Journal of Integer Sequences, Vol. 8 (2005), Article 05.5.2

On the Number of Representations of an Integer by a Linear Form

Gil Alon and Pete L. Clark
1126 Burnside Hall
Department of Mathematics and Statistics
McGill University
805 Sherbrooke West
Montreal, QC H3A 2K6
Canada
alon@math.mcgill.ca
clark@math.mcgill.ca

Abstract

Let $a_1, \ldots, a_k$ be positive integers generating the unit ideal, and j be a residue class modulo $L = \operatorname{lcm}(a_1, \ldots, a_k)$. It is known that the function $r(N)$ that counts solutions to the equation $x_1a_1 + \ldots + x_ka_k = N$ in non-negative integers x_i is a polynomial when restricted to non-negative integers $N \equiv j \pmod{L}$. Here we give, in the case of $k = 3$, exact formulas for these polynomials up to the constant terms, and exact formulas including the constants for $\mathfrak{q} = \gcd(a_1, a_2) \cdot \gcd(a_1, a_3) \cdot \gcd(a_2, a_3)$ of the L residue classes. The case $\mathfrak{q} = L$ plays a special role, and it is studied in more detail.

1 Introduction

We begin with some notation.

For $a \in \mathbb{Z}$ and $n \in \mathbb{Z}^+$, let $\langle a \rangle_n$ denote the unique non-negative integer that is congruent to $a \pmod{n}$ and less than n. We view this as giving a map $\langle\ \rangle : \mathbb{Z}/n\mathbb{Z} \to \{0, \ldots, n-1\}$, and the convention is that operations appearing inside the brackets are performed in the ring $\mathbb{Z}/n\mathbb{Z}$. In particular, if $\gcd(b, n) = 1$, then $\langle b^{-1} \rangle_n$ is the unique integer N in the interval $[0, n-1]$ such that $Nb \equiv 1 \pmod{n}$.

Let $a_1, \ldots, a_k$ be positive integers and $N \in \mathbb{N}$ be a non-negative integer. We are interested in the quantity

$$r(N) = r(a_1, \ldots, a_k; N) = \#\{(x_1, \ldots, x_k) \in \mathbb{N}^k \mid x_1a_1 + \ldots + x_ka_k = N\}.$$

We say that $r(N)$ counts the number of representations of N by the system of weights $(a_1, \ldots, a_k)$.

A system of weights is $(a_1, \ldots, a_k)$ is **primitive** if $\gcd(a_1, \ldots, a_k) = 1$. Given a system $(a_1, \ldots, a_k)$ of weights with $\gcd(a_1, \ldots, a_k) = d$, clearly $r(a_1, \ldots, a_k; N) = 0$ unless N is a multiple of d. And if $N = dM$, say, then we have $r(a_1, \ldots, a_k; dM) = r(\frac{a_1}{d}, \ldots, \frac{a_k}{d}; M)$. So we may, and shall, consider only the primitive case.

Issai Schur showed long ago that there is a simple asymptotic formula for $r(N)$ (Theorem 2.1(b)). We are interested in the problem of giving an "exact formula." When $k = 2$ this is indeed possible: we have **Popoviciu's formula** [7]

$$r(a, b; N) = \frac{N}{ab} - \left\{\frac{\langle a^{-1}\rangle_b N}{b}\right\} - \left\{\frac{\langle b^{-1}\rangle_a N}{a}\right\} + 1, \tag{1}$$

where $\{x\} = x - \lfloor x \rfloor$ denotes the fractional part of x.

Some consequences: it follows that $r(N)$ is the sum of a polynomial $P(N) = \frac{N}{ab}$ and an ab-periodic function $A(N)$. Moreover, for all N, $-1 < A(N) \leq 1$, so

$$\left\lfloor \frac{N}{ab} \right\rfloor \leq r(a, b; N) \leq \left\lfloor \frac{N}{ab} \right\rfloor + 1.$$

More precisely, the minimum value of A is $1 - \frac{b-1}{b} - \frac{a-1}{a}$, attained when

$$\langle a^{-1}N\rangle_b \equiv -1 \pmod{b}, \ \langle b^{-1}N\rangle_a \equiv -1 \pmod{a}.$$

The unique solution in the interval $[0, ab)$ is $N_0 = (a-1)(b-1) - 1$. A straightforward calculation gives $r(N_0) = 0$; alternately, just observe that $A(N_0) < 0$, $\frac{N_0}{ab} < 1$, and $r(N_0) \in \mathbb{N}$. Also, for $N_0 < N < ab$ we have $A(N) > A(N_0)$ and $P(N) > P(N_0)$, so $r(N) > r(N_0) = 0$; and since we clearly have $r(N) \geq 1$ for $N \geq ab$, it follows that N_0 is the largest value of N for which $r(N) = 0$.

For any primitive set $(a_1, \ldots, a_k)$ of weights, one defines the **Frobenius number** $\mathfrak{f}(a_1, \ldots, a_k)$ to be the largest N such that $r(N) = 0$. That such a number exists is not *a priori* obvious, but follows from Theorem 2.1(b) (or from Theorem 7.4). Thus one of the merits of Popoviciu's formula is that from it we can "read off" the Frobenius number for two weights:[1]

$$\mathfrak{f}(a_1, a_2) = (a_1 - 1)(a_2 - 1) - 1. \tag{2}$$

The ideal would be to give a formula for $r(N)$ in the general case that is "as satisfactory" as (1) in the $k = 2$ case. This is probably impossible, and subject to a formalization of "as satisfactory" may even be provably impossible: we propose the heuristic that a "sufficiently satisfactory" formula for $r(a_, \ldots, a_k; N)$ should (as above) lead to an exact formula for the Frobenius number $\mathfrak{f}(a_1, \ldots, a_k)$. However, there is no known exact formula for the Frobenius number when $k \geq 3$, although $\mathfrak{f}$ can be computed in polynomial time for fixed k [5]. When k is included as a parameter, it is known [8] that the problem of computing $\mathfrak{f}(a_1, \ldots, a_k)$ is NP-hard!

[1]Admittedly this formula is easily derived in many other ways; see Section 7.

Still, one observes in the $k = 2$ case that by restricting to values of N in a fixed residue class j modulo $a_1 a_2$, the function $r(a_1, a_2; N)$ is a polynomial in N. Equivalently, for all j, the function $r_j(n) = r(a_1, a_2; j + a_1 a_2 n)$ is a polynomial. This turns out to hold true in the general case: the function $r_j(n) = r(j + a_1 \cdots a_k n)$ is a polynomial function, so we can exchange the one problem of finding a formula for $r(N)$ for the $a_1 \cdots a_k$ problems of finding the coefficients of these polynomials.

