ON STRINGS OF CONSECUTIVE INTEGERS
WITH NO LARGE PRIME FACTORS

ANTAL BALOG and TREVOR D. WOOLEY

Communicated by W. W. L. Chen

Abstract

We investigate conditions which ensure that systems of binomial polynomials with integer coefficients are simultaneously free of large prime factors. In particular, for each positive number ε, we show that there are infinitely many strings of consecutive integers of size about n, free of prime factors exceeding n^ε, with the length of the strings tending to infinity with speed $\log\log\log\log n$.

1. Introduction

In the absence of local conditions imposed by congruential considerations, it seems reasonable to expect that a set of integers having positive density should contain arbitrarily long strings of consecutive integers. A celebrated theorem of Szemerédi [10] shows that such a set must at least contain arbitrarily long arithmetic progressions. Moreover, Hildebrand [8] has established that a set satisfying a certain 'stability' hypothesis contains many strings of k *consecutive* integers provided only that the lower density of the set exceeds $(k - 2)/(k - 1)$. In this note we investigate the existence of long strings of consecutive integers possessing only small prime factors. Integers of the latter type, frequently referred to as 'smooth numbers', have found numerous applications throughout number theory, and have been the subject of extensive investigations by a host of authors (see, for example, Hildebrand and Tenenbaum [9] for an illuminating survey). Thus far, however, it appears that little is known concerning the

---

existence of strings of consecutive smooth numbers.

Our most general conclusions are somewhat complicated to state, and hence we defer their enunciation to Section 2 below. At this point we content ourselves by recording two direct consequences of our methods. Our first result shows that there are arbitrarily long strings of consecutive smooth numbers. Here and henceforth we write $\log_k x$ for the k-fold iterated logarithm, so that $\log_1 x = \log x$, and for each positive integer k we define $\log_{k+1} x = \log(\log_k x)$.

THEOREM 1. *Suppose that u is a fixed real number with $u > 1$. Let*

$$t(n) = [\log_4 n / \log(3u)].$$

Then for infinitely many positive integers, n, none of the prime factors of any of the string of $t(n)$ consecutive integers, $n+1, n+2, \ldots, n+t(n)$, exceed $n^{1/u}$.

It is convenient to refer to an integer whose prime factors are all at most y as being y-smooth. Expressed in this language, Theorem 1 shows that for each positive number ε, there are infinitely many strings of consecutive n^ε-smooth numbers of size about n, with the length of the strings tending to infinity with speed $\log_4 n$. On the other hand, in view of standard conjectures concerning the distribution of prime numbers, it seems inconceivable that permissible values of $t(n)$ could increase with n faster than $(\log n)^2$, and indeed density considerations lead one to believe that growth faster than $\log n$ is unlikely. Any upper bound on permissible $t(n)$ smaller than $\log n$, conditional or not, would therefore appear to be of interest. By way of comparison with Theorem 1, the conclusion of Hildebrand [8] alluded to above shows that for each fixed k, there are infinitely many strings of k consecutive n^{α_k}-smooth numbers of size about n whenever $\alpha_k > \exp(-1/(k-1))$. Although considerably weaker with respect to the smoothness parameter, this consequence of Hildebrand's work has the merit of exhibiting strings of integers, the set of whose elements has positive lower density. By contrast, the integers arising from our methods belong to an exceedingly thin set.

We note that the conclusions of Eggleton and Selfridge [3, § 2] provide a conclusion similar to Theorem 1, save that $t(n)$ is replaced by any fixed number not exceeding 5, and the smoothness parameter $n^{1/u}$ is replaced by $\exp(c \log n / \log_3 n)$ when $t = 2, 3$, and by $\exp(c \log n / \sqrt{\log_3 n})$ when $t = 4, 5$, for a suitable positive constant c. (Here we have corrected a minor oversight in the argument of [3] which arises because primes of half density are used to construct the integers P_0 and P_1.) Moreover, Balog, Erdös and Tenenbaum [1, Theorem 3] have used a construction of Heath-Brown [5] (see Heath-Brown [6] for an important correction to the latter) to show that there are many pairs of integers $n, n+1$ having no prime factors exceeding

$$\exp\left(\frac{8 \log n \log_3 n}{\log_2 n}\right).$$

Although we are unable to improve on the smoothness parameters occurring in these results, we are able to provide similar conclusions for arbitrary numbers of general linear forms.

