ARITHMETIC PROGRESSIONS IN LINEAR COMBINATIONS OF S-UNITS

L. Hajdu

Abstract. M. Pohst asked the following question: is it true that every prime can be written in the form $2^u \pm 3^v$ with some non-negative integers u, v? We put the problem into a general framework, and prove that the length of any arithmetic progression in t-term linear combinations of elements from a multiplicative group of rank r (e.g. of S-units) is bounded in terms of r, t, n, where n is the number of the coefficient t-tuples of the linear combinations. Combining this result with a recent theorem of Green and Tao on arithmetic progressions of primes, we give a negative answer to the problem of M. Pohst.

1. Introduction and results

Linear equations involving elements from a multiplicative group (such as e.g. S-unit equations) play a vital role and have wide and deep applications in several parts of diophantine number theory. For theoretical results and applications of such and related equations we refer to the papers [2–5,8–9], and the references given there. Combining the underlying theory of such equations and a classical result of van der Waerden [10] about arithmetic progressions, we show that the length of any arithmetic progression consisting of t-term linear combinations of elements from a finitely generated multiplicative group of rank r is bounded in terms of r, t, n, where n is the number of the coefficient t-tuples of the linear combinations.

To formulate our results we need some notation. We follow the paper [5], with slight modifications. Let K be an algebraically closed field of characteristic zero. Write K^* for the multiplicative group of the non-zero elements of K, and let Γ be a multiplicative subgroup of K^* having finite rank r. Let t be a positive integer, and let $\mathcal{A}$ be a finite subset of K^t having n elements. Put

$$H_t(\Gamma, \mathcal{A}) = \left\{ \sum_{i=1}^{t} a_i x_i : (a_1, \dots, a_t) \in \mathcal{A},\ (x_1, \dots, x_t) \in \Gamma^t \right\}.$$

The main result of this paper is the following.

2000 *Mathematics Subject Classification:* 11D57 (11B25). *Key words and phrases:* linear equations in variables from a multiplicative group, S-unit equations, arithmetic progressions, primes.
Research supported in part by the János Bolyai Research Fellowship of the Hungarian Academy of Sciences and by the OTKA grants T042985 and T048791.

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

Theorem 1. *There exists a constant $C(r,t,n)$ depending only on r, t and n such that the length of any non-constant arithmetic progression in $H_t(\Gamma,\mathcal{A})$ is at most $C(r,t,n)$.*

Note that in the upper bound $C(r,t,n)$ none of r,t,n could be omitted. This will be demonstrated by a simple example in Remark 1 after the proof of Theorem 1. Further, at the same place we show that the number of arithmetic progressions in $H_t(\Gamma,\mathcal{A})$ can be infinite, in case of any possible length.

Now as an application, we formulate a result concerning primes represented by sums of integers which are rational S-units. This is motivated by the next problem. M. Pohst asked the following question (oral communication): is it true that every prime can be written in the form $2^u \pm 3^v$, with some non-negative integers u,v? As we will see, by a recent, celebrated result of Green and Tao [7] on arithmetic progressions consisting of primes, this question can be reduced to S-unit equations in a natural way. By the help of Theorem 1 we will provide a negative answer to this question, under much more general circumstances. Note that the theorem would be true under even more general conditions, as well. However, we think that it is not natural to use here more general settings.

To formulate this result, let $S=\{p_1,\dots,p_r\}$ be a (nonempty) set of (positive) primes in $\mathbb{Z}$. As usual, let $\mathbb{Z}_S$ denote the set of those integers, which do not have any prime divisors outside S. In particular, we have $\pm 1 \in \mathbb{Z}_S$. Let t be a positive integer and let A be a finite non-empty subset of $\mathbb{Z}^t$. Put

$$H_t(\mathbb{Z}_S,A)=\left\{\sum_{i=1}^{t} a_i s_i : (a_1,\dots,a_t)\in A,\ (s_1,\dots,s_t)\in \mathbb{Z}_S^t\right\}.$$

Theorem 2. *For any S, t and A there are infinitely many primes outside the set $H_t(\mathbb{Z}_S,A)$.*

Taking $S=\{2,3\}$, $t=2$ and $A=\{(1,1)\}$, the above theorem yields a negative answer to the problem of M. Pohst. Note that the smallest prime not of the shape $2^u \pm 3^v$ is 53; this fact is demonstrated in Remark 2 after the proof of Theorem 2. We also mention that it is widely believed that there are infinitely many Mersenne-primes, i.e. primes of the shape 2^u-1 ($u\in\mathbb{N}$). As these primes (would) all belong to $H_2(S,A)$ with $S=\{2\}$, $t=2$ and $A=\{(1,1)\}$, we probably cannot claim that $H_t(S,A)$ contains only finitely many primes in general. Hence the theorem seems to be best possible in the qualitative sense.

