ACTA ARITHMETICA
LXXVII.2 (1996)

Complete sequences of sets of integer powers

by

S. A. BURR (Bronx, N.Y.), P. ERDŐS (Budapest),
R. L. GRAHAM (Murray Hill, N.J.) and
W. WEN-CHING LI (University Park, Penn.)

1. Introduction. For a sequence $S = (s_1, s_2, \ldots)$ of positive integers, define

$$\Sigma(S) := \Big\{ \sum_{i=1}^{\infty} \varepsilon_i s_i : \varepsilon_i = 0 \text{ or } 1,\ \sum_{i=1}^{\infty} \varepsilon_i < \infty \Big\}.$$

Call S *complete* if $\Sigma(S)$ contains all sufficiently large integers.

It has been known for some time (see [B]) that if $\gcd(a,b) = 1$ then the (nondecreasing) sequence formed from the values $a^s b^t$ with $s_0 \le s$, $t_0 \le t \le f(s_0, t_0)$ is complete, where s_0 and t_0 are arbitrary, and $f(s_0, t_0)$ is sufficiently large.

In this note we consider the analogous question for sequences formed from pure powers of integers. Specifically, for a sequence A of integers greater than 1, denote by $\mathrm{Pow}(A; s)$ the (nondecreasing) sequence formed from all the powers a^k where $a \in A$ and $k \ge s \ge 1$. Although we are currently unable to prove it, we believe the following should hold:

CONJECTURE. *For any s, $\mathrm{Pow}(A; s)$ is complete if and only if*

(i) $\sum_{a \in A} 1/(a-1) \ge 1$,
(ii) $\gcd\{a \in A\} = 1$.

The necessity of (ii) is immediate. On the other hand, if (i) fails to hold then standard results in diophantine approximation show (as pointed out by Carl Pomerance) that in fact $\Sigma(\mathrm{Pow}(A; s))$ has upper density less than 1.

2. The main result. For a sequence A, denote by $A(x)$ the number of entries $a \in A$ with $a \le x$.

The research of the fourth author was supported in part by grant from NSA number MDA 904-95-H-1006.

THEOREM 1. *Suppose A is a sequence of integers greater than* 1 *satisfying*:

(i) $\limsup_{n\to\infty} A(n)/n > 0$;
(ii) $\gcd\{a \in A\} = 1$.

Then for any s, there is a finite subset $A' = A'_s$ such that $\mathrm{Pow}(A'; s)$ *is complete.*

Proof. To begin with, we first remove (by (ii)) a finite subsequence $A_0 \subset A$ so that $\gcd\{a \in A_0\} = 1$. We will use A_0 at the end of the proof. Next, we choose (by (i)) a finite increasing subsequence $B = (b_1, \ldots, b_N) \subset A \setminus A_0$ so that $\sum_{i=1}^N b_i^{-1} = \beta > 2$.

Write $\Sigma(\mathrm{Pow}(B; s)) = \{0 = p_0 < p_1 < p_2 < \ldots\}$.

CLAIM 1. *For all $k \geq 0$, $p_{k+1} - p_k \leq 2b_N^{s+1}$.*

Proof of Claim 1. Write $\mathrm{Pow}(B; s) = \{\beta_1 < \beta_2 < \ldots\}$. Observe that for any $k \geq 1$:

(a) The maximum gap size between consecutive terms of $\Sigma(\beta_1, \ldots, \beta_k)$ is at most β_k;

(b) If $\sum_{i=1}^k \beta_i \geq \beta_{k+1}$ then the maximum gap size in $\Sigma(\beta_1, \ldots, \beta_{k+1})$ is less than or equal to the maximum gap size in $\Sigma(\beta_1, \ldots, \beta_k)$.

