On minimal solutions of linear Diophantine equations

Martin Henk Robert Weismantel*

Abstract

This paper investigates the region in which all the minimal solutions of a linear diophantine equation ly. We present best possible inequalities which must be satisfied by these solutions and thereby improve earlier results.

Keywords: Linear Diophantine equations, Hilbert basis, pointed rational cones.

1 Introduction

For two nonnegative integral vectors $a \in \mathbb{N}^n$, $b \in \mathbb{N}^m$, $n, m \geq 1$, let

$$\mathcal{L}(a,b) = \{(x,y) \in \mathbb{N}^n \times \mathbb{N}^m \,:\, a^\intercal x = b^\intercal y\} \tag{1.1}$$

be the set of all nonnegative solutions of the linear Diophantine equation $a^\intercal x = b^\intercal y$. Here we are interested in the *minimal solutions* of this linear Diophantine equation, where $(x,y) \in \mathcal{L}(a,b)$ is called minimal if it can not be written as the sum of two other elements of $\mathcal{L}(a,b)\backslash\{0\}$. The set of all minimal solutions is denoted by $\mathcal{H}(a,b)$. By definition we have

$$\mathcal{L}(a,b) = \left\{\sum\nolimits_{i=1}^{p} q_i h^i \,:\, q_i, p \in \mathbb{N},\, h^i \in \mathcal{H}(a,b)\right\}$$

and $\mathcal{H}(a,b)$ is a minimal subset of $\mathcal{L}(a,b)$ having this generating property.

In other words, $\mathcal{H}(a,b)$ is the *Hilbert basis* of the pointed rational cone

$$C(a,b) = \{(x,y) \in \mathbb{R}^n_{\geq 0} \times \mathbb{R}^m_{\geq 0} \,:\, a^\intercal x = b^\intercal y\}. \tag{1.2}$$

A Hilbert basis of an arbitrary pointed rational polyhedral cone $C \subset \mathbb{R}^n$ is defined as the unique minimal generating system (w.r.t. nonnegative integral combinations) of the semigroup $C \cap \mathbb{Z}^n$. Observe, that $C(a,b) \cap \mathbb{N}^{n\times m} = \mathcal{L}(a,b)$. The existence of such a system of finite cardinality was already shown by Gordan [G1873] for any rational cone. Van der Corput [Cor31] proved the uniqueness for pointed rational cones.

The set $\mathcal{H}(a,b)$ of all minimal solutions of a linear Diophantine equation has been studied for a long time in various contexts, see e.g., [Ehr79], [FT95],

*Supported by a Gerhard-Hess-Forschungsförderpreis of the German Science Foundation (DFG).

[Gre88] and the references within. The purpose of this note is to generalize a result of Lambert [Lam87] and Diaconis, Graham& Sturmfels [DGS94] by proving that the elements of $\mathcal{H}(a,b)$ satisfy a certain system of inequalities.

We assume throughout that $a = (a_1, \ldots, a_n)^T \in \mathbb{N}^n$, $b = (b_1, \ldots, b_m)^T \in \mathbb{N}^m$, $n \geq m \geq 1$, and $a_1 \leq a_2 \leq \cdots \leq a_n$, $b_1 \leq b_2 \leq \cdots \leq b_m$. It is not hard to see that

$$C(a,b) = \text{pos } \left\{ b_j e^i + a_i e^{n+j} : 1 \leq i \leq n,\ 1 \leq j \leq m \right\},$$

where pos denotes the positive hull and $e^i \in \mathbb{R}^{n+m}$ denotes the i-th unit vector. A trivial system of valid inequalities for the elements of $\mathcal{H}(a,b)$ is given by the facet defining hyperplanes of the zonotope

$$\left\{ (x,y) \in \mathbb{R}^{n+m} : (x,y)^\intercal = \sum\nolimits_{i,j} \lambda_{ij} (b_j e^i - a_i e^{n+j}),\ 0 \leq \lambda_{ij} \leq 1 \right\},$$

because it is well known (and easy to see) that the Hilbert basis of a pointed rational cone is contained in the zonotope spanned by the generators of the cone. Stronger inequalities were given by Lambert ([Lam87]) and independently by Diaconis, Graham&Sturmfels [DGS94]. They proved that every $(x,y)^\intercal \in \mathcal{H}(a,b)$ satisfies

$$\sum_{i=1}^{n} x_i \leq b_m \quad \text{and} \quad \sum_{j=1}^{m} y_j \leq a_n. \tag{1.3}$$

