On the Number of Solutions of a Diophantine Equation of Frobenius

Sinan Sertöz
Bilkent University
Department of Mathematics
06533 Ankara, Turkey

1 Introduction

The linear Diophantine equation we are interested in is of the form

\[ x_1a_1 + \cdots + x_na_n = N, \quad (1.1) \]

where \( n > 1 \) is a fixed positive integer and for which we have a fixed set of integers \( 0 < a_1 < \cdots < a_n \) satisfying \((a_1, \ldots, a_n) = 1\). For any nonnegative integer \( N \) we say that a solution exists if there exist nonnegative integers \( x_1, \ldots, x_n \) satisfying equation (1.1). It is well known that for all sufficiently large \( N \) the equation has solutions. In fact for \( n = 2 \) the largest \( N \) for which no solution exists can be explicitly written as \( a_1a_2 - a_1 - a_2 \). However no such formula exists for \( n \geq 3 \) as shown by Curtis in [2]. Several ingenious algorithms exist in the literature for calculating this bound which is known as the conductor. For example the reader may refer to [1, 8, 9, 10, 13] for several algorithms for calculating the conductor. The literature on this subject is vast and for different aspects of the problem we can mention for example [3, 4, 11, 12, 15].
We are interested in calculating how many solutions equation (1.1) has. In particular we want to find a formula, or an algorithm, which is fast in giving the answer. If we denote the number of solutions corresponding to $N$ by $f(N)$, considering that the $a_i$’s are fixed once and for all, then it is known that $f(x) = P(x) + \Delta(x)$ where $P(x)$ is a polynomial in $x$ of degree $n - 1$ with leading coefficient $1/[(a_1 \cdots a_n) \cdot (n - 1)!]$, and $\Delta(x)$ is a periodic function with period $a_1 \cdots a_n$.

In this article we apply an elementary approach to the problem of calculating $\Delta$. We will utilise a roots of unity argument in constructing the desired periodic function. We will write the general form of $f(x)$ for arbitrary $n$, however this involves complicated sums of roots of unity. Therefore we demonstrate the merit of this approach by explicitly working out the $n = 2$ case to obtain a computable expression for $f(x)$. This problem was attacked in [5, 6, 14] from different points of view, but $\Delta$ was not explicitly written.

Apparently the problem of finding the conductor, together with its solution in the $n = 2$ case, first appeared in [16]. Later Frobenius, noting the ubiquity of the concept, raised the general question. Several related questions are also raised in [7]. In fact the problem may look deceptively specialized but nonetheless appears at most unexpected places in mathematics. Arf rings, Gorenstein rings, Weierstrass gaps are examples from geometry where the phenomena surfaces predominantly, and unexpectedly.

Acknowledgements: I thankfully acknowledge several useful discussions with S. Stepanov. His ideas and comments both helped me to overcome difficulties at several stages and to enjoy at a higher level the workings of arithmetic.

2 The General Case

We first set up our notation which we will use throughout. By the set of natural numbers we will mean $\mathbb{N} = \{0, 1, 2, \ldots \}$. We fix a positive integer $n$ and a set of $n$ integers

$$0 < a_1 < a_2 < \cdots < a_n,$$
Diophantine Equation of Frobenius

which we assume to be relatively prime;

\[(a_1, a_2, ..., a_n) = 1.\] (2.2)

Towards the end of the arguments we will assume further that these integers are mutually coprime but that will serve the purpose of achieving a more elegant formulation. We will nevertheless mention how to proceed in the general case of (2.2).

For any \(N \in \mathbb{N}\) we define the counting function

\[f(N) = \#\{ (x_1, x_2, ..., x_n) \in \mathbb{N}^n \mid a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = N \}.\] (2.3)

As a notation we will sometimes use

\[f(a_1, ..., a_n; N) := f(N)\] (2.4)

to emphasize dependence on \(a_i\)’s. The problem we address is to find a computable expression for \(f\).