Let l denote the *least* index such that $\beta_l > 2b_N^{s+1}$. Then by (a), $\Sigma(\beta_1, \ldots, \beta_{k-1})$ has maximum gap size at most $2b_N^{s+1}$ for $k \leq l$. For $k > l$, define $t(i)$, $1 \leq i \leq N$, so that $b_i^{t(i)} < \beta_k \leq b_i^{t(i)+1}$. Then

$$\sum_{i=1}^N \sum_{j=1}^{t(i)} b_i^j = \sum_{i=1}^N \frac{b_i^{t(i)+1} - b_i^{s+1}}{b_i - 1} \geq (\beta_k - b_N^{s+1}) \sum_{i=1}^N \frac{1}{b_i - 1}$$
$$\geq \beta(\beta_k - b_N^{s+1}) \geq \beta_k$$

since $\beta_k \geq 2b_N^{s+1}$ and $\beta > 2$. Thus, by repeated application of (b), $\Sigma(\beta_1, \ldots, \beta_k)$ has maximum gap size bounded by $2b_N^{s+1}$, and consequently, so does $\Sigma(\mathrm{Pow}(B; s))$. ■

Now, let

$$\delta := \frac{1}{2} \limsup_{n\to\infty} \frac{A(n)}{n}.$$

By Szemerédi's theorem [S], there is an integer $R = R(\delta, s)$ such that any subset of R consecutive integers with cardinality at least δR contains an arithmetic progression of length 2^s. By (i), there exist infinitely many m so that the interval $[m, m + R]$ contains at least δR elements of $A' := A \setminus (A_0 \cup B)$. Select an infinite sequence of such disjoint intervals with left-hand endpoints $m_1 < m_2 < \ldots$ Set $A_j := A' \cap [m_j, m_j + R]$. Thus, each A_j

satisfies $|A_j| \geq \delta R$ and consequently, A_j contains an arithmetic progression $a_j + kd_j$, $0 \leq k \leq 2^s - 1$.

CLAIM 2. *For each s it is possible to partition*

$$\{0, 1, \ldots, 2^s - 1\} = C(s) \cup D(s)$$

so that

$$\sum_{c \in C(s)} c^j = \sum_{d \in D(s)} d^j, \quad 0 \leq j \leq s - 1, \tag{1}$$

and

$$\Big| \sum_{c \in C(s)} c^s - \sum_{d \in D(s)} d^s \Big| = s! 2^{\binom{s}{2}}. \tag{2}$$

Proof of Claim 2. To begin, set $C(1) = \{0\}$, $D(1) = \{1\}$. Now, recursively define

$$C(k+1) = C(k) \cup \{2^k + D(k)\},$$
$$D(k+1) = D(k) \cup \{2^k + C(k)\}$$

for $k = 1, 2, \ldots$, so that $C(2) = \{0, 3\}$, $D(2) = \{1, 2\}$, etc. Thus, (1) and (2) hold for $s = 1$.

Now assume that $s \geq 1$ is fixed, and that (1) and (2) hold for s. Then

$$\Big| \sum_{c \in C(s+1)} c^j - \sum_{d \in D(s+1)} d^j \Big|$$
$$= \Big| \sum_{c \in C(s)} c^j + \sum_{d \in D(s)} (2^s + d)^j - \sum_{d \in D(s)} d^j - \sum_{c \in C(s)} (2^s + c)^j \Big|$$
$$= \sum_{i=0}^{j-1} \binom{j}{i} 2^{s(j-i)} \Big| \sum_{d \in D(s)} d^i - \sum_{c \in C(s)} c^i \Big|.$$

By (1) and (2), this reduces to

$$\binom{s+1}{1} 2^s \Big| \sum_{d \in D(s)} d^s - \sum_{c \in C(s)} c^s \Big| = (s+1) 2^s \cdot s! 2^{\binom{s}{2}} = (s+1)! 2^{\binom{s+1}{2}}.$$

Thus, the claim follows by induction. ∎

Since (1) is invariant under the affine transformation $k \mapsto a_j + kd_j$, by Claim 2 we can decompose the set $\{a_j + kd_j : 0 \leq k \leq 2^s - 1\}$ into two disjoint sets P_j and Q_j so that

$$\sum_{p \in P_j} p^i = \sum_{q \in Q_j} q^i, \quad 0 \leq i \leq s - 1,$$

and

$$\left|\sum_{p\in P_j} p^s - \sum_{q\in Q_j} q^s\right| = s!2^{\binom{s}{2}} d_j^s.$$

Of course, there are at most $R \cdot 2^{1-s}$ possible values for d_j, so that one of them, say d, occurs infinitely often. From now on we restrict ourselves to these j, so that we can assume without loss of generality that all $d_j = d$. Let us set $D := s!2^{\binom{s}{2}} d^s$.

