COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
AT YALE UNIVERSITY

Box 2125, Yale University
New Haven, Connecticut 06520

COWLES FOUNDATION DISCUSSION PAPER NO. 917

ON INTEGER POINTS IN POLYHEDRA: A LOWER BOUND

by

Imre Bárány, Roger Howe and László Lovász

May 22, 1989

ON INTEGER POINTS IN POLYHEDRA: A LOWER BOUND

by

Imre Bárány, Roger Howe and László Lovász*

ABSTRACT: Given a polyhedron $P \subset \mathbb{R}^n$ we write P_I for the convex hull of the integral points in P. It is known that P_I can have at most $O(\varphi^{n-1})$ vertices if P is a rational polyhedron with size φ. Here we give an example showing that P_I can have as many as $\Omega(\varphi^{n-1})$ vertices. The construction uses the Dirichlet unit theorem.

*The authors are respectively at the Mathematical Institute, Pf 127, 1364 Budapest, Hungary; the Department of Mathematics, Yale University, New Haven, CT 06520, U.S.A.; and the Departments of Computer Science, Eötvös University, 1088 Budapest, Hungary and Princeton University, Princeton, NJ 08544, U.S.A. The Research of the first author was supported by the Program in Discrete Mathematics at the Cowles Foundation, Yale University and by NSF grant 593A—31—46804, of the second author by NSF grant DMS—8807336 and of the third author by Hungarian National Foundation for Scientific Research grant No. 1812.

1. RESULTS

Given a polyhedron $P \subset \mathbb{R}^n$ write P_I for the convex hull of integral points in P. P is a rational polyhedron if it is given by finitely many inequalities of the form $a^T x \leq \alpha$ where $a \in \mathbb{Q}^n$ and $\alpha \in \mathbb{Q}$. The *size* of this inequality is the number of bits necessary to encode it as a binary string (see Schrijver [S]). The *size of a rational polyhedron* $P \subset \mathbb{R}^n$ is the sum of the sizes of the defining inequalities. Strengthening some earlier results of Shevchenko [Sh] and Hayes and Larman [HL], Cook, Hartmann, Kannan and McDiarmid [CHKM] have proved recently that P_I can have at most $2m^n(12n^2\varphi)^{n-1}$ vertices where m is the number of defining inequalities and φ is the size of P. For some other results and comments see their paper [CHKM]. For $n = 2$ and $n = 3$ there are examples in [R] and in [M] showing that P_I can have as many as $\Omega(\varphi^{n-1})$ vertices. Here we give such a construction for every $n \geq 2$.

THEOREM 1. *For fixed* $n \geq 2$ *and for any* $\varphi \geq 0$ *there exists a rational polyhedron* $P \subset \mathbb{R}^n$ *of size at most* φ *such that the number of vertices of* P_I *is at least* $c\varphi^{n-1}$ *where* c *is a constant depending only on* n. *Moreover, the number of facets of* P *is at most* $2n^2$.

The proof will be based on

THEOREM 2. *There exist* n *by* n, *integral matrices* $A_1, A_2, \ldots, A_{n-1}$ *with the following properties*

(1) *The determinant of each* A_k *equals* 1,

(2) *Every* A_k *has the same set of eigenvectors* $\{s_1, \ldots, s_n\}$,

(3) A_k *has* n *distinct positive eigenvalues* $\lambda_1(A_k), \ldots, \lambda_n(A_k)$, *with* $\lambda_i(A_k)$ *corresponding to the eigenvector* s_i,

(4) *The vectors* $\log \lambda(A_k) \in \mathbb{R}^n$ $(k = 1, \ldots, n-1)$ *are linearly independent over the reals, where the* i–th *component of* $\log \lambda(A_k)$ *is* $\log \lambda_i(A_k)$, $i = 1, \ldots, n$.

Clearly, condition (4) is equivalent to the following

(5) The vectors $\sum_{k=1}^{n-1} \alpha_k \log \lambda(A_k)$ where $\alpha = (\alpha_1, \ldots, \alpha_{n-1})^T \in \mathbb{Z}^{n-1}$ form an (n–1)–dimensional lattice L in $\mathbb{R}^n$.

