so that, since our space is hypermetric,

$$\sum_{i=1}^{n}\sum_{j=1}^{n} \rho_{ij}z_iz_j + 2\sum_{i=1}^{n-1} \rho_{i,n}z_i \leqslant 0. \tag{25}$$

Similarly,

$$\sum_{i=1}^{n-1} z_i + (z_n - 1) = -1,$$

thus

$$\sum_{i=1}^{n}\sum_{j=1}^{n} \rho_{ij}z_iz_j - 2\sum_{i=1}^{n-1} \rho_{i,n}z_i \leqslant 0. \tag{26}$$

Addition of (25) and (26) yields

$$\sum_{i=1}^{n}\sum_{j=1}^{n} \rho_{ij}z_iz_j \leqslant 0. \tag{27}$$

If now $x_1, x_2, \ldots, x_n$ are rational numbers with sum 0, there exists an integer m such that $x_i = z_i/m$, $1 \leqslant i \leqslant n$, where $z_1, z_2, \ldots, z_n$ are integers with sum 0. Then

$$\sum_{i=1}^{n}\sum_{j=1}^{n} \rho_{ij}x_ix_j = \frac{1}{m^2}\sum_{i=1}^{n}\sum_{j=1}^{n} \rho_{ij}z_iz_j \leqslant 0. \tag{28}$$

If $x_1, \ldots, x_n$ are real numbers with sum 0, the left-hand side of (28) is $\leqslant 0$, by continuity. Hence our space is quasi hypermetric.

References

1. J. B. Kelly, Metric inequalities and symmetric differences. *In* "Inequalities—II" (O. Shisha, ed.), pp. 193–212. Academic Press, New York, 1970.
2. J. B. Kelly, Combinatorial inequalities. *In* "Combinatorial Structures and Their Applications" (R. Guy, H. Hanani, N. Sauer, and J. Schonheim, eds.), pp. 201–207. Gordon and Breach, New York, 1970.
3. I. J. Schoenberg, Remarks to Maurice Fréchet's article. *Ann. of Math.* **36** (1935), 724–732.
4. I. J. Schoenberg and J. von Neumann, Fourier integrals and metric geometry. *Trans. Amer. Math. Soc.* **50** (1941), 226–251.
5. G. H. Hardy, J. E. Littlewood, G. Pólya, "Inequalities," 2nd. ed. Cambridge Univ. Press, London and New York, 1952.

How Good Is the Simplex Algorithm?

Victor Klee*

Department of Mathematics, University of Washington, Seattle, Washington

AND

George J. Minty†

Department of Mathematics, Indiana University, Bloomington, Indiana

1. Introduction

By constructing long "increasing" paths on appropriate convex polytopes, we show that the simplex algorithm for linear programs (at least with its most commonly used pivot rule, Dantzig [*1*]) is not a "good algorithm" in the sense of Jack Edmonds. That is, the number of pivots or iterations that may be required is not majorized by any polynomial function of the two parameters that specify the size of the program. In particular, $2^d - 1$ iterations may be required in solving a linear program whose feasible region, defined by d linear inequality constraints in d nonnegative variables or by d linear equality constraints in $2d$ nonnegative variables, is projectively equivalent to a d-dimensional cube. Further, for each d there are positive constants α_d and β_d such that

$$\alpha_d n^{[d/2]} < \Xi(d, n) < \beta_d n^{[d/2]} \qquad \text{for all} \quad n > d, \tag{1}$$

where $\Xi(d, n)$ is the maximum number of iterations required in solving nondegenerate linear programs whose feasible regions are d-dimensional polytopes with n facets. In fact, we show that for each $d \geqslant 2$,

$$\frac{1}{2^{([d/2]^2)}} < \liminf_{n\to\infty} \frac{\Xi(d, n)}{n^{[d/2]}} \leqslant \limsup_{n\to\infty} \frac{\Xi(d, n)}{n^{[d/2]}} \leqslant \frac{2}{[d/2]!}\,. \tag{2}$$

The sharpest lower bound previously established for Ξ [*8*] asserted only that

$$\Xi(d, n) \geqslant (d - 1)(n - d) + 1.$$

Thus result (1) begins to close the "large and embarrassing gap between

* This research was partially supported by the Office of Naval Research and the Boeing Scientific Research Laboratories.

† This research was partially supported by National Science Foundation Grant GP-11878.

what has been observed and what has been proved," a gap which (in the words of Gale [4]) "has stood as a challenge to workers in the field for twenty years now and remains, in my opinion, the principal open question in the theory of linear computation." On the other hand, our results may not be especially significant for the practical aspect of linear programming (see the final section for comments on this point).

2. Preliminaries

Our constructions require some understanding of the geometric interpretation of linear programming. Though the necessary information can be found through Dantzig [2], Grünbaum [5], and Klee [8, 9], it is summarized here as an aid to the reader.

A subset of a real vector space E is called a *polytope* provided that it is the convex hull of a finite set of points or, equivalently, is the bounded intersection of a finite number of closed half spaces. The *faces* of a polytope P are P itself, the intersections of P with its various supporting hyperplanes, and the empty set $\varnothing$. The zero- and the one-dimensional faces are called, respectively, *vertices* and *edges*. Two polytopes are said to be *combinatorially equivalent* provided that there exists a biunique correspondence between the vertex set of one and the vertex set of the other such that vertices determining a face of either polytope correspond to vertices determining a face of the other polytope.

A *d-polytope* is one that is d-dimensional, and the $(d-1)$-dimensional faces of a d-polytope are called its *facets*. A polytope is said to be of *class* (d, n) provided that it is d-dimensional and has precisely n facets. Any bounded subset of R^d, the d-dimensional cartesian space, defined by n linear inequality constraints in d real variables is a polytope of class (c, m) for some $c \leqslant d$ and $m \leqslant n$; of course, its class may be exactly (d, n). A d-polytope is called *simple* provided that each of its vertices is incident to precisely d edges or, equivalently, to precisely d facets. Any simple polytope of class (d, n) is affinely equivalent to the feasible region of a nondegenerate linear program involving n inequality constraints in d real variables, of one involving $n-d$ inequality constraints in d nonnegative variables, and of one involving $n-d$ equality constraints in n nonnegative variables. In fact, there are many such programs of each sort, corresponding to various choices of the objective function.