THEOREM 2. *Let t be a fixed integer with $t \geq 2$, and let a_i and b_i $(1 \leq i \leq t)$ be non-zero integers. Then there are infinitely many integers n for which the linear polynomials n and $a_i n \pm b_i$ $(1 \leq i \leq t)$ are simultaneously y-smooth, where*

$$y = \exp\left(\frac{6 \log n}{(\log_3 n)^{1/t}}\right).$$

The set of integers n constructed in our proof of Theorem 2 is again very thin. For comparison, Balog and Ruzsa [2, Corollary 2] show that whenever $a > 0$ and $c \neq 0$ are fixed integers, and β is a positive number, then there is a set of integers n of positive density for which both n and $an + c$ are n^β-smooth. We note that the latter conclusion generalises a similar earlier conclusion of Hildebrand [7, Corollary 2] for consecutive integers.

The argument we employ to establish the above conclusions is elementary, and can be viewed as a natural extension of that used in Eggleton and Selfridge [3, §2]. The key observation is that the binomial polynomial $x^d - 1$ factors as a product of cyclotomic polynomials of degree at most $\phi(d)$. When a is a positive integer and d is a product of small prime numbers, therefore, it follows that the largest prime factor of $a^d - 1$ is at most $(a+1)^{Ad/\log_2 d}$, for a suitable fixed $A > 0$. On choosing an increasing sequence of such numbers d, we find that the integers $a^d - 1$ become increasingly smooth. In order to apply this observation to obtain the above conclusions, one has only to develop a strategy for choosing integers for which the required linear or binomial polynomials simultaneously assume the desired smooth shape. This we achieve through a simple application of the Chinese Remainder Theorem. The details will be found in the proof of Lemma 2.2 below. We deduce Theorems 1 and 2 from Lemma 2.2 in Section 3 in routine manner.

We adopt the convention throughout that any variable denoted by the letter p is implicitly assumed to be a prime number.

2. The polynomial construction

Before establishing the lemma fundamental to our proof of Theorems 1 and 2, we require a preliminary result needed to partition the prime numbers equitably into a number of classes.

LEMMA 2.1. *Let y be a large real number, and let t and k be positive integers. Then one can partition the prime numbers not exceeding y, and coprime with k, into sets,*

$\mathscr{P}_1, \ldots, \mathscr{P}_t$, *with the property that for each* i,

(2.1)
$$\prod_{p \in \mathscr{P}_i} (1 - 1/p) < 2 \left(\frac{k}{\phi(k) \log y} \right)^{1/t},$$

and

(2.2)
$$\prod_{p \in \mathscr{P}_i} p < y^2 \exp \left(\frac{5y}{4t} \right).$$

PROOF. Let y be a large real number. Then the trivial observation that

$$\prod_{p \mid k} (1 - 1/p) = \phi(k)/k,$$

combined with Mertens' Theorem (see, for example, [4, Theorem 429]), reveals that

(2.3)
$$\prod_{\substack{p \leq y \\ (p,k)=1}} (1 - 1/p) < \frac{k}{\phi(k) \log y}.$$

Also, as a consequence of the Prime Number Theorem (see, for example, [4, Theorem 434]), one has

(2.4)
$$\prod_{\substack{p \leq y \\ (p,k)=1}} p < e^{5y/4}.$$

Let the primes p with $2 \leq p \leq y$ and $(p, k) = 1$ be $p_1 < p_2 < \cdots < p_n$. When $1 \leq i \leq t$ we take $\mathscr{P}_i$ to be the set of prime numbers p_j with $1 \leq j \leq n$ for which $j \equiv i \pmod{t}$. Thus for each i with $1 \leq i \leq t$, and each j with $0 \leq j \leq (n - i)/t$, one has

$$p_{jt+1} \leq p_{jt+i} \leq p_{(j+1)t+1},$$

where we adopt the convention that $p_r = y$ when $r > n$. It follows that

$$\prod_{0 \leq j \leq (n-i)/t} (1 - 1/p_{jt+1}) \leq \prod_{0 \leq j \leq (n-i)/t} (1 - 1/p_{jt+i}) \leq \prod_{1 \leq j \leq (n-i)/t+1} (1 - 1/p_{jt+1}),$$

whence

(2.5)
$$\prod_{p \in \mathscr{P}_1} (1 - 1/p) \leq \prod_{p \in \mathscr{P}_i} (1 - 1/p) \leq 2 \prod_{p \in \mathscr{P}_1} (1 - 1/p).$$

