AN ANALOGUE OF GRIMM'S PROBLEM OF FINDING DISTINCT PRIME FACTORS OF CONSECUTIVE INTEGERS

Paul Erdös and Carl Pomerance*

1. *Introduction.*

In [5] Grimm made the conjecture that if p, p' are consecutive primes, then for each integer m, $p < m < p'$, we can find a prime factor q_m of m such that the q_m's are all different. More generally, if n is a natural number, let $g(n)$ denote the largest number so that for each $m \in \{n+1, n+2, \ldots, n+g(n)\}$ there corresponds a prime factor q_m such that the q_m's are all different. Thus Grimm's conjecture is equivalent to the assertion $p+g(p) \geq p'$ when p, p' are consecutive primes.

It is known that

$$(1) \qquad (\log n/\log\log n)^3 \ll g(n) \ll (n/\log n)^{\frac{1}{2}}.$$

The lower bound is due to Ramachandra, Shorey, and Tijdeman [9]; the upper bound is due to Erdös and Selfridge [3]. From the lower bound, Grimm's conjecture for large primes follows from Cramér's well known conjecture:

$$\lim\sup(p'-p)/(\log p)^2 = 1.$$

From the upper bound it follows that if Grimm's conjecture is true, it must lie very deep. Indeed, Grimm's conjecture and (1) imply

$$(2) \qquad p' - p \ll (p/\log p)^{\frac{1}{2}}.$$

While (2) is undoubtedly true, it is generally recognized as probably hopeless at this time. Even if the Riemann hypothesis is assumed, the best known upper bound result on gaps between consecutive primes is not quite as strong as (2).

As noted in [3], using a result of Ramachandra [8] a better upper bound can be proved for $g(n)$. Indeed from the proof in [8] it follows that there is an $\alpha > 0$ such that for all large n a positive

* Research partially supported by a grant from the National Science Foundation.

proportion of the integers in $(n, n+n^{\frac{1}{2}-\alpha}]$ are divisible by a prime which exceeds $n^{15/26}$. Using this result with the method in [3] gives $g(n) < n^{\frac{1}{2}-c}$ for some fixed $c > 0$ and all large n. It is possible that the methods of Graham [4] will give a further reduction in the exponent, but we have not pursued this issue.

In [7] one of us made the conjecture that there are positive constants c_1, c_2 such that

(3) $$\exp\{c_1(\log n \log\log n)^{\frac{1}{2}}\} \le g(n) \le \exp\{c_2(\log n \log\log n)^{\frac{1}{2}}\}$$

for all large n. It is known that each of the inequalities in (3) separately holds for infinitely many n. (See [3], [7], and [10].)

This paper is addressed to the following question: how does Grimm's problem change if the factors q_m are no longer forced to be prime? Specifically, let $f(n)$ denote the largest integer such that for each composite $m \in \{n+1, n+2, \ldots, n+f(n)\}$ there is a divisor d_m of m with $1 < d_m < m$ and such that the d_m's are all different. We obviously have $f(n) \ge g(n)$ for all n. We prove below that for each $\varepsilon > 0$ we have

(4) $$n^{\frac{1}{2}} \ll f(n) \ll n^{7/12+\varepsilon}.$$

We strengthen the lower bound by showing that

(5) $$\liminf f(n)/\sqrt{n} \ge 4$$

and that there is a certain set A of integers of asymptotic density 1 such that

(6) $$f(n) > 4\sqrt{2n} \text{ for } n \in A, \quad n \text{ large}.$$

We show there are infinitely many twin primes if and only if equality holds in (5). Also if a certain very strong generalization of the twin prime conjecture is true then

(7) $$\limsup f(n)/\sqrt{n} = 4\sqrt{2}.$$

Thus combining our conjectures with our theorems we have $4\sqrt{2}$ as both the maximal order and normal order of $f(n)/\sqrt{n}$, while 4 is the minimal order of $f(n)/\sqrt{n}$.

In Section 5 we consider the function $f(n;c)$ for n a natural number and $c > 1$. This denotes the cardinality of the largest subset of $[n, cn]$ for which we can assign mutually distinct proper divisors. We prove that there is a positive constant $\delta(c)$ such that

(8) $$f(n;c) \sim \delta(c)n \quad \text{as} \quad n \to \infty .$$

The function $\delta(c)$ is continuous and strictly increasing. The fraction $\delta(c)/(c-1)$ is the asymptotic limit of the proportion of integers in $[n,cn]$ that fall in the maximal subset counted by $f(n;c)$. We have

(9) $$\lim_{c\to 1^+} \frac{\delta(c)}{c-1} = 1, \ \lim_{c\to\infty} \frac{\delta(c)}{c-1} = \frac{1}{2}, \ \frac{1}{2} < \frac{\delta(c)}{c-1} < 1 \text{ for all } c > 1.$$

It is probable that $\delta(c)/(c-1)$ is monotonic, but we have not been able to prove this.

We take this opportunity to thank the referee, John L. Selfridge, whose request for more details concerning (6) and (7) led to the discovery of an error in the original version. We also wish to acknowledge a helpful conversation with E. R. Canfield concerning Theorem 3.1.