We start by considering the complex valued function

\[\phi(z) = \prod_{i=1}^{n} (1 - z^{a_i})^{-1}.\] (2.5)

\(\phi\) is a meromorphic function whose only poles are \(e_{a_i}^{k_i}\) for \(i = 1, ..., n\) and \(k_i = 0, ..., a_i - 1\), where we define for any integers \(0 \leq k < a\) the symbol \(e_a^k\) as

\[e_a^k = \exp(2\pi \sqrt{-1} k/a).\]

We will write \(e_a\) for \(e_a^1\), the primitive \(a\)-th root of unity.

The poles of \(\phi\) are located on the unit circle, \(|z| = 1\). The one at \(z = 1\) is a pole of order \(n\). However \(\phi\) is analytic at the origin and has a power series expansion of the form

\[\phi(z) = \sum_{m=0}^{\infty} f(m)z^m, \text{ for } |z| < 1.\] (2.6)
This expansion for $\phi$ implies that for any integer $N \in \mathbb{N}$ we have

$$f(N) = \text{Res}_{z=0} \left( \frac{\phi(z)}{z^{N+1}} \right) = \frac{1}{2\pi \sqrt{-1}} \int_{|z| = \epsilon} \frac{\phi(z)}{z^{N+1}} dz,$$

for any $0 < \epsilon < 1$.

We therefore proceed to find the coefficient of $z^N$ as given in equation (2.6), using equation (2.5). For this the following facts will be useful. We omit the proofs which are straightforward.

**Fact 2.1** For any positive integer $a \in \mathbb{N}$ we have the factorization

$$1 - t^a = \prod_{l=0}^{a-1} (1 - e^{l a} t).$$

As a consequence of this we can rewrite equation (2.5) as

$$\phi(z) = \prod_{j=0}^{n} \prod_{l=1}^{a_j-1} (1 - e^{l a_j} z).$$

From this point on we explicitly assume that

$$(a_i, a_j) = 1 \text{ for } i \neq j. \quad (2.7)$$

This assumption has the advantage of simplifying the partial fractions expansion of the above expression of $\phi$:

$$\phi(z) = \sum_{i=1}^{n} \frac{K_i}{(1 - z)^i} + \sum_{j=1}^{n} \sum_{l=1}^{a_j-1} \frac{\alpha_{j,l}}{(1 - e^{l a_j} z)}, \quad (2.8)$$

where $K_i$ and $\alpha_{j,l}$ for $i, j = 1, ..., n$ and $l = 1, ..., a_j - 1$ are complex constants which can be determined by the theory of partial fractions. In particular to determine the $K_i$’s we will use the following
Fact 2.2 For $i = 1, \ldots, n$ and for all $|z| < 1$ we have

$$\frac{1}{(1 - z)^i} = \sum_{m=0}^{\infty} k_{i,m} z^m$$

(2.9)

where

$$k_{i,m} = \frac{(i + m - 1)!}{(i - 1)! m!} = \binom{i + m - 1}{m}$$

(2.10)

with the usual understanding that $0!$ is replaced by $1$.

Note that $k_{i,m} = (m + 1)(m + 2) \cdots (m + i - 1)/(i - 1)!$, and we can view it as a polynomial in $m$ of degree $i - 1$.

On the other hand since $|e_{a_j}^l z| < 1$ we have

$$\frac{1}{1 - e_{a_j}^l z} = \sum_{m=0}^{\infty} e_{a_j}^{ml} z^m.$$ 

(2.11)

Note here that the coefficients of $z$ are periodic with period $a_j$. Putting equations (2.8), (2.9) and (2.11) together we obtain

$$\phi(z) = \sum_{m=0}^{\infty} \left( \sum_{i=1}^{n} K_i k_{i,m} + \sum_{j=1}^{n} \sum_{l=1}^{a_j-1} \alpha_{j,l} e_{a_j}^{ml} \right) z^m.$$ 