We therefore deduce from (2.3) that

$$\prod_{p \in \mathscr{P}_1} (1 - 1/p) \leq \left(\prod_{i=1}^{t} \prod_{p \in \mathscr{P}_i} (1 - 1/p) \right)^{1/t} < \left(\frac{k}{\phi(k) \log y} \right)^{1/t},$$

whence (2.1) follows from the second inequality of (2.5). Meanwhile, by a similar argument,

$$(2.6) \qquad y^{-1} \prod_{p \in \mathscr{P}_1} p \leq \prod_{p \in \mathscr{P}_i} p \leq y \prod_{p \in \mathscr{P}_1} p.$$

Thus we deduce from (2.4) that

$$y^{-1} \prod_{p \in \mathscr{P}_1} p \leq \left(\prod_{i=1}^{t} \prod_{p \in \mathscr{P}_i} p \right)^{1/t} < \exp \left(\frac{5y}{4t} \right),$$

and (2.2) follows from the second inequality of (2.6). This completes the proof of the lemma.

We are now in a position to establish a rather general conclusion concerning simultaneous smooth values of binomial polynomials. We note that although the conclusions stated in Lemma 2.2 below are restricted to two or more polynomials, such is a matter more of convenience than necessity. One may easily convince oneself that a slight modification of the method presented below will establish the desired conclusion also for a single binomial polynomial.

LEMMA 2.2. *Let* (t_n) *be an increasing (possibly constant) sequence of integers satisfying the property that when* n *is sufficiently large, one has* $t_n \geq 2$. *Also, when* $i \in \mathbb{N}$, *let* k_i, a_i, b_i *be positive integers. When* $n \in \mathbb{N}$, *write*

$$(2.7) \quad \Delta_n = 2 \prod_{i=1}^{t_n} a_i b_i, \qquad K_n = \prod_{i=1}^{t_n} k_i, \qquad \kappa_n = \max_{1 \leq i \leq t_n} k_i, \qquad \alpha_n = \max_{1 \leq i \leq t_n} \frac{a_i}{b_i},$$

and

$$(2.8) \qquad y_n = \tfrac{4}{5} \log(\log n / \log \Delta_n).$$

Then there is an absolute constant C *such that for each natural number* n *for which* $y_n > C \log(2K_n)$, *there is an integer* x *with*

$$(2.9) \qquad \exp(K_n^{-1} (\log n)^{1/11}) \leq x \leq n,$$

which satisfies the property that the largest prime factor of the integer

$$(2.10) \qquad x \prod_{i=1}^{t_n} (a_i x^{k_i} - b_i)$$

is at most

$$\max \left\{ 2, \max_{1 \le i \le t_n} \{a_i, b_i\}, e^{1+\Psi} \right\},$$

where

$$\Psi = 2 \log(\alpha_n x^{\kappa_n}) \left(\frac{\phi(K_n)}{K_n} \log y_n \right)^{-1/t_n}.$$

PROOF. Let t_i, k_i, a_i, b_i ($i \in \mathbb{N}$) satisfy the hypotheses of the statement of the lemma, let n be a large natural number, and define $\Delta_n, K_n, \kappa_n, \alpha_n$ and y_n as in (2.7) and (2.8). For the sake of concision we henceforth omit subscripts n, and write, further, k for K_n. Provided that y is larger than a certain absolute constant y_0, as we now suppose, it follows from Lemma 2.1 that we can partition the prime numbers not exceeding y, and coprime with k, into sets, $\mathscr{P}_1, \ldots, \mathscr{P}_t$, with the property that for each i the inequalities (2.1) and (2.2) are satisfied. When $1 \le i \le t$ we define