The problem of finding a formula for $r_j(n)$ (for any given j) seems not to have received much attention as such. The reasons for this may be as follows: first, "all one has to do" is to compute the coefficients in the partial fraction decomposition the rational function $R(x)$ of equation (5). More precisely, the method of partial fractions leads to a decomposition of $r(N)$ into a finite sum of terms – indexed by the poles of $R(x)$ – in which the term $f_0(n)$ corresponding to the pole at $x = 1$ has the highest order of magnitude. Moreover the computation of $f_0(n)$ is relatively tractable; indeed [1] gives an exact formula for $f_0(n)$ in the general case. Computation of the other terms is more involved, but in the case in which the weights are coprime in pairs, the difference $r(n) - f_0(n)$ is a periodic function, so knowing $f_0(n)$ is enough to compute each polynomial $r_j(n)$ up to a constant. This does not hold in the general case, and here the method of partial fractions, will, in the words of [9], "entail an enormous amount of bookkeeping." The unpleasantness of these calculations, together with the existence of a formula due to Johnson (11) which in the case of $k = 3$ reduces the computation of the Frobenius number $\mathfrak{f}(a_1, a_2, a_3)$ to the pairwise coprime case, seems to have focused attention away from the problem of computation of the $r_j(n)$ in the general case.

In this paper we derive explicit formulas for $r_j(a_1, a_2, a_3; n)$ up to a constant for a general primitive system (a_1, a_2, a_3). In contrast to the methods described above, our approach is mostly independent of the partial fraction decomposition. Moreover, in a sense that will shortly be made precise, our results *improve* as the weights become "less and less pairwise coprime." In particular, for a certain class of weights (called **extremal**), we get exact formulas for $r_j(n)$ for all j.

2 Statement of the Main Results

Let $(a_1, \ldots, a_k)$ be a primitive set of weights, $\mathcal{P} = a_1 \cdots a_k$ and $L = \mathrm{lcm}(a_1, \ldots, a_k)$. For $0 \le j < L$, let $r_j(n) = r(j + Ln)$.

Theorem 2.1. *For any primitive set of weights $(a_1, \ldots, a_k)$, we have*

(a) For each j, the function $n \mapsto r_j(n)$ is the restriction to $\mathbb{N}$ of a degree $k-1$ polynomial with rational coefficients.

(b)

$$r(N) \sim \frac{N^{k-1}}{(k-1)!(a_1 \cdots a_k)}.$$

Remark: Part (b) is due to Issai Schur and has been rediscovered many times since. Part (a) is also well known. Probably the quickest proof is via the method of generating functions, as we shall have reason to recall. On the other hand, we shall also give combinatorial/geometric proofs of part (b), and of part (a) in the case $k = 3$.

The remainder of the results concern only the case $k = 3$. We introduce the following additional notation:

For any ordering (i, j, k) of the set $\{1, 2, 3\}$, let $d_k = \gcd(a_i, a_j)$, and put $\mathfrak{q} = d_1 d_2 d_3$.

The following are all consequences of primitivity:

(i) the d_i's are pairwise coprime;

(ii) $\mathfrak{q}L = \mathcal{P}$;

(iii) $\mathfrak{q}$ divides L.

For $0 \le j < L$, define $R_j(N) = r_j(\frac{N-j}{L})$. The point of this change of variables is so that $R_j(N) = r(N)$ when $N \equiv j \pmod{L}$.

Theorem 2.2.

$$r_0(n) = \frac{L}{2\mathfrak{q}} n^2 + \frac{a_1 d_1 + a_2 d_2 + a_3 d_3}{2\mathfrak{q}} n + 1.$$

Remark: This result appears in several places in the case where the weights are pairwise coprime, but we were not able to find the general case in the literature. Again we give an elementary geometric proof, using a version of Pick's Theorem.

For $0 \le j < L$, define

$$y_i(j) = \langle a_i^{-1} j \rangle_{d_i},$$

$$y(j) = y_1(j)a_1 + y_2(j)a_2 + y_3(j)a_3.$$

(The function $j \mapsto y(j)$ has the effect of projecting a complete system of residues modulo L onto a set of $\mathfrak{q}$ distinct "good" residue classes modulo L, which form a complete system of residues modulo $\mathfrak{q}$.)

Also define polynomials $Q_j(N), P_j(N)$ as follows:

$$Q_j(N) = r_0\left(\frac{N - y(j)}{L}\right),$$

$$P_j(N) = Q_j(N) - Q_j(j) + r(j).$$

The following is our main result:

Theorem 2.3. *For all $0 \le j < L$*

(a) For all $N \equiv y(j) \pmod{L}$, we have $r(N) = R_j(N) = Q_j(N)$.

(b) For all $N \equiv j \pmod{L}$, we have $r(N) = R_j(N) = P_j(N)$.

To illustrate Theorem 2.3 we consider the case of *Chicken McNuggets*. Recall that they are sold in packs of 6, 9 and 20, and a now classic brainteaser asks, "What is the largest number of Chicken McNuggets you *can't* buy?" or more succinctly, "What is $\mathfrak{f}(6, 9, 20)$?" An answer to this question follows directly from some material we shall present in Section 7. Here we are concerned with the problem of computing $r(6, 9, 20; N)$.

Example 1: We will compute the number of ways to buy $1,080,005$ Chicken McNuggets, or $r(6, 9, 20; 1,080,005)$. We have $(d_1, d_2, d_3) = (1, 2, 3)$, $\mathfrak{q} = 6$ and $L = 180$. We shall compute $R_j(N)$ for N congruent to $1080005 \equiv 5 \mod L$. We have $y_1(5) = 0$, $y_2(5) = 1$, $y_3(5) = 1$, so $y(5) = 29$. By Theorem 2.3 we get

$$Q_5(N) = R_{46}(N) = \frac{1}{2160}N^2 + \frac{13}{1080}N + \frac{113}{432}$$

and

$$R_5(N) = Q_5(N) - Q_5(5) + r(5).$$

Since clearly $r(5) = 0$, we get

$$R_5(N) = \frac{1}{2160}N^2 + \frac{13}{1080}N - \frac{31}{432}.$$

Thus there are $r(1,080,005) = R_5(1,080,005) = 540,018,000$ ways to buy $1,080,005$ Chicken McNuggets.

Example 2: We will compute $r(6, 9, 20; 1,000,000)$. We have $1000000 \equiv 100 \pmod L$, so we compute $y_1(100) = 0$, $y_2(100) = 0$, $y_3(100) = 2$, $y(100) = 40$. Applying Theorem 2.3, we get

$$Q_{100}(N) = \frac{1}{2160}N^2 + \frac{1}{540}x + \frac{5}{27}$$

and

$$R_{100}(N) = Q_{100}(N) - Q_{100}(100) + r(100).$$

This last quantity can be evaluated using the identities

$$r(100) = r(80) + r(6, 9; 100) = \ldots = \sum_{i=0}^{5} r(6, 9; 100 - 20i),$$

which divide up the set of all representations of 100 by $(6, 9, 20)$ according to the number of times the weight 20 is used. Clearly $r(6, 9, 100 - 20i) = 0$ unless $100 \equiv 20i \pmod 3$, i.e., unless $i = 2$ or $i = 5$. Certainly $r(6, 9; 0) = 1$, and one sees easily (by Popoviciu's formula, or otherwise) that $r(6, 9; 60) = r(2, 3; 20) = 4$.