The lattice L is orthogonal to the vector $(1,, \ldots, 1) \in \mathbb{R}^n$. This follows from (1). Another way to put (4) or (5) is to say that the matrices $A_1, \ldots, A_{n-1}$ multiplicatively generate an (n–1)–dimensional lattice (a group isomorphic to $\mathbb{Z}^{n-1}$).

As a matter of fact, Theorem 2 is a direct consequence of the Dirichlet Unit Theorem (see, e.g., [BS]). We will explain this in the last section. For the sake of the reader who is not familiar with algebraic number theory a separate and self–contained proof of Theorem 2 will be given in the third section.

2. PROOF OF THEOREM 1

We use Theorem 2. Set $S = \text{cone}\{s_1, \ldots, s_n\}$ and consider $v \in \mathbb{Z}^n \cap \text{int } S$, $v = (v_1, \ldots, v_n)^T$, an integral vector from int S. Define the set

$$V = \{v^a = A_1^{\alpha_1} \cdots A_{n-1}^{\alpha_{n-1}} v \in \mathbb{R}^n : a = (\alpha_1, \ldots, \alpha_{n-1})^T \in \mathbb{Z}^{n-1}\}.$$

Clearly $V \in \mathbb{Z}^n \cap \text{int } S$. For $x = \xi_1 s_1 + \cdots + \xi_n s_n$ define

$$\text{prod}(x) = \prod_{i=1}^{n} \xi_i .$$

Claim 1. For all $w \in V$, $\text{prod}(w) = \text{prod}(v)$.

Indeed, $\operatorname{prod}(A_k v) = \Pi_{i=1}^n \lambda_i(A_k) v_i = \Pi_{i=1}^n \lambda_i(A_k) \operatorname{prod}(v) = \det(A_k)\operatorname{prod}(v) = \operatorname{prod}(v)$ where with $v = \nu_1 s_1 + \cdots + \nu_n s_n$ and the claim follows by an easy induction.

Claim 2. Each $w \in V$ is an extreme point of $\operatorname{conv} V$.

PROOF. The function $f(x) = \log \operatorname{prod}(x)$ is strictly concave on $\operatorname{int} S$:

$$f\left[\frac{x+y}{2}\right] = \log \prod_{i=1}^{n} \frac{\xi_i + \eta_i}{2} \geq \log \prod_{i=1}^{n} \sqrt{\xi_i \eta_i} = \frac{1}{2}(f(x) + f(y))$$

with equality if and only if $x = y$. Then the set $\{x \in S : \operatorname{prod}(x) \geq \operatorname{prod}(v)\}$ is convex with each point $w \in V$ lying on its boundary. Then each $w \in V$ can be strictly separated from the other points of $\operatorname{conv} V$. □

The set $K = \operatorname{conv} V$ is not a polyhedron because it is the convex hull of infinitely many points. However, as we shall see soon, K is "locally" a polytope. More precisely, let Q be the minimal cone having apex v and containing V. Such a minimal cone clearly exists.

Claim 3. Q is a polyhedral cone.

PROOF. We show first that V contains points arbitrarily close to the ray $\{ts_j : t \geq 0\}$ for every $j = 1,, \ldots, n$. For notational convenience we do so only when $j = 1$. Since

$$v^a = A_1^{\alpha_1} \cdots A_{n-1}^{\alpha_{n-1}} v = \prod_{k=1}^{n-1} \lambda_1^{\alpha_k}(A_k)\nu_1 s_1 + \cdots + \prod_{k=1}^{n-1} \lambda_n^{\alpha_k}(A_k)\nu_n s_n ,$$

we have to prove the existence of $a \in \mathbb{Z}^{n-1}$ with

$$\prod_{k=1}^{n-1} \lambda_i^{\alpha_k}(A_k)\nu_i < \epsilon , \quad (i = 2, \ldots, n)$$

for any fixed $\epsilon > 0$. But this is the same as

$$(6) \qquad \sum_{k=1}^{n-1} \alpha_k \log \lambda_i(A_k) + \log \nu_i < \log \epsilon\,, \quad (i = 2, \ldots, n)\,.$$

(Here $\lambda_i(A_k)$ and ν_i are positive by assumption.) The existence of such an $a \in \mathbb{Z}^{n-1}$ is guaranteed by condition (5) and the fact that L is orthogonal to the vector of all ones.

