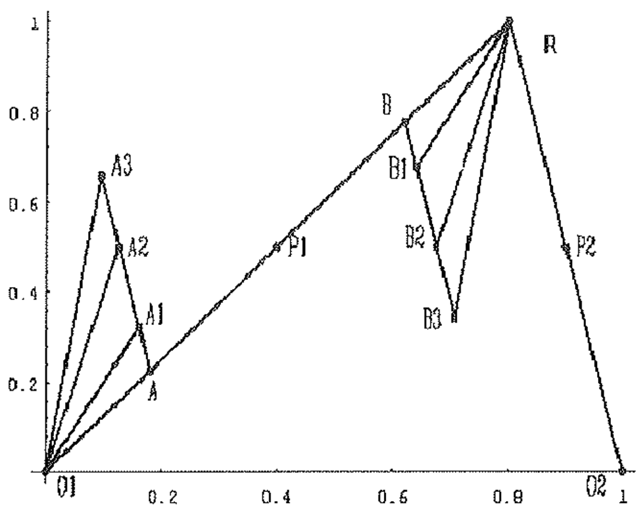


Fig. 8.

Its construction is described in Appendix C where it is shown that $\pi 3(u)$ belongs to a 1-parameter family of facets.

3.6. Mappings

Since our group G is the line (Mod 1), the only automorphism is multiplying every element by -1. However we do have homomorphisms, multiplying each element of G by an integer m is a homomorphism that sends the group elements represented by $1/m, 2/m \ldots$ into $\underline{0}$. These mappings then become the source of still more facets.

Theorem 9. *Mapping Theorem: If $\pi(u)$ is a facet with rhs element u_0, then $\pi_m(u) = \pi(mu)$, where m is any integer, is a facet. The rhs element v_0 of $\pi_m(u)$ can be any one of the m elements v_0 satisfying $mv_0 = u_0$.*

Geometrically, since $\pi_m(u+u(1/m)) = \pi(mu+mu(1/m)) = \pi(mu+\underline{0}) = \pi_m(u)$, $\pi_m(u)$ looks like m copies of $\pi(u)$ stacked next to each other in m successive intervals of length $1/m$. (Figure).

The proof of the Mapping Theorem is given in Appendix D.

Corollary 1. *Mapping Corollary: If $\pi(u)$ is constructed by a process that gives a facet $\pi(u)$ for a range of rhs $u_0 \in [a, b]$, then $\pi_m(u)$ is a facet of the Master Polyhedron $P(G, u_0)$ whenever $mu_0 \in [a, b]$.*

Proof. To make a π_m with rhs u_0, first construct π with rhs mu_0. Then define $\pi_m(u) = \pi(mu)$.

As an example, in Figure 9, a $\pi_3(u)$ facet is created with rhs 0.7, starting from a $\pi(u)$ with rhs 0.1. For facets from constructions like Constructions 1,2, and 3, which are valid for all rhs, this means that they and all their π_m appear in every $P(G, u_0)$.

The Mapping Theorem therefore gives us yet another way to create facets. But it also connects facets and polyhedra $P(G, u_0)$, with different rhs elements u_0.

3.7. Moving from one facet to another

Since, for a fixed rhs element u_0, there is only one polyhedron, we should be able to find a path from one facet to any other facet through a series of intermediate facets. This should be true even if the facets look very different.

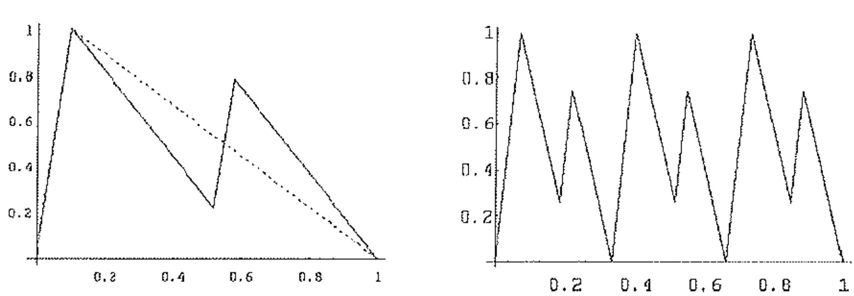


Fig. 9.

We will illustrate this by showing a sequence of facets connecting the three slope facet $\pi 3(u)$ and the standard mixed integer facet replicated twice.

In Figure 10 we start with $\pi 3(u)$. We end in Figure 10 with the figure that describes the last element of the family created by Construction 1. In between we go through some different π's, all of which, with some effort, can be shown to be facets. The Construction 1 family then provides the remainder of the path to the mixed integer facet.

What is different about this infinite dimensional polyhedron is that the movement from $\pi 3(u)$ to the mixed integer cut is along a continuum of facets. We are not moving along an edge to an adjacent facet, but along a curve where each function on the curve is a facet.

4. A merit index and an intersection index

We have shown we can create many different facets. Are some bigger and better than others? What can we say here about size or goodness of the various facets?

One natural figure of merit would be to count the number of paths lying on a facet and take that as a figure of merit. Since the paths lie on a regular grid, more paths will usually mean a bigger facet. One can certainly argue against this by pointing out that a path count would need to be corrected for the alignment of the facet with the regular array of integer points in T-space, or other factors, however, for the moment, we are only at the very crude beginnings of this kind of thinking. The merit index we introduce now is motivated by this idea of counting paths on a facet. However, it also has empirical support, as we will see.

