NUMERICAL SEMIGROUPS: APÉRY SETS AND HILBERT SERIES

JORGE L. RAMÍREZ ALFONSÍN AND ØYSTEIN J. RØDSETH

Dedicated to the memory of Ernst S. Selmer (1920–2006), whose calculations revealed the "Selmer group".

ABSTRACT. Let $a_1, \ldots, a_n$ be relatively prime positive integers, and let S be the semigroup consisting of all non-negative integer linear combinations of $a_1, \ldots, a_n$. In this paper, we focus our attention on AA-semigroups, that is semigroups being generated by almost arithmetic progressions. After some general considerations, we give a characterization of the symmetric AA-semigroups. We also present an efficient method to determine an Apéry set and the Hilbert series of an AA-semigroup.

1. INTRODUCTION

Let $a_1, \ldots, a_n$ be relatively prime positive integers, and let $S = \langle a_1, \ldots, a_n \rangle$ be the set of all non-negative integer linear combinations of $a_1, \ldots, a_n$, that is,

$$S = \{a_1x_1 + \cdots + a_nx_n \mid x_i \in \mathbb{Z}_{\geq 0}\}.$$

Then S is a *numerical semigroup*.

The largest integer *not* in S is called the *Frobenius number* of S, and denoted by $g = g(S)$. The number $g + 1$ is known as the *conductor* of S. The non-negative integers *not* in S are called the *gaps* of S. The number of gaps of S is called the *genus* of S, and denoted $N(S)$.

The numerical semigroup S is *symmetric* if

$$S \cup (g - S) = \mathbb{Z}, \tag{1}$$

where $g - S = \{g - s \mid s \in S\}$. Symmetric semigroups are of special interest because of their rôle in the classification of plane algebraic branches, and because they provide us with answers to questions about monomial curves being complete intersections and also if certain rings are Gorenstein, cf. [21, p. 142].

The *Apéry set* of S with respect to a nonzero $m \in S$ is defined as

$$Ap(S; m) = \{s \in S \mid s - m \notin S\}.$$

Date: December 26, 2008.

2000 *Mathematics Subject Classification.* 05A15, 13P10, 20M14.

Key words and phrases. Numerical semigroup, Hilbert series, Apéry set, Frobenius number, genus.

To determine an Apéry set of a semigroup S is in general very difficult. For, if we know the Apéry set, then, by (2) below, we know $g(S)$. Ramírez Alfonsín [20] has, however, shown that the computational complexity of $g(S)$ is *hard*.

Finally, the *Hilbert series* of S is given by

$$H(S;x) = \sum_{s \in S} x^s.$$

The main intention of this paper is to study the Apéry sets and the Hilbert series of *almost arithmetic semigroups* (AA-semigroups, for short); i.e., semigroups which are generated by *almost arithmetic progressions*, that is, the cases where all but one of the a_i form an ordinary *arithmetic progression*.

In the following section we make some observations about the Apéry sets and the Hilbert series of a general numerical semigroup, and we give simple proofs of some known facts.

In Section 4 we consider AA-semigroups. Rødseth [24] gave an algorithm for the computation of an Apéry set of an AA-semigroup, while Patil and Roberts [19] considered AA-semigroups from a more algebraic point of view. In Section 4.1 we summarize some of Rødseth's results, in terms of which we express the Hilbert series of an AA-semigroup. In Section 4.2 we use Rødseth's results to characterize a symmetric AA-semigroup. Rødseth's algorithm works well on average, but, unfortunately, not from a worst case point of view. For the Frobenius number of an AA-semigroup with $k = 1$, that is, for $g(a, b, c)$, a faster algorithm was given by Greenberg [11]. In Section 4.3 we use Greenberg's idea to speed up Rødseth's algorithm considerably, thus generalizing Greenberg's result.

In the final section we comment on some of the literature on numerical semigroups.

Finite arithmetic progressions of positive integers with a common difference d appear throughout the paper. For simplicity, we set $d > 0$. However, most of the results also hold (or can be adjusted to hold) for d negative; see [24].

2. Apéry Sets and Hilbert Series

The Apéry set $Ap(S;m)$ constitutes a complete set of residues mod m, and we shall write $w(i)$ for the unique $w \in Ap(S;m)$ satisfying $w \equiv i \pmod{m}$. We have

$$g(S) = \max Ap(S;m) - m; \tag{2}$$

a result essentially due to Brauer and Shockley [3] (the case $m = a_1$). We also have

$$N(S) = \frac{1}{m} \sum_{w \in Ap(S;m)} w - \frac{1}{2}(m-1); \tag{3}$$

which is essentially due to Selmer [27] (the case $m = a_1$).

Since

$$S = Ap(S;m) + m\mathbb{Z}_{\geq 0},$$

we have

$$H(S;x) = \sum_{s\in S} x^s = \sum_{w\in Ap(S;m)} \sum_{i=0}^{\infty} x^{w+mi} = \sum_{w\in Ap(S;m)} x^w \sum_{i=0}^{\infty} x^{mi};$$

hence

$$H(S;x) = \frac{1}{1-x^m} \sum_{w\in Ap(S;m)} x^w. \tag{4}$$

If the *degree* of a rational function $\varphi(x)/\psi(x)$ is defined as $\deg\varphi(x) - \deg\psi(x)$ then, by (2) and (4), we have

$$\deg H(S;x) = \max Ap(S;m) - m = g(S). \tag{5}$$

In addition, we have

$$\sum_{s\in S} x^s + \sum_{s \text{ gap}} x^s = \frac{1}{1-x}$$

and thus

$$G(S;x) = \frac{1}{1-x} - H(S;x)$$

for a polynomial $G(S,x)$ of degree $g(S)$ in x, and where $G(S;1) = N(S)$.

By (4), if we know $Ap(S;m)$, then we also know $H(S;x)$. It may, however, require some effort to write $H(S;x)$ as a simple rational function. Conversely, if $H(S;x)$ is known, then $Ap(S;m)$ is completely determined by

$$\sum_{w\in Ap(S;m)} x^w = (1-x^m)H(S;x).$$

So, information about $g(S)$ and $N(S)$ can be obtained from either $H(S;x)$ or $Ap(S;m)$. Many papers about $g(S)$ (and $N(S)$) do, in fact, study $Ap(S;m)$ for some m; usually m is taken as an a_i. Thus results on $H(S;x)$ are implicit in many papers on $g(S)$.