For any linear functional ϕ defined on the space E containing a simple polytope P, a *ϕ-path* on P is a sequence $p_0, p_1, \ldots, p_l$ of vertices of P that are successively adjacent (each of the segments $[p_0, p_1], \ldots, [p_{l-1}, p_l]$ is an edge of P) and such that

$$\phi(p_0) < \phi(p_1) < \cdots < \phi(p_l).$$

The number l is the *length* of the path. The *ϕ-height* of P is the maximum of the lengths of the various ϕ-paths on P, and the *height* $\eta(P)$ is the maximum of P's ϕ-height as ϕ ranges over all linear functionals on E. The *simplex height* $\xi(P)$ and the *symmetric simplex height* $\xi_s(P)$ are later similarly defined with respect to certain special types of ϕ-paths, so that $\eta(P) \geqslant \xi(P) \geqslant \xi_s(P)$.

To maximize ϕ on P by means of the simplex method, one starts with an initial vertex p_0 of P and forms a ϕ-path $p_0, p_1, \ldots, p_l$ leading to a vertex p_l with $\phi(p_l) = \max_P \phi$. In the crudest form of the simplex method, using a pivot rule that permits moving from a vertex of P to any adjacent vertex that provides a larger value of ϕ, any ϕ-path on P may appear as part of some such maximizing ϕ-path. Hence, $\eta(P)$ iterations may be required in solving a linear program whose feasible region is P. The number $\xi(P)$ is similarly related to the most common form of the simplex method [1], in which each pivot maximizes the gradient of ϕ in the space of nonbasic variables [2, Chapter 7; 8]. (This is discussed in more detail later.) The number $\xi_s(P)$ is introduced for technical reasons.

In view of the above facts, it is natural to be interested in $H(d, n)$ [respectively, $\Xi(d, n)$, $\Xi_s(d, n]$, defined as the maximum of $\eta(P)$ [respectively, $\xi(P)$, $\xi_s(P)$] as P ranges over all simple polytopes of class (d, n). Note that

$$H(d, n) \geqslant \Xi(d, n) \geqslant \Xi_s(d, n).$$

Klee [7, 8] once proved $\Xi(d, n) \geqslant (d-1)(n-d)+1$ for all $n > d$ and

$$H(d, n) = \Xi(d, n) = (d-1)(n-d)+1 \qquad \text{when} \quad d \leqslant 3; \tag{3}$$

and was so rash as to conjecture [8, p. 320, 9, p. 150] the equalities always hold in (3). Half of that conjecture is demolished by (1). However, it still seems plausible that $H(d, n) = \Xi(d, n)$, which is to say that the usual simplex algorithm behaves, at its worst, at least as badly as any of its variants in which each pivot improves the value of the objective function. (For practical purposes, the *average behavior* of the algorithm is more important than its *worst possible* behavior. See the final section for additional comments.)

3. Statement of Main Results

The inequalities stated in the Introduction follow from our main results, which are the inequalities

$$H(d+1, n+2) \geqslant 2H(d, n) + 1, \tag{4}$$

$$H(d+2, n+k+1) \geqslant kH(d, n) + k - 1, \tag{5}$$

and the same inequalities with Ξ_s in place of H. These are first proved for H, in order to display the simple geometric ideas that are involved, and then the modifications required for Ξ_s are indicated.

4. Proof That $H(d+1, n+2) \geqslant 2H(d, n) + 1$

If K is a polytope of class (c, m) and P is a polytope of class (d, n), then $K \times P$ is a polytope of class $(c + d, m + n)$. For $0 \leqslant t \leqslant c + d$, the t-faces of $K \times P$ are precisely the sets of the form

$$(r - \text{face of } K) \times (s - \text{face of } P)$$

for $r + s = t$. It follows that $K \times P$ is simple if K and P are, and that $[0,1] \times P$ is a polytope of class $(d + 1, n + 2)$.

Now let E be the vector space containing P, let ϕ be a linear functional on E such that the ϕ-height of P is $\eta(P)$, and let $p_0, p_1, \ldots, p_{\eta(P)}$ be a ϕ-path of length $\eta(P)$ on P. Necessarily, $\phi(p_0) = \min_P \phi$ and $\phi(p_{\eta(P)}) = \max_P \phi$. In the space $R \times E$, let Q be the polytope whose vertices are the points $p^1 = (\phi(p), p)$ and $p^2 = (\sigma - \phi(p), p)$, where σ is fixed with $\sigma > 2 \max_P \phi$ and where p ranges over all vertices of P. For each $(\alpha, x) \in R \times E$, let $\psi(\alpha, x) = \alpha$. Then

$$p_0^1, p_1^1, \ldots, p_{\eta(P)-1}^1, p_{\eta(P)}^1, p_{\eta(P)-1}^2, \ldots, p_1^2, p_0^2$$

is a ψ-path of length $2\eta(P) + 1$. As Q is combinatorially (and even projectively) equivalent to the polytope $[0,1] \times P$, inequality (4) follows.

Plainly $H(2, n) = n - 1$. An immediate consequence of (4) is that $H(3, n) \geqslant 2n - 5$. But then $H(3, n) = 2n - 5$, for it follows from Euler's theorem that any simple polytope of class $(3, n)$ has precisely $2n - 4$ vertices.

For easy visualization of the above construction, suppose that $\phi(p_0) = 0$, $\phi(p_{\eta(P)}) < \frac{1}{2}$, and $\sigma = 1$. Then Q is obtained from the prism $[0,1] \times P$ by tilting the left base $\{0\} \times P$ in one way and the right base $\{1\} \times P$ in the opposite way. (The "tilting" is not effected by rigid motions, but by affine transformations which do not change the E-coordinates of points of $R \times E$.) This is shown in Fig. 1 for the case $d = 1$. The following section describes a similar construction in which R and $[0,1]$ are replaced by R^2 and a convex polygon.