(2.11)
$$\gamma_i = \prod_{p \in \mathscr{P}_i} p \quad \text{and} \quad \Gamma_i = \prod_{\substack{1 \le j \le t \\ j \ne i}} \gamma_j,$$

and write $\Gamma = \gamma_1 \ldots \gamma_t$. Thus it follows from the Prime Number Theorem that

(2.12)
$$k^{-1} e^{3y/4} < \Gamma < e^{5y/4}.$$

The definitions of the γ_j ensure that $(k_j \Gamma_j, \gamma_j) = 1$ ($1 \le j \le t$), and so for each j one may find integers $\tilde{\lambda}_j, \tilde{\mu}_j$ satisfying

(2.13)
$$1 \le \tilde{\lambda}_j, \tilde{\mu}_j < \gamma_j,$$

with the property that

$$k_j \Gamma_j \tilde{\lambda}_j \equiv -1 \pmod{\gamma_j} \quad \text{and} \quad k_j \Gamma_j \tilde{\mu}_j \equiv 1 \pmod{\gamma_j}.$$

Let

(2.14)
$$\lambda_j = \Gamma_j \tilde{\lambda}_j \quad \text{and} \quad \mu_j = \Gamma_j \tilde{\mu}_j \quad (1 \le j \le t),$$

and when $1 \le i, j \le t$ and $i \ne j$, define the integers Λ_{ij} and M_{ij} by

(2.15)
$$\Lambda_{ij} = k_j \lambda_i / \gamma_j \quad \text{and} \quad M_{ij} = k_j \mu_i / \gamma_j,$$

and when $i = j$, by

(2.16)
$$\Lambda_{jj} = (k_j \lambda_j + 1)/\gamma_j \quad \text{and} \quad M_{jj} = (k_j \mu_j - 1)/\gamma_j.$$

We define the integer x satisfying the assertions made in the statement of the lemma by putting

$$(2.17) \qquad x = 2^{\Gamma} \prod_{j=1}^{t} a_j^{\lambda_j} b_j^{\mu_j}.$$

Then on recalling (2.12)–(2.14) and (2.8), we deduce that

$$x \leq 2^{\Gamma} \prod_{j=1}^{t} (a_j b_j)^{\Gamma} < \Delta^{e^{5y/4}} \leq n,$$

and so we obtain the upper bound on x claimed in (2.9). Meanwhile (2.14) and (2.17) together imply that

$$x \geq 2^{\Gamma} \prod_{j=1}^{t} (a_j b_j)^{\Gamma/\tilde{\gamma}},$$

where $\tilde{\gamma} = \max_{1 \leq j \leq t} \gamma_j$, and hence from (2.2), (2.11) and (2.12) we deduce that

$$\log x \geq (ky^2)^{-1} e^{3y/4 - 5y/4t} \log \Delta.$$

On recalling that y is large, it therefore follows from (2.8) that

$$\log x \geq (ky^2)^{-1} (\log n / \log \Delta)^{1/10} \log \Delta \geq k^{-1} (\log n)^{1/11},$$

and the lower bound on x claimed in (2.9) follows immediately.

In order to complete the proof of the lemma we have now only to show that the desired polynomials are simultaneously smooth when evaluated at x. We start by observing that when $1 \leq j \leq t$, one has

$$(2.18) \qquad a_j x^{k_j} - b_j = b_j(z_j^{\gamma_j} - 1),$$

where, on recalling (2.15)–(2.17),

$$(2.19) \qquad z_j = 2^{k_j \Gamma_j} \prod_{i=1}^{t} a_i^{\Lambda_{ij}} b_i^{M_{ij}}.$$

But when d is a positive integer, the polynomial $z^d - 1$ factors as a product of cyclotomic polynomials, the degrees of which are each at most $\phi(d)$. Moreover, plainly, the roots of the latter cyclotomic polynomials all lie on the unit circle. We therefore deduce from (2.15)–(2.17) and (2.19) that the largest prime divisor of $z_j^{\gamma_j} - 1$ is at most

$$(z_j + 1)^{\phi(\gamma_j)} \leq z_j^{\phi(\gamma_j)} e^{\phi(\gamma_j)/z_j} \leq e^{\gamma_j/z_j} \left((a_j/b_j) x^{k_j} \right)^{\phi(\gamma_j)/\gamma_j}.$$

