ACTA ARITHMETICA
138.2 (2009)

Representing integers as linear combinations of S-units

by

Zs. Ádám (Debrecen), L. Hajdu (Debrecen) and F. Luca (Morelia)

1. Introduction. Let A be a finite set of integers, and let $b_1, \dots, b_k$ be positive integers. Motivated by a question of M. Pohst, L. Hajdu [4] proved that there exist infinitely many primes which cannot be represented as $\sum_{i=1}^{k} a_i b_i^{u_i}$ with $a_i \in A$. However, this result is ineffective in the sense that on the one hand it does not give any detail about the distribution of the non-represented primes, and on the other hand it does not provide any information on the set of primes which can be represented in the desired form.

In this paper, we make some progress in this direction. First we prove that only a small fraction of integers can be represented as a linear combination of integral S-units for any fixed S, where the coefficients come from an arbitrary fixed finite set. An immediate consequence is that almost all primes are not representable as linear combinations of S-units. In the special but important case of combinations of pure powers, we get more precise results about the distribution of the numbers which can be represented in that form. The main tools of our proofs are the subspace theorem, and a classical result of Erdős, Pomerance and Schmutz [1] on the small values of the Carmichael function.

2. Results. To formulate our results, we need to introduce some notation. Let $S = \{p_1, \dots, p_t\}$ be some finite set of primes of cardinality t, and put $P = \max_{1 \le i \le t} p_i$. Let $T = \{p_1^{u_1} \cdots p_t^{u_t} : u_i \ge 0\}$ be the set of all positive integers whose prime factors belong to S. Further, let A be some finite set of integers, and write $n = \#A$ and $a^* = \max_{a \in A} |a|$. Finally, let k be a fixed positive integer. Throughout the paper, k will denote the number of terms in the representations by certain linear combinations. Put

$$H_{A,S,k} = \Big\{ v \in \mathbb{N} : v = \sum_{i=1}^{k} a_i s_i \text{ where } a_i \in A,\ s_i \in T\ (i = 1, \dots, k) \Big\}.$$

2000 *Mathematics Subject Classification*: Primary 11D85; Secondary 11A41.
Key words and phrases: S-units, primes, linear combination, representation.

DOI: 10.4064/aa138-2-1

For a real number $x \geq 1$, we put $H_{A,S,k}(x) = H_{A,S,k} \cap [1, x]$. Our first result shows that $H_{A,S,k}(x)$ is very small.

THEOREM 1. *The estimate*

$$\#H_{A,S,k}(x) \ll (\log x)^{tk}$$

holds for all large x. The implied constant depends on n, a^*, P, t and k.

REMARK 1. Under a certain rather strong assumption which is not satisfied in our case, Everest [2] proved a general asymptotic formula which, in particular, applies to count the number of those k-term linear combinations b of S-units which have no vanishing subsum and whose *S-norm* is smaller than x (see [2] for the precise result). The asymptotic formula in [2] is of the size $O((\log x)^{tk})$. Hence, our result is an upper bound of the same strength as the result of Everest but with no extra conditions imposed. It would be interesting to give a common generalization of Everest's result and our Theorem 1.

As a simple and immediate consequence of Theorem 1, we deduce that almost all primes are outside $H_{A,S,k}$. For the precise formulation of this statement, let $\mathbb{P}$ denote the set of all positive primes, and for $L \subseteq \mathbb{P}$ denote by $\mathcal{D}(L)$ the density of L inside $\mathbb{P}$, if it exists.

THEOREM 2. *In the above notation, let $L = \mathbb{P} \cap H_{A,S,k}$. Then $\mathcal{D}(L) = 0$.*

REMARK 2. It is widely believed that the number of Mersenne primes (i.e., primes of the form $2^u - 1$) is infinite. If this conjecture is true, then for the choice $k = 2$, $S = \{2\}$ and $A = \{-1, 1\}$, we find that $\mathbb{P} \cap H_{A,S,k}$ is infinite. Hence, Theorem 2 seems to be best possible in some sense.

Now we give a variant of Theorem 1 concerning representations of integers by linear combinations of pure powers. For this purpose we need some new notation. Let $b_1, \dots, b_k$ be positive integers and write $\underline{b} = (b_1, \dots, b_k)$. Put

$$H_{A,\underline{b}} = \Big\{ v \in \mathbb{N} : v = \sum_{i=1}^{k} a_i b_i^{u_i} \text{ where } a_i \in A, u_i \geq 0 \ (i = 1, \dots, k) \Big\}.$$

Note that if we choose S to be the set of prime divisors of $b_1 \cdots b_k$, then $H_{A,\underline{b}}$ is a subset of $H_{A,S,k}$. For $H \subseteq \mathbb{Z}$ and $m \in \mathbb{Z}$, $m \geq 2$, we put

$$H \pmod m = \{r : 0 \leq r < m,\ d \equiv r \pmod m \text{ for some } d \in H\}.$$

Our main result in this direction is the following.