In the simplest instance, starting with a segment for P, iteration of the above construction leads to a d-polytope which is combinatorially (and even projectively) equivalent to a d-dimensional cube and whose height is $2^d - 1$. Hence $H(d, 2d) \geqslant 2^d - 1$. Further, the polytope in question can be made as

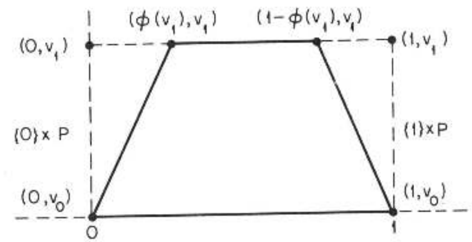

Fig. 1. Construction of a polytope for the case $d = 1$.

close as desired to the cube $[0, 1]^d$. To make this explicit, choose $\epsilon \in]0, \frac{1}{2}[$ and consider the problem

$$\max x_d$$

subject to the constraints

$$\begin{aligned} x_1 &\geqslant 0, \\ x_1 &\leqslant 1, \\ x_2 &\geqslant 0 + \epsilon x_1, \\ x_2 &\leqslant 1 - \epsilon x_1, \\ x_3 &\geqslant 0 + \epsilon x_2, \\ x_3 &\leqslant 1 - \epsilon x_2, \\ &\vdots \\ x_d &\geqslant 0 + \epsilon x_{d-1}, \\ x_d &\leqslant 1 - \epsilon x_{d-1}. \end{aligned}$$

(For notational convenience, we here maximize the last coordinate rather than the first as in Fig. 1.)

For $d = 2$, an x_d-path involving all 2^d vertices is given by

$$(0, 0),$$
$$(1, \epsilon),$$
$$(1, 1 - \epsilon),$$
$$(0, 1).$$

For $d = 3$

$$(0, 0, 0),$$
$$(1, \epsilon, \epsilon^2),$$
$$(1, 1 - \epsilon, \epsilon(1 - \epsilon)),$$
$$(0, 1, \epsilon),$$
$$(0, 1, 1 - \epsilon),$$
$$(1, 1 - \epsilon, 1 - \epsilon(1 - \epsilon)),$$
$$(1, \epsilon, 1 - \epsilon^2),$$
$$(0, 0, 1).$$

For $d = k$, the first $k - 1$ columns of the table are formed by writing down the table for $d = k - 1$ and then repeating it in reverse order. The entries in the last column are obtained in the upper half of the table by multiplying the entries in the $(k - 1)$th column by ϵ, and in the lower half of the table by subtracting ϵ times the $(k - 1)$th column entries from 1. In each case, the last column entries are strictly increasing from 0 to 1.

5. Proof That $H(d + 2, n + k + 1) \geqslant kH(d, n) + k - 1$

Let R^{d+2} denote the $(d + 2)$-dimensional cartesian space, R^2 the subspace of R^{d+2} consisting of all points whose last d coordinates are all 0, and R^d the subspace of R^{d+2} consisting of all points whose first two coordinates are both 0. For each point x of R^{d+2}, let $\alpha(x)$ denote the first coordinate of x and $\omega(x)$ the last one.

Let V_0 be a convex polygon in R^2 whose k vertices $0 = v_0, v_2, ..., v_k$ form an α-path with

$$1 = \alpha(v_2) < \alpha(v_3) < \cdots < \alpha(v_k). \tag{6}$$

Choose δ so that

$$0 < 3\delta < \min\{1, \alpha(v_2), \alpha(v_3) - \alpha(v_2), ..., \alpha(v_k) - \alpha(v_{k-1})\}, \tag{7}$$

let

$$v_1 = \delta v_2, \qquad v_{k+1} = \delta v_k, \tag{8}$$

and set

$$V = \mathrm{con}\{v_1, v_2, ..., v_k, v_{k+1}\}, \tag{9}$$

where "con" denotes the convex hull. An additional restriction is placed on δ later.

Let P be a simple polytope of class (d, n) in R^d admitting an ω-path $0, p_1, ..., p_l$ of length $l = H(d, n)$. In particular, 0 is a vertex of P and

$$0 = \min_P \omega < \max_P \omega = \omega(p_l). \tag{10}$$

We assume without loss of generality that

$$\omega(p_l) = 1. \tag{11}$$

With

$$Q^* = V + P, \tag{12}$$

Q^* is a simple polytope of class $(d + 2, n + k + 1)$, and for $1 \leqslant i \leqslant k + 1$ the polytope $v_i + P$ is a d-face of Q^*. We are going to construct a polytope Q, combinatorially equivalent to Q^*, such that the α-height of Q is at least $kl + k - 1$. It is obtained from Q^* by tilting the d-faces $v_1 + P$ in various

ways and then forming the convex hull of the union of the tilted versions. The tilting of $v_i + P$ is effected by means of an affine transformation which leaves the point v_i invariant, affects only the first two coordinates of any point, and increases [respectively, decreases] the first coordinates of the points of $\{v_i + p: p \in P \text{ with } \omega(p) > 0\}$ when i is odd [respectively, even]. As a first step toward understanding the construction, the reader should attempt with the aid of Fig. 2 to visualize the case in which $d = 1$, $n = 2$, and P is the segment $[(0, 0, 0), (0, 0, 1)]$.

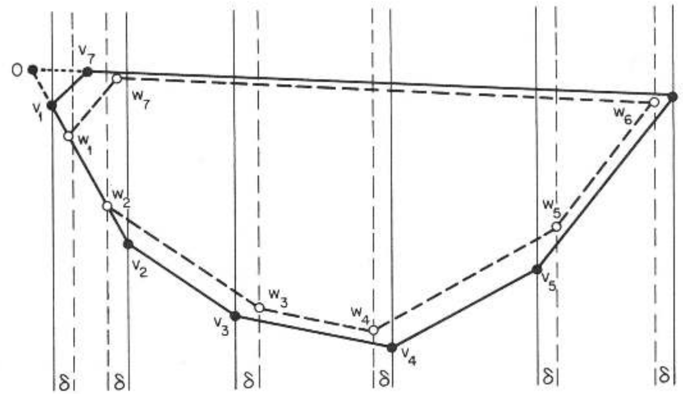

Fig. 2. Construction of a polytope for the case in which $d = 1$, $n = 2$, and P is the segment $[(0, 0, 0), (0, 0, 1)]$.