Thus $r(100) = 5$. In fact $Q_{100}(100) = 5$ as well, so that we have $R_{100}(N) = Q_{100}(N) = R_{40}(N)$. Thus the answer is $r(1,000,000) = 462,964,815$.

As mentioned above, the strength of Theorem 2.3 is inversely proportional to the size of $\frac{L}{\mathfrak{q}}$: we know that $r(N)$ is given by a (possibly) different quadratic polynomial on each residue class of $N \pmod L$, and part (a) gives an exact formula for the $\mathfrak{q}$ "good" residue classes. For the other classes, we are giving the leading term (always $\frac{1}{2\mathcal{P}}$) and the linear term but not the constant term. As the above examples show, the amount of computation needed to compute the constant term depends upon the size of the residue class j: if j is nearly as large as L, then it would require only about twice as much calculation to compute both $r(j)$ and $r(j+L)$ and interpolate, but if j is small the above method is much faster. In particular, we certainly have $r(j)=0$ if $0<j<\min(a_1,a_2,a_3)$ (as in Example 1), so we have rather more exact formulas than was originally advertised.

Note that in the case $\mathfrak{q}=L$ we are getting exact formulas for *every* residue class; in this case (and only in this case) we claim Theorem 2.3 as the analogue of Popoviciu's formula for $k=3$. And indeed it is "sufficiently satisfactory" in the above sense: it can be used to derive an exact formula for $\mathfrak{f}(a_1,a_2,a_3)$. This discussion is carried out in Section 7.

Unfortunately our use of generating functions in the proof of Theorem 2.3(b) seems to be essential, so in order to make our work self-contained, we begin with a review of the generatingfunctionological approach.

3 Quasi-polynomials, generating functions and partial fractions

A function $r:\mathbb{N}\to\mathbb{C}$ for which there exists $L\in\mathbb{Z}^+$ such that for each j, $0\le j<L$, $r_j(n)=r(j+Ln)$ is a polynomial function $P_j(n)$ is called a **quasi-polynomial** with period L. We define the degree of a quasipolynomial as the maximum of the degrees of the P_j's.

Consider the collection $Q=Q(L,d)$ of all quasipolynomials of period L and degree at most $d-1$. Evidently Q is a $\mathbb{C}$-subspace of the vector space of all functions $r:\mathbb{N}\to\mathbb{C}$. Moreover, since each polynomial P_j is uniquely determined by the values of r at $j,j+L,\ldots,j+(d-1)L$, an element of Q is uniquely specified by its first dL values, so $\dim Q=dL$.

In this case, the implied basis is given by L applications of the Lagrange interpolation theorem. Another basis is given as follows: for $0\le i<d$ and $0\le j<L$, we define $\mu_{ij}(n)=\zeta^{nj}n^i$, where ζ is a fixed choice of a primitive Lth root of unity (say $\zeta=e^{\frac{2\pi i}{L}}$). A slight modification of this basis turns out be more convenient: for i and j as above, put $e_{ij}(n)=\binom{n}{i}\zeta^{nj}$. To see why, consider the generating function

$$E_{ij}(x)=\sum_{n\ge 0}e_{ij}(n)x^n.$$

By taking the ith derivative of the identity $\sum_n x^n=\frac{1}{1-x}$, we get:

$$E_{i0}(x)=\sum_{n\ge 0}\binom{n+i}{i}x^n=\frac{1}{(1-x)^{i+1}}.$$

It follows that

$$E_{ij}(x)=E_{i0}(\zeta^j x)=\sum_m\binom{n+i}{i}\zeta^{nj}x^n=\frac{1}{(1-\zeta^j x)^{i+1}}. \tag{3}$$

From this we readily deduce the following basic result:

Proposition 3.1. *For a function $f:\mathbb{N}\to\mathbb{C}$, the following are equivalent:*

(a) $f \in Q(L,d)$ *is a quasi-polynomial of period* L *and degree at most* $d-1$.

(b) *The generating function* $F(x) = \sum_n f(n)x^n$ *of* f *is a rational function of the form* $\frac{P(x)}{(1-x^L)^d}$, *where* $\deg P < Ld$.

Proof. Certainly the space of proper rational functions of the above form has the right dimension, so it suffices to see that the Taylor series coefficients are quasipolynomial functions of n. The key point is that, as for any proper rational function $R(X) = \frac{P(x)}{Q(x)}$ with $Q(0) \neq 0$, we have a **partial fractions decomposition**

$$\frac{P(x)}{Q(x)} = \sum_{j=0}^{M-1} \sum_{i=1}^{m_j} \frac{C_{ij}}{(1-s_j x)^i}. \tag{4}$$

Here $s_1, \ldots, s_M$ are the reciprocals ($s_j = r_j^{-1}$) of the distinct zeros r_j of the denominator $Q(x)$, and m_j denotes the multiplicity of r_j. In the given case, the (reciprocal) zeros of the denominator are Lth roots of unity, so the partial fractions decomposition precisely expresses $\frac{P(x)}{(1-x^L)^d}$ as a $\mathbb{C}$-linear combination of the $E_{ij}(x)$'s. This proves the result. $\square$

Now fix a primitive system of weights $(a_1, \ldots, a_k)$ and consider the generating function $R(x) = \sum r(n)x^n$ for $r(n) = r(a_1, \ldots, a_k; n)$. We have the identity

$$R(x) = (1 + x^{a_1} + x^{2a_1} + \ldots) \cdots (1 + x^{a_k} + x^{2a_k} + \ldots) = \frac{1}{(1-x^{a_1}) \cdots (1-x^{a_k})}. \tag{5}$$

Let $D(x) = (1-x^{a_1}) \cdots (1-x^{a_k})$. Recalling that we have put $L = \operatorname{lcm}(a_1, \ldots, a_k)$, the (reciprocal) zeros of $D(x)$ are all Lth roots of unity, but some zeros may occur with multiplicity greater than one: e.g., $x = 1$ occurs with multiplicity k. Since $1 - x^{a_i}$ has distinct zeros, it is clear that $m_i \leq k$ for all i, so there exists a polynomial $N(x)$ such that $N(x)D(x) = (1-x^L)^k$, whence $R(x) = \frac{N(x)}{(1-x^L)^k}$, and we conclude that $r(n)$ is a quasipolynomial of period L and degree at most $k-1$. Thus we have proved Theorem 2.1(a).

