AUSTRALASIAN JOURNAL OF COMBINATORICS
Volume **32** (2005), Pages 27–45

The Frobenius problem on lattices

Peter Pleasants

Department of Mathematics
The University of Queensland
St. Lucia, QLD 4072
Australia

Harry Ray   Jamie Simpson

Department of Mathematics and Statistics
Curtin University of Technology
GPO Box U1987 Perth, WA 6845
Australia

Abstract

It is widely known that if p and q are relatively prime positive integers then (a) the set of linear combinations of p and q with nonnegative integer coefficients includes all integers greater than $pq-p-q$, (b) exactly half the integers between 0 and $pq-p-q$ belong to this set and (c) an integer m belongs to this set if and only if $pq-p-q-m$ does not. A multidimensional version of statement (a) was recently obtained by Simpson and Tijdeman subject to a geometric condition. Here we unconditionally obtain multidimensional versions of all three statements.

1 Introduction

In a certain far off country there are only two denominations of currency, five dollar notes and eight dollar notes. With these notes the inhabitants can make up amounts of \$5, \$8, \$10, \$13, and so on. The inhabitants find they can't make up \$27 but can make up any amount greater than this. They have also noticed that exactly half the amounts between \$1 and \$26 can be made up and that an amount \$$m$ can be made up if and only if \$$(27-m)$ can't be. There are a number of references in the literature to the amounts that can be made up using denominations \$$p$ and \$$q$. The earliest known to us is a comment by Sylvester [7, p. 134 (p. 620 of *Collected Mathematical Papers, III*)], essentially stating that there are precisely $\frac{1}{2}(p-1)(q-1)$ amounts that cannot be made up when p and q are relatively prime. This was offered as the simplest case of a much more general result and was later presented as a challenge problem in [8][1]. We state the known facts as a theorem.

[1]The volume and page numbers of this problem are commonly wrongly cited in the literature.

Theorem 1.1 *Let p and q be relatively prime positive integers. We say that an integer m is* reachable *if there exist nonnegative integers a and b such that $m = ap + bq$. Then*
(a) $pq - p - q$ is the largest integer that is not reachable,
(b) exactly half the integers m in the range $0 < m < pq - p - q$ are reachable, and
(c) m is reachable if and only if $pq - p - q - m$ is not reachable.

Part (c) says that the set of reachable integers is anti-symmetric about the point $(pq - p - q)/2$, and the other two parts are a consequence of this (part (a) because 0 is clearly the smallest reachable integer). Note that it is necessary for this theorem that p and q are relatively prime, since if they had a common divisor, d say, then integers not divisible by d would be unreachable.

More generally, Frobenius raised the question of the largest amount that cannot be made up from an arbitrary set of denominations \$$p_1$,...,\$$p_k$, where $p_1, \dots, p_k$ have no common factor. There are no such clean results as Theorem 1.1(a) when $k > 2$ but there are various upper bounds for this largest amount, many of them described in [6] which has a brief history of the problem prior to 1981. Selmer and Beyer [4] and Rødseth [3] found different continued fraction algorithms for calculating the largest unreachable number when $k = 3$ and Greenberg [1] combined these approaches to obtain a linear time algorithm for this case. Later Kannan [2] showed that, for each individual k, there is a polynomial time algorithm for calculating the smallest reachable number (though he states that the problem is NP-hard if k is allowed to vary, as well as $p_1, \dots, p_k$).

Here we are concerned with increasing the dimension rather than the number of denominations. A multidimensional analogue of Frobenius' question can be formulated by taking a set S of vectors that generates (as an Abelian group) a lattice L and calling a point of L *S-reachable* if it is a nonnegative integer linear combination of vectors in S. The analogue of Frobenius' question is then to ask for the largest region such that all points of L in the region are reachable. Figure 1 illustrates the situation when $L = \mathbb{Z}^2$ and $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} = \{(1,1),(3,1),(3,6)\}$. The lattice points are the intersections of the grid lines, with the reachable points shown as solid circles. The convex hull of the set of reachable points is the outlined wedge-shaped region with its vertex at the origin (the bottom left corner). Looking like a deep shadow we see a translate of this wedge-shaped region (also outlined) that has all lattice points in its interior reachable. Between these regions is a penumbra in which half the lattice points are reachable and half unreachable.

This paper deals with the cases $|S| = \dim L$ and $\dim L + 1$. Clearly $|S| \geq \dim L$, since S generates L, and in view of what is known when $k \geq 3$ in dimension 1 we cannot expect any sharp, simply-stated results when $|S| \geq \dim L + 2$. In [5] the case $|S| = \dim L + 1$ was considered and, subject to the condition that one vector in S is a positive linear combination of the others,[2] a maximal region in which all lattice points are reachable was identified (a generalization of part (a) of Theorem 1.1). Here we obtain unconditional generalizations of all three parts of the theorem. The

[2]This condition is not emphasized in [5]. It is introduced in the middle of the first page but is not mentioned in the abstract, and Theorem 2 is described as "Sylvester for $k+1$ vectors in $\mathbb{Z}^k$".

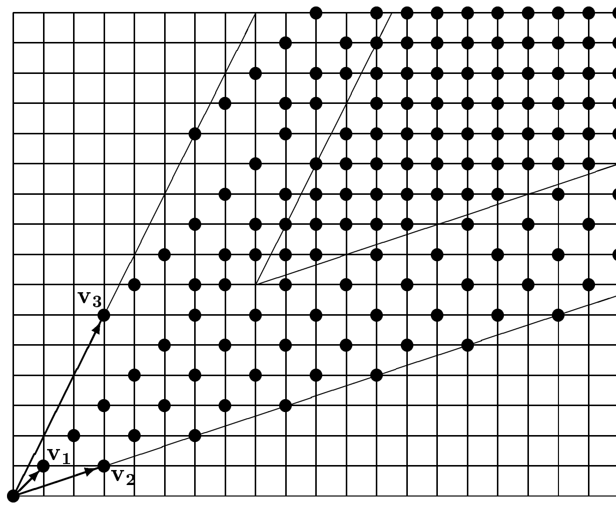

Figure 1: The reachable points (shown as solid circles) when $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} = \{(1,1),(3,1),(3,6)\}$. The lattice is $\mathbb{Z}^2$ with the origin at the lower left corner. All reachable points lie in the wedge-shaped region with its vertex at $\mathbf{0}$ and all lattice points in the interior of the translate of this region with vertex at $(8,7)$ are reachable. A proportion one half of the lattice points between the boundaries of these regions is reachable.

key is to generalize part (c), from which, as in the 1-dimensional case, the other parts can be derived.

As is clear from Figure 1, where less than half the lattice points are reachable, we can no longer expect a single centre of anti-symmetry. We in fact give two different generalizations of part (c). One replaces the global anti-symmetry centre by a network of local anti-symmetry centres, each having limited range but whose ranges cover $\mathbb{R}^n$. This is illustrated in Figure 2 where, with S as in Figure 1, the local centres are shown as two lines of small dots in the centre of the penumbral region between the boundaries of the wedges. The narrow central strip parallel to $\mathbf{v}_1$ is the region of anti-symmetry about the local centre $(4, 3.5)$ (closest to the origin). Widening this strip by replacing its upper boundary by the upper broken line gives the range of anti-symmetry about $(7, 9.5)$, and widening it by replacing its lower boundary by the lower broken line gives the range for the local centre $(8.5, 5)$. The second generalization keeps a single centre ($(4, 3.5)$ in the figures) and gives a simple characterization of the reflections of the unreachable points. Among these reflections are all the reachable points but, in general, other points too. This is illustrated in Figure 3 and will be described in more detail later.

The case $|S| = \dim L$ is trivial, but in Section 4 we formally state the results for this case too, because we use them in treating the case $|S| = \dim L + 1$ and because they provide an introduction to the form of our main result, which is Theorem 6.1.

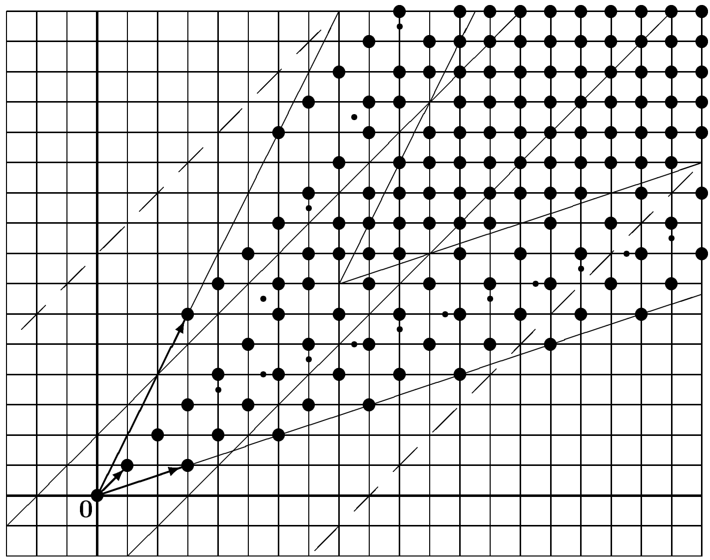

Figure 2: The local antisymmetry centres, given by Theorems 5.5 and 6.2 (with S as in Figure 1), shown as small dots midway between the wedge-shaped regions. The central narrow strip parallel to $\mathbf{v}_1$ is a region of antisymmetry for the centre $(4, 3.5)$, closest to $\mathbf{0}$. Also shown are the region of antisymmetry for the centre $(7, 9.5)$ (the central strip with its upper edge raised to the top broken line) and the region for the centre $(8.5, 5)$ (the central strip with its bottom edge lowered to the bottom broken line). The regions for the centres further from the origin are successively wider strips whose union covers the plane.