We set $m = a_1$ in (4) to get

$$H(S;x) = \frac{1}{1-x^{a_1}} \sum_{w\in Ap(S;a_1)} x^w, \tag{6}$$

so that

$$H(S;x) = \frac{f(S;x)}{(1-x^{a_1})\dots(1-x^{a_n})}$$

where $f(S;x)$ is the polynomial with integer coefficients, given by

$$f(S;x) = (1-x^{a_2})\dots(1-x^{a_n}) \sum_{w\in Ap(S;a_1)} x^w.$$

Notice that $Ap(S;a_1) \subseteq \langle a_2,\dots,a_n\rangle$. For if $w(i) = a_1x_1 + \dots + a_nx_n \in Ap(S;a_1)$, then $x_1 = 0$, since $w(i) - a_1 \notin Ap(S;a_1)$. Let d be a positive common divisor of $a_2,\dots,a_n$, and set $S_d = \langle a_1, a_2/d,\dots,a_n/d\rangle$. If $w_d(i) \in Ap(S_d;a_1)$, then $w_d(i) \mapsto$

$dw_d(i) = w(di)$ defines a bijection $Ap(S_d; a_1) \to Ap(S; a_1)$. This map is injective since $\gcd(a_1, d) = 1$, and thus bijective since both Apéry sets contain exactly a_1 elements. In short,

$$Ap(S; a_1) = dAp(S_d; a_1); \tag{7}$$

that is, $w \in Ap(S_d; a_1)$ if and only if $dw \in Ap(S; a_1)$.

Now, by (4) and (7), we have

$$H(S; x) = \frac{1 - x^{da_1}}{1 - x^{a_1}} H(S_d; x^d), \tag{8}$$

which may be written as

$$f(S; x) = f(S_d; x^d).$$

Moreover, using (5), we get the well-known result

$$g(S) = dg(S_d) + a_1(d - 1), \tag{9}$$

while the quickest way to its companion

$$N(S) = dN(S_d) + \frac{1}{2}(a_1 - 1)(d - 1), \tag{10}$$

is via (3) and (7); cf. [3, 13], [21, pp. 36, 47].

Putting $d = b$ in (9) and (10), we get

$$g(a, b) = ab - a - b \quad \text{and} \quad N(a, b) = \frac{1}{2}(a - 1)(b - 1), \tag{11}$$

two results usually attributed to Sylvester (and partly to Curran Sharp); cf. [21, pp. 31, 103, and 104]. Here we have written $g(a, b)$ for $g(\langle a, b\rangle)$; we will use a similar notation in other cases.

We close this section with a brief discussion of symmetry. In Section 4.2 we shall make use of the following result, which is essentially a lemma due to Apéry [1]; see [21, Lemma 7.2.16]. As before, $m \in S$, $m \neq 0$, and the elements of $Ap(S; m)$ are $w(i)$.

Lemma 1 ([1])**.** *A numerical semigroup S is symmetric if and only if there is an i_0, such that*

$$w(i_0) - w(i) = w(i_0 - i) \tag{12}$$

for all i.

Notice that if (12) holds for some (fixed) i_0 and all i, then $w(i_0) = \max Ap(S; m)$. Also notice that Lemma 1 remains valid if we replace Eq. (12) by the inequality

$$w(i_0) - w(i) \geq w(i_0 - i),$$

since the opposite inequality is always true.

As an easy consequence of Lemma 1 follows the well-known fact that S is symmetric if and only if

$$g(S) + 1 = 2N(S); \tag{13}$$

cf. [21, Lemma 7.2.3].

Fröberg, Gottlieb, and Häggkvist [10, Definition, p. 71] defines the *derived* semigroup S' of S as the semigroup

$$S' = \left\langle \frac{a_1}{\prod_{j\neq 1} d_j}, \ldots, \frac{a_n}{\prod_{j\neq n} d_j} \right\rangle,$$

where $d_i = \gcd(a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n)$ for each $i = 1, \ldots, n$. Notice that in order to determine $H(S; x)$, repeated applications of Proposition (8) leave us with the problem of determining $H(S'; x)$; similarly for Apéry sets, Frobenius number, and genus.

By (11) and (13), $S = \langle a, b\rangle$ is always symmetric. It follows immediately by (9), (10), and (13), that a general S is symmetric if and only if S_d is symmetric. Hence, S is symmetric if and only if the derived semigroup S' is symmetric; see [10, Corollary, p. 71].

3. Generalized Arithmetic Progressions

In this section, we consider a semigroup S_A generated by *a generalized arithmetic progression*; that is, a sequence of relatively prime positive integers $a, ha+d, \ldots, ha+kd$, where also a, h, d, k are positive integers. Since the generators are relatively prime, we have, in particular, $\gcd(a, d) = 1$. The following lemma is due to Roberts [22]; for a short proof see [24, Lemma 1].

Lemma 2. *We have $M \in \langle a, a+d, a+2d, \ldots, a+kd\rangle$ if and only if there exist integers x, y such that*

$$M = ax + dy \quad \textit{with} \quad 0 \leq y \leq kx.$$

The next result follows by adjoining ha to the generators of S_A.

Proposition 1 ([27])**.** *We have*

$$Ap(S_A; a) = \left\{ ha \left\lceil \frac{r}{k} \right\rceil + dr \mid 0 \leq r < a \right\}.$$

Proposition 1 was rediscovered by Matthews [15].

Let s be a non-negative integer, and set $q = \lceil (s-1)/k \rceil$. An easy summation gives

$$\sum_{r=0}^{s-1} x^{a\lceil r/k\rceil + dr} = \frac{F_s(a; x)}{(1-x^d)(1-x^{a+kd})} \tag{14}$$

where

$$F_s(a; x) = (1-x^{a+kd})(1-x^{aq+ds}) - x^d(1-x^a)(1-x^{(a+kd)q}).$$

Using (4) with $m = a$, Proposition 1, and (14), we get the following result.