The tilting of the faces $v_i + P$ is described in terms of a convex polygon

$$W = \mathrm{con}\{w_1, w_2, ..., w_k, w_{k+1}\}, \tag{13}$$

where w_i is close to v_i for $1 \leqslant i \leqslant k + 1$. Specifically,

$$w_1 = 2\delta v_2, \qquad w_2 = (1 - \delta)v_2, \tag{14}$$

the points $w_3, ..., w_k$ are chosen in that order so that

$$\alpha(w_i) = \alpha(v_i) - (-1)^i\delta \qquad (1 \leqslant i \leqslant k) \tag{15}$$

and

$$[w_{i-1}, w_i] \text{ is parallel to } [v_{i-1}, v_i] \qquad (1 \leqslant i \leqslant k), \tag{16}$$

and the point w_{k+1} is then chosen so that

$$[w_k, w_{k+1}] \text{ is parallel to } [v_k, v_{k+1}]$$

and

$$[w_{k+1}, w_1] \text{ is parallel to } [v_{k+1}, v_1]. \tag{17}$$

These choices are all illustrated in Fig. 2. They imply that each w_i is a vertex of W, but of course W need not be contained in V.

When d is 1 and P is the segment $[(0, 0, 0), (0, 0, 1)]$, the desired polytope Q has as its vertices the points v_i and the points $z_i = w_i + (0, 0, 1)$, $1 \leqslant i \leqslant k + 1$. The simple polytope Q is of class $(3, k + 3)$ and the sequence $v_1, z_1, z_2, v_2, v_3, z_3, z_4, \ldots$ (ending with z_k, v_k for even k and with v_k, z_k for odd k) is an α-path of length $2k - 1$ on Q.

Let us now consider the general case, and recall that $\omega P = [0, 1]$. For each $p \in P$, and for $1 \leqslant i \leqslant k + 1$, let

$$p^i = (1 - \omega(p))v_i + \omega(p)w_i \in R^2, \tag{18}$$

whence $0^i = v_i$ and $p_l^{\,i} = w_i$. The "tilted version" of $v_i + P$ is the d-polytope

$$T_i = \{p^i + p : p \in P\}, \tag{19}$$

and the desired $(d + 2)$-polytope is the convex hull

$$Q = \operatorname{con} \bigcup_1^{k+1} T_i . \tag{20}$$

We claim that Q is a simple polytope of class $(d+2, n+k+1)$ and the sequence

$$\begin{aligned} &v_1, \quad p_1^{\,1} + p_1, \ldots, p_l^{\,1} + p_l, \quad p_l^{\,2} + p_l, \ldots, p_1^{\,2} + p_1, \quad v_2, \quad v_3, \quad p_1^{\,3} + p_1, \\ &\ldots, p_l^{\,3} + p_l, \quad p_l^{\,4} + p_l, \ldots, p_1^{\,4} + p_1, \quad v_4, \quad v_5, \ldots \end{aligned} \tag{21}$$

(ending with $p_l^{\,k} + p_l, \ldots, p_1^{\,k} + p_l, v_k$ for even k and with $v_k, p_1^{\,k}, \ldots, p_l^{\,k}$ for odd k) is an α-path of length $kl + k - 1$ on Q. Plainly Q has the claimed dimension, and as, by (18) and (15),

$$\alpha(p_j^{\,i} + p_j) = (1 - \omega(p_j))\, \alpha(v_i) + \omega(p_j)\, \alpha(w_i) = \alpha(v_i) - (-1)^i \delta\omega(p_j),$$

it follows from (6)–(8), (10), (11), (14), and (15) that the function α is strictly increasing for the sequence (21). It remains to show that successive members of (21) are joined by edges of Q, that Q has precisely $n + k + 1$ facets, and that Q is simple. This will all follow from an identification of Q's facets.

With Q^* as in (12), Q^* has n facets of the form

$$V + F = \operatorname{con} \bigcup_{1 \leqslant i \leqslant k+1} (v_i + F), \qquad F \text{ a facet of } P, \tag{22}$$

and $k + 1$ facets of the form

$$E + P - \operatorname{con}((v_i + P) \cup (v_j + P)), \qquad E = [v_i, v_j] \text{ an edge of } V. \tag{23}$$

We are going to show that the facets of Q are simply perturbed versions of those of Q^*. This will be accomplished with the aid of the following result, where the roles of X and Y are played by Q^* and Q, respectively.

LEMMA. *Let X and Y be polytopes having the same number m of vertices, the vertices of X being $x_1, \ldots, x_m$ and those of Y being $y_1, \ldots, y_m$. Suppose that for each index set $I \subset \{1, \ldots, m\}$,*

$$\operatorname{con}\{x_i : i \in I\} \text{ is a facet of } X \Rightarrow \operatorname{con}\{y_i : i \in I\} \text{ is a facet of } Y.$$

Then the reverse implications hold, whence X and Y are combinatorially equivalent.

Proof. For any set $I \subset \{1, \ldots, m\}$, let $X_I = \operatorname{con}\{x_i : i \in I\}$ and $Y_I = \operatorname{con}\{y_i : i \in I\}$. The set I will be called an *X-facet* [respectively, *X-face*] provided that X_I is a *facet* [respectively, *face*] of X. The *Y-facets* and *Y-faces* are similarly defined. The lemma's hypothesis is that every X-facet is a Y-facet, and we want to prove that every Y-facet is an X-facet. As the faces of a polytope are precisely those sets expressible as intersections of its facets, we know every X-face is a Y-face.