2. *The proof of (4).*

The first inequality in (4) is easy. Indeed if we let d_m be the largest proper divisor of the composite number m, then $\sqrt{m} \le d_m < m$. If $d_m = d_k$ where $m < k$, then

$$k - m \ge (k,m) \ge d_k \ge \sqrt{k}.$$

Thus it is not the case that both m and k are in the interval $[n+1,n+\sqrt{n}]$ for any n. We conclude that if m,k are composite and in the interval $[n+1,n+\sqrt{n}]$, then $d_m \ne d_k$. Hence $f(n) \ge [\sqrt{n}]$.

Our proof of the second inequality in (4) relies on some work of Warlimont [11] (also see Cook [1]) concerning the distribution of abnormally large gaps between consecutive primes. First note that if $p_1 < p_2 < q_1 < q_2 < q_3$ are primes with $p_1q_1 > n$, then $n + f(n) < p_2q_3$. Indeed the six integers p_iq_j have collectively only five proper factors larger than 1. Our strategy is thus to find such primes with p_2q_3 as small as possible.

If x is a real number, let $p_i(x)$ denote the i-th prime greater than x. Let $\varepsilon > 0$ be arbitrarily small, but fixed. Let

$$S = \{x: \frac{1}{2}\sqrt{n} \le x \le \frac{3}{4}\sqrt{n}\ ,\ p_2(x)-x \ge \frac{1}{3} n^{1/12+\varepsilon}\}$$

$$T = \{x: \frac{1}{2}\sqrt{n} \le x \le \frac{3}{4}\sqrt{n}\ ,\ p_3(n/x)-n/x \ge \frac{1}{3} n^{1/12+\varepsilon}\}\ .$$

Let p_i denote the i-th prime and let $d_i = p_{i+1}-p_i$. From the estimates of Huxley [6] applied to Warlimont [10], we have a $\delta > 0$ such that for all large x,

$$\sum_{\substack{i\le x\\ d_i>p_i^{1/6+\varepsilon/2}}} d_i < x^{1-\delta}\ . \tag{10}$$

We apply (10) with $x = \sqrt{n}$. If μ denotes Lebesgue measure and if n is large, then (10) implies

$$\mu(S) < 2n^{(1-\delta)/2},\ \mu(T) < 3n^{(1-\delta)/2}.$$

We conclude that there is some x with $\frac{1}{2}\sqrt{n} \le x \le \frac{3}{4}\sqrt{n}$ such that $x \notin S \cup T$. Thus

$$p_2(x) - x \le \frac{1}{3} n^{1/12+\varepsilon}\ ,\quad p_3(n/x) - n/x \le \frac{1}{3} n^{1/12+\varepsilon}\ .$$

Note that $p_2(x) < p_1(n/x)$ (since there are many primes between $\frac{3}{4}\sqrt{n}$ and $\frac{4}{3}\sqrt{n}$) and that $p_1(x)p_1(n/x) > n$. Thus for large n

$$n + f(n) < p_2(x)p_3(n/x) \le (x+\frac{1}{3} n^{1/12+\varepsilon})(n/x+\frac{1}{3} n^{1/12+\varepsilon})$$

$$= n + \frac{1}{3} n^{1/12+\varepsilon}(x+n/x) + \frac{1}{9} n^{1/6+2\varepsilon}$$

$$\le n + \frac{5}{6} n^{7/12+\varepsilon} + \frac{1}{9} n^{1/6+2\varepsilon} \le n + n^{7/12+\varepsilon}\ .$$

We thus have (4) for all large n.

We comment that on the assumption of the Riemann hypothesis, it is known that the "1/6" in (10) can be dropped. Thus the Riemann hypothesis implies $f(n) \le n^{1/2+\varepsilon}$ for all large n.

3. *Blocking configurations.*

From the definition of $f(n)$ there is a set $S = S(n)$ of minimal cardinality such that $S \subset \{n+1, n+2, \ldots, n+f(n)+1\}$, every member of S is composite, and whenever we choose for each $s \in S$ a factor d_s with $1 < d_s < s$, then necessarily $d_s = d_{s'}$ for some pair $s, s' \in S$ with $s \neq s'$. Such a set S is called a *blocking configuration* for n.

If A is a set, denote by $\#A$ the cardinality of A. If S is a set of composite integers, let

$$D(S) = \{d: \text{for some } s \in S,\ d|s,\ 1 < d < s\},$$

$$d(S) = \#D(S).$$

THEOREM 3.1. *(i)* $n + f(n) + 1 \in S(n)$;

(ii) $d(S(n)) = \#S(n) - 1$;

(iii) For all large n, *if* $s \in S(n)$ *then* s *is either the square of a prime or the product of two distinct primes;*

(iv) For each n, $S(n)$ *is unique.*

Proof. (i) This statement is an obvious corollary of the definition of $f(n)$.