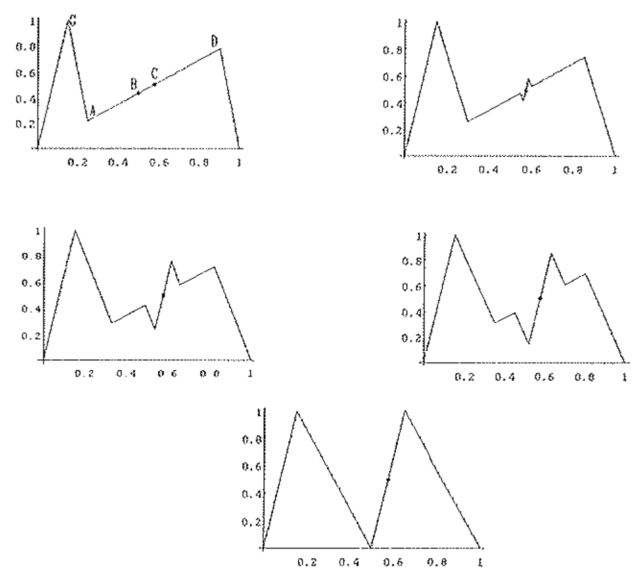


Fig. 10.

4.1. *Merit index*

Consider the unit square C_2 in two dimensions with the x and y taken (Mod 1). Let p be any point (u_1, u_2) in C_2. Then we define the Merit Index $MI(\pi)$ as follows:

Definition 1. *Merit Index Definition: $MI(\pi)$ is twice the area of the set of points $p = (u_1, u_2)$ in C_2 for which $\pi(u_1) + \pi(u_2) = \pi(u_1 + u_2)$ holds.*

Of course it is the area that matters. We take twice the area for convenience as the calculations always turn out simpler with the factor of two, and the largest possible MI becomes 1.0 rather than 0.5.

But how does this definition relate to counting paths? To answer that question let us return to the discussion of paths in Section 1. If we take any path p lying on a face π, we can reduce its length (the total number of group elements in the path) by choosing any pair $(u_1, u_2) \in p$ and replacing them by the single element $u_1 + u_2$. The Pairs Lemma tells us that for any such pair we have $\pi(u_1) + \pi(u_2) = \pi(u_1 + u_2)$ so we have a new shorter path. We can continue doing this until there are only two elements in the path, a complementary pair of elements u and $u_0 - u$ with $\pi(u) + \pi(u_0 - u) = 1$.

If we look at this process in reverse, it shows that every path on a facet is built up from some complementary pair by substituting pairs (u_1, u_2) for single elements $(u_1 + u_2)$ whenever we have $\pi(u_1) + \pi(u_2) = \pi(u_1 + u_2)$. The more such inequalities, the more paths on the face, probably in a very non-linear way. What the merit index measures is the fraction of all possible pairs that are equalities. The larger the index, the more equalities. The more equalities, the more paths.

The second rationale for this index is empirical. It is based on the correlation of the merit index with the results of the shooting experiments discussed in [4]. These experiments were conducted on finite Corner Polyhedra. These Corner Polyhedra are the same concept as the T-space described here, but the mapping of columns is into a finite group instead of into the real line (Mod 1). This results in a finite T-space with as many dimensions as there are non-zero group elements. These finite Corner Polyhedra are the ones discussed in [1].

In the shooting experiments a random direction was chosen at the origin in finite T-space and that direction was pursued until it hit a facet of the Master Polyhedron. That facet was then recorded. The technique used, Gomory's Shooting Theorem, allowed the experiment to be conducted without knowing in advance any of the facets of the polyhedron. The facets were discovered by being hit. The program actually written and used by Evans and Johnson made it possible for them to compute 10,000 hits on each Master Polyhedron used in the experiment, although fewer shots were used on the smaller groups. At the end of each experiment, the number of hits on each facet that was hit was recorded. No knowledge was produced by the shooting experiments about the facets not hit.

What Figure 11 shows is the correlation between the merit index and the number of hits in these shooting experiments. There is a dot for each facet hit. The percent of all shots that hit that facet is plotted vertically, and the merit index of that facet is plotted horizontally. The cyclic group used was the integers, (Mod 17), rhs element 16. A strong

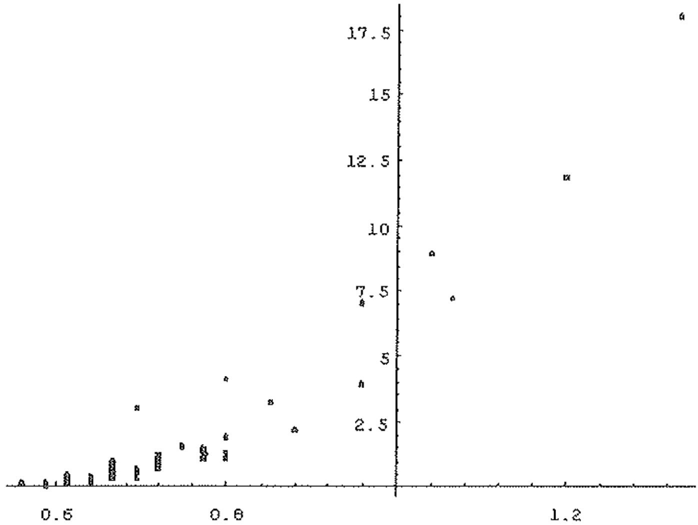


Fig. 11.

non-linear correlation can be seen between the faces that are hit most often, which also has something to do with size, and the merit index.

4.2. *Computing some merit indices for facets*

If we have some face $\pi(u)$ it is quite straightforward to compute the merit index. This can be done exactly for some simple facets and it is very straightforward to get a good numerical approximation for any $\pi(u)$.