The lemma, which is obvious when X is two-dimensional, will be proved by induction on the dimension $\dim X$. Suppose it is known when $\dim X = d - 1$, consider a pair X, Y as described with $\dim X = d$, and suppose there exists a Y-facet J which is not an X-facet. We will show that this contradicts the inductive hypothesis.

Choose an X-facet I. As Y_I and Y_J are both facets of Y, there is a sequence

$$I(0) = I, I(1), \ldots, I(s) = J$$

of Y-facets such that for $1 \leqslant r \leqslant s$ the set $I(r - 1) \cap I(r)$ is a Y-face with

$$\dim Y_{I(r-1) \cap I(r)} = (\dim Y) - 2.$$

It then follows that $I(r - 1)$ and $I(r)$ are the only Y-facets containing $I(r - 1) \cap I(r)$. Let r_0 be the least r for which $I(r)$ is not an X-facet. Then the set $I(r_0 - 1) \cap I(r_0)$, though a Y-face, is not an X-face. For if it were, it would be an intersection of certain X-facets, and each of those X-facets would be a Y-facet. But the only Y-facets containing $I(r_0 - 1) \cap I(r_0)$ are $I(r_0 - 1)$ and $I(r_0)$, and the latter is not an X-facet. We now have a contradiction of the inductive hypothesis as applied to the $(d - 1)$-polytopes $X_{I(r_0-1)}$ and $Y_{I(r_0-1)}$, as every facet of $X_{I(r_0-1)}$ corresponds to a facet of $Y_{I(r_0-1)}$ but not conversely. This completes the proof of the Lemma.

Now note that the vertices of Q^* are precisely the points of the form $v_i + p$, for $1 \leqslant i \leqslant k+1$ and p a vertex of P, and that the number

$$\max\{\|p^i - v_i\|: 1 \leqslant i \leqslant k+1, \quad p \in P\}$$

can be made arbitrarily small by choosing δ small enough. It then follows easily from (19) and (20) that, for all sufficiently small δ, the vertices of Q are precisely the points of the form $p^i + p$, for $1 \leqslant i \leqslant k+1$ and p a vertex of P. This provides a natural correspondence between the vertices of Q^* and those of Q. In describing the correspondence between the facets of Q^* and those of Q, it will be convenient for each $x \in R^{d+2}$ to define x' and x'' by

$$x = x' + x'', \qquad x' \in R^2, \quad x'' \in R^d. \tag{24}$$

Now consider a facet of Q^* of the form $V + F$, as in (22). Let β be a constant and σ a linear functional on R^d such that $\max_P \sigma = \beta$ and $F = \{p \in P: \sigma(p) = \beta\}$. Extend σ to a linear functional μ on R^{d+2} by setting $\mu(x) = \sigma(x'')$ for all $x \in R^{d+2}$. Then the sets

$$\begin{aligned}\{q^* \in Q^*: \mu(q^*) = \beta\} &= V + F \\ &= \operatorname{con}\{v_i + f: 1 \leqslant i \leqslant k+1, \quad f \text{ a vertex of } F\}\end{aligned}$$

and

$$\{q \in Q: \mu(q) = \beta\} = \operatorname{con}\{p^i + f: 1 \leqslant i \leqslant k+1, \quad f \text{ a vertex of } F\}$$

are facets of Q^* and Q, respectively.

Finally, consider a facet of Q^* of the form $E + P$ as in (23), where E is an edge $[v_i, v_j]$ of V. Let β be a constant and σ a linear functional on R^2 such that $\max_V \sigma = \beta$ and $[v_i, v_j] = \{v \in V: \sigma(v) = \beta\}$. As the edges of W are parallel to corresponding edges of V, there is a constant γ such that $\max_W \sigma = \gamma$ and $[w_i, w_j] = \{w \in W: \sigma(w) = \gamma\}$. Extend σ to a linear functional μ on R^{d+2} by setting $\mu(x) = \sigma(x')$ for all $x \in R^{d+2}$, and to another such functional ψ by setting

$$\psi(x) = \sigma(x') + (\beta - \gamma)\,\omega(x). \tag{25}$$

Then the sets

$$\begin{aligned}\{q^* \in Q^*: \mu(q^*) = \beta\} &= E + P \\ &= \operatorname{con}\{v_r + p: r \in \{i, j\}, \quad p \text{ a vertex of } P\}\end{aligned} \tag{26}$$

and

$$\{q \in Q: \psi(q) = \beta\} = \operatorname{con}\{p^r + p: r \in \{i, j\}, \quad p \text{ a vertex of } P\}$$

are facets of Q^* and Q, respectively. To check the equality in (26), use (25), (24), and (18) to see that

$$\begin{aligned}\psi(p^r + p) &= \sigma(p^r) + (\beta - \gamma)\,\omega(p) \\ &= \big(1 - \omega(p)\big)\,\sigma(v_r) + \omega(p)\,\big(\beta - \gamma + \sigma(\omega_r)\big),\end{aligned}$$

whence (for arbitrary $p \in P$) $\psi(p^r + p) = \beta$ if and only if $r \in \{i, j\}$.

Reviewing the facts established since the Lemma was proved, we see from the Lemma that the natural correspondence between the vertices of Q^* and those of Q generates a combinatorial equivalence. Hence Q has all the desired properties and the proof that $H(d+2, n+k+1) \geqslant kH(d, n) + k - 1$ is complete.

6. Proof That $\alpha_d n^{[d/2]} < H(d, n) < \beta_d n^{[d/2]}$

In order to establish the existence of positive constants α_d and β_d such that the stated inequality holds for all $n > d \geqslant 2$, we prove the following more explicit result:

$$\frac{1}{2^{([d/2]^2)}} < \liminf_{n\to\infty} \frac{H(d, n)}{n^{[d/2]}} \leqslant \limsup_{n\to\infty} \frac{H(d, n)}{n^{[d/2]}} \leqslant \frac{2}{[d/2]!}. \tag{27}$$

As will be apparent in the argument below, the upper and lower bounds can be improved by treating separately the cases of odd d and even d.