THEOREM 3. *Let C be an arbitrary positive real number. There exists a positive integer $m = m(k, n, C) > C$ depending only on k, n and C such that*

$$\#(H_{A,\underline{b}} \pmod m) < c_0(k, n)(\log m)^{c_1 k \log\log\log m}.$$

Here, $c_0(k,n)$ is a constant depending only on k and n, and c_1 is an absolute constant.

Roughly speaking, Theorem 3 states that for any A and $\underline{b}$ one can find a "large" modulus m such that "many" residue classes modulo m remain outside the set $H_{A,\underline{b}}$.

As before, for $x \geq 1$, we write $H_{A,\underline{b}}(x) = H_{A,\underline{b}} \cap [1,x]$. As an immediate consequence of Theorem 3, we get the following statement.

COROLLARY 1. *There exists a strictly increasing sequence $X = (x_i)_{i=1}^{\infty}$ of positive integers such that the inequality*

$$\#H_{A,\underline{b}}(x_i) < c_0(k,n)(\log x_i)^{c_1 k \log\log\log x_i}$$

holds for all $i \in \mathbb{N}$, where the constants $c_0(k,n)$ and c_1 are specified in Theorem 3.

REMARK 3. Theorem 1 concerns a more general situation than Corollary 1, and further, the bound in Theorem 1 is better in terms of x. However, a remarkable property of the upper bound in Corollary 1 is that it depends only on k and n, and is independent of $\underline{b}$ and a^*. Further, in view of Theorem 3, one has some extra information about the distribution of the elements of the set $H_{A,\underline{b}}$ (see the remark after Theorem 3). Finally, we note that using the well-known fact that there is some positive constant c^* such that the inequality

$$\varphi(m) > c^* \frac{m}{\log\log m}$$

holds if m is large enough, one can easily prove that $\mathcal{D}(\mathbb{P} \cap H_{A,\underline{b}}) = 0$ for any A and $\underline{b}$.

3. Proofs of the theorems

Proof of Theorem 1. We proceed by induction on t. Assume that $t = 1$. If $k = 1$, then $H_{A,S,1}$ consists of the positive integers of the form ap^l for some $a \in A$ and $l \geq 0$. Clearly, a is non-negative. Obviously, the condition $ap^l \leq x$ leads to $l \leq \log x/\log 2$. Hence,

$$\#H_{A,S,1}(x) \ll n \log x \ll \log x.$$

Assume still that $t = 1$ but $k > 1$. Write $v \in H_{A,S,k}$ as

$$v = a_1 p^{l_1} + \cdots + a_k p^{l_k},$$

where $a_i \in A$ and $l_1 \geq \cdots \geq l_k \geq 0$. Let C_1 be a positive integer such that $p^{C_1} > 2ka^*$. If $l_1 - l_2 > C_1$, then

$$v \geq p^{l_1}\Big(1 - \Big|\sum_{i=2}^{k} a_i p^{l_i - l_1}\Big|\Big) > p^{l_1}\left(1 - \frac{ka^*}{p^{C_1}}\right) > \frac{p^{l_1}}{2},$$

therefore $l_1 \le \log(2x)/\log 2$. Since $l_i \le l_1$ for $i = 2, \dots, k$, we get

$$\#H_{A,S,k} \le n^k (C_2 \log x)^k \ll (\log x)^k,$$

where we can take $C_2 = 2$ provided that x is large. If, on the contrary, $l_1 - l_2 \le C_1$, then we enlarge the set A to the set A' consisting of all elements of the form $a_1 p^\alpha + a_2$, where $a_1, a_2 \in A$ and $\alpha \in \{0, 1, \dots, C_1\}$. Then note that

$$v = (a_1 p^{l_1 - l_2} + a_2) p^{l_2} + \cdots + a_k p^{l_k} \in H_{A',S,k-1},$$

and apply the induction hypothesis.

Assume now that $t > 1$. Again, if $k = 1$, then $v = a s_1$, where $s_1 = p_1^{l_1} \cdots p_t^{l_t}$. Obviously, $l_i \ll \log x$ for $i = 1, \dots, t$, therefore

$$H_{A,S,1} \ll (\log x)^t.$$

Assume now that $t > 1$ and that $k > 1$. Write

$$v = \sum_{i=1}^{k} a_i s_i,$$

and assume that $s_1 \ge \cdots \ge s_k$. One may also assume, because of the induction hypothesis, that $a_i \ne 0$ for all $i = 1, \dots, k$. Let $v \le x$. If $s_1 \le x^2$, then, by the argument used in the case $t > 1$, $k = 1$, we deduce that the number of possibilities for s_1 is $O((\log x)^t)$. Since $s_i \le s_1$ for all $i = 1, \dots, k$, it follows that the number of possibilities for s_i is $O((\log x)^t)$ for each of $i = 1, \dots, k$. Since the coefficients $a_1, \dots, a_k$ can be chosen in at most n^k ways, we see that the number of such v's is $O((\log x)^{kt})$.