The coefficient of $z^N$ is the value we are after:

$$f(N) = \sum_{i=1}^{n} K_N k_{i,N} + \sum_{j=1}^{n} \sum_{l=1}^{a_j-1} \alpha_{j,l} e_{a_j}^{Nl}.$$ 

(2.12)

For notational convenience we introduce the following vectors:

$$\vec{K} := (K_1, \ldots, K_n) \in \mathbb{C}^n$$

$$\vec{k}_N := (k_{1,N}, \ldots, k_{n,N}) \in \mathbb{C}^n$$

$$\vec{\alpha}_j := (\alpha_{j,1}, \ldots, \alpha_{j,a_j-1}) \in \mathbb{C}^{a_j-1}$$

$$\vec{\omega}_j := (e_{a_j}, e_{a_j}^2, \ldots, e_{a_j}^{a_j-1}) \in \mathbb{C}^{a_j-1}$$

$$\vec{\omega}_N := (e_{a_j}^N, e_{a_j}^{2N}, \ldots, e_{a_j}^{(a_j-1)N}) \in \mathbb{C}^{a_j-1}$$
where \( j = 1, \ldots, n \) and \( N \in \mathbb{N} \). Note that for each \( a_j \) and for any integers \( N, M \in \mathbb{N} \)
\[
\vec{\omega}_j^N = \vec{\omega}_j^{N \mod a_j}.
\]
With this notation we can write

**Theorem 2.3 (General Formula)**

\[
f(N) = \vec{K} \cdot \vec{k}_N + \sum_{j=1}^n \vec{\alpha}_j \cdot \vec{\omega}_j^N
\]

where \( \cdot \) denotes the usual inner product.

Observe that \( \vec{K} \cdot \vec{k}_N \) is a polynomial in \( N \) of degree \( n - 1 \). Each \( \vec{\alpha}_j \cdot \vec{\omega}_j^N \) is periodic of period \( a_j \). Since \( (a_1, \ldots, a_n) = 1 \), the total term \( \sum_{j=1}^n \vec{\alpha}_j \cdot \vec{\omega}_j^N \) is periodic with period \( a_1 a_2 \cdots a_n \).

**Calculating \( \vec{K} \) and \( \vec{\alpha}_j \).**

Using elementary means we can obtain the entries of \( \vec{K} = (K_1, \ldots, K_n) \) and \( \vec{\alpha}_j \). For example we can multiply both sides of equation (2.8) by \( (1 - z)^n \), take the \( i \)-th derivative of both sides and evaluate at \( z = 0 \) to obtain

\[
K_i = \frac{1}{(n-i)!} \left( \frac{(1-z)^n}{\prod_{j=1}^{n}(1-z^{a_j})} \right)^{(n-i)}_{z=0}, \text{ for } i = 1, \ldots, n \quad (2.13)
\]

where the superscript denotes the order of the derivative and the subscript denotes the point of evaluation. Similarly multiplying both sides of (2.8) by \( 1 - e_{a_j}^l \) \( z \) and evaluating at \( z = e_{a_j}^{-l} \) we obtain

\[
\alpha_{j,l} = \prod_{i=1, i \neq j}^n (1 - e_{a_j}^{-l a_i})^{-1} \prod_{s=1, s \neq l}^{a_j} (1 - e_{a_j}^{s-l})^{-1}
\]

\[
= \prod_{i=1, i \neq j}^n (1 - e_{a_j}^{-l a_i})^{-1} \prod_{k=1}^{a_j} (1 - e_{a_j}^{k})^{-1}
\]

\[
= \prod_{i=1, i \neq j}^n (1 - e_{a_j}^{-l a_i})^{-1} \cdot \frac{1}{a} \quad (2.14)
\]
The main difficulty of the simplification process lies in finding a expression for (2.14) amenable to summing over the given indices. We will find such an expression for \( n = 2 \). For large \( n \) a computable form for (2.13) is complicated but still seems possible.