Let $\mu(d, n)$ denote the maximum number of vertices of polytopes of class (d, n). It is known (Gale [3], Grünbaum [5]) that

$$\mu(d, n) \geqslant \binom{n - \left[\frac{d+1}{2}\right]}{n - d} + \binom{n - \left[\frac{d+2}{2}\right]}{n - d} \tag{28}$$

for all $n > d$, and (Klee [6], Grünbaum [5]) that equality holds for $n > [d/2]^2$.[1] Plainly $H(d, n) < \mu(d, n)$. Thus for d fixed and $n > [d/2]^2$, the function $H(d, n)$ is majorized by a polynomial of degree $[d/2]$ in n [given by (28)] with leading coefficient $1/([d/2]!)$ when d is even and $2/([d/2]!)$ when d is odd. That justifies the right-hand inequality of (27).

The left-hand inequality of (27) is obvious for $d \leqslant 3$, as $H(2, n) = n - 1$

[1] *Note added in proof*: Equality for all $n > d$ has been proved by Mc Mullen [14].

and $H(3, n) = 2n - 5$. Now suppose the left-hand inequality of (27) holds for a certain value of $d \geqslant 3$, and note that

$$\liminf_{m\to\infty} \frac{H(d+2, 2m+1)}{(2m+1)^{[(d+2)/2]}} \geqslant \liminf_{m\to\infty} \frac{H(d+2, 2m)}{(2m)^{[(d+2)/2]}}$$
$$\geqslant \liminf_{m\to\infty} \frac{(m-1)\,H(d, m) + m - 2}{(2m)\,2^{[d/2]} m^{[d/2]}}$$
$$\geqslant \frac{1}{2^{1+[d/2]}} \liminf_{m\to\infty} \frac{H(d, m)}{m^{[d/2]}}$$
$$> 2^{-(1+[d/2]+[d/2]^2)} > 2^{(-[(d+2)/2]^2)},$$

where the second inequality follows from the result of the preceding section and the fourth inequality follows from the inductive hypothesis. The proof is thus completed by mathematical induction.

7. Replacement of H by Ξ_s in the Inequalities of Previous Sections

As we have seen, $H(d, n)$ is the maximum number of pivots that may be encountered, using a very crude pivot rule, in applying the simplex method to nondegenerate linear programs with bounded feasible region of class (d, n).[2] In practice, however, more refined pivot rules are usually employed. Some of them are discussed by Kuhn and Quandt [*10*] and Dantzig [*2*, Chapter 11]. The most commonly used pivot rule is due to Dantzig [*1*, *2*, Chapter 7], and we want to adapt our earlier construction to apply to it. This will require some preliminary explanation.

In Dantzig's "standard form" [*2*, pp. 86–88, 100–101], a linear programming problem has its feasible region in R^n defined by the nonnegativity constraints $x_j \geqslant 0$ ($j = 1,..., n$), in conjunction with a system of m linear equality constraints,

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1\,, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2\,, \\ &\;\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m\,, \end{aligned}$$

and it requires the maximization of a linear functional

$$c_1x_1 + c_2x_2 + \cdots + c_nx_n = z(x)$$

[2] *Note added in proof*: That the behavior of $H(d, n)$ is "exponentially bad" has also been established by Jack Edmonds, by means of an example associated with an algorithm for finding shortest paths.

over that region. (We maximize rather than minimize, as the former led to more natural notation in the constructions of the preceding sections.) If the problem is nondegenerate and its feasible region P is bounded, then $m \leqslant n$ and P is a simple polytope of class $(n - m, r)$ with $n - m < r \leqslant n$. Dantzig's pivot rule calls, at each stage, for "bringing the most positive column into the basis." In connection with our geometric approach to the construction of examples involving a large number of pivots, we require a coordinate-free description of Dantzig's pivot rule. This is taken from Klee [*8*].

Let X be a flat in a real vector space and let P be a simple d-polytope in X. By a *variable* for the pair (P, X), we mean an affine functional ξ on X such that ξ is not constant on P. For any such ξ, the sets $H(\xi) = \{x \in X: \xi(x) = 0\}$ and $J(\xi) = \{x \in X: \xi(x) \geqslant 0\}$ are, respectively, the *hyperplane* and the *half space associated* with ξ. Consider the problem of maximizing a variable ϕ_0 on P. Let $\mathscr{F}$ denote the set of all facets of P, and for each $F \in \mathscr{F}$ let ϕ_F be a variable such that $F \subset H(\phi_F)$ and $P \subset J(\phi_F)$. Let $\Phi = \{\phi_F: F \in \mathscr{F}\}$, whence P is the intersection of the set $\bigcap_{\phi\in\Phi} J(\phi)$ with the smallest flat containing P. Our discussion is henceforth relative to the system (P, Φ, ϕ_0). For a linear programming problem in standard form, as described in the preceding paragraph, X would be the flat in R^n defined by the equality constraints, P would be the intersection of X with the nonnegative orthant of R^n, Φ would be the set of restrictions to X of those coordinate functionals whose vanishing determines a facet of P, and ϕ_0 would be the restriction to X of the objective function z.

If p is a vertex of the simple d-polytope P, a variable $\phi \in \Phi$ is said to be *basic* or *nonbasic* for p according as $\phi(p) > 0$ or $\phi(p) = 0$; the set of all nonbasic variables is denoted by Φ_p. Each set Φ_p is of cardinality d, and two vertices p and p' of P are adjacent if and only if there is exactly one variable $\phi^{p,p'}$ in $\Phi_p \sim \Phi_{p'}$. The *nonbasic* (p, p') *gradient* of ϕ_0 is then defined as the quotient

$$(\phi_0(p') - \phi_0(p))/(\phi^{p,p'}(p') - \phi^{p,p'}(p)) = (\phi_0(p') - \phi_0(p))/(\phi^{p,p'}(p')),$$

representing the amount of improvement in ϕ_0 achieved per unit increase in $\phi^{p,p'}$ by moving from p along the edge $[p, p']$. The nonbasic (p, p') gradient of ϕ_0 is equal to the coefficient $\gamma(\phi^{p,p'})$, where