Here we show

Theorem 1. *Every $(x,y)^\intercal \in \mathcal{H}(a,b)$ satisfies the $n+m$ inequalities*

$$[J_l] : \quad \sum_{i=1}^{n} x_i + \sum_{j=1}^{l-1} \left\lfloor \frac{b_l - b_j}{a_n} \right\rfloor y_j \leq b_l + \sum_{j=l+1}^{m} \left\lceil \frac{b_j - b_l}{a_1} \right\rceil y_j, \quad l = 1, \ldots, m,$$

$$[I_k] : \quad \sum_{j=1}^{m} y_j + \sum_{i=1}^{k-1} \left\lfloor \frac{a_k - a_i}{b_m} \right\rfloor x_i \leq a_k + \sum_{i=k+1}^{n} \left\lceil \frac{a_i - a_k}{b_1} \right\rceil x_i, \quad k = 1, \ldots, n,$$

where $\lceil x \rceil$ ($\lfloor x \rfloor$) denotes the smallest integer not less than x (the largest integer not greater than x).

Observe, that $[J_m]$ and $[I_n]$ are generalizations of the inequalities stated in (1.3).

2 Proof of Theorem 1

In the following we denote by $\leq$ (respectively by $<$) the usual partial order, i.e., for two vectors x, y we write $x \leq y$ if for each coordinate holds $x_i \leq y_i$ and we write $x < y$ if, in addition, there exists a coordinate with $x_j < y_j$. The proof of Theorem 1 relies on the following observation.

Lemma 1. *Let $(\widehat{x}, \widehat{y})^T \in \mathcal{L}(a,b)$ and let $(x^1, y^1)^T, (x^2, y^2)^T \in \mathbb{N}^{n+m}$ such that $0 < (x^2 - x^1, y^2 - y^1)^T < (\widehat{x}, \widehat{y})^T$ and $a^T x^1 - b^T y^1 = a^T x^2 - b^T y^2$. Then $(\widehat{x}, \widehat{y})^T$ is not an element of $\mathcal{H}(a,b)$.*

Proof. Let $(z_x, z_y) = (x^2 - x^1, y^2 - y^1)$. By assumption we have $(z_x, z_y)^T, (\widehat{x} - z_x, \widehat{y} - z_y)^T \in \mathcal{L}(a,b) \backslash \{0\}$. Thus $(\widehat{x}, \widehat{y}) = (\widehat{x} - z_x, \widehat{y} - z_y) + (z_x, z_y)$ can be written as a non trivial combination of two elements of $\mathcal{L}(a,b) \backslash \{0\}$. □

Proof of Theorem 1. Let $(\tilde{x},\tilde{y})^T \in \mathcal{H}(a,b)$. By symmetry it suffices to consider only the inequalities $[J_l]$, $l = 1,\dots,m$. Let us fix an index $l \in \{1,\dots,m\}$ and let $\xi = \sum_{i=1}^n \tilde{x}_i$, $\upsilon = \sum_{j=1}^m \tilde{y}_j$. We choose a sequence of points $x^i \in \mathbb{N}^n$, $0 \le i \le \xi$, such that