2 Notation and outline of results

For an arbitrary finite set of vectors $S = \{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ in a real vector space W we define

$$C = C(S) = \{x_1\mathbf{v}_1 + \cdots + x_k\mathbf{v}_k : x_1, \dots, x_k \geq 0\}, \tag{1}$$

the *cone* generated by S. It is an unbounded polytope with $\leq \binom{k}{n-1}$ facets. When S spans W we also define $C^\circ = C^\circ(S)$ to be the interior of C, defined by replacing the nonstrict inequalities in (1) by strict ones, and $\partial C = \partial C(S)$ to be the boundary of C. This definition stretches the notion of cone in certain cases. When S does not span W the cone will have dimension less than $\dim W$. And when the convex hull of S contains an open neighbourhood of the origin $C = W$ and $\partial C = \emptyset$. More generally, when the convex hull of S has $\mathbf{0}$ on its boundary the dimension of the lowest dimensional nonempty face of C is ≥ 1 and C ceases to have a vertex, the role of the vertex being played by the unique lowest dimensional face. For a cone C of the full dimension $\dim W$ its *facets* are the faces of codimension 1. Each facet F of C is the cone generated by the vectors of S it contains: $F = C(S \cap F)$. We denote by $H(F)$ the closed half-space containing C with F on its boundary and by $H^\circ(F)$

the corresponding open half-space. Since C is convex, $C = \bigcap H(F)$ over all facets F. Also $\partial C = \bigcup F$. Additionally, we define

$$D(F) = \bigcap_{F' \neq F} H(F'),$$

the enlarged cone got by removing the part F of the boundary of C but retaining all other bounding hyperplanes, which has as its interior

$$D^{\circ}(F) = \bigcap_{F' \neq F} H^{\circ}(F').$$

We call a facet F of $C(S)$ *ordinary* if $S \cap F$ is linearly independent and *exceptional* if $S \cap F$ is linearly dependent.

We further define

$$L = L(S) = \{a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k : a_i \in \mathbb{Z}\}, \tag{2}$$

a $\mathbb{Z}$-module in W that will be a lattice in the cases of interest to us,

$$R = R(S) = \{a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k : a_i \in \mathbb{Z}, a_i \geq 0\}, \tag{3}$$

the set of S-*reachable* points in L, and

$$\mathbf{v} = \mathbf{v}(S) = \mathbf{v}_1 + \cdots + \mathbf{v}_k. \tag{4}$$

We shall use the operators $+$ and $-$, applied to subsets of W, to denote vector addition and subtraction:

$$A \pm B = \{\mathbf{a} \pm \mathbf{b} : \mathbf{a} \in A, \mathbf{b} \in B\}.$$

The following theorem is trivial.

Theorem 2.1 *For any finite set of vectors S, $C(S)$ is the unique minimal cone to contain $R(S)$ in its closure.*

Proof: C is clearly the convex hull of R. □

The kinds of results we shall obtain when $|S| = \dim L$ or $\dim L + 1$ are:

(a) There is a vector $\mathbf{w}$ such that all lattice points in $C^{\circ} + \mathbf{w}$ are reachable but there is a relatively dense[3] set of unreachable lattice points on the boundary of $C + \mathbf{w}$.

(b) The proportion of reachable lattice points in the open region $C^{\circ} \setminus (C + \mathbf{w})$ is a half.

[3] That is, there is a radius ρ such that within a distance ρ of every point of $\partial C + \mathbf{w}$ there is an unreachable lattice point on $\partial C + \mathbf{w}$.

(c1) The sets $L \setminus R$ and $R-(R \cap \partial C)$ are images of each other by reflection in the point $\frac{1}{2}\mathbf{w}$.

(c2) There is a network of local centres of anti-symmetry whose ranges cover W.

These generalize the corresponding parts of Theorem 1.1, where $C=[0,\infty)$ and $\mathbf{w}=pq-p-q$. Part (c) of the theorem, which says that $(pq-p-q)/2$ is a global centre of anti-symmetry, is a particular case of both (c1) and (c2), since in (c1) we have $\partial C=\{0\}$, so $R \cap \partial C=\{0\}$ and $R-(R \cap \partial C)=R$.

3 General lemmas

Before specializing further we give some general lemmas, some of which are almost immediate. The only assumption we make about S in this section is that it spans W, and even that is unnecessary for Lemma 3.2.

Lemma 3.1 *(a) For each facet F of $C(S)$, $C(S)-F=H(F)$.*
(b) $C(S)-\partial C(S)=W \setminus (-C^{\circ}(S))$.

Proof: (a) Being a hyperplane with $\mathbf{0}$ on its boundary, $H(F)=H(F)+H(F)$; hence, since it contains both C and $-F$, $H(F)$ contains $C-F$. Conversely, if $\mathbf{x} \in H(F)$ choose any $\mathbf{y} \in F$ that is not on any lower dimensional face of C. Then $\mathbf{z}=\mathbf{y}+\delta\mathbf{x} \in C$, for $\delta>0$ sufficiently small and $\mathbf{x}=\delta^{-1}\mathbf{z}-\delta^{-1}\mathbf{y} \in C-F$.
(b) $C-\partial C=\bigcup_F(C-F)=\bigcup_F H(F)=W \setminus \bigcap_F(-H^{\circ}(F))=W \setminus (-C^{\circ})$. □

Lemma 3.2 *$R(S)-R(S)=L(S)$.*

Proof: Clearly $R-R \subseteq L$. Conversely, if $\mathbf{l}=a_1\mathbf{v}_1+\cdots+a_k\mathbf{v}_k \in L$ then

$$\mathbf{l}=\sum_{a_i>0} a_i\mathbf{v}_i-\sum_{a_i<0}|a_i|\mathbf{v}_i \in R-R. \quad \square$$

Lemma 3.3 *(a) For each facet F of $C(S)$, $R(S) \cap F=R(S \cap F)$.*
(b) For each ordinary facet F of $C(S)$, $R(S) \cap F=L(S \cap F) \cap C(S)$.
(c) If all facets of $C(S)$ are ordinary then

$$((R(S) \cap \partial C(S))-R(S)) \cap C(S)=R(S) \cap \partial C(S).$$

Proof: (a) Since F is a facet of C, all vectors in $S \setminus F$ lie on the same side of the hyperplane containing F (and not on it). Hence these vectors all have coefficient 0 in any nonnegative linear combination of S that lies on F.
(b) When F is an ordinary facet each point in the hyperplane of F has a unique representation as a linear combination of $S \cap F$. Hence $L(S \cap F) \cap C=R(S \cap F)$.
(c) Clearly $R \cap \partial C \subseteq ((R \cap \partial C)-R) \cap C$. Conversely, suppose $\mathbf{s} \in R \cap \partial C$, $\mathbf{r} \in R$ and F is a facet of C that $\mathbf{s}$ lies on. Then $\mathbf{s}-\mathbf{r} \in -H(F)$, so if $\mathbf{s}-\mathbf{r}$ is also in C then $\mathbf{s}-\mathbf{r} \in F$ and hence $\mathbf{r} \in (F-F) \cap C=F$. Now part (a) gives $\mathbf{s}-\mathbf{r} \in R(S \cap F)-R(S \cap F) \subseteq L(S \cap F)$, and since the facet F is ordinary part (b) gives $L(S \cap F) \cap C=R \cap F \subseteq R \cap \partial C$. □

Lemma 3.4 *If $W = U \oplus V$ is a direct sum of subspaces U and V, S and T are finite subsets of U and V, A_1 and B_1 are arbitrary subsets of U and A_2 and B_2 are arbitrary subsets of V, then*
(a) $\mathbf{v}(S \cup T) = \mathbf{v}(S) \oplus \mathbf{v}(T)$,
(b) $L(S \cup T) = L(S) \oplus L(T)$,
(c) $R(S \cup T) = R(S) \oplus R(T)$,
(d) $C(S \cup T) = C(S) \oplus C(T)$,
(e) $C^\circ(S \cup T) = C^\circ(S) \oplus C^\circ(T)$,
(f) $\partial C(S \cup T) = (\partial C(S) \oplus C(T)) \cup (C(S) \oplus \partial C(T))$,
(g) $(A_1 \oplus A_2) \setminus (B_1 \oplus B_2) = ((A_1 \setminus B_1) \oplus A_2) \cup (A_1 \oplus (A_2 \setminus B_2))$,
where we have used $\oplus$ for vector addition when the first vector is in U and the second in V.