Define now $\epsilon = \frac{1}{2}\min\{\nu_1, \ldots, \nu_n\} > 0$. Let $w_j \in V$ be any point closer than ϵ to the ray $\{te_j : t > 0\}$ in the sense of (6). Define the cone C with apex v as

$$C = v + \operatorname{cone}\{w_1 - v, \ldots, w_n - v\}\,.$$

Clearly $C \in Q$. It is easy to see that the set $S \backslash C$ is bounded. Then (5) implies that $S \backslash C$ contains finitely many points from V, $u_1, \ldots, u_m$, say. Then

$$Q = v + \operatorname{cone}\{w_1 - v, \ldots, w_n - v, u_1 - v, \ldots, u_m - v\}$$

and so Q is a polyhedral cone. □

We define a face, F, of K as a subset $F \subset K$ such that there is a closed halfspace H^+ with bounding hyperplane H such that $K \subset H^+$ and $F = H \cap K$. Then each $w \in V$ is a vertex of K. Moreover, the proof of Claim 3 shows that the facets of K, incident to the vertex v, are all bounded. Thus each face of K incident to v is a polytope. Observe now that V, and consequently $K = \operatorname{conv} V$, is invariant under each linear transformation A_k. This means that the facial structure of K around any one of its vertices is the same. In particular, the boundary of K consists of (n–1)–dimensional polytopes, which we will call facets of K, and for any facet F incident to $w = v^a$ there is a facet F' incident to v such that $A_1^{\alpha_1} \cdots A_{n-1}^{\alpha_{n-1}}(F') = F$. We will use this fact to prove

Claim 4. The function prod assumes a minimum value on the set $\mathbb{Z}^n \cap \operatorname{int} S$.

PROOF. Fix some $v \in \mathbb{Z}^n \cap \operatorname{int} S$ and consider a point $u \in \mathbb{Z}^n \cap \operatorname{int} S$ with $\operatorname{prod}(u) < \operatorname{prod}(v)$. Then the ray $\{tu : t > 0\}$ intersects the boundary of K at a point $u' \in F$ for some facet F incident to some vertex $w = v^a \in V$. Then $\operatorname{prod}(u^{-a}) = \operatorname{prod}(u) < \operatorname{prod}(v)$, and the ray $\{tu^{-a} : t > 0\}$ intersects the boundary of K in the facet $F' = A_1^{-\alpha_1} \cdots A_{n-1}^{-\alpha_{n-1}}(F)$ which is incident to v. This means that $\operatorname{prod}(u) = \operatorname{prod}(u^{-a})$ for some $u^{-a} \in \operatorname{conv}(0 \cup F')$ where F' is a facet incident to v. As the union of the sets $\operatorname{conv}(O \cup F')$ is compact, it contains only finitely many points from $\mathbb{Z}^n$. □

Remark 1. The proof shows further, that the set of values of the function prod is discrete on $\mathbb{Z}^n \cap \operatorname{int} S$. This follows immediately if one uses the Dirichlet unit theorem: it is clear that prod coincides up to a constant multiple with the norm and the norm takes integral values only (see §4, and [BS], [L]).

By virtue of Claim 4, we may assume we have selected $v \in \mathbb{Z}^n \cap \operatorname{int} S$ with $\operatorname{prod}(v)$ minimal in $\mathbb{Z}^n \cap \operatorname{int} S$. Clearly $\operatorname{prod}(v) > 0$. Define V and $K + \operatorname{conv} V$ using this point v. So far we have established that each $w \in V$ is a vertex of the convex hull of $\mathbb{Z}^n \cap \operatorname{int} S$.