Before proceeding further, there is something to be addressed: is it not the case that Proposition 4 already gives the explicit formula for $r(N)$ that we seek? Namely, with C_{ij} equal to the coefficient of $(1-\zeta^{-j}x)^{-i}$, then using (4) and (5), do we not get

$$r(n) = \sum_{j=0}^{M-1} \sum_{i=1}^{m_j} C_{ij} \binom{n+i-1}{i-1} \zeta^{nj} \ ?$$

Certainly we do. However, this formula has its shortcomings. First, it is not really a formula at all until we say what the coefficients C_{ij} are. The total number of such coefficients is equal to $S = a_1 + \ldots + a_k$. We all know how the C_{ij}'s are supposed to be computed: starting with the identity

$$\frac{1}{(1-x^{a_1}) \cdots (1-x^{a_k})} = \sum_{j=0}^{M-1} \sum_{i=1}^{m_j} \frac{C_{ij}}{(1-s_j x)^i}, \tag{6}$$

cross multiplication yields an $S \times S$ linear system to be solved for the C_{ij}'s. But, as any calculus student knows, the amount of computation needed to solve such linear systems quickly gets out of hand. Second, even if we happen to know all the C_{ij}'s, then our explicit formula for $r(n)$ will still be an unwieldy mess. Our difficulties are really linear algebraic in nature: the method of partial fractions computes the coefficients of the basis e_{ij} of $Q(L,d)$, which is the wrong basis for efficient computation of the coefficients of $r_j(n)$.

However, the partial fractions decomposition is a powerful tool for extracting other information about the function $r(n)$. Let us look at it a bit more carefully: first, for any fixed j corresponding to an Lth root of unity ζ^{-j} of order w_j, we get a contribution to $r(n)$ that is a function of the form $f_j(n)\zeta^{-jn}$, where $f_j(n)$ is a polynomial of degree $m_j - 1$. If $j = 0$, $f_0(n)$ is an honest polynomial of degree $k-1$. For a general j, the multiplicity m_j is equal to the number of i's for which $w_j \mid a_i$, so if $w_j > 1$, the primitivity of the system assures that $m_j \le k-1$. (So if the coefficients a_i were coprime in pairs, we would have that $f_j(n)$ is a constant for all $j \neq 0$.) This implies $r(n)$ is asymptotic to the leading term of $f_0(n)$:

$$r(n) \sim C_{0k}\binom{n+k-1}{k-1} = \frac{C_{0k}}{(k-1)!}n^{k-1}.$$

We can compute C_{0k}: just multiply (6) by $(1-x)^k$ and evaluate at $x=1$, getting

$$C_{0k} = \frac{1}{(1+x+\ldots+x^{a_1-1})\cdots(1+x+\ldots+x^{a_k-1})}|_{x=1} = \frac{1}{a_1\cdots a_k}.$$

Thus we get

$$r(n) \sim \frac{n^{k-1}}{(k-1)!(a_1\cdots a_k)},$$

establishing Theorem 2.1(b).

For later use (indeed, the only essential use we shall make of the methods of this section), we record the following result:

Lemma 3.2. *Let $k=3$, and write $r(n) = c_2(n)n^2 + c_1(n)n + c_0(n)$, where the $c_i(n)$'s are periodic modulo L. The linear term c_1 is in fact periodic modulo $\mathfrak{q} = d_1d_2d_3$.*

Proof. In general we have $r(n) = \sum_{j=0}^{M-1} f_j(n)\zeta^{-jn}$, where $f_j(n)$ is a polynomial of degree equal to one less than the number of i's for which a_i divides the order w_j of ζ^{-j}. Thus $f_j(n)$ has a linear term only when w_j divides two of a_1, a_2, a_3, i.e., when $w_j \mid d_1d_2d_3$. □

4 Lattice Geometry

The problem of computing the number of solutions to $x_1a_1+\ldots+x_ka_k = N$ has an evident geometric interpretation: let T_N denote the locus of solutions to this same equation among nonnegative *real numbers*: this forms a simplex in $\mathbb{R}^k$ whose vertices are $V_i = (0,0,\ldots,\frac{N}{a_i},\ldots,0)$ for $1 \le i \le k$. Thus we can look at $r(N)$ as counting the lattice points – i.e., points with integral coordinates, on the simplex T_N.

In other words, even counting lattice points in triangles with rational vertices is a difficult problem! However, it has long been known that if the vertices of the triangle have *integral* coordinates (i.e., are themselves lattice points), then there is a wonderful explicit formula. Theorem 2.2 will follow easily from the following elementary result (for the proof, see e.g. [3]), a version of Pick's Theorem "renormalized" for two-dimensional lattices in higher-dimensional Euclidean spaces.

Theorem 4.1. *(Pick's Theorem) Let Λ be a two-dimensional lattice in $\mathbb{R}^k$ with 2-volume δ. Let P be a lattice polygon containing h interior lattice points and b boundary lattice points. Then the area $A(P)$ of P is equal to $\delta \cdot (h + \frac{b}{2} - 1)$.*

Let $r(P) = h + b$ be the total number of lattice points of P. Then Pick's Theorem is equivalent to the formula

$$r(P) = \frac{A(P)}{\delta} + \frac{b}{2} + 1. \tag{7}$$

We can now prove Theorem 2.2.

Proof. Although we have "the right" to compute $r_0(n) = r(nL)$ directly, we will find it easier to first derive a formula for $r(n\mathcal{P})$ and then change variables to obtain the formula for $r_0(n)$.

To compute $r(n\mathcal{P})$ we must count lattice points on the simplex

$$T_n : x_1a_1 + x_2a_2 + x_3a_3 = n\mathcal{P},$$

whose vertices are $V_1 = (a_2a_3n, 0)$, $V_2 = (0, a_1a_3n, 0)$ $V_3 = (0, 0, a_1a_2n)$. Clearly we have $\frac{A(P)}{\delta} = \alpha n^2$ for some nonzero α, whereas from Section 3 we know that $r(T_n)$ is a quadratic polynomial in n whose leading coefficient is $(\mathcal{P})^2/2(\mathcal{P}) = \frac{\mathcal{P}}{2}$. Thus from our algebraic formalism we get that $\alpha = \frac{a_1a_2a_3}{2}$.