$$\phi_0 = \gamma_0 + \sum_{\phi\in\Phi_p} \gamma(\phi)\phi$$

is the unique expression in that form (the γ's being constants) for the restriction of ϕ_0 to P. A ϕ_0-path $p_0,..., p_l$ is called a *simplex path* for the system (P, Φ, ϕ_0) provided that

$$\gamma(\phi^{p_{i-1},p_i}) = \max \gamma(\phi^{p_{i-1},q}) \qquad \text{for} \quad 1 \leqslant i \leqslant l,$$

the maximum being over all vertices q of P adjacent to p_{i-1}. Thus, a simplex path is one which at each stage maximizes the nonbasic gradient. If the maximum is strict at each stage—that is, if $\gamma(\phi^{p_{i-1}, p_i}) > \gamma(\phi^{p_{i-1}, q})$ for all vertices q of P that are adjacent to p_{i-1} and different from p_i—the simplex path is said to be *unambiguous*. A *symmetric simplex path* for (P, Φ, ϕ_0) is a ϕ_0-path $p_0, ..., p_l$ such that $p_0, ..., p_l$ is an unambiguous simplex path for (P, Φ, ϕ_0), and $p_l, ..., p_0$ is an unambiguous simplex path for $(P, \Phi, -\phi_0)$.

The *simplex height* $\xi(P)$ [respectively, *symmetric simplex height* $\xi_s(P)$] of any simple polytope P is the maximum length of simplex paths [respectively, symmetric simplex paths] in systems of the form (P, Φ, ϕ_0). Finally, $\Xi(d, n)$ [respectively, $\Xi_s(d, n)$] is defined as the maximum of $\xi(P)$ [respectively, $\xi_s(P)$] as P ranges over all simple polytopes of class (d, n). Plainly $\Xi(d, n) \geqslant \Xi_s(d, n)$. It follows from the discussion by Klee [*8*] that $\Xi(d, n)$ is the maximum number of iterations required, using Dantzig's pivot rule, in solving nondegenerate linear programs with bounded feasible region of class (d, n).

It is now a routine, though tedious, matter to verify that the inequalities previously established for the function H are valid for the function Ξ_s as well. The technique is similar, in some respects, to that of Klee [*8*]. We show by way of illustration that

$$\Xi_s(d + 1, n + 2) \geqslant 2\Xi_s(d, n) + 1, \tag{29}$$

presenting the proof of (29) in such a way that the reader will (we hope) see how the same ideas can be used to show

$$\Xi_s(d + 2, n + k + 1) \geqslant k\Xi_s(d, n) + k - 1. \tag{30}$$

Let R^{d+1} denote, as usual, the $(d + 1)$-dimensional cartesian space, R^1 the subspace of R^{d+1} consisting of all points whose last d coordinates are all 0, and R^d the subspace of R^{d+1} consisting of all points whose first coordinate is 0. For each point x of R^{d+1}, let $\alpha(x)$ denote the first coordinate and $\omega(x)$ the last coordinate of x, and let the points x' and x'' be such that

$$x = x' + x'', \qquad x' \in R^1, \quad x'' \in R^d.$$

Let $\bar{\alpha}$ and $\bar{\omega}$ denote, respectively, the restriction of α to R^1 and of ω to R^d.

Let $l = \Xi_s(d, n)$. Then R^d admits a set Φ of n affine functionals such that the following conditions are satisfied:

(a) the intersection $P = \bigcap_{\phi \in \Phi} J(\phi)$ is a simple polytope of class (d, n);

(b) $\bar{\omega}$ does not have the same value at any two vertices of P (convenient, though not essential);

(c) the system $(P, \Phi, \bar{\omega})$ admits a symmetric simplex path $0 = p_0, ..., p_l$ with

$$0 = \min_P \bar{\omega} = \bar{\omega}(p_0) < \cdots < \bar{\omega}(p_l) = \max_P \bar{\omega} = 1.$$

Let V denote the segment $\{x \in R^1: 0 \leqslant \bar{\alpha}(x) \leqslant 1\}$, choose $\delta \in]0, \frac{1}{3}[$, and for each $p \in P$ let

$$p^1 = (\delta\bar{\omega}(p), p), \qquad p^2 = (1 - \delta\bar{\omega}(p), p).$$

Let Q denote the convex hull

$$Q = \operatorname{con}(\{p^1: p \in P\} \cup \{p^2: p \in P\}).$$

Then Q, being combinatorially equivalent to $V + P$, is a simple polytope of class $(d + 1, n + 2)$. The path

$$p_0^1, p_1^1, ..., p_l^1, p_l^2, ..., p_1^2, p_0^2 \tag{31}$$

is an α-path of length $2l + 1$ on Q and we want to show that for a suitable set Ψ of affine functionals on R^{d+1} it is a symmetric simplex path for the system (Q, Ψ, α).

For each $\tau > 0$ let $\Lambda_\tau = \{\tau\bar{\alpha}, \tau - \tau\bar{\alpha}\}$, a pair of affine functionals on R^1. For each such τ,

$$\bigcap_{\lambda \in \Lambda_\tau} J(\lambda) = V, \tag{32}$$

and by choosing τ large enough we may be sure that

the maximum of the absolute value of the nonbasic (v, v') gradient of $\bar{\alpha}$ for the system $(V, \Lambda_\tau, \bar{\alpha})$ is less than δ times the minimum of the absolute value of the nonbasic (p, p') gradient of $\bar{\omega}$ for the system $(P, \Phi\ \bar{\omega})$, where (v, v') [respectively, (p, p')] ranges over all ordered pairs of adjacent vertices of V [respectively, P]. (33)

Having chosen τ for which (33) holds, we then set

$$\Psi = \Phi^\# \cup \Lambda_\tau^\#, \tag{34}$$

where $\Phi^\#$ consists of the functionals $\phi^\#$ given by $\phi^\#(x) = \phi(x')$ (for all $x \in R^{d+1}$, $\phi \in \Phi$), and $\Lambda_\tau^\#$ consists of the two functionals $\tau\alpha^\#$ and $\tau - \tau\alpha^\#$, $\alpha^\#$ being given by

$$\alpha^\#(x) = \bar{\alpha}(x') - \delta\bar{\omega}(x''), \qquad x \in R^{d+1}. \tag{35}$$

Then $Q = \bigcap_{\psi \in \Psi} J(\psi)$ and with the aid of (33) it can be verified that the path (31) is a symmetric simplex path for the system (Q, Ψ, α). This completes the proof of (29).