Proposition 2. *We have*

$$H(S_A; x) = \frac{F_a(ha; x)}{(1-x^a)(1-x^d)(1-x^{ha+kd})}.$$

where

$$F_a(ha;x) = (1-x^{ha+kd})(1-x^{haq+ad}) - x^d(1-x^{ha})(1-x^{(ha+kd)q}).$$

The next theorem tells us when S_A is symmetric; cf. Estrada and López [8], and Matthews [14].

Theorem 1. *The semigroup S_A is symmetric if and only if $a = 1$ or $a \equiv 2 \pmod{k}$.*

Numerical semigroups generated by generalized arithmetic progressions have also been studied by Selmer [27], and Rødseth [24]; see [21] for further references. In addition to Theorem 1, Matthews [14] gave similar results for S_A being *pseudo-symmetric* and also for having the *Arf* property.

4. Almost Arithmetic Progressions

A *numerical* AA-*semigroup* is a semigroup generated by relatively prime positive integers $a, a+d, a+2d, \ldots, a+kd, c$, where also a, d, k, c are positive integers. We shall now consider the Apéry set $Ap(S;a)$ and the Hilbert series $H(S;x)$ of a numerical AA-semigroup. We set $\gcd(a,d) = 1$, which does not imply any loss of generality; cf. Section 2.

4.1. **A Finite Negative-Regular Continued Fraction.** We first recall some notions and results given in [24]. Let $s_{-1} = a$ and determine s_0 by

$$ds_0 \equiv c \pmod{s_{-1}}, \quad 0 \le s_0 < s_{-1}.$$

If $s_0 \neq 0$, we use the Euclidean algorithm with negative division remainders,

$$\begin{aligned}
s_{-1} &= q_1 s_0 - s_1, && 0 \le s_1 < s_0;\\
s_0 &= q_2 s_1 - s_2, && 0 \le s_2 < s_1;\\
s_1 &= q_3 s_2 - s_3, && 0 \le s_3 < s_2;\\
&\cdots\\
s_{m-2} &= q_m s_{m-1} - s_m, && 0 \le s_m < s_{m-1};\\
s_{m-1} &= q_{m+1} s_m, && 0 = s_{m+1} < s_m.
\end{aligned}$$

If $s_0 = 0$, we put $m = -1$. If $m \ge 0$, we have

$$\frac{s_{-1}}{s_0} = q_1 - \cfrac{1}{q_2 - \cfrac{1}{q_3 - \cfrac{1}{\ddots \; - \cfrac{1}{q_m - \cfrac{1}{q_{m+1}}}}}}$$

which is known as the *Jung-Hirzebruch continued fraction* of s_{-1}/s_0.

We have $s_m = \gcd(a, c)$. We define integers P_i by $P_{-1} = 0$, $P_0 = 1$, and (if $m \geq 0$),

$$P_{i+1} = q_{i+1}P_i - P_{i-1}, \quad i = 0, \ldots, m.$$

Then, by induction on i,

$$s_i P_{i+1} - s_{i+1} P_i = a, \qquad i = -1, 0 \ldots, m,$$

and

$$-1 = P_{-1} < 0 = P_0 < \cdots < P_{m+1} = \frac{a}{s_m}.$$

In addition we have,

$$ds_i \equiv cP_i \pmod{a}, \quad i = -1, \ldots, m+1. \tag{15}$$

Putting

$$R_i = \frac{1}{a}\left((a + kd)s_i - kcP_i\right),$$

we then see that all the R_i are integers. Moreover, we have $R_{-1} = a + kd$, $R_0 = \frac{1}{a}\left((a + kd)s_0 - kc\right)$, and

$$R_{i+1} = q_{i+1}R_i - R_{i-1}, \quad i = 0, \ldots, m,$$

and again we see that all the R_i are integers. Furthermore,

$$-\frac{c}{s_m} = R_{m+1} < R_m < \cdots < R_0 < R_{-1} = a + kd,$$

so there is a unique integer v such that

$$R_{v+1} \leq 0 < R_v.$$

For each i, let $\tau(i)$ be the smallest integer for which there exist non-negative integers y, z such that

$$\tau(i) = (a + kd)y + kcz, \quad dy + cz \equiv i \pmod{a}. \tag{16}$$

If there are more than one such pair y, z for some i, choose the one with z minimal. This gives us a unique set L of pairs of nonnegative integers, with $|L| = a$. The set L consists of all lattice points in a closed L-shaped region in the y, z-plane. (Sometimes the L-shape degenerates to a rectangle or an interval.) Rødseth [24] proved that $L = A \cup B$, where

$$A = \{(y, z) \in \mathbb{Z}^2 \mid 0 \leq y < s_v - s_{v+1}, 0 \leq z < P_{v+1}\}, \tag{17}$$

$$B = \{(y, z) \in \mathbb{Z}^2 \mid 0 \leq y < s_v, 0 \leq z < P_{v+1} - P_v\}. \tag{18}$$

Rødseth also proved the following theorem.

Theorem 2 ([24])**.** *For the numerical AA-semigroup S, we have*

$$Ap(S; a) = \left\{a\left\lceil\frac{y}{k}\right\rceil + dy + cz \mid (y, z) \in A \cup B\right\}.$$

Now, we have the following result for the Hilbert series.

Theorem 3. *For the numerical AA-semigroup S, we have*

$$H(S;x)=\frac{F_{s_v}(a;x)(1-x^{c(P_{v+1}-P_v)})+F_{s_v-s_{v+1}}(a;x)(x^{c(P_{v+1}-P_v)}-x^{cP_{v+1}})}{(1-x^a)(1-x^d)(1-x^{a+kd})(1-x^c)}.$$

Proof. By (4) with $m=a$ and Theorem 2, we have

$$H(S;x)=\frac{1}{1-x^a}\sum_{(y,z)\in A\cup B}x^{a\lceil y/k\rceil+dy+cz}.$$

We split the sum like this,

$$\sum_{(y,z)\in A\cup B}=\sum_{y=0}^{s_v-1}\sum_{z=0}^{P_{v+1}-P_v-1}+\sum_{y=0}^{s_v-s_{v+1}-1}\sum_{z=P_{v+1}-P_v}^{P_{v+1}-1}.$$

Then we use (14) on each sum with summation variable y, one in each of the two parts, and the proof is easily completed. □

In particular, Theorem 3 implies the following result.