As for the linear term, the number of lattice points on any line segment joining two integer points $V_i = (x_i, y_i, z_i)$ and $V_j = (x_j, y_j, z_j)$ is equal to $\gcd(x_i - x_j, y_i - y_j, z_i - z_j)$. If for distinct $i,\ j \in \{1, 2, 3\}$ we denote by ℓ_{ij} the line segment from V_i to V_j, we have

$$\#(\mathbb{Z}^3 \cap \ell_{ij}) = a_kN \gcd(a_i, a_j) = a_kd_kN,$$

where k is the remaining index. This gives

$$r(n\mathcal{P}) = \frac{a_1a_2a_3}{2} + \frac{a_1d_1 + a_2d_2 + a_3d_3}{2}n + 1.$$

Finally, observe that $r(nL)$ is also given by a quadratic polynomial in n. On the one hand, this follows immediately from the considerations of the previous section – we know that $r(N)$ is a quasi-polynomial with period L. On the other hand, even without computing the coefficients explicitly, the above argument using Pick's Theorem establishes this fact. Thus $r_0(\mathfrak{q}n) = r(\mathfrak{q}Ln) = r(\mathcal{P}n)$ is an identity of quadratic polynomials, so it can be inverted to give $r_0(n) = r(\mathcal{P}(\frac{n}{\mathfrak{q}}))$. Performing this change of variables gives the formula of Theorem 2.2. ∎

Remark: A less sneaky proof would compute $A(T_n)$ and δ and not merely their ratio. This will be done in Section 6.

5 Recursions and congruences

Let $(a_1, \ldots, a_k)$ be any primitive set of weights. We will get a lot of mileage out of the following innocuous identity:

$$r(n + a_1) = r(n) + r(a_2, \ldots, a_k; n + a_1). \tag{8}$$

The proof could hardly be simpler: the representations of $n+a_1$ by weights $a_1, \ldots, a_k$ that use a_1 at least once are in one-to-one correspondence with all representations of n by $a_1, \ldots, a_k$, and the ones that do not use a_1 at all are actually representations by $a_2, \ldots, a_k$.

There are of course analogous formulas for the other weights, which we use without comment.

We now return to the case of $k = 3$.

Proposition 5.1. *Let $0 \leq y_i < d_i$. Then*

$$r(y_1a_1 + y_2a_2 + y_3a_3 + nL) = r(nL).$$

Proof. Assume that $d_1 > 1$. We have $r(a_1+nL) = r(nL)+r(a_2, a_3; a_1+nL)$. Since $\mathfrak{q}$ divides L and $(a_1, d_1) = 1$, we have that d_1 does not divide a_1+nL, so necessarily $r(a_2, a_3; a_1+nL) = 0$. Repeated application of this argument gives the result. $\square$

Recall that we have defined numbers $y_i(j) = \langle a_i^{-1} j\rangle_{d_i}$ and $y_j = y_1a_1 + y_2a_2 + y_3a_3$. By construction $y(j)$ is congruent to j modulo L and satisfies the hypothesis of Proposition 5.1. Thus we have

$$r_{y(j)}(n) = r(y_1a_1 + y_2a_2 + y_3a_3 + nL) = r(nL) = r_0(n).$$

The change of variables $n = \frac{N-y(j)}{L}$ gives us the expression for $r(N)$ valid on the congruence class $y(j)$, which gives Theorem 2.3(a). Since $y(j) \equiv j \pmod{L}$, applying Lemma 3.2, the expression $Q_j(N) = r_0(\frac{N-y(j)}{L})$ differs from the correct expression for $r(N)$ on the class j by at most a constant. The expression for $P_j(N)$ corrects for this (in a rather tautological way) by ensuring that $P_j(j) = r(j)$. This completes the proof of Theorem 2.3.

6 Geometric and Combinatorial Proofs

In this section we give proofs of the results of this paper (except Theorem 2.3(b)) that do not require the theory of generating functions.

6.1 More lattice geometry

First, recall that $r(N) = r(a_1, \ldots, a_k; N)$ counts lattice points on the simplex T_N given by the intersection of the hyperplane $x_1a_1 + \ldots + x_ka_k = N$ with the positive orthant. As N varies, the lattices of integral solutions are isometric – they are all principal homogeneous spaces for the rank $k - 1$ free abelian group of integral solutions to $x_1a_1 + \ldots x_ka_k = 0$. Thus these lattices have a common 2-volume δ. Because the simplices T_N are similar, as N approaches infinity we must have $r(N) \sim \frac{A(T_N)}{\delta}$ (more precisely, the geometric picture gives $|r(N) - \frac{A(T_N)}{\delta}| = O(N^{k-2})$.) We saw earlier that the ratio of these two quantities is $\frac{N^{k-1}}{(k-1)!(a_1 \cdots a_k)}$, but we shall now compute both quantities individually.

Let $\Lambda \subset \mathbb{R}^k$ be a d-dimensional sublattice of $\mathbb{Z}^k$, i.e., the $\mathbb{Z}$-span of d $\mathbb{R}$-linearly independent vectors $\{v_1, \ldots, v_k\}$ with integral coordinates. The volume of Λ is the d-dimensional volume of any fundamental region for Λ, e.g. of the paralleletope formed by $\{x_1 v_1 + \ldots x_d v_d \mid 0 \leq x_i \leq 1\}$.

Given an orthonormal basis $e_1, \ldots, e_d$ of the vector space V spanned by Λ, we can write $v_i = \sum_{ij} \alpha_{ij} e_j$, and we have $\mathrm{Vol}(\Lambda) = \det(\alpha_{ij})$. Alternately, for any $\mathbb{Z}$-basis $\{v_1, \ldots, v_k\}$ for Λ, $\mathrm{Vol}(\Lambda)$ is the square root of the determinant of the **Gram matrix** $M_{ij} = v_i \cdot v_j$.

The problem is that our lattice

$$\Lambda_0 = \{(z_1, \ldots, z_k) \in \mathbb{Z}^k \mid z_1 a_1 + \ldots + z_k a_k = 0\},$$

a $k-1$-dimensional lattice in $\mathbb{R}^k$, does not come equipped with any natural basis. Nevertheless we present the following method to compute the volume of any lattice $\Lambda_{\mathbf{a}}$ defined by the hyperplane equation $(z_1, \ldots, z_k) \cdot \mathbf{a} = 0$, for $\mathbf{a} = (a_1, \ldots, a_k)$ a primitive set of weights (rebaptized here as a **primitive vector**).

Let $\Lambda = \mathbb{Z}v_1 \oplus \ldots \oplus \mathbb{Z}v_{k-1} \subset \mathbb{Z}^k$. With $e_1, \ldots, e_k$ an orthonormal basis of $\mathbb{R}^k$, define the generalized cross-product

$$\mathrm{cross}(v_1, \ldots, v_{k-1}) = \begin{bmatrix} e_1 & \ldots & e_k \\ \longleftarrow & v_1 & \longrightarrow \\ & \vdots & \\ \longleftarrow & v_{k-1} & \longrightarrow \end{bmatrix}$$

By expanding out the determinant, one checks that $\mathrm{cross}(v_1, \ldots, v_{k-1})$ is orthogonal to Λ and that $\mathrm{Vol}(\Lambda) = ||\,\mathrm{cross}(v_1, \ldots, v_{k-1})||$.