The merit index (MI) of the mixed integer facet is $MI = \eta(u_0)^2 + (1 - \eta(u_0))^2$. For this facet the merit index depends strongly on the choice of u_0 and ranges from 0.5 in the middle ($u_0 = 0.5$) to a value that approaches 1 as u_0 approaches 0 (or 1). We can sometimes exploit this dependence on right-hand side by using the Merit Index Mapping Theorem.

Theorem 10. *Merit Index Mapping Theorem: Mappings $\pi^*(u) = \pi(nu)$ of a face always have the same merit index as the original $\pi(u)$.*

This theorem seems intuitively correct and the proof is not difficult.

Of course the facet $\pi^*(u)$ created by the mapping will have a rhs element that is different from the rhs element of the original $\pi(u)$. Mapping thus makes it possible to move the mapping of a family member with higher merit index into a position where the family member has lower merit index. For example, the mixed integer cut has $MI = 0.52$ for rhs element 0.6. We can form a new facet with rhs 0.6 by using as our $\pi(u)$ the mixed integer cut with rhs element 0.8, and generating from it the new facet $\pi^*(u) = \pi(2u)$. Since $\pi(u)$ has $MI = 0.68$, the new facet π^* also has $MI = 0.68$. So we have created a new facet with the desired rhs 0.6, but with a larger merit index.

In Figure 12 we also show a minimal, but non-facet, inequality. It has a low MI namely $MI = 0.1075$ as we would expect. There is always a facet that satisfies all the equalities of a non-facet and others in addition as well, and therefore we should expect a lower score on non-facets. In particular, the mixed integer facet dominates it and has a larger $MI = 0.68$.

So far we have discussed the merit index for different facets or for the same facets with different rhs element. There is also the question of the variation of the merit index within a family of facets all having the same rhs. Figure 13, based on Construction 1, gives an example. In the merit curve in Figure 13, there is a discontinuous drop from $MI = 0.52$ to $MI = 0.43$ as the λ of Construction 1 becomes non-zero. Then, as λ grows, the merit index first decreases to a low value slightly less than 0.38, and then rises smoothly to 0.68. This example shows that the choice of the parameter value can make a difference of more than 0.3 in the Merit Index.

4.3. The intersection index

It would be good to know about two faces how many of the paths on each face they have in common. It gives information for how far apart they are on the polyhedron, or perhaps more accurately, how large their intersection is. For practical cutting planes one might not want to add two cutting planes that have almost the same set of paths on them.

Note that in T-space the intersection of two facets is never empty. The two element paths $\{u, (u - u_0)\}$ lie on every facet since every facet satisfies the minimality condition $\pi(u) + \pi(u - u_0)) = \pi(u_0) = 1$.

Again it is difficult to obtain the actual paths. However, as we remarked above, all paths can be built up by starting with a complementary pair $\{u, (u - u_0)\}$, and then using equalities to build longer paths. So if we take the set of equalities that two facets have in common, all the paths that those equalities can generate from the complementary pairs

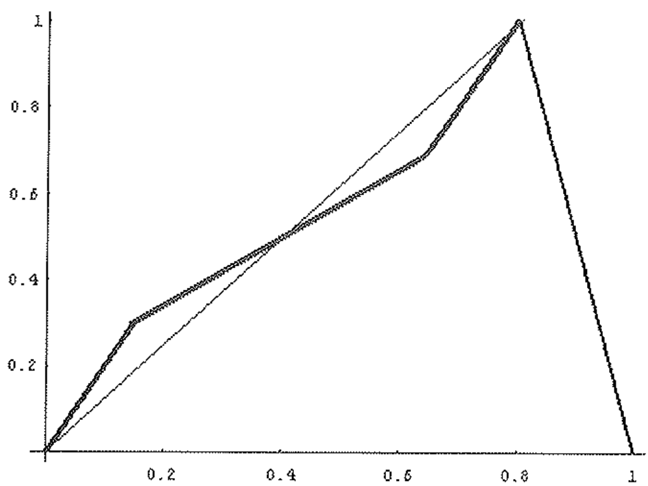
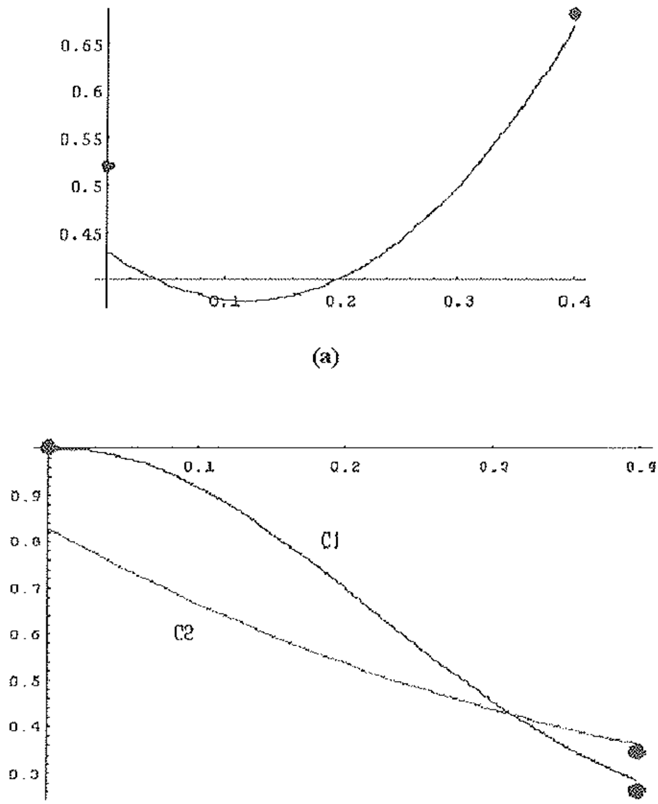


Fig. 12.