Theorem 4. *Let a,b,c be positive integers with* $\gcd(a,b)=1$. *Then,*

$$H(a,b,c;x)=\frac{1-x^{bs_v}-x^{cP_{v+1}}-x^{a(R_v-R_{v+1})}+x^{aR_v+cP_{v+1}}+x^{bs_v-aR_{v+1}}}{(1-x^a)(1-x^b)(1-x^c)}.$$

Proof. We apply Theorem 3 with $k=1$, $b=a+d$. In particular, we have $F_s(a;x)=(1-x^d)(1-x^{bs})$. □

By Theorem 4, we obtain

$$g(a,b,c)=\deg H(a,b,c;x)=\max\{aR_v+cP_{v+1},bs_v-aR_{v+1}\}-a-b-c;$$

a result due to Rødseth [23, Theorem 1].

Several authors have proved Theorem 4 in one form or another. We say more about this in Section 5.

4.2. **Symmetry.** We shall now see how to decide if an AA-semigroup S is symmetric or not. We continue to use the notation and results from the previous section. In particular, we still assume that $\gcd(a,d)=1$, which represents no restriction.

However, to gain some extra insight before we state our theorem and present its proof, we first consider the case $k=1$; that is $S=\langle a,b,c\rangle$ (with $b=a+d$). In this case the set L was given in [23]. By Lemma 1, S is symmetric if the L-shaped plane region degenerates to a rectangle (or an interval); that is, if $v=-1$ or $v=m$. Otherwise, if S is symmetric or not, depends on the two right upper corners of the L-shaped region; that is, the lattice points $(s_v-1,P_{v+1}-P_v-1)$ and $(s_v-s_{v+1}-1,P_{v+1}-1)$. The difference between the corresponding elements in $Ap(S;a)$ is

$$b(s_v-1)+c(P_{v+1}-P_v-1)-(b(s_v-s_{v+1}-1)+c(P_{v+1}-1))$$
$$=aR_{v+1}+c(P_{v+1}-P_v),$$

so that S is symmetric if $R_{v+1} = 0$. It is easy to see that this is the only possibility for symmetry in the non-degenerate case.

A simple check reveals that in each of the three cases $v = -1$, $v = m$, $R_{v+1} = 0$, the derived semigroup S' is generated by (at most) two of the three generators. Thus $S = \langle a, b, c\rangle$ is symmetric if and only if its derived semigroup is generated by two elements, a result due to Fröberg, Gottlieb, and Häggkvist [10, Corollary p. 77].

We now turn to general $k \geq 1$.

Theorem 5. *The numerical AA-semigroup S is symmetric if and only if one of the following conditions is satisfied.*

(i) $s_v = 1$,
(ii) $s_v \equiv 2 \pmod{k}$ *and* $s_{v+1} = 0$,
(iii) $s_v \equiv 2 \pmod{k}$ *and* $s_v = a$,
(iv) $s_v - s_{v+1} = 1$ *and* $R_{v+1} = 1 - k$,
(v) $s_v \equiv 2 \pmod{k}$ *and* $s_v - s_{v+1} > 1$ *and* $R_{v+1} = 0$,
(vi) $k \geq 2$ *and* $s_{v+1} = k - 1$ *and* $R_v = 1$.

A straightforward way to prove Theorem 5 would be to use Theorem 2 to determine $g(S)$ and $N(S)$, and then find necessary and sufficient conditions for (13) to hold. However, this turns out to be rather complicated. We will avoid this by using Lemma 1.

Proof of Theorem 5. We shall apply Lemma 1 with $m = a$. We set $w(i) = w(y, z)$ if $i \equiv dy + cz \pmod{a}$ and $(y, z) \in L = A \cup B$, for A, B given by (17) and (18). In particular, we write $w(i_0) = w(y_0, z_0)$. By Theorem 2, we then have

$$w(y, z) = a\left\lceil \frac{y}{k} \right\rceil + dy + cz.$$

Also notice that

$$R_i = s_i + k\frac{1}{a}\left(ds_i - cP_i\right) \equiv s_i \pmod{k}.$$

We consider four cases. In the first two cases, the L-shape of L degenerates, and in each case there is only one possible choice of (y_0, z_0) such that i_0 may satisfy (12). Otherwise we have two possible choices for (y_0, z_0); one choice is covered by Case 3, and the other is covered by Case 4.

Case 1. Suppose that $s_{v+1} = 0$. Then $v = m$, $P_{v+1} = a/s_m$, and

$$L = \{(y, z) \mid 0 \leq y < s_m,\ 0 \leq z < a/s_m\}.$$

Let $w(i_0) = w(s_m - 1, a/s_m - 1)$. Now, $w(i_0 - i) = w(s_m - 1 - y, a/s_m - 1 - z)$, and S is symmetric if and only if

$$\left\lceil \frac{s_m - 1}{k} \right\rceil - \left\lceil \frac{y}{k} \right\rceil = \left\lceil \frac{s_m - 1 - y}{k} \right\rceil \qquad \text{for } 0 \leq y < s_m;$$

that is, if and only if (i) or (ii) holds. (Notice that (i) implies $s_{v+1} = 0$.)

Case 2. Suppose that $s_v = a$ and $s_{v+1} \neq 0$. Then $L = \{(y, 0) \mid 0 \leq y < a\}$. Let $(y_0, z_0) = (a-1, 0)$. Then $w(i_0 - i) = w(a - 1 - y, 0)$, and, by Lemma 1, S is symmetric if and only if

$$\left\lceil \frac{a-1}{k} \right\rceil - \left\lceil \frac{y}{k} \right\rceil = \left\lceil \frac{a-1-y}{k} \right\rceil \qquad \text{for } 0 \leq y < a,$$

or equivalently, if and only if $a \equiv 2 \pmod{k}$. (We have $a = s_v \neq 1$, since $s_{v+1} \neq 0$.)

Case 3. Let $s_v \neq a$ and $s_{v+1} \neq 0$. Assume that $(y_0, z_0) = (s_v - 1, P_{v+1} - P_v - 1)$.