We have just shown that each sequence $\mathrm{Pow}(A_j; s)$ contains two terms which differ by D. Since the A_j's are mutually disjoint, we conclude:

CLAIM 3. *For any $u \geq 1$, $\Sigma(\mathrm{Pow}(A_1; s)) + \ldots + \Sigma(\mathrm{Pow}(A_u; s))$ contains an arithmetic progression of length $u+1$ and step size D.*

Finally, we will need:

CLAIM 4. *$\Sigma(\mathrm{Pow}(A_0; s))$ contains a complete residue system modulo D.*

Proof of Claim 4. Let $q_1 < \ldots < q_r$ be the distinct primes dividing D. By hypothesis, for some $a(i) \in A_0$, we have $\gcd(a(i), q_i) = 1$, $1 \leq i \leq r$. Thus, there exist $t_i(1) < t_i(2) < t_i(3) < \ldots$ so that $a(i)^{t_i(k)} \bmod D$ does not depend on k, say

$$a(i)^{t_i(k)} \equiv c(i) \pmod D, \qquad k = 1, 2, \ldots,$$

where, of course, $\gcd(c(i), q_i) = 1$. Define $Q := q_1 \ldots q_r$. Then $\Sigma(\mathrm{Pow}(A_0; s))$ certainly contains integers $M(j)$ so that

$$M(j) \equiv \frac{Q}{q_1} c(1) + \ldots + \frac{Q}{q_r} c(r) := M \pmod D$$

for $1 \leq j \leq D$. Note that $\gcd(M, D) = 1$. Finally, since

$$\sum_{j=1}^{k} M(j) \equiv kM \pmod D, \qquad 1 \leq k \leq D,$$

it follows that $\Sigma(\mathrm{Pow}(A_0; s))$ contains a complete residue system modulo D as claimed. ∎

To conclude the proof of Theorem 1, we observe by Claims 3 and 4 that

$$\begin{aligned}\Sigma(\mathrm{Pow}(A_0; s)) + \Sigma(\mathrm{Pow}(A_1; s)) + \ldots + \Sigma(\mathrm{Pow}(A_u; s))\\ = \Sigma(\mathrm{Pow}((A_0, A_1, \ldots, A_u); s))\end{aligned}$$

must contain at least $2b_N^{s+1}$ *consecutive* integers, provided u is taken sufficiently large. However, by Claim 1, it follows at once that

$$\Sigma(\mathrm{Pow}((B, A_0, A_1, \ldots, A_u); s)) \subset \Sigma(\mathrm{Pow}(A; s))$$

contains all sufficiently large integers. This proves the theorem. ∎

3. Concluding remarks. We remark here that with very similar arguments, one can prove somewhat sharper forms of Theorem 1 when the initial set A has a special structure.

THEOREM 2. *For any $\varepsilon > 0$, there is an integer $n_0(\varepsilon)$ so that if $n > n_0(\varepsilon)$ and $N > (e+\varepsilon)n$ then* $\mathrm{Pow}(\{n, n+1, \ldots, N\}; 1)$ *is complete (and, in fact, contains all integers $\geq n$).*

Note that the bound on N is essentially best possible because of the necessity of condition (i) in the conjecture for any set A to have $\mathrm{Pow}(A; s)$ complete.