(b)

Fig. 13.

will be paths lying on both faces. Again we can expect the number of common equalities to be positively but non-linearly correlated to the number of common paths.

With this in mind we return to the unit square C_2 in two dimensions used in the Merit Index. Again let p be any point (u_1, u_2) in C_2. For any facet π, let $S(\pi)$ be the set of points p in C_2 for which $\pi(u_1) + \pi(u_2) = \pi(u_1 + u_2)$ holds. We define the Intersection Index $II(\pi_1, \pi_2)$ of two facets π_1 and π_2 by:

Definition 2. *Intersection Index Definition: $II(\pi_1, \pi_2)$ is the area of $S(\pi_1) \cap S(\pi_2)$ divided by the area of $S(\pi_1)$.*

Note that $II(\pi_1, \pi_2)$ is not symmetric in π_1 and π_2. $II(\pi_1, \pi_2)$ and $II(\pi_2, \pi_1)$ are, however, closely related through $II(\pi_1, \pi_2)/II(\pi_2, \pi_1) = MI(\pi_2)/MI(\pi_2)$.

Figure 13(a) shows a plot of the merit index for the family of Construction 1 where $u_0 = 0.6$. Figure 14 shows in black the values of u_1, u_2 for which $\pi(u_1) + \pi(u_2) = \pi(u_1 + u_2)$ with u_1 on the horizontal axis and u_2 on the vertical axis. Figure 14(a) is for $\lambda = 0$, i.e. the mixed integer facet. Figure 14(b) is for $\lambda = \epsilon > 0$ and shows the reason for the discontinuity in the merit index seen in Figure 13(a). The merit index drops from 0.52 to a value just larger than 0.43. However, the two triangles including the points (0.3,1) and (1,0.3) increase in size as λ increases. These two triangles are not in the areas where equality holds for the mixed-integer facet. There is a small discontinuity at $\lambda = 1$ when the small triangle above (0.5,0.5) in Figure 14(d) comes into the equality region. This triangle results from the fact that, e.g., $u_1 = 0.5$ and $u_2 = 0.55$ have $\pi(u_1) + \pi(u_2) = 0 + 0.5 = \pi(u_1 + u_2) = \pi(0.05) = 0.5$. Thus, as λ approaches 1, the merit index jumps from 0.67 to 0.68. In Figure 13(b) we show the intersection indices for the family of Construction 1. Both $II(\pi_1, \pi_2)$ and $II(\pi_2, \pi_1)$ are shown. In Figure 13(b), π_1 is the family member created by increasing λ, while π_2 does not change and is always the facet $\lambda = 0$, i.e. the mixed integer facet. In the Figure 13(b) C_1 marks the curve of $II(\pi_1, \pi_2)$ values and C_2 marks the curve of $II(\pi_2, \pi_1)$ values. The horizontal axis gives the values of a parameter which ranges from 0 to 0.4 as λ ranges from 0 to 1.

For very small values of the parameter, when π_1 is first separating from π_2, there is a sudden drop in C_2 but C_1 moves continuously and so remains near 1 (see Figure 14 (a) and (b)). Since in this parameter range we also have a discontinuous change in the merit index of π_1, it appears that we are moving from a larger face abruptly to a somewhat smaller face. The fact that C_1 remains near 1 shows that the smaller face π_1 still has almost all of its paths in common with the larger one, but the drop in C_2 shows that π_2 now has a significant collection of paths no longer in common with π_1.

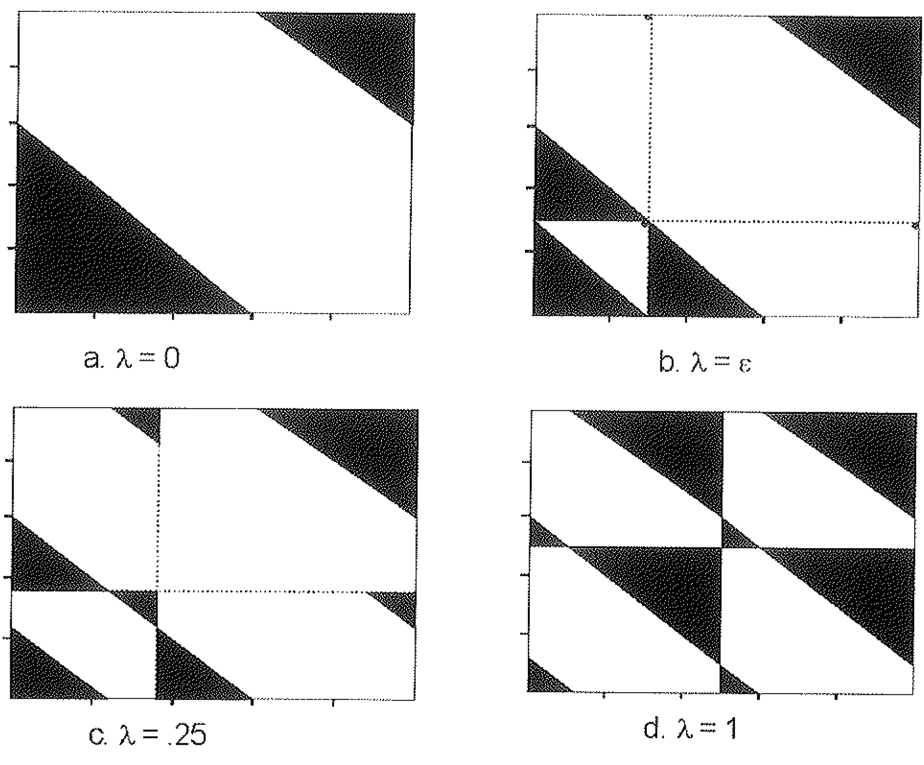


Fig. 14. Two-slope Family

A more extreme example can be obtained from the three slope family of Construction 3 (Figure 8), with $u_0 = 0.8$. Choose a family member with $\lambda_2 = \varepsilon$ very near 0. This means that in Figure 8 we choose an A_1 arbitrarily close to A, and a B_1 arbitrarily close to B.