$$0 = x^0 < x^1 < x^2 < \cdots < x^\xi = \tilde{x}. \tag{2.1}$$

Next we define recursively a sequence of points $y^j \in \mathbb{N}^m$, $0 \le j \le \upsilon$, by $y^0 = 0$ and $y^j = y^{j-1} + e^{d(j)}$, $j \ge 1$, where the index $d(j)$ is given by $d(j) = \min\{1 \le d \le m : y_d^{j-1} + e^d \le \tilde{y}_d\}$. Observe that here e^d denotes the d-th unit vector in $\mathbb{R}^m$. Obviously, we have

$$0 = y^0 < y^1 < y^2 < \cdots < y^\upsilon = \tilde{y}. \tag{2.2}$$

For two points $x \in \mathbb{N}^n$, $y \in \mathbb{N}^m$ let $r(x,y) = a^T x - b^T y$ and for a given point x^i let $y^{\mu(i)}$ be the unique point such that

$$r(x^i, y^{\mu(i)}) = \min\left\{ r(x^i,y^j) : r(x^i,y^j) \ge 0,\ 0 \le j \le \upsilon \right\}.$$

For abbreviation we set $r(i) = r(x^i, y^{\mu(i)})$. It is easy to see that $r(i) \in \{0,\dots, b_m - 1\}$ and

$$0 = y^{\mu(0)} \le y^{\mu(1)} \le \cdots \le y^{\mu(\xi)} = \tilde{y}. \tag{2.3}$$

Moreover, by definition of y^j we have the relation

$$r(i) \ge b_t \Longrightarrow y_j^{\mu(i)} = \tilde{y}_j,\ 1 \le j \le t. \tag{2.4}$$

So we have assigned to each $i \in \{0,\dots,\xi-1\}$ its residue $r(i)$ and now we count the number of different residues which may occur. To this end let

$$R_l = \{ i \in \{0,\dots,\xi-1\} : r(i) < b_l \},$$

and for $l+1 \le j \le m$ let

$$R_j = \left\{ i \in \{0,\dots,\xi-1\} : b_l \le r(i) < b_j,\ y_{j-1}^{\mu(i)} = \tilde{y}_{j-1},\ y_j^{\mu(i)} < \tilde{y}_j \right\}.$$

Since $\{0,\dots,\xi-1\} = \bigcup_{j=l}^m R_j$ we have

$$\sum_{i=1}^n \tilde{x}_i \le \#R_l + \sum_{j=l+1}^m \#R_j. \tag{2.5}$$

By Lemma 1, (2.1), (2.2) we have

$$\#R_l = \#\{r(i) : i \in R_l\} \le b_l. \tag{2.6}$$

We claim that for $j = l+1,\dots,m$

$$\#R_j \le \left\lceil \frac{b_j - b_l}{a_1} \right\rceil \tilde{y}_j. \tag{2.7}$$

To show this let $\zeta \in \{0,\dots,\tilde{y}_j - 1\}$ and let $x^{i_1} < \cdots < x^{i_\tau}$ be all vectors of the x-sequence (cf. (2.1)) satisfying $y_j^{\mu(i)} = \zeta$ and $i \in R_j$. By construction we have $y^{\mu(i_1)} = y^{\mu(i_2)} = \cdots = y^{\mu(i_\tau)}$ and so

$$(\tau - 1)a_1 \le a^T x^{i_\tau} - a^T x^{i_1} = r(i_\tau) - r(i_1) \le (b_j - 1) - b_l.$$

Hence $\tau \le \lceil (b_j - b_l)/a_1 \rceil$ and we get (2.7).

So far we have proved (cf. (2.5), (2.7))

$$\sum_{i=1}^{n} \tilde{x}_i \le \#R_l + \sum_{j=l+1}^{m} \left\lceil \frac{b_j - b_l}{a_1} \right\rceil \tilde{y}_j. \tag{2.8}$$

In the following we estimate the number of residues in $\{0, \dots, b_l - 1\}$ which are not contained in $\{r(i) : i \in R_l\}$.