2. Proofs of the theorems

To prove our theorems, we need several tools. The first one is a deep and general finiteness result for the number of solutions of linear equations involving elements of Γ, due to Evertse, Schlickewei and Schmidt [5].

Keeping the notation from the previous section, consider the equation

$$a_1x_1+\ldots+a_tx_t=1 \tag{1}$$

in $\underline{x}=(x_1,\dots,x_t)\in\Gamma^t$, where $\underline{a}=(a_1,\dots,a_t)\in(K^*)^t$. A solution $\underline{x}$ is called non-degenerate, if no subsum of the left hand side of (1) vanishes, that is $\sum\limits_{i\in I} a_ix_i \neq 0$ for any nonempty subset I of $\{1,\dots,t\}$. The next statement is a simple and immediate consequence of Theorem 1.1 from [5].

Theorem A. *There exists a constant $c_1(r,t)$ depending only on r and t (independent of $\underline{a}$) such that equation* (1) *has at most $c_1(r,t)$ non-degenerate solutions $\underline{x} \in \Gamma^t$.*

We will also need the following simple and well-known corollary of the above theorem.

Corollary 1. *There exists a constant $c_2(r,t)$ depending only on r and t with the following property. If $(x_1,\dots,x_t) \in \Gamma^t$ is a solution to* (1) *then $x_i = \alpha_{P(i)} x_i^*$ ($i = 1,\dots,t$) with some $\alpha_{P(i)}, x_i^* \in \Gamma$, where $(x_1^*,\dots,x_t^*)$ belongs to a set of cardinality at most $c_2(r,t)$. Further, here $P_1,\dots,P_s,P_{s+1}$ is a partition of $\{1,\dots,t\}$, $P(i)$ denotes the class P_l for which $i \in P_l$, and $\alpha_{P_{s+1}} = 1$.*

Proof. Partitioning the sum at the left hand side of (1) into vanishing subsums (the indices in the subsums compose the classes $P_1,\dots,P_s$, respectively) and a subsum yielding 1 (the indices in this subsum compose P_{s+1}) such that none of these subsums has a vanishing subsum, the statement follows from Theorem A by a simple inductive argument. □

The next well-known result from Ramsey theory is due to van der Waerden (cf. [10]). This theorem will be very helpful in taking care of the vanishing subsums in the occurring linear equations of the shape (1).

Theorem B. *For every positive integers k and h there exists a positive integer $W = W(k,h)$ such that for any coloring of the set $\{1,\dots,W\}$ using k colors, we get a non-constant monochromatic arithmetic progression, having at least h terms.*

Finally, in the proof of Theorem 2 we also make use of the following recent deep and celebrated theorem of Green and Tao [7] about arithmetic progressions of primes.

Theorem C. *There are arbitrarily long arithmetic progressions of primes.*

Now we are ready to prove our results.

Proof of Theorem 1. We proceed by induction on t. Let $t = 1$ and take an arbitrary non-empty subset $\mathcal{A}$ of K having n elements. Let $q_1,\dots,q_L$ be a non-constant arithmetic progression in $H_1(\Gamma,\mathcal{A})$; write $q_j = a^{(j)} x^{(j)}$ ($a^{(j)} \in \mathcal{A}$, $x^{(j)} \in \Gamma$, $j = 1,\dots,L$). Without loss of generality we may assume that $0 \notin \mathcal{A}$; otherwise we can give bounds for the lengths of the positive and negative parts of the progression independently, and then simply combine them. Let $d := q_2 - q_1 \neq 0$ denote the common difference of the progression. Subtracting the consecutive terms, we get the equalities

$$(a^{(j+1)}/d) x^{(j+1)} - (a^{(j)}/d) x^{(j)} = 1 \quad (j = 1,\dots,L-1).$$

If $L - 1 > n^2 c_1(r,2)$ then by $|\mathcal{A}| = n$ and the box principle we get that for some $j \in \{1,\dots,L-1\}$ the equation

$$(a^{(j+1)}/d) x_1 - (a^{(j)}/d) x_2 = 1$$

has more than $c_1(r,2)$ solutions in $(x_1,x_2) \in \Gamma^2$. However, by Theorem A this is a contradiction. Hence $L \leq C(r,1,n) := n^2 c_1(r,2) + 1$, and the theorem follows for $t = 1$.