Example: Take $k = 2$ and $\mathbf{a} = (a_1, a_2)$ a primitive vector. Then $\Lambda_{\mathbf{a}} = \mathbb{Z}(-a_2, -a_1)$, i.e. the orthogonal complement of v is obtained by rotating $\mathbb{Z}\mathbf{a}$ through a 90 degree angle. Notice that in this case the 1-volume of $\mathbf{a}$ is equal to the $(k-1)$-volume of $\Lambda_{\mathbf{a}}$.

That this remains true in higher dimensions is the content of the following result:

Lemma 6.1. *Let* $\mathbf{a} = (a_1, \ldots, a_k)$ *be a primitive vector. Then*

$$\mathrm{Vol}(\Lambda_{\mathbf{a}}) = \mathrm{Vol}(\mathbf{a}) = \sqrt{a_1^2 + \ldots + a_k^2}.$$

Proof. For any $v_1, \ldots, v_{k-1}$ is a $\mathbb{Z}$-basis of $\Lambda_{\mathbf{a}}$, $\mathrm{Vol}(\Lambda_{\mathbf{a}}) = ||\,\mathrm{cross}(v_1, \ldots, v_{k-1})||$. On the other hand, the cross product of integer vectors remains an integer vector, so $\mathrm{cross}(v_1, \ldots, v_{k-1})$ and $\mathbf{a}$ are two integer vectors spanning the same 1-dimensional $\mathbb{R}$-subspace. Since $\mathbf{a}$ is assumed to be primitive, we must have $\mathrm{cross}(v_1, \ldots, v_{k-1}) = n\mathbf{a}$, for some integer n, so that $\mathrm{Vol}(\Lambda_{\mathbf{a}}) \geq \mathrm{Vol}(\mathbf{a})$. We claim that $n = 1$, which will finish the proof.

To establish the claim, consider the set $\mathcal{C} = \{\mathrm{cross}(w_1, \ldots, w_{k-1}) \mid w_i \in \Lambda_{\mathbf{a}}\}$ of vectors that are cross-products from $\Lambda_{\mathbf{a}}$. Writing $w_i = \Sigma \alpha_{ij} v_j$ with α_{ij} in $\mathbb{Z}$, we get

$$\mathrm{cross}(w_1, \ldots, w_{k-1}) = \det(\alpha_{ij})\,\mathrm{cross}(v_1, \ldots, v_{k-1}),$$

which shows that $\mathcal{C}$ is just the 1-dimensional lattice generated by $\mathrm{cross}(v_1, \ldots, v_{k_1})$. In particular, we can exploit the fact that $\mathcal{C}$ is a $\mathbb{Z}$-module as follows: note that $\Lambda_{\mathbf{a}}$ is precisely the set of integer solutions of

$$a_1 z_1 + \ldots + a_k z_k = 0.$$

Thus

$$v_1 = (-a_2, a_1, 0, \ldots, 0),\ v_2 = (-a_3, 0, a_1, 0, \ldots, 0), \ldots,\ v_{k-1} = (-a_k, 0, \ldots, a_1)$$

are all elements of $\Lambda_{\mathbf{a}}$, so that

$$\mathrm{cross}(v_1, \ldots, v_{k-1}) = a_1^{k-2}(a_1, \ldots, a_{k-1})$$

is a vector in $\mathcal{C}$. By symmetry, we see that $a_2^{k-2}(a_1, \ldots, a_{k-1}), \ldots, a_k^{k-2}(a_1, \ldots, a_k)$ are also vectors of $\mathcal{C}$. Since $\mathcal{C}$ is a $\mathbb{Z}$-module, we know that for any integers $z_1, \ldots, z_k$, we have

$$(z_1 a_1^{k-2} + z_2 a_2^{k-2} + \ldots z_k a_k^{k-2})(a_1, \ldots, a_k) \in \mathcal{C}.$$

Since $(a_1, \ldots, a_k)$ is primitive, so is $(a_1^{k-2}, \ldots, a_k^{k-2})$, so that we can choose $z_1, \ldots, z_k$ such that $z_1 a_1^{k-2} + \ldots + z_k a_k^{k-2} = 1$. But this implies that $\mathbf{a} = (a_1, \ldots, a_k)$ is in $\mathcal{C}$, and $n = 1$. This establishes the claim and completes the proof of Lemma 6.1. $\square$

Coming back to our lattice Λ_0, we have $\mathrm{Vol}(\Lambda_0) = \sqrt{a_1^2 + \ldots + a_k^2}$. It remains to compute the volume of T_N. First consider the $(k-1)$-paralleletope P_N whose sides are the vectors $v_1 - v_2, v_1 - v_3, \ldots, v_1 - v_k$, where $v_i = \frac{N}{a_i} e_i$ gives the intersection points of the simplex $H_N \cap \mathbb{R}^{k+}$ with the coordinate axes. Using the cross-product, we calculate

$$\mathrm{Vol}(P_N) = \frac{\sqrt{a_1^2 + \ldots + a_k^2} N^{k-1}}{a_1 \cdots a_k}.$$

Noting that the cone on an r-dimensional base has volume equal to $1/(r+1)$ times the corresponding cylinder, induction then gives that

$$\mathrm{Vol}(T_N) = \frac{1}{(k-1)!} \mathrm{Vol}(P_N) = \frac{\sqrt{a_1^2 + \ldots + a_k^2}}{a_1 \cdots a_k} \frac{N^{k-1}}{(k-1)!}.$$

Thus we conclude once again that

$$r(a_1, \ldots, a_k, N) \sim \frac{N^{k-1}}{a_1 \cdots a_k (k-1)!}.$$

6.2 More recursions and congruences

Using the results of the last section, it is easy to rederive the quasipolynomiality of $r(N) = r(a_1, a_2, a_3; N)$. We need only use our basic recursion formula

$$r(a_1 + N) = r(N) + r(a_2, a_3; a_1 + N)$$

"in reverse." Namely, fix a class $j \pmod{L}$. Since, for all positive integers N we have $r(N) = r(N + a_1) - r(a_2, a_3; N + a_1)$, taking $N = j + nL$, we get

$$r(j + nL) = r(a_1 + j + nL) - r(a_2, a_3; a_1 + j + nL). \tag{9}$$

A similar formula holds for a_2 and a_3. By Theorem 1(b), every sufficiently large integer N is representable by the system (x_1, x_2, x_3); in particular, there exist x_1, x_2, x_3 such that $L \mid x_1a_1 + x_2a_2 + x_3a_3 + j$. Repeated application of (8) gives that $r(j+nL) - r(x_1a_1 + x_2a_2 + x_3a_3 + j + nL)$ is equal to a sum of $x_1 + x_2 + x_3$ terms of the form $r(a_i, a_j; a_k + c_{ijk} + nL)$ for various positive integers c_{ijk}. Now it is easy to see that each of these terms is either zero or a linear polynomial in n: e.g. this can be seen by contemplating the family of planar line segments

$$L_N = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1a_1 + x_2a_2 = N,\ x_1,\ x_2 \geq 0\},$$

the two-dimensional analogue of the family of simplices T_N considered above. The geometric picture makes clear that for any positive integers a, b, j, $r(a, b; j + \mathrm{lcm}(a, b)) - r(a, b; j)$ is equal to 1 if $\gcd(a, b) \mid j$ and 0 otherwise. Details are left to the reader.