To do this we have to extend our x-sequence. For $v \in \mathbb{N}$ let $p_v, q_v \in \mathbb{N}$ be the uniquely determined numbers with $v = p_v \xi + q_v$, $0 \le q_v < \xi$, and let

$$\overline{x}^v = p_v x^\xi + x^{q_v}.$$

Observe that $r(\overline{x}^v, y) = p_v b^T \tilde{y} - b^T y + a^T x^{q_v}$. For $s \in \{1, \dots, l-1\}$ and $t \in \{0, \dots, \tilde{y}_s - 1\}$ let $y^{s,t}$ be the point of the y-sequence (cf. (2.2)) with coordinates

$$y_s^{s,t} = t, \quad y_j^{s,t} = \tilde{y}_j,\ 1 \le j \le s-1, \quad \text{and} \quad y_j^{s,t} = 0,\ s+1 \le j \le m.$$

For such a vector $y^{s,t}$ let $\overline{x}^{\delta(s,t)}$ be the point of the $\overline{x}$-sequence such that

$$r(\overline{x}^{\delta(s,t)}, y^{s,t}) = \min \left\{ r(\overline{x}^i, y^{s,t}) : r(\overline{x}^i, y^{s,t}) \ge b_s,\ i \in \{0, \dots, \xi\} \right\}.$$

Observe that such a point $\overline{x}^{\delta(s,t)}$ exists, because $t \in \{0, \dots, \tilde{y}_s - 1\}$. Moreover, $\overline{x}^{\delta(s,t)}$ belongs to the "original" x-sequence. In particular, we have

$$b_s \le r(x^{\delta(s,t)}, y^{s,t}) < b_s + a_n. \tag{2.9}$$

Let $r_{s,t} = \{\overline{x}^i : b_s \le r(\overline{x}^i, y^{s,t}) < b_l\}$. Obviously, by (2.9) we have

$$\#r_{s,t} \ge \lfloor (b_l - b_s)/a_n \rfloor. \tag{2.10}$$

Now we study the cardinality of

$$\overline{R} = \bigcup_{s=1}^{l-1} \left\{ \bigcup_{t=0}^{\tilde{y}_s - 1} \left\{ r(\overline{x}^i, y^{s,t}) : b_s \le r(\overline{x}^i, y^{s,t}) < b_l \right\} \right\}$$

and we show

$$\#\overline{R} \ge \sum_{s=1}^{l-1} \left\lfloor \frac{b_l - b_s}{a_n} \right\rfloor \tilde{y}_s. \tag{2.11}$$

Suppose the contrary. Then, by (2.10), we can find $s, s' \in \{1, \dots, l-1\}$, $t \in \{0, \dots, \tilde{y}_s - 1\}$, $t' \in \{0, \dots, \tilde{y}_{s'} - 1\}$ and vectors $\overline{x}^v$, $\overline{x}^w$ of the $\overline{x}$-sequence such that $r(\overline{x}^v, y^{s,t}) = r(\overline{x}^w, y^{s',t'})$. We may assume $y^{s,t} < y^{s',t'}$ and therefore $\overline{x}^v < \overline{x}^w$, i.e., $v \le w$. Since

$$r(\overline{x}^v, y^{s,t}) = p_v b^T \tilde{y} - b^T y^{s,t} + a^T x^{q_v} = p_w b^T \tilde{y} - b^T y^{s',t'} + a^T x^{q_w} = r(\overline{x}^w, y^{s',t'})$$

we get $p_w \in \{p_v, p_v + 1\}$.

a) If $p_w = p_v$ then $0 < \overline{x}^w - \overline{x}^v = x^{q_w} - x^{q_v} < x^\xi$ and we can apply Lemma 1 to $(\overline{x}^v, y^{s,t})^T$, $(\overline{x}^w, y^{s',t'})^T$ which yields the contradiction $(\tilde{x}, \tilde{y}) \notin \mathcal{H}(a, b)$.

b) If $p_w = p_v + 1$ then $0 < \overline{x}^w - \overline{x}^v = x^\xi + x^{q_w} - x^{q_v}$. Since

$$a^T (x^{q_v} - x^{q_w}) = b^T \tilde{y} + b^T y^{s,t} - b^T y^{s',t'} > 0$$

we have $x^{q_w} < x^{q_v}$ and thus $0 < \overline{x}^w - \overline{x}^v < x^\xi$. Hence, also in this case we can apply Lemma 1 and obtain a contradiction.