Consider now $\varphi \in \mathbb{R}$, large enough, and the set

$$V(\varphi) = \{w = \omega_1 s_1 + \cdots + \omega_1 s_n \in V : w_i \leq 2^{\varphi}, i = 1, \ldots, n\}.$$

The cardinality of $V(\varphi)$ is the same as the number of points $a \in \mathbb{Z}^{n-1}$ with

$$\sum_{k=1}^{n-1} \alpha_k \log \lambda_i(A_k) + \log \nu_i \leq \varphi, \quad i = 1, \ldots, n. \tag{7}$$

In view of (5), this number is essentially the same as the (n–1)–dimensional volume of the set

$$\left\{x = \sum_{k=1}^{n-1} \alpha_k \log \lambda(_k) + \log v : a = (\alpha_1, \ldots, \alpha_{n-1}) \in \mathbb{R}^n, x_i \le \varphi, i = 1, \ldots, n\right\}.$$

As this set is a simplex, its volume is equal to $\text{const}\cdot\varphi^{n-1}$ with the constant depending only on $A_1, \ldots, A_{n-1}$. Thus

$$|V(\varphi)| \ge \text{const}\cdot\varphi^{n-1}. \tag{8}$$

Now we are going to define the polyhedron P whose existence is claimed in the theorem. Let the point $x = \xi_1 s_1 + \cdots + \xi_n s_n \in \mathbb{R}^n$ have components $x_1, \ldots, x_n$ in the standard coordinates on $\mathbb{R}^n$. Define $Z(\varphi)$ as the set of points $x \in \mathbb{Z}^n$ with $0 < \xi_i \le 2\varphi$, $i = 1, \ldots, n$. Let $v_i \in Z(\varphi)$ be the point with minimal i–th component in the basis $s_1, \ldots, s_n$ $(i = 1, \ldots, n)$, and let m_i be the i–th component of v_i. Then the inequality $\xi_i \ge m_i$ is implied by n inequalities that define facets of $\operatorname{conv} Z(\varphi)$. These inequalities have the form

$$0 \le \text{ or } \ge \det\begin{bmatrix} 1 & \cdots & 1 & 1 \\ w_1 & \cdots & w_n & x \end{bmatrix} = b_0 + b_1x_1 + \cdots + b_nx_n, \tag{9}$$

where $w_i \in Z(\varphi)$. We may assume the s_i are unit vectors in the standard Euclidean norm. Then as the Euclidean distance of w_i from the origin is at most $n2^{\varphi}$, its components in the standard basis are at most $n2^{\varphi}$ in absolute value. So b_i is equal to the value of an integral n by n determinant all of whose entries are at most $n2^{\varphi}$ in absolute value. Thus the size of the inequality (9) is at most $\text{const}\cdot\varphi$ for some constant, depending only on n. The number of such inequalities is n for each v_i and so it is n^2 altogether. Similarly, let $u_i \in Z(\varphi)$ be the point with maximal i–th component $(i = 1, \ldots, n)$, and let the i–th component of u_i be equal to M_i. Then the inequality $\xi_i \le M_i$ is implied by n inequalities that define facets of $\operatorname{conv} Z(\varphi)$. These latter inequalities are of the form (9) and their number is at most n^2. Let now P be the polyhedron defined by these $2n^2$ inequalities. Then P is rational, has size at most $\text{const}\cdot\varphi$ with the constant depending only on n. Moreover, $V(\varphi) \subset P$ and $0 < \xi_i < 2\varphi$ for any $x \in P$. This implies that every $w \in V(\varphi)$ is a vertex of P_I. This proves the theorem. □

Remark 2. We have shown that $V(\varphi) \subset \text{vert } P_I$. This and the symmetry of V imply that the number of k–dimensional faces $(k = 0, 1, \ldots, n-1)$ of P_I is at least $\text{const} \cdot \varphi^{n-1}$. It would be interesting to extend the results of [CKHM] by showing that P_I has at most $O(\varphi^{n-1})$ k–dimensional faces for any polytope P of size φ.

Remark 3. Using the above construction one can find highly regular triangulations of $\mathbb{R}^{n-1}$ that are perhaps new and interesting. Consider the convex set K defined above and assume each facet of K is a simplex. This gives rise to a simplicial complex K with (infinite) vertex set V where vertices $w_1 = v^{a_1}, \ldots, w_d = v^{a_d}$ form a simplex if their convex hull is a face of K. K is $(n-1)$–dimensional and can be represented as a triangulation T of $\mathbb{R}^{n-1}$ with vertex set $\mathbb{Z}^{n-1}$ in the following way. The points $a_1, \ldots, a_d \in \mathbb{Z}^{n-1}$ form a simplex if the convex hull of the points $v^{a_1}, \ldots, v^{a_d}$ is a face of K. The triangulation T is invariant under translations from $\mathbb{Z}^{n-1}$. The geometric properties of T could be deduced from those of the cone Q. When one uses the Dirichlet unit theorem for the construction, the triangulation comes from an irreducible polynomial. So most probably, there are many different triangulations of this type.