Proof: These are all straightforward. □

4 $|S| = \dim L$

Let $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a set of $n = \dim W$ linearly independent vectors in W. This case is particularly simple because every vector in W has a unique representation as a linear combination of S. The following lemma is immediate.

Lemma 4.1 *When $|S| = \dim L$, C has n facets given by $C(S\backslash\{\mathbf{v}_i\})$, $i = 1, \ldots, n$, all of which are ordinary, and the section of C by the hyperplane through $\lambda_1\mathbf{x}_1, \ldots, \lambda_n\mathbf{x}_n$, where $\lambda_1, \ldots, \lambda_n$ are any positive real numbers, is an $(n-1)$-dimensional simplex.*

To state the main result of this section we need some notation to describe the regions of anti-symmetry. For a vector $\mathbf{s} = s_1\mathbf{v}_1 + \cdots + s_n\mathbf{v}_n \in R$ we define

$$P_{\mathbf{v}}(\mathbf{s}) = \{x_1\mathbf{v}_1 + \cdots + x_n\mathbf{v}_n : x_i - x_j < s_i + 1 \text{ for } i, j = 1, \ldots, n\}.$$

It is an infinite open prism parallel to $\mathbf{v}$ containing a neighbourhood of the origin and it is easily checked that it is invariant under reflection in the point $\frac{1}{2}(\mathbf{s} - \mathbf{v})$. The prisms for different $\mathbf{s}$ are nested in a manner corresponding to the componentwise partial ordering of R: if $\mathbf{s} < \mathbf{t}$ componentwise then $P_{\mathbf{v}}(\mathbf{s}) \subset P_{\mathbf{v}}(\mathbf{t})$.

Theorem 4.2 *Let C, L, R and $\mathbf{v}$ be as in (1), (2), (3) and (4). When S is a basis of W we have*
(a) all points of L in the interior of $C - \mathbf{v}$ are S-reachable, but no points of L on the boundary of $C - \mathbf{v}$ are S-reachable,
(c1) the sets $L \setminus R$ and $R - (R \cap \partial C)$ are images of each other by reflection in the point $-\frac{1}{2}\mathbf{v}$, and
(c2) for each $\mathbf{s} \in R \cap \partial C$, $\frac{1}{2}(\mathbf{s} - \mathbf{v})$ is a centre of anti-symmetry for the points of R in the prism $P_{\mathbf{v}}(\mathbf{s})$ and these prisms are nested in a manner corresponding to the componentwise partial ordering of $R \cap \partial C$ and cover W.

Proof: (a) If $\mathbf{l} = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n \in L \cap (C^\circ - \mathbf{v})$ then $a_i > -1$ for $i = 1, \ldots, n$ and, since the a_i's are integers, $\mathbf{l}$ is reachable. Points of L on the boundary of $C^\circ - \mathbf{v}$ have at least one a_i equal to -1, so are unreachable.
(c1) A vector $\mathbf{l} \in L$ is unreachable if and only if $a_i < 0$ for some i and is in $R - (L \cap \partial C)$ if and only if $a_i \geq 0$ for some i. Reflection in $-\frac{1}{2}\mathbf{v}$ changes each a_i to $-1 - a_i$, which has the opposite sign (counting 0 among the positives), so $L \setminus R$ and $R - (L \cap \partial C)$ are images of each other by this reflection.
(c2) Since $\mathbf{s} \in \partial C$, $s_i = 0$ for some i, so if $\mathbf{l}$ is reachable the reflection of $\mathbf{l}$ in $\frac{1}{2}(\mathbf{s} - \mathbf{v})$ has a negative coefficient of $\mathbf{v}_i$ and is unreachable (whether or not $\mathbf{l}$ is in $P_{\mathbf{v}}(\mathbf{s})$). For the other direction, we use the trivial fact that if $e_i, e_j > 0$ then

$$e_j a_i - e_i a_j < b e_j + c e_i \Rightarrow (a_j \leq -c \Rightarrow a_i < b) \tag{5}$$

(presented in this general form because we shall make further use of it later). With $e_i = e_j = c = 1$ and $b = s_i$ it shows that if $\mathbf{l} \in P_{\mathbf{v}}(\mathbf{s})$ then $a_1, \ldots, a_n$ are either all ≥ 0 or all $< s_i$. Consequently the unreachable lattice points in $P_{\mathbf{v}}(\mathbf{s})$ reflect to reachable points (not only for $\mathbf{s} \in R \cap \partial C$ but, more generally, for $\mathbf{s} \in R$). The nesting of the prisms is clear from their definition. For the covering of W, any $\mathbf{x} \in W$ can be written uniquely as

$$\mathbf{x} = x_1\mathbf{v}_1 + \cdots + x_n\mathbf{v}_n + \lambda\mathbf{v}$$

with $\lambda \in \mathbb{R}$, $x_1, \ldots, x_n \geq 0$ and $x_i = 0$ for some i. Now the point

$$\mathbf{s} = \lceil x_1 \rceil \mathbf{v}_1 + \cdots + \lceil x_n \rceil \mathbf{v}_n$$

is a reachable lattice point on the facet $C(S \setminus \{\mathbf{v}_i\})$ of C and $\mathbf{x} \in P_{\mathbf{v}}(\mathbf{s})$. □

5 $|S| = \dim L + 1$: nondegenerate case

Let $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_{n+1}\}$ be a set of $n + 1$ vectors that generate an n-dimensional lattice L in a real vector space W. There is a single linear relation between the $\mathbf{v}$'s which is rational and which, after reordering the vectors and multiplying by a scalar, can be put in the form

$$d_1\mathbf{v}_1 + \cdots + d_r\mathbf{v}_r = d_{r+1}\mathbf{v}_{r+1} + \cdots + d_{r+s}\mathbf{v}_{r+s} \tag{6}$$

with $r, s \geq 0$, $1 \leq r + s \leq n + 1$ and $d_1, \ldots, d_{r+s}$ positive integers having no common factor. (When r or s is 0 the corresponding side of the equation is to be interpreted as $\mathbf{0}$.) The ordering of the $\mathbf{v}$'s is determined up to permuting $\mathbf{v}_1, \ldots, \mathbf{v}_r$, permuting $\mathbf{v}_{r+1}, \ldots, \mathbf{v}_{r+s}$ and interchanging these two sets of vectors. Once an ordering is chosen the relation (6) is unique. We denote by

$$\mathbf{u} = \mathbf{u}(S) = d_1\mathbf{v}_1 + \cdots + d_r\mathbf{v}_r = d_{r+1}\mathbf{v}_{r+1} + \cdots + d_{r+s}\mathbf{v}_{r+s} \tag{7}$$

the common value of both sides of (6).

Having $r + s < n + 1$ in (6) is an exceptional situation, unlikely to occur if $\mathbf{v}_1, \ldots, \mathbf{v}_{n+1}$ is a randomly chosen generating set for L. We shall call such a set S *degenerate*. In this section we consider only nondegenerate sets of vectors, with $r + s = n + 1$.

Lemma 5.1 *When $|S| = \dim L + 1$ and S is nondegenate, C has rs facets all of which are ordinary. More precisely, the sets $S \cap F$, for the facets F of C, are the sets $S \setminus \{\mathbf{v}_i, \mathbf{v}_j\}$ with $1 \le i \le r$, $r+1 \le j \le n+1$.*

Proof: In the nondegerate case no n $\mathbf{v}_k$'s are linearly dependent so every facet F of C contains exactly $n-1$ $\mathbf{v}_k$'s which are linearly independent: $S \cap F = S \setminus \{\mathbf{v}_i, \mathbf{v}_j\}$, say, with $i < j$. These $\mathbf{v}_k$'s span a hyperplane that contains no positive linear combination of $\mathbf{v}_i$ and $\mathbf{v}_j$ (since these vectors are on the same side of the hyperplane) so $\mathbf{v}_i$ and $\mathbf{v}_j$ are on opposite sides of (6) and $1 \le i \le r$, $r+1 \le j \le n+1$. Conversely, for each such pair $\{\mathbf{v}_i, \mathbf{v}_j\}$, $S \setminus \{\mathbf{v}_i, \mathbf{v}_j\}$ is linearly independent and spans a hyperplane which contains $d_i\mathbf{v}_i - d_j\mathbf{v}_j$ but not $\mathbf{v}_i$ or $\mathbf{v}_j$. Hence $d_i\mathbf{v}_i$ and $d_j\mathbf{v}_j$ (and therefore the whole of C) are on the same side of the hyperplane and consequently $C(S \setminus \{\mathbf{v}_i, \mathbf{v}_j\})$ is a facet of C not containing $\mathbf{v}_i$ or $\mathbf{v}_j$. □

Lemma 5.2 *The sets $R - (R \cap \partial C)$ and $L \setminus R$ are the images of each other by reflection in the point $\frac{1}{2}(\mathbf{u} - \mathbf{v})$.*

This is illustrated in Figure 3 where, with S as in Figure 1, the solid circles are again the reachable points, the points of $R - (R \cap \partial C)$ not in R are hollow circles and the reflection point is the small dot at $(4, 3.5)$.