Next we claim that

$$\overline{R} \cap \{r(i) : i \in R_l\} = \emptyset. \tag{2.12}$$

Otherwise there exist $\overline{x}^v$, $y^{s,t}$ with $b_s \leq r(\overline{x}^v, y^{s,t}) < b_l$ and $\overline{x}^i$, $y^{\mu(i)}$, $0 \leq i \leq \xi - 1$, such that $r(\overline{x}^v, y^{s,t}) = r(\overline{x}^i, y^{\mu(i)})$. Since $r(\overline{x}^v, y^{s,t}) \geq b_s$ but $y_s^{s,t} < \tilde{y}_s$ we have $y^{s,t} \neq y^{\mu(i)}$ (cf. (2.4)). Hence, we may assume $y^{s,t} < y^{\mu(i)}$ or $y^{\mu(i)} < y^{s,t}$.
a) If $y^{s,t} < y^{\mu(i)}$ then $\overline{x}^v < \overline{x}^i$ and thus $v < i < \xi$. Again, by Lemma 1 we find $(\tilde{x}, \tilde{y}) \notin \mathcal{H}(a,b)$.
b) If $y^{\mu(i)} < y^{s,t}$ then $\overline{x}^i < \overline{x}^v$. As above, it is easy to see that $p_v \in \{0,1\}$ and that in both cases Lemma 1 can be applied in order to get a contradiction.

Finally, we note that (2.6), (2.12) and (2.11) imply

$$\#R_l \leq b_l - \sum_{s=1}^{l-1} \left\lfloor \frac{b_l - b_s}{a_n} \right\rfloor \tilde{y}_s,$$

which proves inequality $[J_l]$ (cf. (2.8)). □

3 Remarks

Theorem 1 shows that the minimal solutions of a linear Diophantine equation ly in the region that one obtains from intersecting the zonotope associated with the generators of $C(a,b)$ with all the halfspaces induced by the inequalities $[I_k]$, $k = 1, \ldots, n$ and $[J_l]$, $l = 1, \ldots, m$. We believe that a stronger statement is true: every element of $\mathcal{H}(a,b)$ is a convex combination of 0 and the generators $b_j e^i + a_i e^{n+j}$ of $C(a,b)$. More formally, let

$$P(a,b) = \text{conv}\left\{0, b_j e^i + a_i e^{n+j} : 1 \leq i \leq n,\ 1 \leq j \leq m\right\}.$$

We conjecture that

Conjecture 1. $\mathcal{H}(a,b) \subset P(a,b)$.[1]

We remark that there is an example by Hosten and Sturmfels showing that if one replaces $P(a,b)$ by the "smaller" polytope $\widetilde{P}(a,b) = \text{conv}\{0, (b_j e^i + a_i e^{n+j})/\gcd(b_j, a_i) : 1 \leq i \leq n,\ 1 \leq j \leq m\}$, then $\mathcal{H}(a,b) \not\subset \widetilde{P}(a,b)$.

For $m = 1$ Theorem 1 implies the inclusion $\mathcal{H}(a,b) \subset P(a,b)$. This can easily be read off from the representation

$$P(a,b) = \left\{(x,y)^T \in \mathbb{R}^n \times \mathbb{R} :\ a^T x = b_1 y,\quad x, y \geq 0,\ \sum_{i=1}^n x_i \leq b_1\right\}.$$

It is not difficult to check that the inequalities $[I_k]$ and $[J_l]$ of Theorem 1 "without rounding" define facets of $P(a,b)$.