3. PROOF OF THEOREM 2

Define the polynomial

$$p(\lambda) = (\lambda-2)(\lambda-4) \cdots (\lambda-2n) + 1 = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_0 .$$

Clearly, $a_{n-1}, \ldots, a_0$ are integers. Computing p at $\lambda = 1, 3, \ldots, 2n+1$ we see that p has n real roots $\lambda_1 < \lambda_2 < \cdots < \lambda_n$. The root λ_i is close to $2i$, more precisely:

(10) $$|\lambda_i - 2i| < 1 \text{ and so } |\lambda_i - 2j| > 1 \text{ when } i \neq j .$$

Define the n by n integral matrix A as

$$A = \begin{bmatrix} 1 & 1 & 0 \cdot \cdot & 0 & 0 \\ 0 & 0 & 1 \cdot \cdot & 0 & 0 \\ \cdot & \cdot & \cdot \cdot \cdot & \cdot & \cdot \\ 0 & 0 & 0 \cdot \cdot & 0 & 1 \\ -a_0 & -a_1 & \cdot \cdot \cdot & -a_{n-2} & -a_{n-1} \end{bmatrix} .$$

Then, as it is well–known and actually easy to check

$$\det(A - \lambda I) = (-1)^n p(\lambda) .$$

A has n (real) eigenvectors $s_1, \ldots, s_n$ with $As_i = \lambda_i s_i$. Define now

$$A_k = A - 2kI , \; k = 1, 2, \ldots, n .$$

Then $A_k s_i = (A - 2kI)s_i = (\lambda_i - 2k)s_i$ and so A_k has the same set of eigenvectors as A with eigenvalues $\lambda_i(A_i) = \lambda_i - 2k$. Then $\det(A_k) = (-1)^n p(\lambda - 2k) = (-1)^n$ and $A_1 \cdots A_n = -I$, because $A_1 \cdots A_n s_i = \Pi_{k=1}^n \lambda_i(A_k)s_i = \Pi_{k=1}^n (\lambda_i - 2k)s_i = -s_i$. Next, we prove property (4) for the vectors $\log|\lambda(A_k)|$, $k = 1, \ldots, n-1$. Then the theorem will follow for the matrices $A_1^2, A_2^2, \ldots, A_{n-1}^2$ (instead of $A_1, \ldots, A_{n-1}$; but the A_k^2 serve just as well). So assume

$$\sum_{k=1}^{n-1} \alpha_k \log|\lambda(A_k)| = 0$$

for some real numbers $\alpha_1, \ldots, \alpha_{n-1}$. Defining $\alpha_n = 0$ we have

$$\sum_{k=1}^{n} \alpha_k \log|\lambda(A_k)| = 0 .$$

Set $|\alpha_j| = \max\{|\alpha_k| : k = 1, \ldots, n\}$. If $j = n$ then we are done. So assume $j \neq n$

and consider the j–th component of the above equation.

$$\sum_{k=1}^{n} \alpha_k \log|\lambda_j(A_k)| = 0 .$$

Then, using (10)

$$\begin{aligned}
|\alpha_j \log|\lambda_j(A_j)|| &= \Big|\sum_{k\neq j} \alpha_k \log|\lambda_j(A_k)|\Big| \\
&\le \sum_{k\neq j} |\alpha_k|\,|\log|\lambda_j(A_k)|| \\
&\le \sum_{k\neq j} |\alpha_k| \log|\lambda_j(A_k)| \\
&\le |\alpha_j| \sum_{k\neq j} \log|\lambda_j(A_k)| \\
&= |\alpha_j|\,|\log|\lambda_j(A_j)||
\end{aligned}$$

because $A_1 \cdots A_n = -I$ implies $\Sigma_{k=1}^{n} \log|\lambda(A_k)| = 0$. But then equality holds throughout and so $|\alpha_j| = |\alpha_k| = 0$. □