The merit index of the mixed integer facet, $\lambda_2 = 0$, is 0.68 but the merit index with $\lambda_2 = \varepsilon$ is almost exactly 0.20 (see Figure 15(b)). This very large change occurs because in π_2 the segments from the origin to A_1 and from B_1 to R are no longer in line with the segment AR, while the new equalities that π_1 has gained from the arbitrarily small new segments AA_1 and BB_1 can make only an arbitrarily small contribution to the merit index.

For the same reasons the intersection index of π_1 with π_2 remains almost 1, but the index $II(\pi_2, \pi_1)$ becomes 0.20/0.68 which is roughly 0.29. This suggests that π_1 has become much smaller but remains firmly attached to π_2, while the larger π_2 has lots of paths that have nothing to do with π_1. Since the λ_1 in Construction 3 can be chosen freely from the range $0 \leq \lambda \leq 1$, there is not one, but rather a whole family of these much smaller three-slope facets firmly attached to different parts of the basic mixed integer facet. Despite the small merit index, all of these 3-slope facets are facets. They include

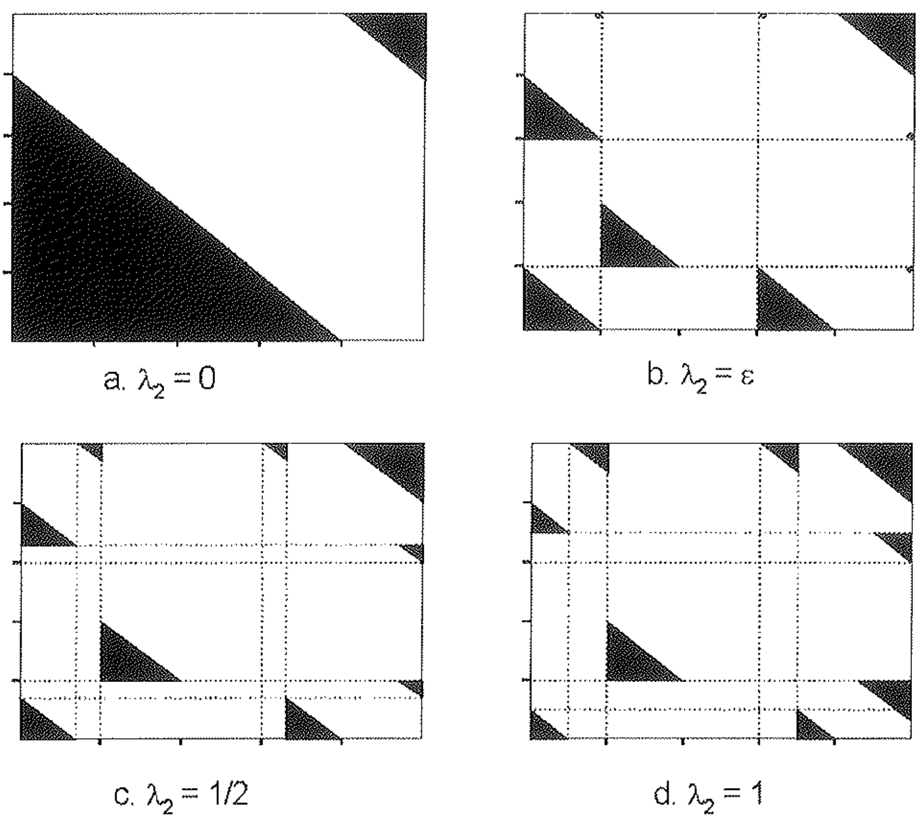


Fig. 15. Construction 3 Family

four triangles where equality holds and does not hold for the mixed integer facets. These four triangles (Figure 15(b), (c), (d)) have corners at (.2,1), (.6,1), (1,.6), and (1,.2) and come from the line OP2 with the line segments A, $A1$ and $B1$, B.

With the merit index and the intersection index we have done some very preliminary exploring of the polyhedron associated with G. Exploring the polyhedron reminds us that cutting planes are not isolated things, e.g. "the Gomory cut," rather they are all facets of this huge polyhedron and it is possible to move from one to another, no matter how different they look, through a continuous succession of intermediate facets, and these too are facets and give cuts.

Since all these cuts are facets, they are all independent. No facet (or cut) is implied by any other facet or by any combination of other facets (or cuts).

5. Two dimensions (or more)

A similar theory can be developed for mappings into a two dimensional group. The group is the unit square in two dimensions with the x and y taken (Mod 1). The mapping consists of taking two rows rather than one and sending the pair of fractional parts associated with each non-basic variable into the unit square. The π now has to be a function of two arguments (f_1, f_2) the fractional parts associated with one non-basic variable in the two rows. Again the continuous variables are linked to the slopes at the corners, but now these slopes are more varied and more accurately reflect the role of the continuous variables.

Here is a minimal inequality in this setup (Figure 16). However it seems likely that functions like the one pictured in Figure 16 with a single central peak can be minimal but can not be facets. Figure 17 shows what we believe to be a facet. These tentative statements reflect the fact that with the exception of [5] very little work has been done in this area. It looks like a promising area both for theory and for practice.

There are reasons to think that such inequalities would be stronger since they deal with the properties of two rows, not one. They can also much more accurately reflect the structure of the continuous variables. There is also an argument that says that not too much can be expected from more rows in the case of a pure integer problem with integer starting data. For in that case the fractional parts of a single row can represent the structure of the fractional parts of all rows.