(ii) Let S be a blocking configuration for n. If $T \subset S$, $T \neq S$, by the minimality of S it follows that for each $t \in T$ there is a divisor d_t with $1 < d_t < t$ and the d_t are distinct. Thus $d(T) \geq \#T$. Consider the bipartite graph from S to $D(S)$ where s and d are connected by an edge if $d|s$, $1 < d < s$. Since by assumption this graph does not contain a matching of S into $D(S)$, it follows from the "Marriage Theorem" of P. Hall that there is some $T \subset S$ with $\#T > d(T)$. But we have just seen that this inequality fails if $T \neq S$. Thus the guaranteed set T must be S itself. Let $m \in S$. Then

$$\#S - 1 = \#(S-\{m\}) \leq d(S-\{m\}) \leq d(S) < \#S,$$

so $d(S) = \#S - 1$.

(iii) Note that if n is large, then (4) implies $f(n) < n^{2/3}$. Thus if $s, s' \in S(n)$, $s \neq s'$, then $(s,s') < n^{2/3}$. If $s_o \in S(n)$ has three prime factors, then it has a factor d_o with $s_o^{2/3} \leq d_o < s_o$. For each $s \in S(n) - \{s_o\}$, let d_s be a factor with $1 < d_s < s$ and such that the d_s are distinct. But $d_o \neq d_s$ for all $s \in S(n) - \{s_o\}$ for otherwise

$(s,s_o) > n^{2/3}$. This contradicts the definition of S(n). Thus there is no $s_o \in S(n)$ with three prime factors.

(iv) Let $T \subset \{n+1,n+2,\ldots,n+f(n)+1\}$ be any set of composite numbers. Then $d(T) \geq \#T - 1$, for if not then $n + f(n) + 1 = m \in T$ and

$$d(T-\{m\}) \leq d(T) < \#(T-\{m\}),$$

contradicting the definition of f(n). Now suppose S_1, S_2 are two different subsets of $\{n+1,\ldots,n+f(n)+1\}$ such that every member of S_1, S_2 is composite, $d(S_i) < \#S_i$, but if $T \subset S_i$, $T \neq S_i$, then $d(T) \geq \#T$. From the proof of (ii) it follows that a blocking configuration has these properties. Thus to prove (iv) it suffices to show that these properties force $S_1 = S_2$. Note that neither of S_1, S_2 is a subset of the other. Thus

$$\begin{aligned} d(S_1 \cup S_2) &\leq d(S_1) + d(S_2) - d(S_1 \cap S_2) \\ &\leq \#S_1 - 1 + \#S_2 - 1 - \#(S_1 \cap S_2) \\ &= \#(S_1 \cup S_2) - 2 < d(S_1 \cup S_2), \end{aligned}$$

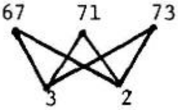
a contradiction.

4. *The normal size of f(n).*

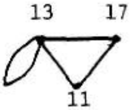
If n is large we can represent the blocking configuration S(n) of n as a graph whose vertices are the prime factors of the members of S(n) and two primes are joined by an edge if and only if their product is in S(n). If the square of a prime is in S(n), we represent this as a loop. We call this graph G(n). Each vertex of G(n) has valence at least 2 and there is exactly one more edge in G(n) than vertices. We give examples of a few G(n).

G(132):

G(133):



G(134):



Note that (4) shows that if n is large, then the largest prime in $G(n)$ is at most $n^{7/12+\varepsilon}$, so that the smallest prime in $G(n)$ is at least $n^{5/12-\varepsilon}$.

We now show (5). Let p_1 denote the largest prime in $G(n)$. Say it is connected to p_i and p_j where $p_i > p_j$. (Note that we allow the possibility $p_1 = p_i$.) Now p_j is connected to some $p_k \neq p_1$. Since we may assume all of these primes are odd, we have

$$f(n) \geq p_1 p_i - p_k p_j \geq p_1(p_j+2) - (p_1-2)p_j$$

$$= 2p_1 + 2p_j > 4\sqrt{p_1 p_j} > 4\sqrt{n}.$$

Note that if $f(n) < (4+\varepsilon)\sqrt{n}$, then the above argument shows that $p_i = p_j + 2$. On the other hand if $p, p+2$ are both prime, then the configuration $p^2, p(p+2), (p+2)^2$ shows that $f(p^2-1) \leq 4p+4$. Thus equality in (5) is equivalent to the existence of infinitely many twin primes.

For most integers n, the interval $[n+1, n+10\sqrt{n}]$ is free of squares of primes. If A denotes the set of such n, then A has asymptotic density 1. In fact, the number of $n \leq x$ with $n \notin A$ is $O(x/\log x)$. We now show (6) for the set A.

THEOREM 4.1. *For all sufficiently large* $n \in A$ *we have* $f(n) > 4\sqrt{2n}$.

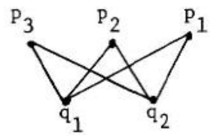
Proof. Suppose $n \in A$ and that $f(n) \le 4\sqrt{2n}$. Also suppose n is large enough so that (iii) of Theorem 3.1 holds. Thus every member of $S(n)$ is of the form pq where p, q are primes. Since the smallest prime in $G(n)$ is at least $n^{5/12-\varepsilon}$, we may assume all of the primes in $G(n)$ exceed 5.