Let now t be an arbitrary integer with $t \geq 2$, and assume that the statement is true for $t-1$. That is, the length of any arithmetic progression in $H_{t-1}(\Gamma, \mathcal{B})$ with any non-empty $\mathcal{B} \subseteq K^{t-1}$, $|\mathcal{B}| = m$ is at most $C(r, t-1, m)$ for some constant $C(r, t-1, m)$ depending only on $r, t-1, m$. Further, let $\mathcal{A}$ be a non-empty subset of K^t having n elements, and let $q_1, \dots, q_L$ be a non-constant arithmetic progression in $H_t(\Gamma, \mathcal{A})$. Assume first that $n = 1$. Let $\mathcal{A} = \{(a_1, \dots, a_t)\}$, and put

$$q_j = \sum_{i=1}^{t} a_i x_i^{(j)} \quad (j = 1, \dots, L)$$

where $(x_1^{(j)}, \dots, x_t^{(j)}) \in \Gamma^t$. We have

$$\sum_{i=1}^{t} (a_i/d) x_i^{(j+1)} - \sum_{i=1}^{t} (a_i/d) x_i^{(j)} = 1 \quad (j = 1, \dots, L-1)$$

where $d := q_2 - q_1 \neq 0$ is the common difference of the progression. Note that if $a_1 \dots a_t = 0$ then by the induction step we immediately have $L \leq C(r, t-1, 1)$ and the theorem follows in this case. Otherwise, Corollary 1 implies that for each $j \in \{1, \dots, L-1\}$, $x_i^{(j)}$ is of the form $x_i^{(j)} = \alpha_{P(i)} x_i^*$ with certain $(x_1^*, \dots, x_t^*)$ coming from a finite subset of Γ^t of cardinality bounded by some $c_2(r, t)$ and certain $\alpha_{P(i)} \in \Gamma$ $(i = 1, \dots, t)$. Here $P_1, \dots, P_s, P_{s+1}$ is some partition of the set $\{1, \dots, t\}$, and $P(i)$ denotes the class P_l $(1 \leq l \leq s+1)$ for which $i \in P_l$. Further, P_{s+1} is possibly empty, but otherwise $\alpha_{P_{s+1}} = 1$. Obviously, we have $1 \leq s+1 \leq t$, further $1 \leq s \leq t$ if P_{s+1} is empty. Now we paint the terms q_j $(j = 1, \dots, L-1)$ of the arithmetic progression. We code the colors in the following way. Those q_j will get the same color, where in the above representation the very same partition of the indices $\{1, \dots, t\}$ occurs, moreover, the "parameter t-tuples" $(x_1^*, \dots, x_t^*)$ also coincide. That is, q_{j_1} and q_{j_2} will get the same color if and only if we have

$$(x_1^{(j_1)}, \dots, x_t^{(j_1)}) = (\alpha_{P(1)} x_1^*, \dots, \alpha_{P(t)} x_t^*)$$

and

$$(x_1^{(j_2)}, \dots, x_t^{(j_2)}) = (\alpha'_{P(1)} x_1^*, \dots, \alpha'_{P(t)} x_t^*)$$