If $0 \leq z < P_{v+1} - P_v$ and $0 \leq y < s_v$, then $w(i_0 - i) = w(s_v - 1 - y, P_{v+1} - P_v - 1 - z)$, and (12) holds if and only if

$$\left\lceil \frac{s_v - 1}{k} \right\rceil - \left\lceil \frac{y}{k} \right\rceil = \left\lceil \frac{s_v - 1 - y}{k} \right\rceil;$$

that is, if and only if $s_v \equiv 2 \pmod{k}$. (We have $s_v > s_{v+1} \geq 1$.)

If $P_{v+1} - P_v \leq z < P_{v+1}$, then $0 \leq y < s_v - s_{v+1}$. (Notice that $P_v \neq 0$ since $s_v \neq a$.) It is easily seen that $w(i_0 - i) = w(s_v - s_{v+1} - 1 - y, 2P_{v+1} - P_v - 1 - z)$, and for $s_v \equiv 2 \pmod{k}$, Eq. (12) holds if and only if

$$1 - \left\lceil \frac{y}{k} \right\rceil = \left\lceil \frac{1 - y - R_{v+1}}{k} \right\rceil \quad \text{for all } 0 \leq y < s_v - s_{v+1}. \tag{19}$$

If $s_v - s_{v+1} = 1$, then $y = 0$, and (19) holds if and only if $R_{v+1} \geq 1 - k$. Since

$$R_{v+1} \equiv s_{v+1} \equiv s_v - 1 \equiv 1 \pmod{k},$$

Eq. (19) holds if and only if $R_{v+1} = 1 - k$.

If $s_v - s_{v+1} > 1$, we first set $y = 1$, and find that $R_{v+1} = 0$. Then it is easily seen, that Eq. (19) holds if and only if $R_{v+1} = 0$. Hence, in this case, S is symmetric if and only if (iv) or (v) holds.

Case 4. Let $s_v \neq a$, $s_{v+1} \neq 0$, and assume that $(y_0, z_0) = (s_v - s_{v+1} - 1, P_{v+1} - 1)$.

If $0 \leq y < s_v - s_{v+1}$, then we have $w(i_0 - i) = w(s_v - s_{v+1} - 1 - y, P_{v+1} - 1 - z)$, and (12) holds if and only if

$$\left\lceil \frac{s_v - s_{v+1} - 1}{k} \right\rceil - \left\lceil \frac{y}{k} \right\rceil = \left\lceil \frac{s_v - s_{v+1} - 1 - y}{k} \right\rceil.$$

This is if and only if $s_v - s_{v+1} = 1$ or $s_v - s_{v+1} \equiv 2 \pmod{k}$.

Moreover, if $s_v - s_{v+1} \leq y < s_v$, then $w(i_0 - i) = w(2s_v - s_{v+1} - 1 - y, P_{v+1} - P_v - 1 - z)$. For (12) to be satisfied, we must have

$$\left\lceil \frac{s_v - s_{v+1} - 1}{k} \right\rceil - \left\lceil \frac{y}{k} \right\rceil = \left\lceil \frac{s_v - s_{v+1} - 1 - y + R_v}{k} \right\rceil. \tag{20}$$

Suppose that $s_v - s_{v+1} = 1$. We have $s_v > s_{v+1} \geq 1$, and (20) gives for $y = 1$,

$$-1 = \left\lceil \frac{-1 + R_v}{k} \right\rceil,$$

so that $R_v \leq 0$, which is impossible.

If $s_v - s_{v+1} > 1$, we consider the case $s_v - s_{v+1} \equiv 2 \pmod{k}$. Then (20) becomes

$$1 - \left\lceil \frac{y}{k} \right\rceil = \left\lceil \frac{1 - y + R_v}{k} \right\rceil. \tag{21}$$

For S to be symmetric, this identity must be satisfied for $s_v - s_{v+1} \leq y < s_v$. For $y = s_v - s_{v+1}$, Eq. (21) holds if and only if $k \geq 2$ and $R_v = 1$. (Notice that $R_v = 1$ implies $s_v \equiv 1 \pmod{k}$.) If the conditions $k \geq 2$ and $R_v = 1$ are satisfied, Eq. (21) implies that $y \not\equiv 1 \pmod{k}$. Now it follows easily that S is symmetric if and only if (vi) holds. □

If S is symmetric, then $g(S)$ is odd by (13). For $g(S)$ even, we have the corresponding notion of S being a *pseudo*-symmetric semigroup, which holds if and only if

$$S \cup (g - S) = \mathbb{Z} \setminus \{g/2\}.$$

For pseudo-symmetric semigroups, we have a result corresponding to Lemma 1, which can be used as a starting point for a study of pseudo-symmetric AA-semigroups, using the machinery above.

4.3. **Speeding up the AA-Algorithm.** We continue our considerations of the numerical AA-semigroup S generated by $a, a + d, \ldots, a + kd, c$. Also in this section we use the notation of Sections 4.1 and 4.2.

If $\gcd(a, d) = 1$, then, as soon as we know the values of $s_v, s_{v+1}, P_v, P_{v+1}$, we can plug these into the theorems in Sections 4.1 and 4.2, and get out the Apéry set $Ap(S; a)$ and the Hilbert series $H(S; x)$. Moreover, we get the Frobenius number, the genus, and we can easily decide whether S is symmetric or not.

So, we first apply the Euclidean algorithm (with positive division remainders) to remove a possible $\gcd(a, d) > 1$ to make a and d coprime, and at the same time we compute s_0.

With the goal of determining $Ap(S; a)$ and $H(S; x)$, we now need to know the values of $s_v, s_{v+1}, P_v, P_{v+1}$. Note that we do not need to know v. We may proceed as described in Section 4.1, where we applied the Euclidean algorithm with negative division remainders to the quotient s_{-1}/s_0 until we got the stop order $R_{v+1} \leq 0$. This algorithm works well on average; cf. [18, p. 423]. Unfortunately, the number m of steps in the Euclidean algorithm with negative division remainders can be as large as $a - 2$. We now avoid this by using the Euclidean algorithm with *positive* division remainders instead of negative, thus limiting the number of steps in this Euclidean algorithm to $O(\log a)$. This approach was first used by Greenberg [11] for the computation of $g(a, b, c)$.