We first note that no prime $p > \sqrt{n}$ in $G(n)$ is connected to two primes differing by 6 or more for otherwise

$$f(n) \ge 6p > 6\sqrt{n} > 4\sqrt{2n}.$$

Since there are no squares of primes in $n+1, n+10\sqrt{n}$, we conclude that no two primes exceeding $\sqrt{n}$ are connected in $G(n)$. Denote the primes exceeding $\sqrt{n}$ in $G(n)$ by $p_1 < p_2 < \dots < p_\ell$ and the remaining primes by $q_1 < q_2 < \dots < q_k$. Each p is connected to exactly two q's (which are necessarily consecutive primes differing by 2 or 4) and each q is connected to at least two p's.

If $k = 2$, then $\ell \ge 3$ and $G(n)$ must contain the subgraph



and so this subgraph must in fact be $G(n)$. But $p_3 \ge p_1 + 6$, so that

$$f(n) \ge p_3 q_2 - p_1 q_1 \ge (p_1 + 6)(q_1 + 2) - p_1 q_1$$

$$> 2p_1 + 6q_1 > 2p_1 + 6n/p_1 \ge 4\sqrt{3n} > 4\sqrt{2n},$$

a contradiction. We conclude that $k \ge 3$.

Say for some i, j we have $q_j - q_i \ge 4$ and that q_j is connected to p_a and q_i is connected to p_b. Then we must have $p_a \le p_b$. For if not, then

$$f(n) \ge p_a q_j - p_b q_i \ge (p_b + 2)(q_i + 4) - p_b q_i$$

$$> 4p_b + 2q_i > 4p_b + 2n/p_b > 6\sqrt{n} > 4\sqrt{2n},$$

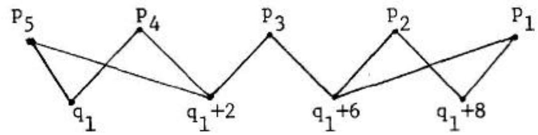
a contradiction.

We conclude that $q_2 - q_1 = 2$. For if $q_2 - q_1 = 4$, q_1 is connected to $p_b < p_\ell$ and p_ℓ is connected to $q_j > q_1$, then $q_j - q_1 \geq 4$ and we are in the situation just covered. Similarly we have $q_k - q_{k-1} = 2$.

Putting together what we have learned about $q_1,\ldots,q_k$ with the fact that the q's are coprime to 30, we have

$$\{q_1,\ldots,q_k\} = \{q_1, q_1+2, q_1+6, q_1+8\}.$$

We conclude that G(n) must contain the subgraph



and so this must be G(n).

Now note that $p_5 - p_4 = 2$. For if $p_5 - p_4 \geq 4$, then

$$f(n) \geq p_5(q_1+2) - p_4q_1 \geq (p_4+4)(q_1+2) - p_4q_1$$
$$> 2p_4 + 4q_1 > 2p_4 + 4n/p_4 \geq 4\sqrt{2n}.$$

Similarly we have $p_2 - p_1 = 2$.

We have $p_3 \leq \sqrt{2n}$, for if not, then

$$f(n) \geq p_3(q_1+6) - p_3(q_1+2) = 4p_3 > 4\sqrt{2n}.$$

Thus $q_1 + 2 > \sqrt{n/2}$. But then $p_5 - p_3 \leq 6$, for if not,

$$f(n) \geq p_5(q_1+2) - p_3(q_1+2) \geq 8(q_1+2) > 4\sqrt{2n}.$$

Similarly $p_3 - p_1 \leq 6$. We conclude that $p_4 - p_3 = p_3 - p_2 = 4$, so that one of these primes is divisible by 3, a contradiction. Thus the theorem is established.

We now give a heuristic argument for (7). However, the first part of the argument is rigorous. Let $1 > \varepsilon > 0$ be arbitrary, but fixed. Let h(x) denote the function $2x + \frac{n}{x} - 9$. For each large integer n we can find an integer x_0 with the following properties:

$$(11)\qquad \begin{cases} \sqrt{n/2} < x_0 < \sqrt{n/2} + 9n^{1/4}, \\ 1 - \frac{2}{3}\,\varepsilon < h(x_0) - [h(x_0)] < 1 - \frac{1}{3}\,\varepsilon, \\ [h(x_0)] \equiv 39 \pmod{210}. \end{cases}$$

To see that x_0 exists, note that for $0 \le \alpha \le 9$

$$h(\sqrt{n/2} + \alpha n^{1/4}) = 2\sqrt{2n} - 9 + 2\sqrt{2}\,\alpha^2 + O(n^{-1/4}).$$

Clearly there is a real number α_0, $1 < \alpha_0 < 9$, such that

$$2\sqrt{2n} - 9 + 2\sqrt{2}\,\alpha_0^2 \stackrel{\text{def}}{=} m_0 \equiv 40 \pmod{210}.$$

Thus there are positive quantities $1 > \delta_2 > \delta_1 > 0$ such that if x_0 is any integer in the interval

$$\sqrt{n/2} + (\alpha_0-\delta_2)n^{1/4} < x_0 < \sqrt{n/2} + (\alpha_0-\delta_1)n^{1/4},$$

then

$$m_0 - \frac{2}{3}\,\varepsilon < h(x_0) < m_0 - \frac{1}{3}\,\varepsilon.$$

This number x_0 satisfies the conditions in (11).