**Remark:** If the \( a_i \)'s are not mutually coprime then the partial fraction expansion of \( f \) will involve terms of the form \( 1/(1 - e_i^k)^r \), and the determination of the coefficient with which it appears will be similar to determining the \( K_i \)'s. Clearly this will entail an enormous amount of bookkeeping.

**The \( n = 3 \) Case:** In [5, 6, 14] the polynomial part of \( f \) for \( n = 3 \) is determined by different approaches. Here we give the result from our approach for completeness. Certainly for the polynomial part we need to determine \( K_1, K_2 \) and \( K_3 \), which are found as:

\[
K_1 = \frac{\sigma_1^2 + \sigma_2 - 6\sigma_1 + 6}{12\sigma_3} \\
K_2 = \frac{\sigma_1 - 3}{2\sigma_3} \\
K_3 = \frac{1}{\sigma_3}
\]

where the \( \sigma_i \)'s are the elementary symmetric polynomials of \( a_1, a_2, a_3 \) defined as usual as \( \sigma_1 = a_1 + a_2 + a_3 \), \( \sigma_2 = a_1a_2 + a_1a_3 + a_2a_3 \) and \( \sigma_3 = a_1a_2a_3 \).

### 3 The \( n = 2 \) Case

In the \( n = 2 \) case it is well known that for any integer \( N \in \mathbb{N} \) we have \( f(N) = \lceil N/(a_1a_2) \rceil + f(N - \lceil N/(a_1a_2) \rceil : a_1a_2) \), where \( \lceil \cdot \rceil \) denotes the greatest integer function. In fact \( f(N) = 1 \) for all \( a_1a_2 - a_1 - a_2 < N < a_1a_2 \), so it suffices to find \( f \) for \( 0 \leq N \leq a_1a_2 - a_1 - a_2 \). For this we define some auxiliary functions:

**Definition 3.1** \( \varepsilon_{a|n} = \begin{cases} 1 & \text{if } a|n \\ 0 & \text{if } a \not{|} n \end{cases} \)
**Definition 3.2** \( r_{a,b,n} \) is defined for some positive integers \( a, b > 0 \) with \( (a,b) = 1 \) and \( n \geq 0 \), as the unique solution of the equation \( bx + n \equiv 0 \mod a \), with \( 0 \leq x \leq a - 1 \).

With these definitions we state the main result of this work:

**Theorem 3.3** For two positive integers \( a, b \) with \( (a,b) = 1 \), if we define for any \( N \in \mathbb{N} \) the function

\[
 f(N) = \# \{ (x, y) \in \mathbb{N}^2 \mid ax + by = N \}
\]

then we have

\[
 f(N) = \frac{N + br_{a,b,N} + ar_{b,a,N}}{ab} + \varepsilon_{a|N} + \varepsilon_{b|N} - 1.
\]

Before we prove the theorem we state two elementary lemmas whose proofs are briefly indicated.

**Lemma 3.4** For any two integers \( a > 0 \) and \( N \geq 0 \), we have

\[
 \sum_{k=1}^{a-1} e_a^{kN} = a\varepsilon_{a|N} - 1.
\]

**Proof:** The proof is straightforward once it is observed that either \( e_a^N = 1 \) or \( e_a^N = e_r \) for some positive integer \( r \) dividing \( a \).

**Lemma 3.5** If \( (a,b) = 1 \), then for any integer \( 1 \leq k \leq a - 1 \), we have

\[
 \frac{t^{a-1} + t^{a-2} + \cdots + t + 1}{t - e_a^{kb}} = \sum_{i=0}^{a-2} C_{k,b,a}(i) t^{a-2-i} \tag{3.15}
\]

where

\[
 C_{m,a}(i) = 1 + e_a^m + e_a^{2m} + \cdots + e_a^{im}.
\]
Proof: Observe that
\[ C_{kb,a}(i+1) - e_a^{kb} C_{kb,a}(i) = 1, \]
for \( i = 0, ..., a - 3 \). Also observe that \( e_a^{kb} C_{kb,a}(a - 2) = e_a^{b} + e_a^{2b} + \cdots + e_a^{(a-1)b} = -1 \) as shown by lemma 3.4. Now multiply both sides of (3.15) by \( t - e_a^{b} \) and simplify.