Then in view of (2.17) and (2.18), the expression (2.10) is π-smooth, where

$$(2.20) \qquad \pi = \max \left\{ 2, \max_{1 \le i \le t}\{a_i, b_i\}, e^{\psi} \right\},$$

and

$$(2.21) \qquad \psi = \max_{1 \le j \le t} \left(\frac{\phi(\gamma_j)}{\gamma_j} \log(\alpha x^\kappa) + \frac{\gamma_j}{z_j} \right).$$

Next, on recalling (2.11) and (2.1) we deduce that when $1 \le j \le t$,

$$(2.22) \qquad \frac{\phi(\gamma_j)}{\gamma_j} = \prod_{p \in \mathscr{P}_j} (1 - 1/p) < 2 \left(\frac{k}{\phi(k) \log y} \right)^{1/t}.$$

Moreover, from (2.2), (2.11), (2.12) and (2.19) one has for $1 \le j \le t$,

$$z_j \ge 2^{\Gamma_j} \ge \exp\left((ky^2)^{-1} e^{3y/4 - 5y/4t} \log 2 \right) > \exp\left((2ky^2)^{-1} e^{y/8} \right),$$

and hence from (2.2) and (2.11),

$$\frac{\gamma_j}{z_j} < y^2 \exp\left(\frac{5y}{4t} - \frac{e^{y/8}}{2ky^2} \right).$$

When $y/\log(2k)$ is sufficiently large we may therefore conclude that $\gamma_j/z_j < 1$, and thus the proof of the lemma is completed by collecting together (2.20)–(2.22).

3. Simultaneous smooth values of linear polynomials

It is now a simple matter, by inserting suitable choices for the parameters, to deduce Theorems 1 and 2 from Lemma 2.2.

PROOF OF THEOREM 1. Let u be a fixed real number with $u > 1$, and let $t(n)$ be defined as in the statement of Theorem 1. We note, in particular, that when n is sufficiently large in terms of u, one has $t(n) \ge 2$. We apply Lemma 2.2 with $t_n = t(n)$ ($n \in \mathbb{N}$), and with

$$k_i = 1, \qquad a_i = 1, \qquad b_i = i \qquad (i \in \mathbb{N}).$$

In the notation defined by (2.7), we then have

$$\Delta_n = 2(t_n!), \qquad K_n = 1, \qquad \kappa_n = 1, \qquad \alpha_n = 1 \qquad (n \in \mathbb{N}).$$

In particular, when n is sufficiently large in terms of u one has

$$\Delta_n < 2t_n^{t_n} < \log_2 n,$$

whence, in the notation defined by (2.8),

$$y_n > \tfrac{4}{5} \log(\log n / \log_3 n).$$

It follows that y_n satisfies the hypothesis necessary to apply Lemma 2.2 provided only that n is sufficiently large in terms of u. Under such conditions, therefore, we may conclude from the latter lemma that there is an integer x with

$$\exp((\log n)^{1/11}) \le x \le n,$$

satisfying the condition that the largest prime factor of $(x - 1)(x - 2) \cdots (x - t_n)$ is at most $\max\{2, t_n, ex^\beta\}$, where

$$\beta = 2 \exp\left(-\frac{1}{t_n} \log_2 y_n \right) < \frac{1}{u}.$$

Consequently there is a strictly increasing sequence (n_k) of positive integers satisfying the property that for each k, the integers $n_k + 1, n_k + 2, \ldots, n_k + t(n_k)$ are all $n_k^{1/u}$-smooth. This completes the proof of the theorem.