Now let $w_0 = [\frac{n}{x_0}] - 9$, so that

$$[h(x_0)] = 2x_0 + w_0 \equiv 39 \pmod{210}.$$

Thus there are infinitely many integers y such that the nine numbers

$$(12)\qquad \begin{cases} x_0 + y & w_0 - 2y \\ x_0 + y + 2 & w_0 - 2y + 2 \\ x_0 + y + 6 & w_0 - 2y + 6 \\ x_0 + y + 8 & w_0 - 2y + 12 \\ & w_0 - 2y + 14 \end{cases}$$

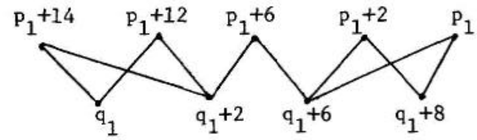
are simultaneously coprime to 210. Indeed, if $y \equiv 11 - x_0 \pmod{210}$, then the first column modulo 210 is 11, 13, 17, 19. Since

$$w_0 - 2y \equiv w_0 + 2x_0 - 22 \equiv 17 \pmod{210},$$

the second column modulo 210 is 17, 19, 23, 29, 31.

Thus from the prime k-tuples conjecture there are infinitely many values of y for which the integers in (12) are all prime. We now make an even stronger conjecture. Namely, we assert that for all sufficiently large n there is a value of y with $|y| < n^{1/5}$ and such that each number in (12) is prime.

With such a value of y, let $q_1 = x_0 + y$, $p_1 = w_0 - 2y$ and consider the graph



The three largest integers represented by edges in this graph are

$$
(13)\quad \begin{cases}
(q_1+8)(p_1+2) = x_0w_0 + y(w_0-2x_0) + 8w_0 + 2x_0 + (y+8)(-2y+2), \\
(q_1+6)(p_1+6) = x_0w_0 + y(w_0-2x_0) + 6w_0 + 6x_0 + (y+6)(-2y+6), \\
(q_1+2)(p_1+14) = x_0w_0 + y(w_0-2x_0) + 2w_0 + 14x_0 + (y+2)(-2y+14),
\end{cases}
$$

while the three smallest integers represented by edges in the graph are

$$
\begin{aligned}
q_1(p_1+12) &= x_0w_0 + y(w_0-2x_0) + 12x_0 + y(-2y+12), \\
(q_1+2)(p_1+6) &= x_0w_0 + y(w_0-2x_0) + 2w_0 + 6x_0 + (y+2)(-2y+6), \\
(q_1+6)p_1 &= x_0w_0 + y(w_0-2x_0) + 6w_0 + (y+6)(-2y).
\end{aligned}
$$

Since

$$
x_0 = \sqrt{n/2} + O(n^{1/4}), \quad w_0 = \sqrt{2n} + O(n^{1/4}), \quad y^2 = O(n^{2/5}),
$$

it follows that the least of the 3 smallest numbers is the middle one for all large n. Moreover, since $w_0 - 2x_0 = O(n^{1/4})$, it follows that

$$
\begin{aligned}
(q_1+2)(p_1+6) &= x_0w_0 + O(n^{9/20}) + 2w_0 + 6x_0 + O(n^{2/5}) \\
&= x_0(w_0+10) + O(n^{9/20}) \\
&= x_0([\tfrac{n}{x_0}]+1) + O(n^{9/20}) \\
&= n + x_0 - x_0(\tfrac{n}{x_0} - [\tfrac{n}{x_0}]) + O(n^{9/20}) .
\end{aligned}
$$

Thus from (11),

$$n + \frac{1}{3}\varepsilon x_0 + O(n^{9/20}) \le (q_1+2)(p_1+6) \le n + \frac{2}{3}\varepsilon x_0 + O(n^{9/20}) .$$

We conclude that for all large n, we have

$$n < (q_1+2)(p_1+6) \le n + \varepsilon\sqrt{n}.$$

Since each of the numbers in (13) is $(4\sqrt{2} + o(1))\sqrt{n}$ more than $(q_1+2)(p_1+6)$, it follows that

$$f(n) \le (4\sqrt{2} + \varepsilon + o(1))\sqrt{n}$$

for all large n. Since $\varepsilon > 0$ is arbitrary, we have

$$\limsup f(n)/\sqrt{n} \le 4\sqrt{2}.$$

This inequality (which depends on a strong form of the prime k-tuples conjecture) with Theorem 4.1 implies (7).

5. *A related problem.*

We say that a set of natural numbers S has the distinct divisor property if for each $s \in S$ we can find a divisor $d_s | s$, $1 \le d_s < s$, such that the d_s are all distinct. If $c > 1$, let $f(n;c)$ denote the cardinality of the largest subset of $[n,cn]$ which has the distinct divisor property.

Put another way, we can let $G(n;c)$ denote the bipartite graph from the integers in $[n,cn]$ to the set of their proper divisors, where edges connect numbers in $[n,cn]$ to their proper divisors. Then $f(n;c)$ is the cardinality of the largest matching in $G(n;c)$.

THEOREM 5.1. *For each* $c > 1$ *there is a constant* $\delta(c)$ *such that* $f(n;c) \sim \delta(c)n$ *as* $n \to \infty$.