with the same partition $P_1, \dots, P_s, P_{s+1}$, the same $(x_1^*, \dots, x_t^*) \in \Gamma^t$, and some $\alpha_{P(1)}, \dots, \alpha_{P(t)}, \alpha'_{P(1)}, \dots, \alpha'_{P(t)} \in \Gamma$. Observe that by Corollary 1 and elementary combinatorics, the number of colors is bounded by some constant $c_3(r, t)$ depending only on r and t. Take $k = c_3(r, t)$ and $h = C(r, t-1, 1) + 1$. Suppose that $L - 1 \geq W(k, h)$. Then by Theorem B we find that there exists a monochromatic arithmetic progression in $H_t(\Gamma, \mathcal{A})$ corresponding to the above coloring, of length $C(r, t-1, 1) + 1$. If this subprogression corresponds to a case where P_{s+1} is non-empty, then observe that in each corresponding q_j the very same constant $\sum\limits_{P(i)=P_{s+1}} a_i x_i^*$ occurs. Cancelling this constant from each term of the subprogression, we get an arithmetic progression in $H_{t-1}(\Gamma, \mathcal{A}')$ (with the appropriate one-elemented $\mathcal{A}'$) of length $C(r, t-1, 1) + 1$, which is a contradiction. Suppose now that P_{s+1} is empty. Observe that in this case $s < t$ must be valid. Hence there exists a class, say P_1

with at least two members. However, then writing $b_l = \sum\limits_{P(i)=P_l} a_i x_i^*$ $(l = 1, \dots, s)$ the representation

$$q_j = \sum_{l=1}^{s} b_l \alpha_{P_l}$$

belongs to $H_{t-1}(\Gamma, \{\underline{b}\})$, with $\underline{b} = (b_1, \dots, b_s, 0 \dots, 0) \in K^{t-1}$. Hence we get an arithmetic progression in the latter set, of length $C(r, t-1, 1) + 1$, which is a contradiction again. As there are now more cases to distinguish, we get that $L \leq C(r, t, 1) := W(k, h)$ must be valid. Hence the theorem follows in this case.

Finally, consider the general case, i.e. with a non-empty $\mathcal{A} \subseteq K^t$, $|\mathcal{A}| = n$, and let $q_1, \dots, q_L$ be a non-constant arithmetic progression in $H_t(\Gamma, \mathcal{A})$. Paint q_j $(j = 1, \dots, L)$ with a color corresponding to that $\underline{a} \in \mathcal{A}$ which belongs to the representation of q_j. Let $k = n$ and $h = C(r, t, 1) + 1$. Applying Theorem B we get that if $L \geq W(k, h)$, then there exists a monochromatic subprogression of the original arithmetic progression of length at least $C(r, t, 1) + 1$. As in this subprogression the terms correspond to the same $\underline{a} \in \mathcal{A}$, this is a contradiction. Hence $L \leq C(r, t, n) := W(k, h) - 1$, and the theorem follows. □

Remark 1. As we mentioned in the introduction, in the upper bound $C(r, t, n)$ none of r, t, n could be omitted. To see this, for simplicity take $K = \mathbb{Q}$. First let t be arbitrary but fixed, take $\Gamma = \{-1, 1\}$ and let $\mathcal{A} = \{(1, \dots, 1)\}$. As the arithmetic progression $-t, -t+2, \dots, t-2, t$ belongs to $H_t(\Gamma, \mathcal{A})$, the dependence on t is necessary. Let now $t = 1$, and take an arbitrary positive integer k. Choosing either $\Gamma = \{1\}$ and $\mathcal{A} = \{1, \dots, k\}$ or $\Gamma = U_S$ with $S = \{p : p \text{ is prime and } p \mid k!\}$ (for the notation see the proof of Theorem 2 below) and $\mathcal{A} = \{1\}$, in both cases we get that the arithmetic progression $1, \dots, k$ belongs to $H_t(\Gamma, \mathcal{A})$. This shows that the dependence on both r and n is necessary, as well.

Further, in general it is not possible to give a bound for the *number* of progressions in $H_t(\Gamma, \mathcal{A})$. Indeed, take $K = \mathbb{Q}$, $S = \{2\}$ and let $\Gamma = U_S$. Setting $\mathcal{A} = \{0, 1\}$ we see that $0, 2^u, 2^{u+1}$ is an arithmetic progression in $H_1(\Gamma, \mathcal{A})$ for any $u \in \mathbb{N}$. To get a "non-trivial" example, observe that $1, 2^u + 1, 2^{u+1} + 1$ is an arithmetic progression consisting of pairwise relatively prime terms in $H_2(\Gamma, \mathcal{A})$, for any $u \in \mathbb{N}$. In general, take arbitrary K, Γ, t and $\mathcal{A}$, and suppose that $q_1, \dots, q_L$ is an arithmetic progression in $H_t(\Gamma, \mathcal{A})$. Then $q_1 + x, \dots, q_L + x$ is an arithmetic progression in $H_{t+1}(\Gamma, \mathcal{A}')$ with any $x \in \Gamma$, where $\mathcal{A}'$ is chosen accordingly. This shows that $H_t(\Gamma, \mathcal{A})$ can contain infinitely many arithmetic progressions in general.