As it stands, Theorem 1 is stated in such a way that the parameter u is assumed to be a fixed real number with $u > 1$. At the cost of complicating the statement of our result, this assumption in the statement of Theorem 1 may be somewhat weakened. For example, one may establish through an almost identical argument that with $v(n) = [\log_4 n / \log_5 n]$, there are infinitely many positive integers, n, for which none of the prime factors of the string of $v(n)$ consecutive integers $n + 1, n + 2, \ldots, n + v(n)$, exceed

$$\exp\left(3 \log n / \log_4 n\right).$$

PROOF OF THEOREM 2. Let t be a fixed integer with $t \ge 2$, and let A_i and B_i $(1 \le i \le t)$ be non-zero integers. We apply Lemma 2.2 with $t_n = t$ $(n \in \mathbb{N})$, and with

$$k_i = 2, \qquad a_i = A_i^2, \qquad b_i = B_i^2 \qquad (1 \le i \le t).$$

In the notation defined by (2.7), we then have

$$\Delta_n = 2 \prod_{i=1}^{t} (A_i B_i)^2, \qquad K_n = 2^t, \qquad \kappa_n = 2, \qquad \alpha_n = \max_{1 \le i \le t} \frac{a_i}{b_i} \qquad (n \in \mathbb{N}).$$

It follows, in the notation defined by (2.8), that y_n satisfies the hypothesis necessary to apply Lemma 2.2 provided only that n is sufficiently large in terms of t, A and B. Under such conditions, therefore, we may conclude that there is an integer x with

$$\exp\left(2^{-t}(\log n)^{1/11}\right) \le x \le n,$$

satisfying the condition that the largest prime factor of $x \prod_{i=1}^{t}(A_i^2 x^2 - B_i^2)$ is at most

$$\max\left\{2, \max_{1 \le i \le t}\{a_i, b_i\}, e(\alpha_n x^2)^{\beta}\right\},$$

where

$$\beta = 2\exp\left(-\frac{1}{t}\log(\tfrac{1}{2}\log y_n)\right).$$

On recalling (2.8), we deduce that there is a strictly increasing sequence (n_k) of positive integers satisfying the property that for each k, the integers n_k and $A_i n_k \pm B_i$ $(1 \le i \le t)$ are simultaneously z-smooth, where

$$z = \exp\left(6\log n_k/(\log_3 n_k)^{1/t}\right).$$

This completes the proof of the theorem.

References

[1] A. Balog, P. Erdös and G. Tenenbaum, 'On arithmetic functions involving consecutive divisors', in: *Analytic number theory (Allerton Park, IL, 1989)* Progr. Math. 85, (Birkhäuser, Boston, 1990) pp. 77–90.

[2] A. Balog and I. Z. Ruzsa, 'On an additive property of stable sets', in: *Sieve methods, exponential sums and their applications in number theory, Cardiff, 1995*, London Mathematical Society Lecture Notes No. 237 (Cambridge University Press, Cambridge, 1997), pp. 55–63.

[3] R. B. Eggleton and J. L. Selfridge, 'Consecutive integers with no large prime factors', *J. Austral. Math. Soc. Ser. A* **22** (1976), 1–11.

[4] G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers*, 5th edition (Clarendon Press, Oxford, 1989).

[5] D. R. Heath-Brown, 'Consecutive almost-primes', *J. Indian Math. Soc.* **52** (1987), 39–49.

[6] ———, 'Correction and Footnote to "Consecutive almost-primes" ', *(J. Indian Math. Soc.* **52** (1987), 39–49), *J. Indian Math. Soc.*, to appear.

[7] A. Hildebrand, 'On a conjecture of Balog', *Proc. Amer. Math. Soc.* **95** (1985), 517–523.

[8] ———, 'On integer sets containing strings of consecutive integers', *Mathematika* **36** (1989), 60–70.

[9] A. Hildebrand and G. Tenenbaum, 'Integers without large prime factors', *J. Théor. Nombres Bordeaux* **5** (1993), 411–484.

[10] E. Szemerédi, 'On sets of integers containing no k terms in arithmetic progression', *Acta Arith.* **27** (1975), 199–245.

Mathematical Institute
Budapest 1364
Hungary
e-mail: balog@math-inst.hu

Department of Mathematics
University of Michigan
East Hall
525 East University Avenue
Ann Arbor, Michigan 48109-1109
USA
e-mail: wooley@math.lsa.umich.edu