Proof: Without loss of generality we can suppose that $r \neq 0$. The key idea is that, by using (6), each point $\mathbf{l}$ of L can be put in the canonical form

$$\mathbf{l} = a_1\mathbf{v}_1 + \cdots + a_{n+1}\mathbf{v}_{n+1} : \begin{array}{l} a_i \ge 0 \text{ for all } i \text{ with } 1 \le i \le r, \\ a_i < d_i \text{ for some } i \text{ with } 1 \le i \le r, \end{array}$$

and that then $\mathbf{l} \in R$ if and only if $a_j \ge 0$ for $r+1 \le j \le n+1$, since by the uniqueness of (6) these coefficients can be increased only by decreasing each a_i with $i \le r$ by a multiple of d_i.

Suppose that $\mathbf{l} \in L \setminus R$ and i, j are such that $1 \le i \le r$, $0 \le a_i < d_i$, $r+1 \le j \le n+1$, $a_j < 0$. Then $\mathbf{u} - \mathbf{v} - \mathbf{l} + \mathbf{s} \in R$, where

$$\mathbf{s} = \sum_{k \neq i,j} (|a_k| + 1)\mathbf{v}_k \in R \cap F,$$

with F the facet $C(S \setminus \{\mathbf{v}_i, \mathbf{v}_j\})$, and we have used the representation of $\mathbf{u}$ on the left of (6). So the reflection, $\mathbf{u} - \mathbf{v} - \mathbf{l}$, of $\mathbf{l}$ in $\frac{1}{2}(\mathbf{u} - \mathbf{v})$ is in $R - R \cap \partial C$.

Conversely, suppose that $\mathbf{l} = \mathbf{r} - \mathbf{s}$, with $\mathbf{r} \in R$ and $\mathbf{s} \in R \cap F$ for some facet $F = C(S \setminus \{\mathbf{v}_i, \mathbf{v}_j\})$ of C. Then $a_i, a_j \ge 0$ and the contribution of $\mathbf{v}_i$ and $\mathbf{v}_j$ to $\mathbf{u} - \mathbf{v} - \mathbf{l}$ is $(d_i - a_i - 1)\mathbf{v}_i - (a_j + 1)\mathbf{v}_j$, with $d_i - a_i - 1 < d_i$ and $-(a_j + 1) < 0$. Since the coefficient of $\mathbf{v}_i$ will not be positive if it is decreased by a multiple of d_i, the coefficient of $\mathbf{v}_j$ must be negative in the canonical representation of the reflection of $\mathbf{l}$, which is therefore not in R. □

For a facet F of C and a coset $\Lambda = L(S \cap F) + \mathbf{l}$ of $L(S \cap F)$ in L, the reflection $L(S \cap F) + \mathbf{u} - \mathbf{v} - \mathbf{l}$ of Λ in $\frac{1}{2}(\mathbf{u} - \mathbf{v})$ is also a coset of $L(S \cap F)$ in L. Lemma 5.2 leads to the following useful property of such reflection pairs of cosets:

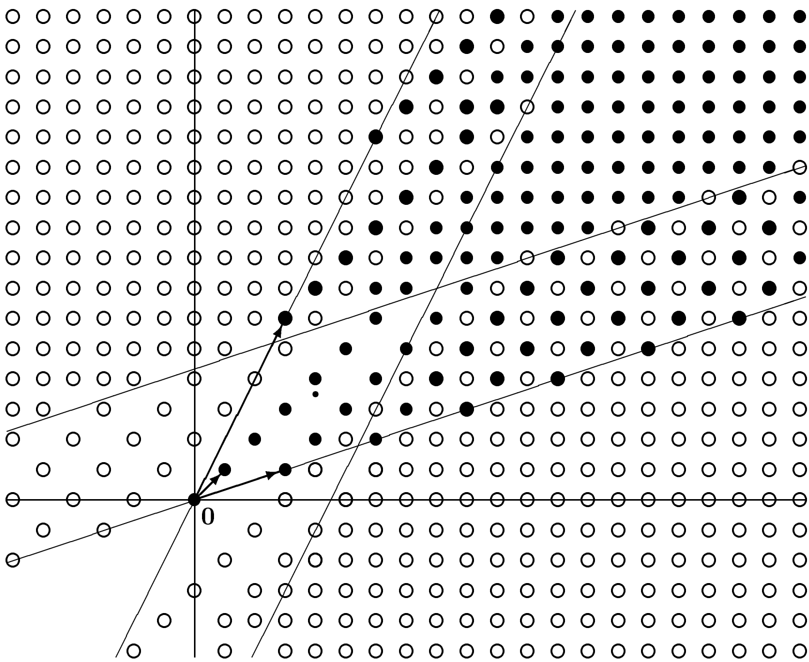

Figure 3: The global reflection centre, represented by the small dot at $(4, 3.5)$. The set S is again as in Figures 1 and 2, and the solid circles are the reachable points R, but grid lines (other than the axes) have been omitted for clarity. The hollow circles are the unreachable points in $R - (R \cap \partial C)$. So circles (solid and hollow) reflect to hollow circles or empty spaces. The edges of the cones C and $C + \mathbf{u} - \mathbf{v}$ have been extended backwards to form a parallelogram with the reflection centre at its centre.

Lemma 5.3 *For any facet F of C and any pair of cosets of $L(S \cap F)$ in L that are reflections of each other in $\frac{1}{2}(\mathbf{u} - \mathbf{v})$, one coset contains no reachable points and the other has every point which lies in the interior of $D(F) + \mathbf{u} - \mathbf{v}$ reachable.*

We note that every coset of $L(S \cap F)$ has points in $D^{\circ}(F) + \mathbf{u} - \mathbf{v}$ by an argument like that in the proof of Lemma 3.1(a): for any point $\mathbf{f} \in F$ that is not in any lower dimensional face of C there is a ball B with centre $\mathbf{f}$ contained in $D^{\circ}(F)$; then $\lambda B \subset D^{\circ}(F)$ for all $\lambda > 0$ and λB contains points of the coset when λ is large enough. So the cosets in each reflection pair are indeed of two quite distinct types.

Proof: By Lemma 3.2 $L(S \cap F) = R(S \cap F) - R(S \cap F)$, so if $\mathbf{l} \in R$ then $L(S \cap F) + \mathbf{l} \subseteq R + R(S \cap F) - R(S \cap F) \subseteq R - (R \cap \partial C)$. Hence all points in $L(S \cap F) + \mathbf{u} - \mathbf{v} - \mathbf{l}$ are unreachable, by Lemma 5.2. Conversely, if $\mathbf{l} \in D^{\circ}(F) + \mathbf{u} - \mathbf{v}$ is unreachable then, by Lemma 5.2, $\mathbf{u} - \mathbf{v} - \mathbf{l} \in -D^{\circ}(F) \cap (R - R \cap \partial C)$. But $-D^{\circ}(F) \cap (R - \partial D(F)) = \emptyset$, by Lemma 3.1(b) applied to D, and clearly $\partial C \setminus \partial D(F) \subset F$, so $\mathbf{u} - \mathbf{v} - \mathbf{l} \in R - (R \cap F)$ and, in view of Lemma 3.3(a), there are reachable points in $L(S \cap F) + \mathbf{u} - \mathbf{v} - \mathbf{l}$. $\square$

Theorem 5.4 *Let C, L, R and $\mathbf{v}$ be as in (1), (2), (3) and (4). When $|S| = \dim W + 1$, L is an n-dimensional lattice in W and $\mathbf{u}$ is as in (7) with $r + s = |S|$ we have*
(a) all points of L in the interior of $C + \mathbf{u} - \mathbf{v}$ are reachable and a point $\mathbf{l} + \mathbf{u} - \mathbf{v}$ of L on the boundary of $C + \mathbf{u} - \mathbf{v}$ is unreachable if and only if the corresponding point $\mathbf{l}$ on the boundary of C is reachable, and
(b) the proportion of points of L in C° but not in $C + \mathbf{u} - \mathbf{v}$ that are reachable is a half, in the sense that the difference between the numbers of reachable and unreachable points of L in $(C^\circ \cap B_\rho) \setminus (C + \mathbf{u} - \mathbf{v})$ is $O(\rho^{n-2})$, where B_ρ is the ball with radius ρ centred at the origin.