Thus $r(j + nL)$ differs from $r(x_1a_1 + x_2a_2 + x_3a_3 + j + nL)$ by a linear polynomial. But since $x_1a_1 + x_2a_2 + x_3a_3 + j = ML$, we have $r(x_1a_1 + x_2a_2 + x_3a_3 + j + nL) = r((M + n)L) = r_0(M + n)$, which by Theorem 2 is a quadratic polynomial. Thus $r(j+nL)$ is itself a quadratic polynomial, which was to be shown.

7 Extremal systems

Suppose (a_1, a_2, a_3) is a primitive system of weights with $\mathfrak{q} = L$; we say the system is **extremal**. This implies that $a_i = \frac{\mathfrak{q}}{d_i}$. Conversely, given any triple (d_1, d_2, d_3) of pairwise coprime positive integers, by putting $a_i = \frac{d_1d_2d_3}{d_i}$, we get an extremal system.

The extremal systems are, in many ways, the three-dimensional systems whose behavior most closely parallels that of two-dimensional systems. In particular, their Frobenius number can be computed:[2]

Proposition 7.1. *Let d_1, d_2, d_3 pairwise coprime positive integers. Then*

$$\mathfrak{f}(d_1d_2, d_1d_3, d_2d_3) = 2d_1d_2d_3 - d_1d_2 - d_1d_3 - d_2d_3.$$

This result appeared as Problem A3 on the 1983 International Mathematical Olympiad. We shall discuss four proofs. First, we begin with the following general observation.

Lemma 7.2. *Let (a_1, a_2, a_3) be any primitive system of weights. For $1 \leq i \leq 3$, the Frobenius number $\mathfrak{f} = \mathfrak{f}(a_1, a_2, a_3)$ satisfies $\mathfrak{f} \equiv -a_i \pmod{d_i}$.*

Proof. By symmetry, it suffices to consider the case $i = 1$. Since $\mathfrak{f} + a_1 > \mathfrak{f}$, we can write $\mathfrak{f} + a_1$ as $x_1a_1 + x_2a_2 + x_3a_3$, with nonnegative x_i's. But x_1 cannot be positive, since that

[2]Another property enjoyed by all extremal systems is the so-called Gorenstein condition: namely, that exactly half of the positive integers $N \leq \mathfrak{f}(a_1, \ldots, a_k) + 1$ can be represented [6].

would lead to a representation of $\mathfrak{f}$ by a_1, a_2 and a_3. So $\mathfrak{f} + a_1$ is an integral combination of a_2 and a_3, and is therefore divisible by d_1.

Compiling the three congruences, we get

$$\mathfrak{f} \equiv (d_1 - 1)a_1 + (d_2 - 1)a_2 + (d_3 - 1)a_3 \pmod{\mathfrak{q}}. \tag{10}$$

The right-hand side of (10) is equal to $3d_1d_2d_3 - d_1d_2 - d_1d_3 - d_2d_3 = j_0$, say.

By Theorem 2.3, we have $r_{j_0}(n) = r_0(n)$ for all n, so that we have an *explicit formula* for r on the class containing $\mathfrak{f}$, namely $\mathfrak{q}^2 R_{j_0}(N) =$

$$\frac{1}{2}N^2 + \frac{2a_1 + 2a_2 + 2a_3 - 3\mathfrak{q}}{2}N + \mathfrak{q}^2 - \frac{3}{2}(a_1 + a_2 + a_3)\mathfrak{q} + \frac{1}{2}(a_1^2 + a_2^2 + a_3^2) + \mathfrak{q}(d_1 + d_2 + d_3).$$

Hence $R_{j_0}(j_0 - L) = R_{j_0}(j_0 - 2L) = 0$. Since $R_{j_0}(N)$ is a quadratic polynomial, it can have no further zeros, so $\mathfrak{f}(a_1, a_2, a_3) = j_0 - L$. □

Here is a second, more conceptual proof: we know that $r(j_0) = r_{j_0}(0) = r_0(0) = 1$, so $\mathfrak{f} \neq j_0$. In general, since $r(L) > 0$, the function $r(N)$ is nondecreasing on any given congruence class modulo L. Here $L = \mathfrak{q} = a_id_i$, so we conclude that $\mathfrak{f} < j_0$ and it is enough to show that $r(j_0 - L) = 0$. But, by our usual recursion we have

$$1 = r(j_0) = r(j_0 - d_1a_1) + \sum_{i=0}^{d_1-1} r(a_2, a_3; j_0 - ia_i).$$

Clearly $r(j_0 - L) = r(j_0 - d_1a_1) \in \{0, 1\}$ and is zero if and only if any one of the terms of the sum is nonzero. But taking $i = d_1 - 1$, we have

$$r(a_2, a_3; j_0 - (d_1 - 1)a_i) = r(a_2, a_3; (d_2 - 1)a_2 + (d_3 - 1)a_3) > 0.$$

Third proof: Let us recall the following formula, due to Johnson [4]:

Lemma 7.3. *For any primitive system* (a_1, a_2, a_3), *one has*

$$\mathfrak{f}(a_1, a_2, a_3) = d_3\mathfrak{f}(\frac{a_1}{d_3}, \frac{a_2}{d_3}, a_3) + (d_3 - 1)a_3. \tag{11}$$

Proof. Let $\mathfrak{f} = \mathfrak{f}(a_1, a_2, a_3)$. By lemma 7.2, we have $\mathfrak{f} \equiv -a_3 \pmod{d_3}$. For any number N satisfying $N \equiv -a_3 \pmod{d_3}$, in any representation of N as $x_1a_1 + x_2a_2 + x_3a_3$ we must have $x_3 \equiv -1 \pmod{d_3}$, hence N is representable by a_1, a_2 and a_3 if and only if $N - (d_3 - 1)a_3$ is representable by a_1, a_2 and d_3a_3. This is equivalent to $(N - (d_3 - 1)a_3)/d_3$ being representable by $\frac{a_1}{d_3}, \frac{a_2}{d_3}$ and a_3, whence the result. □

This leads to an explicit formula for $\mathfrak{f}$ whenever $a_1 = d_2d_3$:[3]

$$\mathfrak{f}(d_2d_3, a_2, a_3) = d_3\mathfrak{f}(d_2, \frac{a_2}{d_3}, a_3) + (d_3 - 1)a_3 = d_3\mathfrak{f}(d_2, \frac{a_2}{d_3}) + (d_3 - 1)a_3 =$$

[3]This holds in particular for the "McNuggets" system $(6, 9, 20)$, and so can be used to answer the brain-teaser.

$$d_2a_2 - d_2d_3 - a_2 + d_3a_3 - a_3 = a_2d_2 + a_3d_3 - a_1 - a_2 - a_3,$$

where the second equality uses the fact that, since the weight a_3 is a multiple of the weight d_2, it can be omitted without changing the Frobenius number.