6. Challenges

The exploration of the these polyhedra has barely begun. There is much to be done. We mention some of those things here.

6.1. A facet conjecture

Every single facet that we have dealt with has been piecewise linear. Are there facets that are not piecewise linear? We have not been able to construct a curved π that is a facet, so we state the following conjecture:

Conjecture 1. Facet Conjecture: If π is a facet, π is piecewise linear.

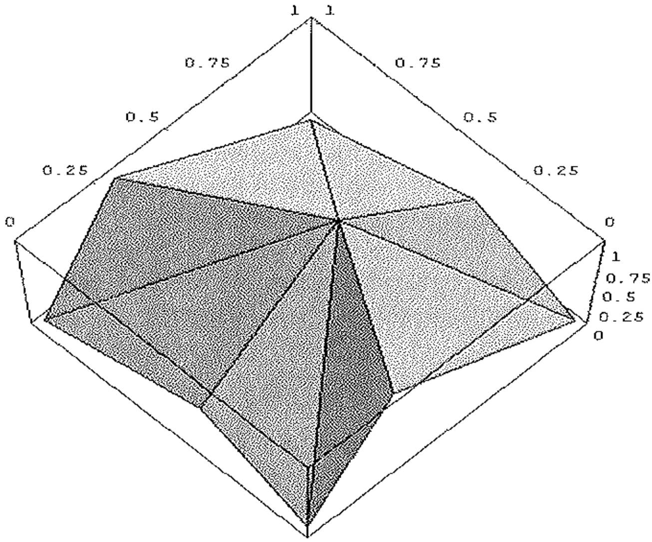

1
1
0.75
0.75
0.5
0.5
0.25
0.25
0
0
1
0.75
0.5
0.25
0

Fig. 16.

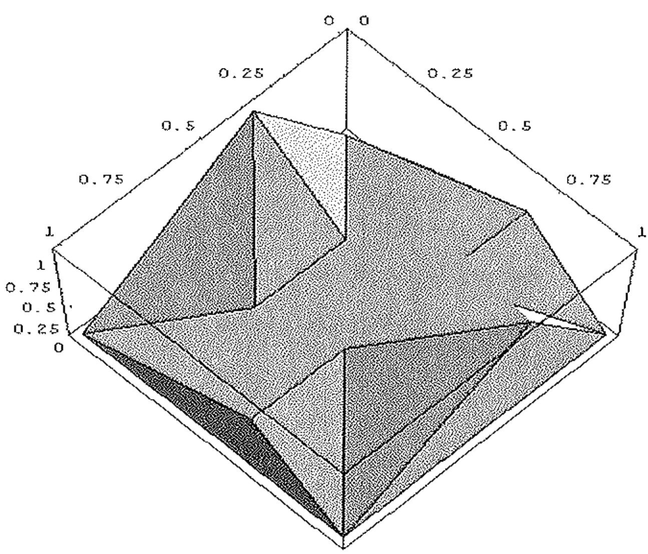

0
0
0.25
0.25
0.5
0.5
0.75
0.75
1
1
1
0.75
0.5
0.25
0

Fig. 17.

6.2. *Better methods*

While the Interval Lemma and Facet Theorem and the various lemmas about subadditivity give us useful tools, their application to generate facets is still somewhat ad hoc. Also we don't have general theorems for three or more slope facets analogous to the Gomory-Johnson Two Slope Theorem.

6.3. *Smaller facets*

There has been an inherent bias in much of what we have done so far. Because of our interest in cutting planes, we have discussed mainly large facets, families of two-slope facets or three-slope facets. What about much smaller facets? It seems extremely likely that there are much smaller more complex facets all around the big ones. Can we develop a better understanding of this and through that an understanding of the fundamental limits to what can be accomplished by cutting planes?

6.4. *More dimensions*

Although the equivalent of the Interval Lemma and the Facet Theorem seem to work with group of dimension higher than 1, and these were the methods used to produce the examples in Section 5, we know very little about 2 or more dimensions. Or what about m dimensions where m is the dimension of the original columns?

6.5. *Using the cutting planes*

We have an endless supply of facets. All of them are cutting planes and none are implied by the others. How are they to be used? One possibility is to keep a list of the n highest ranked facets and just use them all, all the time. The list could be developed from the merit index or from experiments, while the number on the list is determined by practical computational considerations.

In the opposite direction is developing methods to exploit the wealth of cutting planes to form cutting planes that reflect what variables are actually present or the actual numerical structure of the row being worked. For example, $\pi(u)$ that have small values for certain integer variables can be developed. These are good cuts for those variables. In addition, we also want cuts that have small s^+ and s^- for the continuous variables. The Gomory mixed integer cut has the smallest s^+ and s^-.

Can we transform *families* of facets and carry them along in integer programming rather than dealing with one facet at a time?

6.6. *A better index*

While the merit index helps, it does not take explicitly into account the effect of steep slopes s^+ and s^- at the origin which weaken cuts in mixed integer programming. The polyhedra $P_{=}^{+}(U, u_0)$ used in [2] are analogous to the polyhedra we have been dealing with, but they explicitly take into account the role of the continuous variables. Perhaps the theory and the indices can be redone in this space.

6.7. Beyond cutting planes

Perhaps we should shift our thinking away from the notion of cutting plane and think instead of adding, not a cutting plane, but rather an approximation to the corner polyhedron. We could iterate on successive approximate Corner Polyhedra, rather than adding cutting planes and scooting off to a new corner. The practical tendency in computation to add many cutting planes at once is a step in that direction, and with this theory we are now in a position to use in actual computations a very large coherent set of inequalities that approximates the corner polyhedron.