Proof: **(a)** Since $C^\circ + \mathbf{u} - \mathbf{v}$ is the reflection of $-C^\circ$ in $\frac{1}{2}(\mathbf{u} - \mathbf{v})$ the first part is immediate from Lemmas 5.2 and 3.1(b). For the second part, the reflection of $\partial C + \mathbf{u} - \mathbf{v}$ in $\frac{1}{2}(\mathbf{u} - \mathbf{v})$ is $-\partial C$. If $\mathbf{l} + \mathbf{u} - \mathbf{v} \in \partial C + \mathbf{u} - \mathbf{v}$ then, by Lemmas 5.1 and 3.3(c), its reflection $-\mathbf{l} \in -\partial C$ is in $R - R \cap \partial C$ if and only if $\mathbf{l}$ is reachable. So, by Lemma 5.2, this is a necessary and sufficient condition for $\mathbf{l} + \mathbf{u} - \mathbf{v}$ to be unreachable.
(b) For each facet F of C the slice $K(F) = H^\circ(F) \setminus (H(F) + \mathbf{u} - \mathbf{v})$ of W maps into itself under reflection in $\frac{1}{2}(\mathbf{u} - \mathbf{v})$ and contains only finitely many cosets of $L(S \cap F)$ in L. The difference between the numbers of points of $K(F) \cap (D^\circ(F) + \mathbf{u} - \mathbf{v}) \cap B_\rho$ in the cosets of any reflection pair in $K(F)$ is $O(\rho^{n-2})$, since discrepancies occur only within a bounded distance of the $(n-2)$-dimensional boundary of $F \cap B_\rho$. By Lemma 5.3, in one coset all these points are reachable and in the other none of them are. The regions $K(F) \cap (D^\circ(F) + \mathbf{u} - \mathbf{v})$, for the different facets F, are disjoint (because each lies outside only one $H(F) + \mathbf{u} - \mathbf{v}$) and cover the whole of $C^\circ \setminus (C + \mathbf{u} - \mathbf{v})$ with the exception of certain areas that lie within a bounded distance of the $(n-2)$-dimensional faces of C. Since these areas contain $O(\rho^{n-2})$ points of $L \cap B_\rho$, this establishes (b). □

Before giving a result of type (c2) we again need some notation to describe the regions of anti-symmetry. By the proof of Lemma 3.3(a) and the fact that all facets of F are ordinary, any $\mathbf{s} \in R \cap \partial C$ has a unique representation as $\mathbf{s} = s_1\mathbf{v}_1 + \cdots + s_{n+1}\mathbf{v}_{n+1}$, with $s_1, \ldots, s_{n+1}$ nonnegative integers. For such a point $\mathbf{s}$ we define

$$\Gamma_{ij} = \Gamma_{ij}(\mathbf{s}) = (d_j x_i - d_i x_j) - (d_j s_i + d_i)$$

and

$$P_{\mathbf{u}}(\mathbf{s}) = \{x_1\mathbf{v}_1 + \cdots + x_{n+1}\mathbf{v}_{n+1} : \Gamma_{ij} < 0 \text{ for } 1 \le i,j \le r \text{ and } r+1 \le i,j \le n+1\}.$$

Since Γ_{ij} is unchanged when multiples of $\mathbf{u}$ are added to $\mathbf{x} = x_1\mathbf{v}_1 + \cdots + x_{n+1}\mathbf{v}_{n+1}$, $P_{\mathbf{u}}(\mathbf{s})$ is an infinite open prism parallel to $\mathbf{u}$ with parallel pairs of opposite facets defined by $\Gamma_{ij}(\mathbf{s}) = 0$, $\Gamma_{ji}(\mathbf{s}) = 0$, for certain pairs, i, j. (The constraints $\Gamma_{ii}(\mathbf{s}) < 0$ are trivially satisfied and can be ignored.) This prism contains a neighbourhood of the origin and it is easily checked that it is invariant under reflection in the point $\frac{1}{2}(\mathbf{u} - \mathbf{v} + \mathbf{s})$. The prisms for different $\mathbf{s}$ are nested in a manner corresponding to the componentwise partial ordering of $R \cap \partial C$.

Theorem 5.5 *With the assumptions of Theorem 5.4, if r and s are both nonzero then, for each $\mathbf{s} \in R \cap \partial C$, $\frac{1}{2}(\mathbf{u} - \mathbf{v} + \mathbf{s})$ is a centre of anti-symmetry for the points of R in the prism $P_{\mathbf{u}}(\mathbf{s})$. These prisms are nested in a manner corresponding to the componentwise partial ordering of $R \cap \partial C$ and cover W.*

Proof: By Lemmas 3.3(a) and 5.1 any $\mathbf{s} \in R \cap \partial C$ can be represented as $\mathbf{s} = s_1\mathbf{v}_1 + \cdots + s_{n+1}\mathbf{v}_{n+1}$, with $s_1, \ldots, s_{n+1}$ nonnegative integers, $s_I = 0$ for some I with $1 \le I \le r$ and $s_J = 0$ for some J with $r+1 \le J \le n+1$. If $\mathbf{l} = a_1\mathbf{v}_1 + \cdots + a_{n+1}\mathbf{v}_{n+1}$ is a reachable point of L in canonical form then the coefficient of $\mathbf{v}_I$ in the reflection of $\mathbf{l}$ in $\frac{1}{2}(\mathbf{u} - \mathbf{v} + \mathbf{s})$ (using the representation of $\mathbf{u}$ on the left of (6)) is $< d_I$ and the coefficient of $\mathbf{v}_J$ is < 0. Hence the canonical form of this reflection has a negative coefficient of $\mathbf{v}_J$ and the reflection is unreachable (whether or not $\mathbf{l}$ is in $P_{\mathbf{u}}(\mathbf{s})$).

Conversely, (5) with $e_i = d_i$, $e_j = d_j$, $b = d_i + s_i$, $c = 1 - d_j$ shows that if $\mathbf{l} \in P_{\mathbf{u}}(\mathbf{s})$ is in canonical form then $a_k < d_k + s_k$ for $1 \le k \le r$ and hence its reflection has nonnegative coefficients of $\mathbf{v}_1, \ldots, \mathbf{v}_r$. If, further, $\mathbf{l}$ is unreachable then (5) with $e_i = d_i$, $e_j = d_j$, $b = s_i$, $c = 1$ shows that $a_k < s_k$ for $r+1 \le k \le n+1$, and hence that the reflection of $\mathbf{l}$ also has nonnegative coefficients of $\mathbf{v}_{r+1}, \ldots, \mathbf{v}_{n+1}$ and so is reachable. Again, the conclusion in this direction holds more generally for $\mathbf{s} \in R$, not just $\mathbf{s} \in R \cap \partial C$.

The nesting of the prisms is clear from their definition. For the covering of W, since $r, s \ne 0$ we can use (7) to express any $\mathbf{x} \in W$ uniquely in the form

$$\mathbf{x} = x_1\mathbf{v}_1 + \cdots + x_{n+1}\mathbf{v}_{n+1} + \lambda\mathbf{u}$$

with $\lambda \in \mathbb{R}$, $x_1, \ldots, x_{n+1} \ge 0$, $x_i = 0$ for some i with $1 \le i \le r$ and $x_j = 0$ for some j with $r+1 \le j \le n+1$. Now the point

$$\mathbf{s} = \lceil x_1 \rceil \mathbf{v}_1 + \cdots + \lceil x_{n+1} \rceil \mathbf{v}_{n+1}$$

is a reachable point on the facet of C spanned by $S \setminus \{\mathbf{v}_i, \mathbf{v}_j\}$ and $\mathbf{x} \in P_{\mathbf{u}}(\mathbf{s})$. □

6 $|S| = \dim L + 1$: general case

For the general case of a set $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_{n+1}\}$ of $n+1$ vectors in an n-dimensional real space W that generate a lattice L of the full dimension n, and hence satisfy an essentially unique relation (6) when suitably ordered, we can write W as the direct sum

$$W = U \oplus V,$$

where U and V are spanned by the sets of vectors $S_U = \{\mathbf{v}_1, \ldots, \mathbf{v}_{r+s}\}$ and $S_V = \{\mathbf{v}_{r+s+1}, \ldots, \mathbf{v}_{n+1}\}$. Then S_U is a nondegenerate set of $r+s$ vectors in the $(r+s-1)$-dimensional space U, to which the results of Section 5 apply, and S_V is a linearly independent set of vectors, to which the results of Secton 4 apply. Lemma 3.4 can then be used to combine the two.

For example, by Lemmas 3.4(f), 4.1 and 5.1, the cone C has precisely $rs+n+1-r-s$ facets: rs ordinary facets of the type $F = F_U \oplus C_V$ (for which $S \cap F$ consists of

$n-1$ linearly independent vectors) and $n+1-r-s$ exceptional facets of the type $F = C_U \oplus F_V$ (for which $S \cap F$ consists of n vectors that satisfy a linear relation).