Proposition 1. *For $l = 1, \ldots, m$ let*

$$J_l = \left\{(x,y) \in \mathbb{R}^n \times \mathbb{R}^m : \sum_{i=1}^n x_i + \sum_{j=1}^{l-1} \frac{b_l - b_j}{a_n} y_j \leq b_l + \sum_{j=l+1}^m \frac{b_j - b_l}{a_1} y_j\right\}$$

[1] This conjecture was independently made by Hosten and Sturmfels, private communication.

and for $k = 1, \ldots, n$ let

$$I_k = \left\{ (x,y) \in \mathbb{R}^n \times \mathbb{R}^m : \sum_{j=1}^{m} y_j + \sum_{i=1}^{k-1} \frac{a_k - a_i}{b_m} x_i \le a_k + \sum_{i=k+1}^{n} \frac{a_i - a_k}{b_1} x_i \right\}.$$

Then we have $P(a,b) \subset J_l$, $P(a,b) \subset I_k$. Moreover, $P(a,b) \cap J_l$ and $P(a,b) \cap I_k$ are facets of $P(a,b)$, $1 \le l \le m$, $1 \le k \le n$.

Proof. It is quite easy to check that all vectors $b_j e^i + a_i e^{n+j}$, $1 \le i \le n$, $1 \le j \le m$, are contained in J_l, $l = 1, \ldots, m$. Moreover, the inequality corresponding to J_l is satisfied with equality by the $n + m - 1$ linearly independent points $b^l e^i + a_i e^{n+l}$, $1 \le i \le n$, $b_j e^n + a_n e^{n+j}$, $1 \le j \le l-1$, $b_j e^1 + a_1 e^{n+j}$, $l+1 \le j \le m$. The halfspaces I_k can be treated in the same way. □

Elementary considerations show that for $m = 2$ the polytope $P(a,b)$ can be written as $P(a,b) = \{(x,y)^T \in \mathbb{R}^n \times \mathbb{R}^2 : a^T x = b^T y; x, y \ge 0, (x,y)^T \in I_k, 1 \le k \le n\}$, and thus Theorem 1 and Proposition 1 imply that the conjecture is "almost true" when $m = 2$ (or respectively, for $n = 2$).

Acknowledgements. We would like to thank Robert T. Firla and Bianca Spille for helpful comments.

References

[Cor31] J.G. van der Corput, *Über Systeme von linear-homogenen Gleichungen und Ungleichungen*, Proceedings Koninklijke Akademie van Wetenschappen te Amsterdam 34, 368 - 371 (1931).

[DGS94] P. Diaconis, R. Graham, and B. Sturmfels, *Primitive partition identities,* Paul Erdös is 80, Vol. II, Janos Bolyai Society, Budapest, 1-20 (1995).

[Ehr79] E. Ehrhart, *Sur les équations diophantiennes linéaires*, C. R. Acad. Sc. Paris, t. 288, Série A, 785 – 787 (1979).

[FT95] M. Filgueiras and A.P. Tomás, *A fast method for finding the basis of non-negative solutions to a linear Diophantine equation*, J. Symbolic Computation **19**, 507 – 526 (1995).

[G1873] P. Gordan, *Über die Auflösung linearer Gleichungen mit reellen Coefficienten*, Math. Ann. 6, 23 - 28 (1873).

[Gre88] H. Greenberg, *Solution to a linear Diophantine equation for nonnegative integers*, J. of Algorithms **9**, 343 – 353 (1988).

[Lam87] J.L. Lambert, *Une borne pour les générateurs des solutions entiéres postives d'une équation diophantienne linéaire*, C.R. Acad. Sci. Paris **305**, Série I, 1987, 39–40.

Otto-von-Guericke-Universität Magdeburg, FB Mathematik / IMO, Universitätsplatz 2, D-39106 Magdeburg, Germany, {henk,weismantel}@imo.math.uni-magdeburg.de