Proof. To prove the theorem we break the graph $G(n;c)$ into little pieces and then put them back together. Perhaps there is a more direct proof, but we have not been able to find it.

Let $B > 0$ be a fixed but arbitrarily large integer. Let $G(n;c,B)$ denote the subgraph of $G(n;c)$ where each of the edges m, d (where $m \in [n,cn]$, $d|m$, $1 \le d < m$) satisfies $m/d \le B$. Let $f(n;c,B)$ denote the cardinality of the largest matching in $G(n;c,B)$. We shall show below that there is a constant $\delta(c,B)$ such that

(14) $$f(n;c,B) \sim \delta(c,B)n \quad \text{as} \quad n \to \infty .$$

This result implies the theorem. Indeed, if $B_1 < B_2$, then clearly $f(n;c,B_1) \le f(n;c,B_2)$ so that $\delta(c,B_1) \le \delta(c,B_2)$. Thus

$$\delta(c) \overset{\text{def}}{=} \lim_{B\to\infty} \delta(c,B)$$

exists. Now the number of d such that $d|m$ for some $m \in [n,cn]$ and $m/d > B$ is at most cn/B. Thus

$$f(n;c,B) \le f(n;c) \le f(n;c,B) + cn/B,$$

and so

$$\delta(c,B) \le \underline{\lim}\, \frac{1}{n} f(n;c) \le \overline{\lim}\, \frac{1}{n} f(n;c) \le \delta(c,B) + c/B.$$

Letting $B \to \infty$, we have $\delta(c) = \lim \frac{1}{n} f(n;c)$, which was to be proved.

Let $B' > B$ be arbitrarily large but fixed. If m is any positive integer, we can factor $m = a(m)b(m)$ where $a(m)$ is the largest divisor of m that is coprime to $B!$. Let $G(n;c,B,B')$ denote the subgraph of $G(n;c,B)$ where we take only those $m \in [n,cn]$ with $b(m) \le B'$. We shall show below that there is a constant $\delta(c,B,B')$ such that if $f(n;c,B,B')$ is the cardinality of the largest matching in $G(n;c,B,B')$, then

(15) $$f(n;c,B,B') \sim \delta(c,B,B')n \text{ as } n \to \infty .$$

In the same way as the theorem follows from (14), we can show (14) follows from (15). Indeed, if $B_1' < B_2'$, then $f(n;c,B,B_1') \le f(n;c,B,B_2')$. Thus $\delta(c,B,B_1') \le \delta(c,B,B_2')$ so that

$$\delta(c,B) \overset{\text{def}}{=} \lim_{B'\to\infty} \delta(c,B,B')$$

exists. Now the number of $m \in [n,cn]$ with $b(m) > B'$ is at most

$$\sum_{b>B'}{}' \; cn/b$$

where Σ' denotes the sum over those b free of primes exceeding B. Since $\Sigma' \; 1/b < \infty$, it follows that

$$\lim_{B'\to\infty} \sum_{b>B'}{}' \; 1/b = 0.$$

But

$$f(n;c,B,B') \le f(n;c,B) \le f(n;c,B,B') + \sum_{b>B'}' cn/b,$$

so that

$$\delta(c,B,B') \le \underline{\lim}\, \frac{1}{n} f(n;c,B) \le \overline{\lim}\, \frac{1}{n} f(n;c,B) \le \delta(c,B,B') + \sum_{b>B'}' c/b.$$

Letting $B' \to \infty$, we have (14).

If a is an integer coprime to $B!$, let $G(n;c,B,B',a)$ denote the subgraph of $G(n;c,B,B')$ where we take only those $m \in [n,cn]$ with $a(m) = a$. Note that if m, d is an edge in $G(n;c,B,B',a)$, then $a|d$. Indeed, $m/d \le B$ and $(a,B!) = 1$, so $a|d$. A corollary is that if $a_1 \ne a_2$, then any connected component of $G(n;c,B,B')$ does not intersect both $G(n;c,B,B',a_1)$ and $G(n,c,B,B',a_2)$. Indeed, if m, d and m', d are two edges in $G(n;c,B,B')$, then $a(m)|d|m'$ and $a(m')|d|m$, so that $a(m) = a(m')$.

It thus follows that a maximal matching in $G(n;c,B,B')$ corresponds to a union of maximal matchings in the $G(n;c,B,B',a)$'s. If $f(n;c,B,B',a)$ is the cardinality of the largest matching in $G(n;c,B,B',a)$, then

$$f(n;c,B,B') = \sum_a f(n;c,B,B',a). \tag{16}$$

Although there are many terms in this sum, we note that up to isomorphism there are really only a bounded number of different graphs $G(n;c,B,B',a)$. Indeed, we can list the numbers $b \le B'$ composed solely of the primes up to B in increasing order:

$$b_1 < b_2 < \ldots < b_k.$$

Then for each a for which $G(n;c,B,B',a) \ne \emptyset$ there is an i, j with $1 \le i \le j \le k$ and such that the vertices in $[n,cn]$ that are in $G(n;c,B,B',a)$ are $b_i a, b_{i+1} a, \ldots, b_j a$. In addition, the edges in $G(n;c,B,B',a)$ connect a number ba to a divisor da where $1 < b/d \le B$. For each pair i, j with $1 \le i \le j \le k$, let $G_{i,j}$ denote the bipartite graph from $\{b_i, b_{i+1}, \ldots, b_j\}$ to factors where b, d is an edge if $d|b$ and $1 < b/d \le B$. Thus we have seen that each $G(n;c,B,B',a) \ne \emptyset$ is canonically isomorphic to a $G_{i,j}$; and so if $f_{i,j}$ is the cardinality of the largest matching in $G_{i,j}$, then

$$f_{i,j} = f(n;c,B,B',a).$$

Let $g_{i,j}(n;c,B,B')$ denote the number of values of a with $G(n;c,B,B',a)$ canonically isomorphic to $G_{i,j}$. Then from (16), we have

$$f(n;c,B,B') = \sum_{1\le i\le j\le k} f_{i,j}g_{i,j}(n;c,B,B').$$

Thus to prove (15) and ultimately the theorem it is sufficient to show there are constants $\delta_{i,j}(c,B,B')$ with

$$(17)\qquad \begin{aligned} &g_{i,j}(n;c,B,B') \sim \delta_{i,j}(c,B,B')n \quad \text{as} \quad n \to \infty \\ \text{or} \quad &g_{i,j}(n;c,B,B') = O(1) \quad \text{as} \quad n \to \infty . \end{aligned}$$

For $G(n;c,B,B',a)$ to be canonically isomorphic to $G_{i,j}$ it is necessary and sufficient that

$$(a,B!) = 1,\ b_{i-1}a < n \le b_i a, \text{ and } b_j a \le cn < b_{j+1}a$$

(where we let $b_0 = 0$, $b_{k+1} = \infty$). Let

$$\alpha = \max\left\{\frac{1}{b_i}, \frac{c}{b_{j+1}}\right\} \quad , \quad \beta = \min\left\{\frac{1}{b_{i-1}}, \frac{c}{b_j}\right\} .$$

Then the difference between $g_{i,j}(n;c,B,B')$ and the number of $a \in [\alpha n, \beta n]$ with $(a,B!) = 1$ is at most 2. This possible error is caused by the ambiguity of the 2 possible extreme values for a. Thus if $\alpha < \beta$,

$$g_{i,j}(n;c,B,B') \sim (\beta-\alpha)\frac{\phi(B!)}{B!}n \text{ as } n \to \infty ,$$

while if $\alpha \ge \beta$, then

$$g_{i,j}(n;c,B,B') = O(1) \quad \text{as} \quad n \to \infty .$$

This proves (17) and thus the theorem.

We now collect together some results about the function $\delta(c)$.

THEOREM 5.2. *(i) The function* $\delta(c)$ *is continuous and strictly increasing,*
(ii) $1/2 < \delta(c)/(c-1) < 1$, $\delta(c) < c/2$,
(iii) $\lim_{c\to 1^+} \delta(c)/(c-1) = 1$, $\lim_{c\to\infty} \delta(c)/(c-1) = 1/2$.

Proof. Let S be a subset of $[n,cn]$ with the distinct divisor property and let $\varepsilon > 0$ be arbitrary. If S' denotes the set of even numbers in $(cn,(c+\varepsilon)n]$, then $S \cup S'$ has the distinct divisor property. Indeed, the members of S' can be mapped to their halves; this does not interfere with the divisors of members of S. We thus have

$$f(n;c) + \varepsilon n/2 - 1 < f(n;c+\varepsilon) \le f(n;c) + \varepsilon n,$$

so that

$$\delta(c) + \varepsilon/2 \le \delta(c+\varepsilon) \le \delta(c) + \varepsilon .$$

This proves (*i*).

The even numbers in $[n,cn]$ have the distinct divisor property. A proper divisor d of any number in $[n,cn]$ satisfies $d \le cn/2$. These two observations immediately give

$$(18) \qquad (c-1)/2 \le \delta(c) \le c/2 \text{ for all } c > 1.$$

Therefore $\lim_{c\to\infty} \delta(c)/(c-1) = 1/2$, which is part of (*iii*). To see that we can make the first inequality in (18) strict, note that the small odd multiples of 3 in $[n,cn]$ can be mapped to 1/3 of themselves and this will not interfere with mapping evens in $[n,cn]$ to their halves. Specifically we have

$$\delta(c) \ge \begin{cases} 2(c-1)/3, & 1 < c \le 3/2 \\ (c-1)/2 + 1/12, & c \ge 3/2, \end{cases}$$

so that $\delta(c)/(c-1) > 1/2$ for all $c > 1$, proving part of (*ii*).

To see that the second inequality in (18) is strict, suppose not, so $\delta(c) = c/2$ for some c. If $c' > c$, the argument that gives (*i*) shows that

$$\delta(c') \ge \delta(c) + (c'-c)/2 = c'/2$$

so that $\delta(c') = c'/2$. Thus we may assume the value of c with $\delta(c) = c/2$ also satisfies $c \ge 2$. If $S_n \subset [n,cn]$ is a maximal set with the distinct divisor property, then $\#S_n = cn/2 + o(n)$. But each proper divisor of a member of S_n does not exceed $cn/2$. Therefore, but for $o(n)$ exceptions, we can map the integers in $[1,cn/2]$ to distinct multiples in $[n,cn]$. Since $c/2 \ge 1$, we thus have a subset $T_n \subset [1,n]$ with $\#T_n = n + o(n)$ such that the members of T_n can be mapped to distinct multiples in $[n,cn]$.