4. RELATION TO TOTALLY REAL NUMBER FIELDS

The above construction is a particularly transparent case of a general phenomenon of algebraic number theory. Precisely, given $A_1, \ldots, A_{n-1}$ as in Theorem 2, let $\mathcal{A}_{\mathbb{Q}}$ be the set of all linear combinations, with rational coefficients, of the products of the A_k's. Let $\mathcal{A}_{\mathbb{R}}$ be defined similarly, but allow real coefficients. And let $\mathcal{J}$ be what you get when you restrict yourself to integer coefficients. Then $\mathcal{A}_{\mathbb{R}}$ will be an n–dimensional vector space and will be an algebra, i.e., closed under multiplication. Also $\mathcal{J}$ will be a lattice in $\mathcal{A}_{\mathbb{R}}$, and will also be closed under multiplication. Each matrix A_k will be a unit of $\mathcal{J}$, in the sense that A_k^{-1} will also be in $\mathcal{J}$. This is so because, since A_k is integral with determinant 1, it satisfies an equation

$$A^n + c_1 A^{n-1} + \cdots + c_{n-1} A + I = 0$$

where the c_i are integers. Hence

$$A_k^{-1} = -(c_{n-1} I + c_{n-2} A_k + \cdots + c_1 A_k^{n-2} + A_k^{n-1})$$

and the right hand side is obviously in $\mathcal{J}$.

The entity $\mathcal{A}_{\mathbb{Q}}$ is a vector space of dimension n over $\mathbb{Q}$, and is closed under multiplication. In fact, $\mathcal{A}_{\mathbb{Q}}$ is a field: every element in it is invertible. It is a type of field known as *totally real number field.* Precisely, a totally real number field is a field generated by the rational numbers $\mathbb{Q}$ together with an element x which satisfies an equation

$$p(x) = x^n + c_1 x^{n-1} + \cdots + c_1 x + c_n = 0$$

with all c_i's in $\mathbb{Q}$. The polynomial p should be irreducible over $\mathbb{Q}$, but should have n distinct real roots.

Given a totally real number field F, there is a distinguished spanning lattice I_F in F, called the *ring of integers* of F. It consists of all elements of F which satisfy polynomials with coefficients in $\mathbb{Z}$ and main coefficient 1. It is closed under multiplication. Let U be the group of *units* of I_F, i.e., elements A of I_F such that A^{-1} is also in I_F. Then the Dirichlet unit theorem [BS], [L] guarantees that U contains n–1 elements A_k as required by Theorem 2. Other objects of the discussion can also be interpreted as appurtenances of a totally real number field.

ACKNOWLEDGMENT

The authors thank Herb Scarf, Bill Cook, Ravi Kannan and David Shallcross for fruitful discussions and the Cowles Foundation for hospitality.

REFERENCES

[BS] Z. I. BOREVICH AND I. R. SAFAREVICH, *Number Theory*. New York and London: Academic Press (1966).

[CHKM] W. COOK, M. HARTMANN, R. KANNAN AND C. MCDIARMID, On integer points in polyhedra, *Combinatorica*, submitted (1989).

[HL] A. C. HAYES AND D. G. LARMAN, The vertices of the knapsack polytope, *Discrete Applied Math.*, **6** 1983, 135–138.

[L] S. LANG, *Algebraic Number Theory*, Graduate Texts in Mathematics 100. New York: Springer Verlag (1986).

[M] D. A. MORGAN, The set of vertices of the convex hull of integer points in regions defined by particular linear inequalities, submitted to *Mathematika*.

[R] D. S. RUBIN, On the unlimited number of faces in integer hulls of linear programs with a single constraint, *Operations Research*, **18**, 1970, 940–946.

[S] A. SCHRIJVER, *Theory of Linear and Integer Programming*. Chichester: Wiley (1987).

[Sh] V. N. SCHEVCHENKO, On the number of extreme points in integer programming, *Kibernetika*, **2** 1981, 133–134.