We now present the algorithm for the determination of $s_v, s_{v+1}, P_v, P_{v+1}$, using the Euclidean algorithm with *positive* division remainders.

Algorithm Apéry

Input: a, d, c, k, s_0
Output: $s_v, s_{v+1}, P_v, P_{v+1}$

1. $r_{-1} = a, r_0 = s_0$

2. $r_{i-1} = \kappa_{i+1} r_i + r_{i+1}, \quad \kappa_{i+1} = \lfloor r_{i-1}/r_i \rfloor, \quad 0 = r_{\mu+1} < r_\mu < \cdots < r_{-1}$

3. $p_{i+1} = \kappa_{i+1} p_i + p_{i-1}, \quad p_{-1} = 0, \quad p_0 = 1$

4. $T_{i+1} = -\kappa_{i+1} T_i + T_{i-1}, \quad T_{-1} = a + kd, \quad T_0 = \dfrac{1}{a}((a+kd)r_0 - kc)$

5. IF there is a minimal u such that $T_{2u+2} \le 0$, THEN

$$\begin{pmatrix} s_v & P_v \\ s_{v+1} & P_{v+1} \end{pmatrix} = \begin{pmatrix} \gamma & 1 \\ \gamma - 1 & 1 \end{pmatrix} \begin{pmatrix} r_{2u+1} & -p_{2u+1} \\ r_{2u+2} & p_{2u+2} \end{pmatrix}, \quad \gamma = \left\lfloor \frac{-T_{2u+2}}{T_{2u+1}} \right\rfloor + 1$$

6. ELSE $s_v = r_\mu, s_{v+1} = 0, P_v = p_\mu, P_{v+1} = p_{\mu+1}$.

Usually we do not need to know the value of P_v in step 6.

We now prove the correctness of Algorithm Apéry.

Theorem 6. *Algorithm Apéry determines precisely the quantities $s_v, s_{v+1}, P_v, P_{v+1}$, as defined in Section 4.1, in $O(\log a)$ steps.*

Proof. From steps 1–4 in the algorithm and induction on i, we get

$$r_i p_{i+1} + r_{i+1} p_i = a, \tag{22}$$

$$dr_i - (-1)^i c p_i \equiv 0 \pmod{a}, \tag{23}$$

$$T_i = \frac{1}{a}\left((a+kd)r_i - (-1)^i k c p_i\right). \tag{24}$$

Thus $T_{2i+1} > 0$. We have two cases.

Case 1. Assume $T_i \le 0$ for some $-1 \le i \le \mu + 1$. Then, by (24), i is even. Let u be minimal such that

$$T_{2u+2} \le 0.$$

Since $T_{2i+1} > 0$, we have, by step 4,

$$T_{2i+2} - T_{2i} = -\kappa_{2i+2} T_{2i+1} < 0,$$
$$T_{2i+1} - T_{2i-1} = -\kappa_{2i+1} T_{2i},$$

so that

$$\cdots < T_{2u+4} < T_{2u+2} \le 0 < T_{2u} < T_{2u-2} < \ldots,$$
$$\cdots > T_{2u+5} > T_{2u+3} \ge T_{2u+1} < T_{2u-1} < T_{2u-3} < \ldots$$

Let

$$-T_{2u+2} = (\gamma - 1) T_{2u+1} + T^*, \qquad 0 \le T^* < T_{2u+1}, \tag{25}$$

with

$$\gamma = \left\lfloor \frac{-T_{2u+2}}{T_{2u+1}} \right\rfloor + 1.$$

If $u \geq 0$, then by step 4,

$$\kappa_{2u+2}T_{2u+1} = T_{2u} - T_{2u+2} > -T_{2u+2} = (\gamma-1)T_{2u+1} + T^* \geq (\gamma-1)T_{2u+1}.$$

Since $T_{2u+1} > 0$, we thus have $\kappa_{2u+2} > \gamma - 1$; that is, $1 \leq \gamma \leq \kappa_{2u+2}$. It follows by step 3 that

$$p_{2u+2} - \gamma p_{2u+1} \geq p_{2u+2} - \kappa_{2u+2}p_{2u+1} = p_{2u} > 0.$$

Thus $p_{2u+2} > \gamma p_{2u+1}$, which is trivially true also if $u = -1$.

Now recall the definition of $w(i)$ given at the beginning of Section 2. Also recall that $\tau(i)$ is the smallest integer for which there exist nonnegative integers y, z satisfying (16). If there are more than one such pair for some i, we chose the one with z minimal. The set of such pairs y, z is $L = A \cup B$ for A, B given by (17), (18).

By (16), (25), and (24), we have

$$\tau(i) - aT^* = (a+kd)(y + (\gamma-1)r_{2u+1} + r_{2u+2}) + kc(z + (\gamma-1)p_{2u+1} - p_{2u+2}),$$

and, by (23),

$$d(y + (\gamma-1)r_{2u+1} + r_{2u+2}) + c(z + (\gamma-1)p_{2u+1} - p_{2u+2}) \equiv i \pmod{a}.$$

By the minimality of $\tau(i)$, we thus have if $T^* > 0$,

$$z < p_{2u+2} - (\gamma-1)p_{2u+1},$$

which is also true if $T^* = 0$, because of the minimality of z.

Similarly,

$$\begin{aligned}\tau(i) - a(T_{2u+1} - T^*) &= \tau(i) - a(\gamma T_{2u+1} + T_{2u+2}) \\ &= (a+kd)(y - \gamma r_{2u+1} - r_{2u+2}) + kc(z - \gamma p_{2u+1} + p_{2u+2}),\end{aligned}$$

and it follows that

$$y < \gamma r_{2u+1} + r_{2u+2}.$$

Finally, we have

$$\tau(i) - aT_{2u+1} = (a+kd)(y - r_{2u+1}) + kc(z - p_{2u+1}),$$

so that

$$y < r_{2u+1} \quad \text{or} \quad z < p_{2u+1}.$$

Let us set

$$\begin{aligned}A_1 &= \{(y,z) \mid 0 \leq y < r_{2u+1},\ 0 \leq z < p_{2u+2} - (\gamma-1)p_{2u+1}\}, \\ B_1 &= \{(y,z) \mid 0 \leq y < \gamma r_{2u+1} + r_{2u+2},\ 0 \leq z < p_{2u+1}\}.\end{aligned}$$

Then $L \subseteq A_1 \cup B_1$. Now, $|A_1 \cup B_1| = a$ by (22), and it follows by the uniqueness of L, that $A_1 = A$, $B_1 = B$. Hence we have the results in step 5.