We can now proceed with the proof of the theorem.

Proof of theorem 3.3: From equation (2.12) we write
\[
 f(N) = K_1k_{1,N} + K_2k_{2,N} + \sum_{k=1}^{a-1} \alpha_{1,k} e_a^{kN} + \sum_{k=1}^{b-1} \alpha_{2,k} e_b^{kN}. \tag{3.16}
\]
From equations (2.10), (2.13) and (2.14) we have
\[
 K_1 = \frac{a + b - 2}{2ab} \tag{3.17}
\]
\[
 K_2 = \frac{1}{ab} \tag{3.18}
\]
\[
 k_{1,N} = 1 \tag{3.19}
\]
\[
 k_{2,N} = N + 1 \tag{3.20}
\]
\[
 \alpha_{1,k} = \frac{1}{1 - e_a^{-kb}} \tag{3.21}
\]
\[
 \alpha_{2,k} = \frac{1}{1 - e_b^{-ka}} \tag{3.22}
\]
Therefore it remains to evaluate the sum
\[
 \sum_{k=1}^{a-1} \alpha_{1,k} e_a^{kN} = \frac{1}{a} \sum_{k=1}^{a-1} \frac{e_a^{kN}}{1 - e_a^{-kb}} \]
\[
 = \frac{1}{a} \sum_{k=1}^{a-1} e_a^{kN} \left( 1 - \frac{1}{1 - e_a^{-kb}} \right) \]
\[
 = \frac{1}{a} \sum_{k=1}^{a-1} e_a^{kN} - \frac{1}{a} \sum_{k=1}^{a-1} \frac{e_a^{kN}}{1 - e_a^{-kb}} \]
\[
 = \frac{1}{a} (a e_a|N - 1) - \frac{1}{a} \sum_{k=1}^{a-1} \frac{e_a^{kN}}{1 - e_a^{-kb}}. \tag{3.21}
\]
where we have written the last line from lemma 3.4. We now proceed to evaluate the sum of roots of unity that appears in the above sum:

\[
\sum_{k=1}^{a-1} \frac{e_{kN}}{1 - e_{ka}^{kb}} = \sum_{k=1}^{a-1} e_{ka}^{kN} \cdot \frac{1}{1 - e_{ka}^{kb}} = \sum_{k=1}^{a-1} e_{ka}^{kN} \cdot \frac{(1 - e_{ka}^b) \cdots (1 - e_{ka}^{(a-1)b})}{(1 - e_{ka}^b) \cdots (1 - e_{ka}^{(a-1)b})} = \sum_{k=1}^{a-1} e_{ka}^{kN} \cdot \frac{1}{a} \cdot \prod_{j=1, j \neq k}^{a-1} (1 - e_{ka}^j) = \frac{1}{a} \sum_{k=1}^{a-1} e_{ka}^{kN} \cdot \sum_{i=0}^{a-2} C_{kb,a}(i) \tag{3.22}
\]

where we write the last line from (3.15). We now calculate this last sum. For this first observe that

\[
C_{kb,a}(0) = 1 \quad C_{kb,a}(1) = 1 + e_{ka}^{kb} \quad C_{kb,a}(2) = 1 + e_{ka}^{kb} + e_{ka}^{2kb} \\
\vdots \quad \vdots \quad \vdots \\
C_{kb,a}(a-2) = 1 + e_{ka}^{kb} + e_{ka}^{2kb} + \cdots + e_{ka}^{(a-2)kb} \\
\sum_{i=0}^{a-2} C_{kb,a}(i) = (a-1) + (a-2)e_{ka}^{kb} + \cdots + (a-1-j)e_{ka}^{jb} + \cdots + e_{ka}^{(a-2)kb} = \sum_{j=0}^{a-2} (a-1-j)e_{ka}^{jb}.
\]