THEOREM 3. *There exists a function $f : \mathbb{Z}^+ \to \mathbb{Z}^+$ such that for any $s \geq 1$, if $A = \{n, n+1, \ldots, N\}$ with $N > f(s)n^s$ then* $\mathrm{Pow}(A; s)$ *is complete. Moreover, for any $\varepsilon > 0$,*

$$f(s) = o(2^{s^3/2+\varepsilon}) \quad \textit{as } s \to \infty.$$

The results we have described have all had an asymptotic flavor. That is, the sets A for which $\mathrm{Pow}(A; s)$ was proved complete were large. One might well ask for similar results for specific small sets A (indeed, this was our original motivation). The first nontrivial example is probably the set $\{3, 4, 7\}$ (since $\frac{1}{3-1} + \frac{1}{4-1} + \frac{1}{7-1} = 1$). Using fairly recent estimates in diophantine approximation, such as the inequality

$$|3^p - 4^q| > \exp\{\ln 3(p - 500 \ln 4 (8 + \ln p)^2)\}$$

of Mignotte and Waldschmidt [MW, Corollary 10.1], we can show that the largest integer *not* in $\Sigma(\mathrm{Pow}(\{3, 4, 7\}; 1))$ is 581. Similarly, the largest missing integer in $\Sigma(\mathrm{Pow}\{3, 5, 7, 13\}; 1)$ is 111, and the largest missing integer in $\Sigma(\mathrm{Pow}\{3, 6, 7, 13, 21\}; 1)$ is 16. Of course, when $\sum_{a \in A} 1/(a-1)$ is larger than 1, then one would expect it to be easier to show completeness of $\mathrm{Pow}(A; s)$, and our limited computational experience confirms this. For example, the largest missing integer in $\Sigma(\mathrm{Pow}(\{3, 4, 5\}; 1))$ is 78.

We are still fairly far from being able to prove the conjecture stated at the beginning. A related problem one could look at is the following. Suppose $0 < a_1 < \ldots < a_k$ satisfy

$$\sum_i \frac{1}{\log a_i} > \frac{1}{\log 2}.$$

Must $\Sigma(\mathrm{Pow}\{a_1, \ldots, a_k\}; s)$ have positive density? positive upper density? For example, what about set $\{3, 4\}$?

We close by remarking that our investigations grew out of the following conjecture of Erdős and Lewin [EL]. Suppose $\{a_1, \ldots, a_k\}$ is a set of $k \geq 2$ positive integers so that $\gcd(a_1, \ldots, a_k) = 1$. Prove that every sufficiently large integer is a sum of terms $a_1^{r_1} a_2^{r_2} \ldots a_k^{r_k}$ with all $r_i \geq 1$ so that no term in the sums divides any other. This was shown to hold for $\{2,3\}$ by Selfridge and Lewin (independently), and for $\{2,5,7\}$ (and several other sets) by Erdős and Lewin [EL].

References

[B] B. J. Birch, *Note on a problem of Erdős*, Proc. Cambridge Philos. Soc. 55 (1959), 370–373.

[EL] P. Erdős and M. Lewin, *d-complete sequences of integers*, Math. Comp. 65 (1996), 837–840.

[MW] M. Mignotte and M. Waldschmidt, *Linear forms in two logarithms and Schneider's method, II*, Acta Arith. 53 (1989), 251–287.

[S] E. Szemerédi, *On sets containing no k elements in arithmetic progression*, ibid. 27 (1975), 199–245.

Department of Mathematics
Lehman College (CUNY)
Bronx, New York 10468
U.S.A.
E-mail: stefan@bellcore.com

Mathematical Institute
Hungarian Academy of Sciences
Reáltanoda U. 13-15
H-1053, Budapest, Hungary

AT&T Bell Laboratories
Murray Hill, New Jersey 07974
U.S.A.
E-mail: rlg@research.att.com

Department of Mathematics
Pennsylvania State University
University Park, Pennsylvania 16802
U.S.A.
E-mail: wli@leibniz.math.psu.edu

Received on 3.2.1995
and in revised form on 9.4.1996 (2738)