Case 2. If $T_i > 0$ for $i = -1, 0, \ldots, \mu+1$ then by putting $i = \mu$ in (22), we see that $p_{\mu+1} = a/r_\mu$. Moreover, by (24) with $i = \mu + 1$ (recall that $r_{\mu+1} = 0$), we get

$$T_{\mu+1} = (-1)^\mu \frac{kc}{r_\mu}.$$

Hence, μ is even in this case, and (24) with $i = \mu$ gives us

$$T_\mu = \frac{1}{a}\left((a+kd)r_\mu - kcp_\mu\right).$$

Now,

$$w(i) - aT_\mu = (a+kd)(y - r_\mu) + kc(z + p_\mu),$$

where

$$d(y - r_\mu) + c(z + p_\mu) \equiv i \pmod{a},$$

because of (23) and the fact that μ is even. By the minimality of $\tau(i)$, we thus have

$$y < r_\mu.$$

Similarly,

$$\tau(i) - aT_{\mu+1} = (a+kd)\, y + kc\left(z - \frac{a}{r_\mu}\right)$$

gives

$$z < \frac{a}{r_\mu}.$$

Thus we have a pairs y, z satisfying (16). Since $dy + cz$ form a complete residue system modulo a, we have

$$L = \{(y,z) \mid 0 \le y < r_\mu,\ 0 \le z < \frac{a}{r_\mu}\}.$$

By the uniqueness of L, we now have $s_v = r_\mu$, so that $v = m$. Moreover, $s_{v+1} = 0$ and $P_{v+1} = a/r_\mu = p_{\mu+1}$. In addition, by (15) and (23), both $x = P_v$ and $x = p_\mu$ satisfy

$$\frac{c}{r_\mu}x \equiv d \pmod{\frac{a}{r_\mu}}, \quad 0 \le x < \frac{a}{r_\mu},$$

so that $P_v = p_\mu$, and we have the results in step 6.

Finally, it is well known how to use Fibonacci numbers to show that the index μ in step 2 satisfies $\mu \le \log a / \log \alpha$, where $\alpha = (1+\sqrt{5})/2$. Therefore, the index u and all the intermediate quantities needed to determine $s_v, s_{v+1}, P_v, P_{v+1}$ can be found in at most $O(\log a)$ steps. □

As an alternative to what we have done in this section, we could instead try to use a procedure for translating a negative-regular continued fraction into a regular one; cf. [18, Lemma 1]. Here may be something interesting hiding, since Dedekind sums appear in [18] and could perhaps be connected to numerical semigroups in this way. On the other hand, we already know from the book of Beck and Robins [2] that there is a connection to Dedekind sums.

5. Closing Comments

We have used combinatorial methods to study the numerical semigroup $S = \langle a_1, \ldots, a_n \rangle$. However, in the study of S, many authors use algebraic methods in the following way:

For a field $\Bbbk$, let $R = \Bbbk[x_1, \ldots, x_n]$ be the polynomial ring graded by $\deg x_i = a_i$ for $i = 1, \ldots, n$. Then $H(S; x)$ is the Hilbert series of this graded ring R. Let $A[S] = \Bbbk[t^{a_1}, \ldots, t^{a_n}]$ be the semigroup algebra associated with S. Let the map $\pi : R \longrightarrow A[S]$ be induced by $\pi(x_i) = t^{a_i}$, and let $I = \ker \pi$. Then $A[S] \cong R/I$. The *toric ideal* I is principal if $n = 2$. Herzog [12] showed that for $n = 3$, I has at most three generators. Then the situation changes dramatically, as Bresinsky [4] shows that for any (fixed) $n \geq 4$ and any $m \geq 1$, there exist $a_1, \ldots, a_n$ for which I requires at least m generators. Bresinsky's result implies that, for any (fixed) $n \geq 4$ and any $m \geq 1$, there exist $a_1, \ldots, a_n$ such that $f(S; x)$ contains at least m terms.

In between the nice situation $n = 2$ and the unwieldy situation $n \geq 4$, we have the case $n = 3$. The case $S = \langle a_1, a_2, a_3 \rangle$ has been the subject of many papers, with much overlap. The results are mainly expressed in two ways: Either *implicitly*, that is, in terms of quantities L_i and s_{ij}, where each L_i is the minimum positive integer for which the equation $a_i L_i = \sum_{j \neq i} a_j s_{ij}$ admits a solution in non-negative integers s_{ij}, or *semi-explicitly*[1], that is, algorithmically.

In the semi-explicit case the solution contains an algorithm for the computation of the L_i and the s_{ij}. For example, three integers always form an almost arithmetic progression. So, by putting $b = a + d$ and $k = 1$, Section 4.1 (and Section 4.3) contain semi-explicit results for $S = \langle a, b, c \rangle$. The following proposition, which is essentially [21, Claim 8.4.3], shows that the algorithms in Section 4.1 and in Section 4.3 determine the L_i and the s_{ij}.

Proposition 3. *For* $\gcd(a, b) = 1$, *we have*

$$\begin{aligned} a(R_v - R_{v+1}) &= b(s_v - s_{v+1}) + c(P_{v+1} - P_v), \\ bs_v &= aR_v + cP_v \qquad \text{if} \quad R_{v+1} \neq 0, \\ bs_{v+1} &= cP_{v+1} \qquad \text{if} \quad R_{v+1} = 0, \\ cP_{v+1} &= a(-R_{v+1}) + bs_{v+1}, \end{aligned}$$

where the left hand sides are the least positive multiples of a, b, b, c *belonging to* $\langle b, c \rangle$, $\langle a, c \rangle$, $\langle a, c \rangle$, $\langle a, b \rangle$, *respectively.*

Implicit results for $n = 3$ were given by Johnson [13], Brauer and Shockley [3], Fel [9], Herzog [12], Rosales and García-Sánchez [26], and others. A result by Selmer and Beyer [28] was the forerunner for the semi-explicit results of Rødseth [23] and Greenberg [11]. Various variants of Greenberg's result have been given in several papers, including [5, 6, 25]. Several results on the case $n = 3$, both implicit and semi-explicit, can be found in the monograph [21].