We shall use the convention that $\mathbf{u}(T) = \mathbf{0}$ when T is linearly independent. The convenience of this is that then

$$\mathbf{u}(S) = \mathbf{u}(S_U) \oplus \mathbf{u}(S_V) = \mathbf{u}_U \oplus \mathbf{u}_V, \tag{8}$$

but it also sits well with the fact that $\mathbf{u}(T \cup \{\mathbf{0}\}) = \mathbf{0}$ when T is a linearly independent set of n vectors. Also, for any $\mathbf{x} \in W$ we shall, as in (8), write its U and V components as $\mathbf{x}_U$ and $\mathbf{x}_V$, and for any entity $E(S)$ that is a function of a set of vectors S we shall use E_U and E_V as abbreviations for $E(S_U)$ and $E(S_V)$ (as we have already been doing above). When we need to deal with two possible symmetric situations without specifying which, we shall use X to denote one of U, V and Y to denote the other.

We can now give our most general form of results of types (a), (b) and (c1).

Theorem 6.1 *Let C, L, R and $\mathbf{v}$ be as in (1), (2), (3) and (4). When $|S| = \dim W + 1$, L is an n-dimensional lattice in W and $\mathbf{u}$ is as in (7) we have*
(a) all points of L in the interior of $C + \mathbf{u} - \mathbf{v}$ are reachable and a point $\mathbf{l} + \mathbf{u} - \mathbf{v}$ of L on the boundary of $C + \mathbf{u} - \mathbf{v}$ is unreachable if and only if the corresponding point $\mathbf{l}$ on the boundary of C is in $L(S \cap F)$ for some facet F of C,
(b) the proportion of points of L in C° but not in $C + \mathbf{u} - \mathbf{v}$ that are reachable is a half (in the sense of Theorem 5.4) and
(c1) the sets $R - (R \cap \partial C)$ and $L \setminus R$ are the images of each other by reflection in the point $\frac{1}{2}(\mathbf{u} - \mathbf{v})$.

We note that (a) agrees with Theorem 5.4(a), since $L(S \cap F) \cap C = R \cap F$ when F is ordinary, by Lemma 3.3(b), and all facets are ordinary in the nondegenerate case. In general, however, $L(S \cap F) \cap C \neq R \cap F$ for exceptional facets. These facets have $|S \cap F| = \dim F + 1$ and identifying the reachable points on them is just the problem of identifying the reachable points in W with the dimension reduced by 1.

Proof: (a) The first part follows from Theorems 4.2(a), 5.4(a), Lemma 3.4(a)–(e) and (8). For the second part, let $\mathbf{m} = \mathbf{l} + \mathbf{u} - \mathbf{v}$ be a point of L on the boundary of $C + \mathbf{u} - \mathbf{v}$. Then

$$\begin{aligned}
\mathbf{m} \notin R \quad &\Longleftrightarrow \quad \mathbf{m}_X \notin R_X \text{ for } X = U \text{ or } X = V \text{ (Lemma 3.4(b)(c))}\\
&\Longleftrightarrow \quad \mathbf{l}_X \in R_X \cap \partial C_X \text{ for some } X \text{ (Theorems 4.2(a), 5.4(a))}\\
&\Longleftrightarrow \quad \mathbf{l}_X \in R_X \cap F_X \text{ for some } X \text{ and some facet } F_X \text{ of } C_X\\
&\Longleftrightarrow \quad \mathbf{l}_X \in L(S_X \cap F_X) \text{ for some } X \text{ and } F_X \text{ (Lemmas 3.3(b), 3.4(d))}\\
&\Longleftrightarrow \quad \mathbf{l} \in L((S_X \cap F_X) \cup S_Y) \text{ for some } X \text{ and } F_X \text{ (Lemma 3.4(b))}\\
&\Longleftrightarrow \quad \mathbf{l} \in L(S \cap F) \text{ for some facet } F \text{ of } C,
\end{aligned}$$

where the last step uses the fact that, by Lemma 3.4(f), the facets of C are the sets $F_X \oplus C_Y$ with $X = U$ or V.

(b) Since $C_V^\circ \subset C_V - \mathbf{v}_V$, the lattice points in $C^\circ \setminus (C + \mathbf{u} - \mathbf{v})$ are those in

$$(C_U^\circ \setminus (C_U + \mathbf{u}_U - \mathbf{v}_U)) \oplus C_V^\circ,$$

by Lemma 3.4 and (8). These points are reachable if and only if their U-components are reachable (since all points in C_V° are reachable), so by Theorem 5.4(b) and the fact that there are $O(\rho^{n+1-r-s})$ lattice points in $B_\rho \cap C_V^\circ$ the difference between the numbers of reachable and unreachable points of L in $(C^\circ \cap B_\rho) \setminus (C + \mathbf{u} - \mathbf{v})$ is $O(\rho^{r+s-3+n+1-r-s}) = O(\rho^{n-2})$.
(c1) Let $\mathbf{l}$ be a point of L and $\bar{\mathbf{l}} = \mathbf{u} - \mathbf{v} - \mathbf{l}$ its reflection in $\frac{1}{2}(\mathbf{u} - \mathbf{v})$. Then

$$\begin{aligned} \bar{\mathbf{l}} \notin R &\iff \bar{\mathbf{l}}_X \notin R_X \text{ for } X = U \text{ or } X = V \text{ (Lemma 3.4(b)(c))} \\ &\iff \mathbf{l}_X \in R_X - (R_X \cap \partial C_X) \text{ for some } X \text{ (Theorem 4.2(c1), Lemma 5.2)} \\ &\iff \mathbf{l} \in R - (R \cap \partial C) \text{ (Lemmas 3.4(b)(c)(f) and 3.2).} \quad \square \end{aligned}$$

Again we separate off the type (c2) result, as it is more complicated to describe. By Lemma 3.4(c), any $\mathbf{s} \in R$ can be written as

$$\mathbf{s} = \mathbf{s}_U \oplus \mathbf{s}_V = (s_1\mathbf{v}_1 + \cdots + s_{r+s}\mathbf{v}_{r+s}) \oplus (s_{r+s+1}\mathbf{v}_{r+s+1} + \cdots + s_{n+1}\mathbf{v}_{n+1}), \qquad (9)$$

with $s_1, \ldots, s_{n+1}$ nonnegative integers. When $\mathbf{s}$ lies on an ordinary facet of C (i.e. $\mathbf{s}_U \in R_U \cap \partial C_U$) this representation is unique, but in general there are several representations arising from different representations of $\mathbf{s}_U$. Writing an arbitrary $\mathbf{x} \in W$ as

$$\mathbf{x} = x_1\mathbf{v}_1 + \cdots + x_{n+1}\mathbf{v}_{n+1},$$

we define $P_1(\mathbf{s})$ to be the set of all points $\mathbf{x} \in P_{\mathbf{u}}(\mathbf{s}_U) \oplus P_{\mathbf{v}}(\mathbf{s}_V)$ that satisfy

$$-s_j d_i - 1 < x_i - d_i x_j < s_i + d_i \text{ for } r < i \leq r+s \text{ and } r+s < j \leq n+1 \qquad (10)$$

(inequalities similar to $\Gamma_{ij} < 0$, but with 1 in place of d_j when $j > r+s$). When $\mathbf{s}$ has more than one representation (9) we take for $P_1(\mathbf{s})$ the union of these regions over all such representations. Then $P_1(\mathbf{s})$ is an infinite prism parallel to $\mathbf{u}_U \oplus \mathbf{v}_V$ containing a neighbourhood of the origin. We also define $P_2(\mathbf{s})$ in the same way except for changing the range of i in (10) to $1 \leq i \leq r$. Finally we put

$$P(\mathbf{s}) = P_1(\mathbf{s}) \cup P_2(\mathbf{s}),$$

an infinite prism parallel to $\mathbf{u}_U \oplus \mathbf{v}_V$ containing a neighbourhood of the origin and invariant under reflection in the point $\frac{1}{2}(\mathbf{u} - \mathbf{v} + \mathbf{s})$. Then

Theorem 6.2 *With the assumptions of Theorem 6.1, if $\{r, s\} \neq \{n+1, 0\}$ then, for each $\mathbf{s} \in R \cap \partial C$, $\frac{1}{2}(\mathbf{u} - \mathbf{v} + \mathbf{s})$ is a centre of anti-symmetry for the points of R in the prism $P(\mathbf{s})$. These prisms are nested in a manner corresponding to the componentwise partial ordering of $R \cap \partial C$. If r and s are both nonzero then W is covered by the prisms $P(\mathbf{s})$ corresponding to the reachable points $\mathbf{s}$ on the ordinary facets of C and if $r + s < n+1$ it is covered by the prisms $P(\mathbf{s})$ corresponding to the reachable points $\mathbf{s}$ on the exceptional facets of C.*

Proof: When $\mathbf{s} \in R \cap \partial C$ we have $\mathbf{s}_X \in R_X \cap \partial C_X$ with X one of U or V, by Lemma 3.4(c)(f). So, by Lemma 3.4 and the proof of Theorem 4.2(c2) or 5.5,

$$\mathbf{l} \in R \Rightarrow \mathbf{l}_X \in R_X \Rightarrow \mathbf{u}_X - \mathbf{v}_X + \mathbf{s}_X - \mathbf{l}_X \in L_X \setminus R_X \Rightarrow \mathbf{u} - \mathbf{v} + \mathbf{s} - \mathbf{l} \in L \setminus R$$

(whether or not $\mathbf{l}$ is in $P(\mathbf{s})$).