Now we multiply this with \( e_{ka}^{kN} \) and sum over \( k \):

\[
\sum_{k=1}^{a-1} e_{ka}^{kN} \cdot \sum_{i=0}^{a-2} C_{kb,a}(i) = \sum_{k=1}^{a-1} \sum_{j=0}^{a-2} (a-1-j)e_{ka}^{k(N+jb)}
\]
Diophantine Equation of Frobenius

\[
\left( a - 1 \right) \sum_{k=1}^{a-1} e_a^{k(N+jb)} = \sum_{j=0}^{a-2} (a-1-j)(a \varepsilon_{a|(N+jb)} - 1),
\]  
(3.23)

where the last equation follows from lemma 3.4. Note that since \(0 \leq j \leq a-1\) and \((a,b) = 1\), there is only one value of \(j\) which satisfies \(N+jb \equiv 0 \mod a\). In definition 3.2 we denoted this unique solution by \(r_{a,b,N}\). Hence we have

\[
a \varepsilon_{a|(N+jb)} - 1 = \begin{cases} 
-1 & \text{if } j \neq r_{a,b,N} \\
 a-1 & \text{if } j = r_{a,b,N}
\end{cases}. 
\]  
(3.24)

Putting this in (3.23) we have

\[
\sum_{k=1}^{a-1} e_a^{kN} \cdot \sum_{i=0}^{a-2} C_{kb,a}(i) = a(a - 1 - r_{a,b,N}) - \sum_{j=0}^{a-2} (a - 1 - j) 
\]
\[
= a \left( \frac{a-1}{2} - r_{a,b,N} \right).
\]

Inserting this into (3.22) gives

\[
\sum_{k=1}^{a-1} \frac{e_a^{kN}}{1 - e_a^{kb}} = \left( \frac{a-1}{2} - r_{a,b,N} \right).
\]

Further inserting this into (3.21) gives

\[
\sum_{k=1}^{a-1} \alpha_{1,k} e_a^{kN} = \epsilon_{a|N} - \frac{a+1}{2a} + \frac{r_{a,b,N}}{a}. 
\]  
(3.25)

Similarly we find

\[
\sum_{k=1}^{b-1} \alpha_{2,k} e_b^{kN} = \epsilon_{b|N} - \frac{b+1}{2b} + \frac{r_{b,a,N}}{b}. 
\]  
(3.26)

Finally we substitute equations (3.25), (3.26), (3.17), (3.18), (3.19) and (3.20) into equation (3.16) and obtain

\[
f(N) = \frac{N + b r_{a,b,N} + a r_{b,a,N}}{ab} + \epsilon_{a|N} + \epsilon_{b|N} - 1.
\]
as claimed.

Remark: The fraction \( \frac{N + br_{a,b,N} + ar_{b,a,N}}{ab} \) is always an integer; \( a|(N + br_{a,b,N}) \) by definition, so \( a|(N + br_{a,b,N} + ar_{b,a,N}) \). Similarly \( b|(N + br_{a,b,N} + ar_{b,a,N}) \). But \( (a, b) = 1 \), so \( ab|(N + br_{a,b,N} + ar_{b,a,N}) \).

Remark: That we have obtained a computable formula can be checked easily; for example a simple Maple V program can calculate immediately that

\[
    f(1234567, 12345678; 123456789012345) = 8.
\]

Recall equation (2.4) for notation.

References


Diophantine Equation of Frobenius


Bilkent University
Department of Mathematics
06533 Ankara, Turkey

sertoz@fen.bilkent.edu.tr