[1] We have borrowed this term from [2, p. 16].

Herzog [12] showed that the monomial curve $C = \Bbbk[t^a, t^b, t^c]$, considered as an R-module (with $n = 3$), has resolution

$$0 \to R^2 \xrightarrow{M} R^3 \to R \to C \to 0.$$

Herzog expressed the matrix M implicitly, and so did Denham [7], while Morales [16] gave M on a semi-explicit form, using the results of [23]. The Hilbert series of the graded ring C is then easily computed from M. The results show that $f(S;x)$ has at most four terms if $S = \langle a, b, c\rangle$ is symmetric, and at most six terms otherwise. (This is also clear from Theorem 4 and the beginning of Section 4.2.)

Finally, let us mention that Morales [17] used his results in [16] to construct a large class of monomial curves, where each curve is defined by an ideal P in R such that the symbolic Rees ring $R^{(P)}$ is Noetherian.

Acknowledgment

The authors thank the referee for careful reading of the first version of this paper, for valuable comments and suggestions, and for making us aware of the interesting reference [18].

References

[1] R. Apéry, Sur les branches superlinéaires des courbes algébriques, *C.R. Acad. Sci. Paris* **222** (1946), 1198-1200.
[2] M. Beck and S. Robins, *Computing the Continuos Discretely*, UTM, Springer, New York, 2007.
[3] A. Brauer and J. E. Shockley, On a problem of Frobenius, *J. Reine Angew. Math.* **211** (1962), 215–220.
[4] H. Bresinsky, On prime ideals with generic zero $x_i = t^{n_i}$, *Proc. Amer. Math. Soc.* **47** (1975), 329–332.
[5] Y. Cheng and F. K. Hwang, Diameters of weighted double loop networks, *J. Algorithms* **9** (1988), 401–410.
[6] J. L. Davison, On the linear diopantine problem of Frobenius, *J. Number Theory* **48** (1994), 353–363.
[7] G. Denham, Short generating functions for some semigroup algebras, *Electron. J. Combin.* **10** (2003), #R36, 7 pp. (electronic).
[8] M. Estrada and A. López, A note on symmetric semigroups and almost arithmetic sequences, *Comm. Algebra* **22** (10) (1994), 3903–3905.
[9] L. G. Fel, Frobenius problem for semigroups $S(d_1, d_2, d_3)$, *Func. Anal. Other Math.* **1** (2006), 119–157.
[10] R. Fröberg, C. Gottlieb and R. Häggkvist, On numerical semigroups, Semigroup Forum **35** (1987), 63–83.
[11] H. Greenberg, Solution to a diophantine equation for nonnegative integers, *J. Algorithms* **9** (3) (1988), 343–353.
[12] J. Herzog, Generators and relations of abelian semigroup rings, *Manuscripta Math.* **3** (1970), 175–193.
[13] S. M. Johnson, A linear diophantine problem, *Canad. J. Math.* **12** (1960), 390–398.
[14] G. L. Matthews, On numerical semigroups generated by generalized arithmetic sequences, *Comm. Algebra* **32** (9) (2004), 3459–3469.

[15] G. L. Matthews, On integers nonrepresentable by a generalized arithmetic sequences, *Integers* **5** (2) (2005), # A12, 6 pp. (electronic).
[16] M. Morales, *Syzygies of monomial curves and a linear diophantine problem of Frobenius*, Internal report, Max Planck Institut für Mathematik, Bonn, 1987.
[17] M. Morales, Noetherian symbolic blow-ups, *J. Algebra* **140** (1) (1991), 12–25.
[18] G. Myerson, On semi-regular finite continued fractions, *Arch. Math.* **48** (1987), 420–425.
[19] D. P. Patil and L. G. Roberts, Hilbert functions of monomial curves, *J. Pure Appl. Algebra* **183** (1–3) (2003), 275–292.
[20] J. L. Ramírez Alfonsín, Complexity of the Frobenius problem, *Combinatorica* **16** (1) (1996), 143-147.
[21] J. L. Ramírez Alfonsín, *The Diophantine Frobenius Problem*, Oxford Lectures Series in Mathematics and its Applications **30**, Oxford University Press (Oxford, 2005).
[22] J. B. Roberts, Note on linear forms, *Proc. Amer. Math. Soc.* **7** (1956), 465–469.
[23] Ø. J. Rødseth, On a linear diophantine problem of Frobenius, *J. Reine Angew. Math.* **301** (1978), 171–178.
[24] Ø. J. Rødseth, On a linear diophantine problem of Frobenius II, *J. Reine Angew. Math.* **307/308** (1979), 431–440.
[25] Ø. J. Rødseth, *Progression bases for finite cyclic groups*, in: Number Theory: New York Seminar 1991–1995, D. V. Chudnovsky, G. V. Chudnovsky, and M. B. Nathanson, eds., Springer, 1996, 269–279.
[26] J. C. Rosales and P. A. García-Sánchez, Numerical semigroups with embedding dimension three, *Arch. Math.* **83** (2004), 488–496.
[27] E. S. Selmer, On the linear diophantine problem of Frobenius, *J. Reine Angew. Math.* bf 293/294 (1977), 1–17.
[28] E. S. Selmer and Ø. Beyer, On the linear diophantine problem of Frobenius in three variables, *J. Reine Angew. Math.* **301** (1978), 161–170.

Equipe Combinatoire et Optimisation, Université Pierre et Marie Curie, Paris 6, 4 Place Jussieu, 75252 Paris Cedex 05
E-mail address: ramirez@math.jussieu.fr
URL: http://www.ecp6.math.jussieu.fr/pageperso/ramirez/ramirez.html

Department of Mathematics, University of Bergen, Johs. Brunsgt. 12, N-5008 Bergen, Norway
E-mail address: rodseth@math.uib.no
URL: http://math.uib.no/People/nmaoy