In the converse direction, if $\mathbf{s} \in R$ and $\mathbf{l} \in L$ is in canonical form with $\mathbf{l}_U \in P_{\mathbf{u}}(\mathbf{s}_U)$ then the second paragraph of the proof of Theorem 5.5 shows that $\mathbf{u} - \mathbf{v} + \mathbf{s} - \mathbf{l}$ has nonnegative coefficients of $\mathbf{v}_1, \ldots, \mathbf{v}_r$. Now (5) with $e_k = d_k$ for $r < k \le r+s$, $e_k = 1$ for $r + s < k \le n+1$, $b = s_i$ or s_j and $c = 1$ shows that if $\mathbf{l} \in P_1(\mathbf{s})$ then

$$\mathbf{l}_X \in L_X \setminus R_X \Rightarrow \mathbf{u}_Y - \mathbf{v}_Y + \mathbf{s}_Y - \mathbf{l}_Y \in R_Y$$

when X is either of U or V (and Y is the other one). Also the proofs of Theorems 4.2(c2) and 5.5 show that for $\mathbf{l}_X \in P_{\mathbf{x}}(\mathbf{s}_X)$ (with $\mathbf{x} = \mathbf{u}$ or $\mathbf{v}$)

$$\mathbf{l}_X \in L_X \setminus R_X \Rightarrow \mathbf{u}_X - \mathbf{v}_X + \mathbf{s}_X - \mathbf{l}_X \in R_X,$$

since the implication in this direction requires only that $\mathbf{s}_X \in R_X$. So, by Lemma 3.4, if $\mathbf{s} \in R$ then the reflection in $\frac{1}{2}(\mathbf{u} - \mathbf{v} + \mathbf{s})$ of any unreachable point of L in $P_1(\mathbf{s})$ is reachable (whether or not $\mathbf{s} \in \partial C$). The same holds for unreachable points $\mathbf{l} \in P_2(\mathbf{s})$, by interchanging the roles of the indices $1, \ldots, r$ and $r+1, \ldots, r+s$ in the definition of canonical form.

The nesting of the prisms is clear, provided we define $\mathbf{s} \le \mathbf{t}$ to mean that there exist representations (9) of $\mathbf{s}$ and $\mathbf{t}$ for which the inequality holds componentwise.

For the coverings of W, let

$$\mathbf{x} = (x_1\mathbf{v}_1 + \cdots + x_{r+s}\mathbf{v}_{r+s}) \oplus (x_{r+s+1}\mathbf{v}_{r+s+1} + \cdots + x_{n+1}\mathbf{v}_{n+1}) = \mathbf{x}_U \oplus \mathbf{x}_V$$

be an arbitrary point of W. If r and s are both nonzero then by Theorem 5.5 we can find an $\mathbf{s}_U \in R_U \cap \partial C_U$ such that $\mathbf{x}_U \in P_{\mathbf{u}}(\mathbf{s}_U)$. Now use (6) to decrease $x_{r+1}, \ldots, x_{r+s}$ (at the expense of increasing $x_1, \ldots, x_r$) so that

$$x_i < d_i \min(x_{r+s+1}, \ldots, x_{n+1}) \text{ for } i = r+1, \ldots, r+s,$$

making the expression $x_i - d_i x_j$ in the middle of (10) negative for all relevant i and j so that the right hand inequality is automatically satisfied. Then any $\mathbf{s}_V \in R_V$ with sufficiently large coefficients has $\mathbf{x}_V \in P_{\mathbf{v}}(\mathbf{s}_V)$ and satisfies the left hand inequality of (10), so $\mathbf{x} \in P_1(\mathbf{s}_U \oplus \mathbf{s}_V)$ and $\mathbf{s}_U \oplus \mathbf{s}_V$ is a reachable point on an ordinary facet of C. Similarly, if $r + s \ne n+1$ then by Theorem 4.2(c2) there is an $\mathbf{s}_V \in R_V \cap \partial C_V$ with $\mathbf{x}_V \in P_{\mathbf{v}}(\mathbf{s}_V)$ and we can use (6) to arrange that

$$x_i > d_i \max(x_{r+s+1}, \ldots, x_{n+1}) \text{ for } i = r+1, \ldots, r+s,$$

so that the expression $x_i - d_i x_j$ in the middle of (10) is positive for all relevant i and j and the left hand inequality is automatically satisfied. Now any $\mathbf{s}_U \in R_U$ with large enough coefficients has $\mathbf{x}_U \in P_{\mathbf{u}}(\mathbf{s}_U)$ and satisfies the right hand inequality of (10). Then $\mathbf{s}_U \oplus \mathbf{s}_V$ is a reachable point on an exceptional facet of C and $\mathbf{x} \in P_1(\mathbf{s}_U \oplus \mathbf{s}_V)$. Since $P_1(\mathbf{s}) \subseteq P(\mathbf{s})$ these are enough to establish the coverings, but the argument also applies with P_2 in place of P_1, of course. □

We note that the anti-symmetry centres in this theorem all lie on the boundary of the cone $C + \frac{1}{2}(\mathbf{u} - \mathbf{v})$, midway between C and $C + (\mathbf{u} - \mathbf{v})$.

7 Loose ends

Here we comment on some matters arising from the main body of the paper.

Interconnections

Naturally Theorem 5.4 and Lemma 5.2 are a special case of Theorem 6.1 and Theorem 5.5 is a special case of Theorem 6.2.

Also when (in the case $|S| = \dim L + 1$) $\mathbf{v}_1 = \mathbf{0}$, we have $\mathbf{u} = \mathbf{0}$, so Theorem 6.1(a)(c1) gives Theorem 4.2(a)(c1) as a special case and Theorem 6.2 gives Theorem 4.2(c2) as a special case. (From this point of view the case $|S| = \dim L$ is extremely degenerate.)

When $\{r, s\} = \{n+1, 0\}$ (6) shows that $\mathbf{0}$ is in the interior of the convex hull of S and hence that $C = W$. In that case Theorem 5.4(a) tells us that every lattice point is reachable. When $n = 1$ this even adds to Theorem 1.1, giving the not quite trivial result that when one of the two denominations is negative then every (positive or negative) integral sum of money can be made up.

With $\{r, s\} = \{1, n\}$ Theorem 5.4(a) gives Theorem 2 of [5].

More vectors

In the light of what is known about the ordinary 1-dimensional Frobenius problem it would not seem possible to obtain any such simply stated result as Theorem 6.1 for sets of more than $n + 1$ vectors in an n-dimensional lattice. Instead one might look for extensive, but non-optimal, regions in which all lattice points are reachable. It might also be possible to use methods like those of [2] to obtain, for fixed n and k, polynomial time algorithms to compute the maximal translate of C in which all lattice points are reachable.

Comparing cones

Theorem 2.1 identifies the minimal cone that contains all reachable lattice points and Theorem 6.1(a) identifies a cone that contains no unreachable lattice points in its interior but has unreachable lattice points on its boundary. Both this fact and inspection of Figure 1 suggest that the cone of Theorem 6.1 is in some sense maximal. This sense can be made precise by putting a suitable partial ordering on the set of all cones in W. The partial ordering should extend to all cones (open, closed or neither) with vertices anywhere in W. The partial ordering we chose is to order primarily on the solid angle at the vertex and secondarily, for cones with the same solid angle, by inclusion. With this definition we have

Theorem 7.1 *The cone $C + \mathbf{u} - \mathbf{v}$ of Theorem 6.1 is the unique maximal cone such that all points of L in its interior are S-reachable.*

Proof: Two elementary facts are, first, that if the solid angle of a cone C_1 is less than that of a cone C_2, or the solid angles are equal but C_2 is not a translate of C_1, then a positive proportion of all lattice points lies in $C_2^\circ \setminus C_1$ (because the solid angle of a cone measures the proportion of space it covers) and, second, that if C_2 is a translate of C_1 but is not contained in C_1 then it contains in its interior an infinite segment of some edge of C_1. By the first of these facts, any cone that is not a translate of C but whose solid angle is greater than or equal to that of C contains in its interior lattice points not in C, which by Theorem 2.1 are unreachable. And by the second fact, a cone that is a translate of C but not contained in $C + \mathbf{u} - \mathbf{v}$ contains in its interior an infinite segment of some edge $\{\lambda \mathbf{v}_i + \mathbf{u} - \mathbf{v} : \lambda \geq 0\}$ of $C + \mathbf{u} - \mathbf{v}$. This segment contains lattice points $m\mathbf{v}_i + \mathbf{u} - \mathbf{v}$, for sufficiently large m, which by Theorem 6.1(a) are also unreachable. Hence $C + \mathbf{u} - \mathbf{v}$ is the unique maximal cone containing no unreachable lattice points in its interior. □

Our reason for introducing solid angles is that if cones are ordered by inclusion alone then maximal cones containing no unreachable lattice points in their interiors are not unique. For example, in Figure 1 the wedge $C(\{\mathbf{v}_1, \mathbf{v}_2\}) + (3.5, 5.5)$ contains no unreachable lattice points in its interior yet is not contained in $C(\{\mathbf{v}_2, \mathbf{v}_3\}) + \mathbf{u} - \mathbf{v}$, the maximal such wedge. However, it sacrifices infinitely many reachable points inside the maximal wedge to gain only 8 reachable points outside, and does not detract from our feeling that the maximal wedge is the significant one.