It remains to deal with the case when $s_1 > x^2$. For this, we use the subspace theorem. Let $\overline{S} = S \cup \{\infty\}$ regarded as a set of valuations of $\mathbb{Q}$ containing the infinite one. For each $\mu \in \overline{S}$, let $L_{i,\mu}(\mathbf{x}) \in \mathbb{Z}$ be the linear form of k variables given by $L_{i,\mu}(\mathbf{x}) = x_i$ if μ is finite, or $\mu = \infty$ and $i > 1$, and $L_{1,\infty} = \sum_{i=1}^{k} a_i x_i$. We evaluate the double product

$$\prod_{i=1}^{k} \prod_{\mu \in \overline{S}} |L_{i,\mu}(\mathbf{s})|_\mu, \tag{1}$$

where $\mathbf{s} = (s_1, \dots, s_k)$. Since $s_2, \dots, s_k$ are all S-units, we get

$$\prod_{\mu \in \overline{S}} |L_{i,\mu}(\mathbf{s})|_\mu = 1 \quad \text{for all } i = 2, \dots, k. \tag{2}$$

When $i = 1$, again since s_1 is an S-unit, we get

$$\prod_{\mu \in \overline{S} \setminus \{\infty\}} |L_{i,\mu}(\mathbf{s})|_\mu = s_1^{-1}. \tag{3}$$

Finally,

$$|L_{1,\infty}(\mathbf{s})|_\infty = |a_1 s_1 + \cdots + a_k s_k| < x < s_1^{1/2}. \tag{4}$$

Estimates (2)–(4) show that for the double product given by (1), we have

$$\prod_{i=1}^{k} \prod_{\mu \in \overline{S}} |L_{i,\mu}(\mathbf{s})|_\mu < s_1^{-1/2}. \tag{5}$$

By the subspace theorem, there exist finitely many hyperplanes such that all solutions $\mathbf{s}$ of equation (5) are contained in these hyperplanes. More precisely, there exist non-zero integral vectors $\mathbf{d}^{(1)}, \ldots, \mathbf{d}^{(\lambda)}$ with

$$\lambda < C_3(t, k, n, a^*, P),$$

where $C_3(t, k, n, a^*, P)$ is an explicitly computable constant depending on the specified parameters (see e.g. [3]) such that each solution $\mathbf{s}$ of (5) satisfies

$$\sum_{i=1}^{k} d_i^{(j)} s_i = 0 \quad \text{for some } j = 1, \ldots, \lambda. \tag{6}$$

For each $j = 1, \ldots, \lambda$, let i_j denote the minimal index for which $d_{i_j}^{(j)} \neq 0$. Put

$$A' = A \cup \{a - a' d_i^{(j)} / d_{i_j}^{(j)} : a, a' \in A,\ i = 1, \ldots, k,\ j = 1, \ldots, \lambda,\ i \neq i_j\}.$$

Observe that $\#A' \leq C_4(k, t, n, a^*, P)$, where $C_4(k, t, n, a^*, P)$ is also a constant depending on k, t, n, a^* and P.

Let now $v \in H_{A,S,k}(x)$, and write

$$v = \sum_{i=1}^{k} a_i s_i$$

with $a_i \in A$. Then, as $\mathbf{s}$ satisfies (6) with the appropriate $\mathbf{d}^{(j)}$, we get

$$s_{i_j} = -\sum_{\substack{i=1 \\ i \neq i_j}}^{k} (d_i^{(j)} / d_{i_j}^{(j)}) s_i.$$

Then

$$v = \sum_{\substack{i=1 \\ i \neq i_j}}^{k} a_i' s_i,$$

where $a_i' = a_i - a_{i_j}(d_i^{(j)} / d_{i_j}^{(j)})$. In particular, $a_i' \in A'$. By the induction hypothesis for A', the number of such v's is $\ll (\log x)^{(k-1)t} \ll (\log x)^{kt}$, which finishes the argument. ∎

Proof of Theorem 2. As is well-known by the Prime Number Theorem, $\#(\mathbb{P} \cap [1, x]) = \pi(x)$ is asymptotically $x/\log x$ as $x \to \infty$. Thus, the desired statement follows immediately from Theorem 1. ∎

We continue with the proof of Theorem 3. For this purpose, we need some new notation and two lemmas.