We hope that many will be interested in exploring these and still other possibilities.

Appendices

Appendix A – Valid Inequalities and cutting planes – Non-integer variables

We now extend this analysis to include non-integer variables along with the integer ones.

If t_i is a non-integer variable we still obtain the inequality in (4), but with non-integer t_i we can not take the next step and get the inequality in (5) which is the cutting plane. Since in our discussion of this case we will go from group element to real number and back we need some notation: If u is a group element, then the corresponding element in the real interval $0 \leq x < 1$ we denote by $\eta(u)$. Going the other direction if we have a real number x then the corresponding group element is $g(x)$.

We will discuss only the simplest case: using the mapping χ that maps the vector $c_i t_i$ into the group element $a = c_{i,k} t_i$ (Mod 1). For the non-integer variable t_i we reason as follows: Denote by a/n the group element $g(R(a)/n)$, we have from subadditivity for any integer n:

$$\pi(a) \leq n\pi(a/n), \text{ then for sufficiently large } n\pi(a/n) \leq (s^{+} + \varepsilon)\eta(a/n)$$

$$= (s^{+} + \varepsilon)\eta(a)/n \text{ so } \pi(a) \leq (s^{+} + \varepsilon)\eta(a) \qquad \text{(A1)}$$

In (6) s^{+} is lim sup $\pi(u)/\eta(u)$, as u approaches $\underline{0}$ from the right. We assume π is piecewise linear, with s^{+} denoting the slope of the first segment of $\pi(u)$. Since (6) holds for all n, we can make ε arbitrarily small. So:

$\pi(a) \leq n\pi(a/n) \leq s^{+}\eta(a)$ which implies when $c_{i,k}$ is positive that

$$\pi(a) \leq s^{+}\eta(a) \leq s^{+}c_{i,k}t_i. \qquad \text{(A2)}$$

We can repeat this argument when s^{-} is the (negative) slope of the last segment of $\pi(u)$ to obtain:

$$\pi(a) \leq s^{-}c_{i,k}t_i. \qquad \text{(A3)}$$

Putting (A2), (A3), and (5) together gives us

$$\textstyle\sum_{i\varepsilon I}\pi(F(c_{i,k}))t_i+\sum_{i\varepsilon T^+}(s^+c_{i,k})t_i+\sum_{i\varepsilon T^-}(s^-c_{i,k})t_i\geq 1. \qquad \text{(A4)}$$

In (A4) I is the set of variables that are integer, T^+ is the set of variables that are non-integer and have $c_{i,k}>0$ and T^- is the set of variables that are non-integer and have $c_{i,k}<0$. s^+ and s^- are the slopes of the first and last linear segments of $\pi(u)$. So we have proved:

Theorem 11. *Mixed Integer Cutting Plane Theorem: If $\pi(u)$ is subadditive and piecewise linear, then the inequality (A4) on the non-basic variables t is a valid cutting plane for the mixed integer programming problem whose constraints are $Bx+Nt=b$.*

This theorem justifies the complete process of cutting plane generation described at the beginning of Section 1 above.

Appendix B – Proof of the Gomory-Johnson two slope theorem

Theorem 12. *Gomory-Johnson Two Slope Theorem: If $\pi(u)$ is subadditive, minimal, and has only two slopes, then it is a facet.*

Proof. If we have a $\pi(u)$ satisfying the conditions of the theorem, let us take any one of its line segments $S=[p_1,p_2]$ and the corresponding group interval $I=[u_1,u_2]$. Assume that S has slope s^+, the right hand slope at $\underline{0}$. For the interval U of the Interval Lemma, choose $[\underline{0},\varepsilon]$ where ε is smaller than the length of any interval of S. For the interval V choose $[u_1,u_2-\varepsilon]$. $U+V$ then becomes I. Since for these intervals we have $\pi(u)+\pi(v)=\pi(u+v)$, the Interval Lemma asserts that any other $\pi^*(u)$ satisfying them must be linear on U, V, and $U+V$ and have the same slope s_1^* on all three segments. Repeating this reasoning on each s^+ segment shows that π^* will be linear on all of them and that they all have slope s_1^* because they all have the slope that π^* has on $U=[\underline{0},\varepsilon]$. Similarly we can show for all segments of $\pi(u)$ with slope s^-, that $\pi^*(u)$ would have to be a straight line on these with slope s_2^*.

If we can show that s_1^* and s_2^* must in fact be the same values as s^+ and s^-, we will have shown that any $\pi^*(u)$ satisfying the same equations as $\pi(u)$ is in fact identical with $\pi(u)$, so that $\pi(u)$ must be a facet. To do this we use the same basic conditions as in Construction 1, namely that $\pi^*(\underline{0})=0$, and $\pi^*(u_0)=1$. The analysis is slightly more complicated, but the result is provided by the Conditions Lemma.

Lemma 5. *Conditions Lemma: Consider two two-dimensional non-zero row vectors X_1 and X_2 that point into the first quadrant. Then the equations $s_1X_1+s_2X_2=X_3$, where s_1 and s_2 are scalars and X_3 points into the second or fourth quadrant, have either one solution or none.*

Proof. If the rank of (X_1,X_2) is 2, there is exactly one solution. If the rank is one, X_1 and X_2 are scalar multiples of each other so that $s_1X_1+s_2X_2=s_3X_1$. But s_3X_1 can only point into the first or third quadrants, so there is no solution. Since the vectors are non-zero, rank 0 is not possible.