Finally, Proposition 7.1 is a consequence of the following beautiful result of Alfred Brauer [2].

Theorem 7.4. *Given a primitive set of weights $(a_1, \dots, a_k)$, for $1 \leq i \leq k$ let $e_i = \gcd(a_1, \dots, a_i)$. Then*

(a) $\mathfrak{f}(a_1, \dots, a_k) \leq \sum_{i=2}^{k} a_i \frac{e_{i-1}}{e_i} - \sum_{i=1}^{k} a_i + 1.$

(b) One has equality in part (a) if and only if, for all $i \geq 2$, $\frac{e_{i-1}}{e_i} a_i$ can be represented by the system $(a_1, \dots, a_i)$.

There is an analogue of Proposition 7.1 for arbitrary $k \geq 2$: let $d_1, \dots, d_k$ be a set of pairwise coprime positive integers, define $\mathcal{D} = \prod_j d_j$, and for $1 \leq i \leq k$ put $a_i = \frac{\mathcal{D}}{d_i}$. Then we say $(a_1, \dots, a_k)$ is an **extremal** system of weights.

Proposition 7.5. *Let $(a_1, \dots, a_k)$ be the extremal system formed from the pairwise coprime system $(d_1, \dots, d_k)$. Then:*

$$\mathfrak{f}(a_1, \dots, a_k) = (k-1)(d_1 \cdots d_k) - \sum_{i=1}^{k} a_i. \tag{12}$$

The result is again a consequence of Brauer's Theorem 7.4. We leave to the interested reader the task of investigating which of the other proofs of Proposition 7.1 can be made to go through.

We end with an explanation of the term "extremal." For this, let $r : \mathbb{N} \to \mathbb{C}$ be any function given by a quasipolynomial mod L, of total degree at most d, and which is not identically zero on any residue class. Then we can define $\mathfrak{f}(r)$ as the largest N for which $r(N) = 0$. Indeed, since r can have at most d zeros on any given residue class, we have $\mathfrak{f}(r) \leq dL - 1$. By Theorem 2.1, we can apply this to any primitive system of weights $(a_1, \dots, a_k)$, getting the bound

$$\mathfrak{f}(a_1, \dots, a_k) \leq (k-1)L - 1.$$

In fact this bound is never sharp, as can be seen by comparing to Brauer's bound. This is not so surprising, since the quasipolynomial associated with a set of weights has many special properties.

Notice however that this bound differs from the right hand side of (12) by precisely $a_1 + \dots + a_k - 1$. This can be explained as follows: for a primitive system $(a_1, \dots, a_k)$, define $\mathfrak{f}^+(a_1, \dots, a_k)$ to be the largest positive integer N such that the equation $x_1a_1 + \dots + x_ka_k = N$ has no solution in *positive integers* x_i. This is closely related to the Frobenius number defined above; indeed, we have

$$\mathfrak{f}^+(a_1, \dots, a_k) = \mathfrak{f}(a_1, \dots, a_k) + a_1 + \dots + a_k.$$

Moreover, we can define

$$r^+(a_1,\ldots,a_k;N) = \#\{(x_1,\ldots,x_k)\in(\mathbb{Z}^+)^k \mid x_1a_1+\ldots+x_ka_k = N\}.$$

The associated generating function $R^+(x)$ is easily computed:

$$\sum_{n\geq 0} r^+(n)x^n = (x^{a_1}+x^{2a_1}+\ldots)\cdots(x^{a_k}+x^{2a_k}+\ldots) = \frac{x^{a_1+\ldots+a_k}}{(1-x^{a_1})\cdots(1-x^{a_k})}.$$

This is very nearly the generating function of a quasipolynomial: the only problem is that the degree of the numerator is equal to the degree of the denominator. Thus

$$x^{-1}R^+(x) = \sum_{n\geq 0} a_{n+1}x^n$$

is a quasipolynomial, so that $r^+(n)$ is a quasipolynomial modulo L when restricted to positive integral values, of degree $k-1$ on each residue class. From this information alone we deduce as above the bound

$$\mathfrak{f}^+(a_1,\ldots,a_k) \leq (k-1)L$$

and thus the bound

$$\mathfrak{f}(a_1,\ldots,a_k) \leq (k-1)L - \sum_{i=1}^{k} a_i.$$

It is remarkable that this bound *is* attained, by all extremal systems. Comparing with Brauer's bound, it is easily seen that the bound is *only* attained for extremal systems, whence the name.

References

[1] M. Beck, I. Gessel and T. Komatsu, The polynomial part of a restricted partition function related to the Frobenius problem, *Electronic J. Comb.* **8** (2001), N7.

[2] A. Brauer, On a problem of partitions, *Amer. J. Math.* **64** (1942), 299–312.

[3] P. Erdos and J. Suranyi, *Topics in the theory of numbers* , Springer, 2003.

[4] S. M. Johnson, A linear diophantine problem, *Canad. J. Math.* **12** (1960), 390–398.

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

[6] A. Nijenhuis and H. Wilf, Representations of integers by linear forms in non-negative integers, *J. Number Theory* **4** (1972), 98–106.

[7] T. Popoviciu, Asupra unei probleme de patitie a numerelor, *Acad. Republicii Populare Romane, Filiala Cluj, Studii si cercetari stiintifice* **4** (1953), 7–58.

[8] J. L. Ramirez-Alfonsin, Complexity of the Frobenius problem, *Combinatorica* **16** (1996), 143–147.

[9] S. Sertoz, On the number of solutions of the Diophantine equation of Frobenius, *Diskret. Mat.* **10** (1998), 62–71.

2000 *Mathematics Subject Classification*: Primary 05A15; Secondary 52C07.
Keywords: Frobenius problem, quasi-polynomial, representation numbers, Pick's theorem.

Received June 23 2005; revised version received October 19 2005. Published in *Journal of Integer Sequences*, October 20 2005.

Return to Journal of Integer Sequences home page.