Proof of Theorem 2. Let t and S be fixed, and let A be a non-empty subset of $\mathbb{Z}^t$ with $|A| = n$. As is well-known, taking $K = \mathbb{Q}$ and

$$U_S = \{p/q : p, q \in \mathbb{Z} \setminus \{0\},\ \gcd(p, q) = 1,\ pq \in \mathbb{Z}_S\},$$

U_S is a finitely generated multiplicative subgroup of $\mathbb{Q}^*$ (with $\mathbb{Z}_S \subseteq U_S$), of rank $r = |S|$. Further, Theorem C obviously implies that there are infinitely many pairwise disjoint arithmetic progressions of primes of length $C(r, t, n) + 1$ (where $C(r, t, n)$ is specified in Theorem 1). As by Theorem 1 each such progression contains a prime outside $H_t(U_S, A)$, the statement follows. □

Remark 2. The smallest prime yielding a negative answer to the problem of M. Pohst is 53. This can be seen as follows. On the one hand, it is easy to check that all the smaller primes can be represented in the desired form, with "small" u, v. (The "largest" decomposition is given by $2^7 - 3^4 = 47$.) On the other hand, if 53 is of the shape $2^u \pm 3^v$, then we have $2^\alpha y^2 = 3^\beta x^3 + 53$ with $\alpha \in \{0, 1\}$ and $\beta \in \{0, 1, 2\}$ where $\pm x$ and y are powers of 3 and 2, respectively. However, a simple computation with Magma (see [1]) gives that these elliptic equations have no solutions of the required shape, and our claim follows. Note that as these equations can be easily transformed into Mordell equations, their solutions are already known from [6].

3. Acknowledgement

The author is grateful to A. Pethő for letting him know about the problem of M. Pohst, and for his encouragement.

References

[1] J. Cannon et al., *The Magma computational algebra system*, http://magma.maths.usyd.edu.au.

[2] J.-H. Evertse, K. Győry, *On unit equations and decomposable form equations*, J. Reine Angew. Math. **358** (1985), 6–19.

[3] J.-H. Evertse, K. Győry, C. Stewart, R. Tijdeman, *S-unit equations and their applications*, New Advances in Transcendence Theory (A. Baker, ed.), Cambridge University Press, Cambridge, 1988, pp. 110–174.

[4] J.-H. Evertse, H. P. Schlickewei, *The absolute subspace theorem and linear equations with unknowns from a multiplicative group*, Number theory in progress, Vol. 1 (Zakopane-Kościelisko, 1997), de Gruyter, Berlin, 1999, pp. 121–142.

[5] J.-H. Evertse, H. P. Schlickewei, W. M. Schmidt, *Linear equations in variables which lie in a multiplicative group*, Annals Math. **155** (2002), 807–836.

[6] J. Gebel, A. Pető, H. G. Zimmer, *On Mordell's equation*, Compositio Math. **110** (1998), 335-367.

[7] B. Green, T. Tao, *The primes contain arbitrarily long arithmetic progressions*, arXiv:math.NT/0404188 **v5** (9 Feb 2006), 56 pp.

[8] K. Győry, *Some recent applications of S-unit equations*, Astérisque **209** (1992), 17–38.

[9] K. Győry, *Solving Diophantine equations by Baker's theory*, A panorama of number theory or the view from Baker's garden (Zürich, 1999), Cambridge Univ. Press, Cambridge, 2002, pp. 38–72.

[10] B. L. van der Waerden, *Beweis einer Baudetschen Vermutung*, Nieuw Archief voor Wiskunde **19** (1927), 212–216.

L. Hajdu
Number Theory Research Group
of the Hungarian Academy of Sciences, and
Institute of Mathematics
University of Debrecen
P.O. Box 12
4010 Debrecen
Hungary

E-mail address:
`hajdul@math.klte.hu`