Now to determine s_1^* and s_2^*, we use the equations $s_1 X_1 + s_2 X_2 = X_3$ where $X_1 = (L_1, L_2)$ with L_1 the total length of intervals of slope s^+ to the left of u_0 in $\pi(u)$ and L_2 the total length of intervals of slope s^- to the right. The components of X_2 are obtained in the same way using s^- in place of s^+. For X_3 we use $(1, -1)$. The equations then represent the condition that the sum of the rise in each segment to the left of u_0 should equal 1, and the sum of the rises to the right should equal -1. The equations satisfy the conditions of the Conditions Lemma. Since $\pi(u)$ does satisfy these equations, the Conditions Lemma tells us there is exactly one solution. So s_1^* and s_2^* must in fact be the same values as $s+$ and s Applying the Facet Theorem ends the proof.

Appendix C – Construction of $\pi 3(u)$

To construct $\pi 3$ (see the first figure in Figure 10): (1) Select a rhs element u_0 with $\eta(u_0)$ in the open interval [0, 0.25]. (2) Find the midpoint C of the line segment $[O, R]$ with slope $1/(1 + \eta(u_0))$. On this $[O, R]$, which is not the direct segment $[O, R]$, (a) find the midpoint $C = (0.5(1 + \eta(u_0)), 0.5)$, (b) select a point A located to the right of u_0 but less than halfway from O to C. Because u_0 is in $[0, 0.25]$ and C is to the right of 0.5, there always is such an A. $\pi 3(u)$ is then defined as the successive direct segments O, R, A, C, D, O, where D is the complementary point to A. Since there is a range of possible A's, this process does not produce one, but rather a family of facets, depending on the choice of A.

Appendix D – Proof of mapping theorem

Theorem 13. *Mapping Theorem: If $\pi(u)$ is a facet with rhs element u_0, then $\pi_m(u) = \pi(mu)$, where m is any integer, is a facet. The rhs element v_0 of $\pi_m(u)$ can be any one of the m elements v_0 satisfying $mv_0 = u_0$.*

Proof. There is a very tight relation between π_m and π. Because of the sequence $\pi_m(u_1) + \pi_m(u_2) = \pi(mu_1 + mu_2) = \beta + \pi(m(u_1 + u_2)) = \beta + \pi_m(u_1 + u_2)$ the equalities ($\beta = 0$) of π translate into equalities of π_m as do the inequalities where $\beta > 0$. The subadditivity of π_m follows from the subadditivity of π because the β that appears in any sequence will be non-negative because of the subadditivity of π.

To prove uniqueness we first show that any π^* satisfying the equalities that π_m does, must, like π_m, be made up of a single pattern that is repeated in each interval of length $1/m$. If we use $u(1/m)$ to indicate the group element u whose $u) = 1/m$, then $\pi_m(u(1/m)) = \pi(mu(1/m)) = \pi(\underline{0}) = 0$. Therefore π_m satisfies the equality $\pi_m(u(1/m)) + \pi_m(u(1/m)) = \pi_m(u(1/m))$. π^* must also π satisfy this equality which implies that $\pi^*(u(1/m)) = 0$. Next, since $\pi_m(u) + \pi_m(u(1/m) = \pi_m(u + u(1/m))$, π^* must also satisfy $\pi^*(u) = \pi^*(u + u(1/m))$. Therefore π^* too is made up of a single pattern that is repeated in each interval of length $1/m$.

There is an addition preserving $1 - 1$ mapping φ between the first of these intervals, I regarded as a group (Mod 1/n) and the group G. We define $\varphi(u) = mu$ for any element u in I.

Therefore any π defined on I produces a new π, π_I defined on G by $(u) = \pi(\varphi^{-1}(u))$. We can verify that, as one would expect, that if we use as our π, πm, we get $\pi(u) = \pi_m(\varphi^{-1}(u) = \pi(m\varphi^{-1}(u)) = \pi(u)$.

Since π^* is now repetitive, like π_m, we can use π^* to produce a function $\pi^*(\varphi^{-1}(u))$ which has the same tight relation to π^* as π does to π_m. The same sequence of equalities written above with β holds with this new pair. In particular if π^* satisfied some triple $\pi^*(u_1) + \pi^*(u_2) = \pi^*(u_1 + u_2)$ which πm did not, there would be a triple that $\pi^*(\varphi^{-1}(u))$ satisfies that π did not. But this contradicts the assumption that π was a facet. This contradiction ends the proof.

References

[1] Gomory, R.E.: Some Polyhedra Related to Combinatorial Problems. Journal of Linear Algebra and Its Applications **2**(4), 451–558 (1969)

[2] Gomory, R.E., Johnson, E.L.: Some Continuous Functions Related to Corner Polyhedra, Part I. Mathematical Programming **3**(1), North-Holland, 23–85 (1972)

[3] Gomory, R.E., Johnson, E.L.: Some Continuous Functions Related to Corner Polyhedra, Part II. Mathematical Programming **3**(3), North-Holland, 359–389 (1972)

[4] Gomory, R.E., Johnson, E.L., Evans, L.: Corner Polyhedra and their Connection with Cutting Planes. To appear in Mathematical Programming

[5] Gomory, R.E., Johnson, E.L.: On the Group Problem for Mixed Integer Programming. Mathematical Programming Study **2**, 137–179 (1974)

[6] Zhang, Z.: Subadditive Functions and Integer Programming Polyhedra. Ph.D. Thesis SUNY Stony Brook, Department of Applied Mathematics and Statistics, 1992

[7] Araoz, J.: Polyhedral Neopolarities. Ph.D. Thesis, University of Waterloo, Faculty of Mathematics, 1974

[8] Balas, E., Ceria, S., Cornuejols, G., Natraj, N.: Gomory Cuts Revisited. Operations Research Letters **19**, 1–9 (1996)