Let $\lambda(m)$ be the Carmichael function of the positive integer m. Recall that for $m \in \mathbb{N}$, $\lambda(m)$ denotes the least positive integer for which

$$b^{\lambda(m)} \equiv 1 \pmod{m}$$

for all $b \in \mathbb{Z}$ with $\gcd(b, m) = 1$.

LEMMA 1. *Let $m = q_1^{\alpha_1} \cdots q_z^{\alpha_z}$ where $q_1, \ldots, q_z$ are distinct primes, and let $b \in \mathbb{Z}$. Then*

$$\#\{b^u \pmod{m} : u \geq 0\} \leq \lambda(m) + \max_{1 \leq i \leq z} \alpha_i.$$

Proof. Let $\beta = \max_{1 \leq i \leq z} \alpha_i$. Write $m = m_1 m_2$, where m_2 is maximal with the property that $\gcd(m_2, b) = 1$. Consequently, the powers b^u are congruent to 0 (mod m_1) if $u \geq \beta$. So, for $u \geq \beta$, the class $b^u \pmod{m}$ "varies" only modulo m_2. Since $(b, m_2) = 1$, the length of the period of the powers b^u, with $u \geq \beta$, is at most $\lambda(m_2)$ modulo m_2. As $(m_1, m_2) = 1$, the length of the period modulo m is at most $\lambda(m_2)$ as well by the Chinese remainder theorem. Since $\lambda(m_2) \leq \lambda(m)$, the period length modulo m is at most $\lambda(m)$, while the length of the preperiod is at most β. The statement now follows. ∎

The next lemma is a nice result of Erdős, Pomerance and Schmutz [1] concerning "small" values of the Carmichael function.

LEMMA 2. *There exists a positive constant c_1 and a strictly increasing sequence $N = (n_i)_{i=1}^{\infty}$ of natural numbers such that for all $i \in \mathbb{N}$ we have*

$$\lambda(n_i) < (\log n_i)^{c_1 \log\log\log n_i}.$$

We are now ready to prove Theorem 3.

Proof of Theorem 3. Let k, n and C be fixed, and choose an element $m = n_i$ from the sequence N in Lemma 2 such that $m > C$. Write $m = q_1^{\alpha_1} \cdots q_z^{\alpha_z}$ in the usual manner, where $q_1, \ldots, q_z$ are distinct primes and $\alpha_1, \ldots, \alpha_z$ are positive integers. Lemma 1 implies that for all $b \in \{b_1, \ldots, b_k\}$,

$$\#\{a \cdot b^u \pmod{m} : a \in A, u \geq 0\} \leq n(\lambda(m) + \max_{1 \leq i \leq z} \alpha_i). \tag{7}$$

On the other hand, a simple calculation shows that

$$n(\lambda(m) + \max_{1 \leq i \leq z} \alpha_i) \leq n\left(\frac{\log m}{\log 2} + (\log m)^{c_1 \log\log\log m}\right), \tag{8}$$

where c_1 is specified in Lemma 2. The statement is an immediate consequence of inequalities (7) and (8). ∎

Proof of Corollary 1. Arranging in increasing order the resulting values of the moduli m corresponding to different values of C provided by Theorem 3, we get an infinite sequence $X = (x_i)_{i=1}^{\infty}$ which has the desired property. ∎

Acknowledgements. The authors thank the referees for their useful remarks.

Research of L. Hajdu was supported in part by the Hungarian Academy of Sciences, by the grants T48791 and T67580 of the Hungarian National Foundation for Scientific Research.

Research of F. Luca was supported in part by grants SEP-CONACyT 79685 and PAPIIT 100508.

References

[1] P. Erdős, C. Pomerance and E. Schmutz, *Carmichael's lambda function*, Acta Arith. 58 (1991), 363–385.

[2] G. R. Everest, *Counting the values taken by sums of S-units*, J. Number Theory 35 (1990), 269–286.

[3] J.-H. Evertse and H. P. Schlickewei, *A quantitative version of the absolute subspace theorem*, J. Reine Angew. Math. 548 (2002), 21–127.

[4] L. Hajdu, *Arithmetic progressions in linear combinations of S-units*, Period. Math. Hungar. 54 (2007), 175–181.

Zs. Ádám
Institute of Mathematics
University of Debrecen
P.O. Box 12
H-4010 Debrecen, Hungary
E-mail: adzs@freemail.hu

L. Hajdu
Institute of Mathematics
University of Debrecen
and the Number Theory Research Group
of the Hungarian Academy of Sciences
P.O. Box 12
H-4010 Debrecen, Hungary
E-mail: hajdul@math.klte.hu

F. Luca
Mathematical Institute, UNAM
Ap. Postal 61–3 (Xangari), CP 58 089
Morelia, Michoacán, Mexico
E-mail: fluca@matmor.unam.mx

Received on 11.1.2008
and in revised form on 6.2.2009 (5615)