In the argument just completed, the purpose of (33) was to ensure that, whenever a move is made along an edge of Q so as to maximize the nonbasic gradient of α for the system (Q, Ψ, α), the edge in question will correspond (under the natural combinatorial equivalence between Q and $V+P$) to an edge of P rather than to an edge of V, provided that any improvement in α can be achieved at that stage by traveling along an edge of Q that corresponds to one of P.

The inequality (30) is established by combining the ideas of the above construction with those used earlier in proving (5). With V as in the earlier section (a convex polygon rather than a segment) and P as above, Q is defined by means of (20) and Λ_τ is chosen subject to (32) and (33). Comparing the roles of (25) and (35), one is led to the proper analog of (34) and the proof of (30) is completed. The details are omitted because they are tedious and not particularly instructive.

8. Final Comments

We have discussed Dantzig's pivot rule explicitly because it is the one most commonly used in practice. However, our methods could probably be used to exhibit the same bad behavior for many other pivot rules. Indeed, we do not believe there exists a pivot rule that turns the simplex method into a "good algorithm" in the sense of Edmonds, though the rule calling at each stage for greatest possible improvement (rather than gradient) of the objective function would seem to merit further study (see Dantzig's comments on this [2, p. 240]).[3]

We have of course been discussing the *worst* rather than the *average* behavior of the simplex algorithm, and it should be emphasized that the number of iterations required in our examples is much greater than the number usually encountered in practice ([2, p. 160]) or even in formal experimental studies of the simplex method (Kuhn and Quandt [10]). Inequality (5) led to (1) and (2), which are rather satisfactory from a theoretical viewpoint. However, so far as linear programs of common size are concerned, it is most instructive to compare our inequality,

$$\Xi(d, 2d) \geqslant 2^d - 1, \tag{36}$$

with Dantzig's summary [2, p. 160] of "empirical experience with thousands of practical programs." In a statement applying to nondegenerate linear programs with feasible regions, defined by m linear equality constraints in n nonnegative variables, he reports that "the number of iterations may run anywhere from m as a minimum, to $2m$ and rarely to $3m$. The number is

[3] *Note added in proof*: This rule has recently been studied by Jeroslow [13], with results similar to those obtained here for Dantzig's pivot rule.

usually less than $3m/2$ when there are less than 50 equations and 200 variables (to judge from informal empirical observations)." For a nondegenerate problem with $m = 40$ and $n = 80$, the feasible region could be of class (40, 80) and our $\Xi(40, 80) \geqslant 2^{40} - 1$ is to be compared with Dantzig's $3m = 120$. Any mathematical explanation of this contrast must be in the realm of geometric probability, and the results of Rényi and Sulanke [11] and Schmidt [12] may be relevant.

Dantzig [2, p. 160] also reports: "Some believe that for a randomly chosen problem with fixed m, the number of iterations grows in proportion to n." As interpreted in terms of *worst* rather than *average* behavior, this would imply that for each m there is a constant $\gamma(m)$ such that

$$\Xi(n - m, n) \leqslant \gamma(m)n \qquad \text{for all} \quad n > m + 1.$$

We do not know of any results contradicting this. However, it follows from (36) that no such function $\gamma(m)$ is majorized by any polynomial in m.

References

1. G. B. Dantzig, Maximization of a linear function of variables subject to linear in equalities. *In* "Activity Analysis of Production and Allocation" (Cowles Commission Monograph No. 13) (T. C. Koopmans, ed.), pp. 339–347. Wiley, New York, 1951.
2. G. B. Dantzig, "Linear Programming and Extensions." Princeton Univ. Press, Princeton, New Jersey, 1963.
3. D. Gale, Neighborly and cyclic polytopes. *In* "Convexity" (V. L. Klee, ed.), *Proc. Symp. Pure Math.*, Vol. VII, pp. 225–232. Amer. Math. Soc., Providence, Rhode Island, 1963.
4. D. Gale, How to solve linear inequalities. *Amer. Math. Monthly* **76** (1969), 589–599.
5. B. Grünbaum, "Convex Polytopes." Wiley, New York, 1967.
6. V. Klee, On the number of vertices of a convex polytope. *Canad. J. Math.* **16** (1964), 701–720.
7. V. Klee, Heights of convex polytopes. *J. Math. Anal. Appl.* **11** (1965), 176–190.
8. V. Klee, A class of linear programs requiring a large number of iterations. *Numer. Math.* **7** (1965), 313–321.
9. V. Klee, Convex polytopes and linear programming. *In Proc. IBM Sci. Comput. Symp. Combinatorial Problems, IBM Data Processing Div., White Plains, New York, 1966*, pp. 123–158.
10. H. Kuhn and R. E. Quandt, An experimental study of the simplex method. *In* "Experimental Arithmetic, High-Speed Computing and Mathematics" (N. C. Metropolis *et al.* eds.), *Proc. Symp. Appl. Math.*, Vol. XV, pp. 107–124. Amer. Math. Soc., Providence, Rhode Island, 1963.
11. A. Rényi and R. Sulanke, Zufällige konvexe Polygone in einem Ringgebiet. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **9** (1968), 146–157.
12. W. M. Schmidt, Some results in probabilistic geometry. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **9** (1968), 158–162.
13. R. Jeroslow, The simplex algorithm with the pivot rule of maximizing criterion. (To be published.)
14. P. McMullen, The maximum numbers of faces of a convex polytope. *Mathematika* **17** (1970), 179–184.