The relationship between (c1) and (c2)

The fact that Theorems 5.5 and 6.2 do not apply when $\{r, s\} = \{n+1, 0\}$ is inevitable, as in that case all points of L are reachable, as already mentioned, and there can be no centres of local anti-symmetry. In this respect Theorem 6.2 is slightly less general than Theorem 6.1(c1).

From the proofs of Theorems 4.2(c2), 5.5 and 6.2 we see that the restricted ranges are necessary in the direction unreachable-to-reachable only. In the other direction we have, without restriction: *if* $\mathbf{l} \in R$ *then for every* $\mathbf{s} \in R \cap \partial C$ *the reflection of* $\mathbf{l}$ *in* $\frac{1}{2}(\mathbf{u} - \mathbf{v} + \mathbf{s})$ *is unreachable.* In the unreachable-to-reachable direction, since the prisms in these theorems cover W we have: *if* $\mathbf{l} \in L \setminus R$ *then there is some* $\mathbf{s} \in R \cap \partial C$ *such that the reflection of* $\mathbf{l}$ *in* $\frac{1}{2}(\mathbf{u} - \mathbf{v} + \mathbf{s})$ *is reachable.* The combination of these two italicized statements is equivalent to (c1) (at least for the cases with $\{r, s\} \neq \{n + 1, 0\}$).

We should mention here that the prisms $P(\mathbf{s})$ are not maximal regions of anti-symmetry. For example, $(C^\circ + \mathbf{u} - \mathbf{v}) \cup P(\mathbf{0}) \cup (-C^\circ)$ is a region of anti-symmetry about $\frac{1}{2}(\mathbf{u} - \mathbf{v})$ larger than $P(\mathbf{0})$, and there are similar regions (consisting of two opposing cones joined by a prismatic neck) for the other centres. This shows in particular that there is no uniquely determined direction for the axes of prisms that are regions of anti-symmetry: there are non-empty prisms that are regions of anti-symmetry in any direction close enough to $\mathbf{u}_U \oplus \mathbf{v}_V$. Even if we restrict attention to regions that are convex prisms, it is not easy to determine whether the prisms $P(\mathbf{s})$ are in some sense maximal.

The case $\{r, s\} = \{1, n\}$

The requirement of [5] that one $\mathbf{v}_i$ is a positive linear combination of the others is, with our present notation, equivalent to $\{r, s\} = \{1, n\}$. How restrictive is this? For one thing, it restricts C to having only n facets (its cross-section is a simplex) whereas in general it can have as many as $\lfloor (n+1)^2/4 \rfloor$ facets when r and s are nearly equal. We can also calculate the probability of the condition holding when the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_{n+1}$ are randomly chosen in an isotropic manner (so that the vectors $\mathbf{v}$ and $-\mathbf{v}$ have an equal chance of being picked). We could either pick vectors from a preassigned lattice L, discarding nongenerating sets, or pick arbitrary vectors subject to the requirement that at every stage $\mathbf{v}_1, \ldots, \mathbf{v}_i$ span a subspace of dimension at least $i-1$ and that whenever $\mathbf{v}_i$ is chosen in the subspace spanned by $\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}$ then it is chosen to be a rational linear combination of them. In either case the resulting set satisfies a nontrivial rational linear relation

$$c_1\mathbf{v}_1 + \cdots + c_{n+1}\mathbf{v}_{n+1} = \mathbf{0}.$$

Disregarding the degenerate case when some c_i's vanish (which has probability zero), we can normalize so that $c_1 = 1$, and then all patterns of signs for $c_2, \ldots, c_{n+1}$ are equally likely. There are 2^n possible allocations of sign in all and, for $n \geq 2$, $n+1$ of them correspond to the $\{1, n\}$ case, which therefore has probability $(n+1)/2^n$.

When $n = 1$ there are two nondegenerate possibilities for $\{r, s\}$: $\{1, 1\}$, when p and q have the same sign, and $\{2, 0\}$, when p and q have opposite signs so that $C = \mathbb{R}$ and all integers are reachable. Each has probability $\frac{1}{2}$. When $n = 2$ there are also two possibilities: $\{1, 2\}$ and $\{3, 0\}$ (when $C = W$), with probabilities 3/4 and 1/4. But when $n = 3$ there are the three possibilities $\{1, 3\}$, $\{2, 2\}$, $\{4, 0\}$, of which the last is the case $C = W$. A typical example of the second is when $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ lie along the generators of a square pyramid. The probabilities of these three cases are 1/2, 3/8, 1/8. The corresponding probabilities for $n = 4$ are 5/16, 5/8, 1/16, and the case $\{r, s\} = \{1, 4\}$ is no longer among the most probable.

Validating generating sets

Mention above of discarding sets of $n+1$ vectors that do not generate L raises the question of how to determine whether a set of vectors in a lattice generates the lattice. Our point of view throughout has been that we start with a set of vectors S, which we are told generate a lattice, and we investigate the reachable points of that lattice. But what if we are interested in a specific lattice Λ ($\mathbb{Z}^n$ say) and are given a set S of vectors in Λ? How can we decide whether $L(S) = \Lambda$ so that our results apply? The following theorem gives a general criterion for a set of vectors in a lattice to generate it.

Theorem 7.2 *Let S be a subset of an n-dimensional lattice Λ. A necessary and sufficient condition for S to generate Λ is that the set of integer indices*[4]

$$\{[\Lambda : L(T)] = \det L(T)/\det\Lambda : T \subseteq S,\ |T| = n\} \tag{11}$$

[4] We use the convention that $[\Lambda : L(T)] = 0$ when T is contained in a subspace of dimension less than n.

has no common divisor.

Here $\det \Lambda$ is the volume of the fundamental region[5] of Λ. Note that this theorem does not require S to be finite. Nor does it require $|S| \geq n$, if we adopt the natural convention that the empty set has common divisor 0.

Proof: If $T \subseteq S$ then $L(T) \subseteq L(S)$ and $[\Lambda : L(T)] = [\Lambda : L(S)][L(S) : L(T)]$. So the integers (11) are all divisible by $[\Lambda : L(S)]$, and if S does not generate Λ then $[\Lambda : L(S)] > 1$. Conversely, suppose the integers (11) are all divisible by a prime p. The quotient Abelian group $\Lambda/p\Lambda$ is isomorphic to $\mathbb{F}_p^n$, where $\mathbb{F}_p$ is the finite field with p elements, and the fact that each $\det L(T)/\det \Lambda$ is divisible by p means that the image in $\mathbb{F}_p^n$ of every n-element subset of S is linearly dependent. There is certainly one subset of S whose image in $\mathbb{F}_p^n$ is linearly independent: the empty set. (When all vectors of S are in $p\Lambda$ this is the only such subset.) Let T be a maximal such subset. Then $|T| < n$ and the image of T spans a proper subspace Z of $\mathbb{F}_p^n$ which contains the images of all vectors in S. Now $[\Lambda : L(S)] \geq [\mathbb{F}_p^n : Z] = p^{n-\dim Z} > 1$, so $L(S) \neq \Lambda$. $\square$

References

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

[2] R. Kannan, Lattice translates of a polytope and the Frobenius problem, *Combinatorica* **12** (1992), 161–177.

[3] Ø. Rødseth, On a linear diophantine problem of Frobenius, *J. reine angew. Math.* **301** (1978), 171–178.

[4] E. S. Selmer and Ö. Beyer, On the linear diophantine problem of Frobenius in three variables, *J. reine angew. Math.* **301** (1978), 161–170.

[5] R. J. Simpson and R. Tijdeman, Multi-dimensional versions of a theorem of Fine and Wilf and a formula of Sylvester, *Proc. Amer. Math. Soc.* **131** (2003), 1661–1671.

[6] C. Smoryński, Skolem's solution to a problem of Frobenius, *Math. Intelligencer* **3** (3) (1981), 123–132.

[7] J. J. Sylvester, On subinvariants, i.e. semi-invariants to binary quantics of an unlimited order, *Amer. J. Math.* **5** (1882), 79–136 (*Collected Mathematical Papers, Vol. III*, Cambridge U. P., London, 1909, no. 67, pp. 568–622).

[8] J. J. Sylvester, *Educational Times* **37** (1884), 26.

(Received 4 Mar 2003; revised 20 Jan 2005)

[5] Some authors define $\det \Lambda$ to be the square of this volume.