Let t denote an arbitrarily large, but fixed integer. Consider all of the integers

$$k, 2k, \ldots, tk$$

where k runs over the integers in $(\frac{n}{t+1}, \frac{n}{t}]$. These integers are all different, for if $ik = jk'$ with $k < k'$, then

$$\frac{t}{t-1} \le \frac{i}{j} = \frac{k'}{k} < \frac{t+1}{t},$$

a contradiction. Then for some $k \in (\frac{n}{t+1}, \frac{n}{t}]$ and for each n bigger than some $n_0(t)$, the set T_n contains all of $k, 2k, \ldots, tk$. For if not, then

$$\#T_n \le (1 - \frac{1}{t^2+t})\, n + 1,$$

contradicting $\#T_n = n + o(n)$. Thus if $n \ge n_0(t)$, there is some $k \in (\frac{n}{t+1}, \frac{n}{t}]$ such that $k, 2k, \ldots, tk$ all have distinct multiples in $[n, cn]$. It then follows from Theorem 2 in [2] that

$$c \ge (\frac{2}{\sqrt{e}} + o(1))\sqrt{\log t/\log\log t}$$

where the "$o(1)$" tends to 0 as $t \to \infty$. But this inequality fails for large t. This contradiction shows that $\delta(c) < c/2$, proving another part of *(ii)*.

The second inequality in (18) shows that $\delta(c)/(c-1) < 1$ for $c > 2$. Suppose now $1 < c \le 2$. Let T denote the set of integers in $[n, cn]$ not divisible by any prime up to $c/(c-1)$. Then

$$\#T \sim \{(c-1)\Pi_1\, (1 - 1/p)\}n \quad \text{as} \quad n \to \infty$$

where Π_1 denotes the product over primes $p \le c/(c-1)$. If $d|t$, $d < t$, $t \in T$, then $d \le cn/p_0$ where p_0 is the first prime exceeding $c/(c-1)$. Moreover, d is not divisible by any prime $p \le c/(c-1)$. Thus the number of proper divisors of members of T is

$$\le (1 + o(1))\{(c/p_0)\Pi_1\, (1-1/p)\}n \quad \text{as} \quad n \to \infty.$$

Since $c/p_0 < c - 1$, it follows that

$$(19) \qquad \delta(c) \le c - 1 - \{(c-1)\Pi_1\, (1 - 1/p) - (c/p_0)\Pi_1\, (1 - 1/p)\} < c - 1.$$

This completes the proof of *(ii)*.

It remains to show $\lim_{c\to 1^+} \delta(c)/(c-1) = 1$. Let $p_1 < p_2$ denote the first two primes with $p_2/p_1 \le c$. Let U denote the set of $m \in [n, cn]$ divisible by some prime below p_2. Then

$$\# U \sim \{(c-1)(1-\Pi_2(1-1/p))\} n \quad \text{as} \quad n \to \infty$$

where Π_2 denotes the product over the primes $p < p_2$. If $m \in U$, let $p(m)$ denote the least prime factor of m. Then the mapping $m \to m/p(m)$ is one-to-one on U. For if $m_1/p(m_1) = m_2/p(m_2)$ where $p(m_1) \le p(m_2)$, then

$$1 \le p(m_2)/p(m_1) = m_2/m_1 \le c.$$

Since $p(m_2) < p_2$, we have $p(m_2) = p(m_1)$ so that $m_2 = m_1$. Thus the mapping is one-to-one as claimed. We conclude that $f(n;c) \ge \# U$, so that

$$\delta(c) \ge (c-1)(1-\Pi_2(1-1/p)).$$

By the prime number theorem $p_2 \to \infty$ as $c \to 1^+$. Thus $\Pi_2(1-1/p)$ diverges to 0 as $c \to 1^+$. We conclude that

$$\liminf_{c \to 1^+} \delta(c)/(c-1) \ge 1.$$

Combined with (19), we have $\lim_{c \to 1^+} \delta(c)/(c-1) = 1$.

6. *Further comments.*

In section 3 we proved that if n is large, then every member of n's blocking configuration is a product of two primes. We have computed the blocking configurations for all $n \le 436$ and we found that in each case every member is the product of two primes. We thus conjecture that there are no exceptions, that for every n, $S(n)$ consists solely of integers the product of two primes.

Is $\# S(n)$ bounded? In particular, can this be seen to follow from our other conjectures?

Let $f_1(n)$ be the corresponding function to $f(n)$, but now we allow the divisor 1 to be used (but only once, of course). The function $f_1(n)$ behaves very much like $f(n)$. The only change is that the numbers in (5), (6), (7) are different.

Suppose in the definition of $f(n)$, instead of asking that $m \in [n+1, n+f(n)]$ be composite, let us ask that m has at least three (or r), not necessarily distinct prime factors. The blocking sets get much more complicated (in fact, how large is the smallest blocking set for r prime factors?) and it seems that instead of $f(n) \asymp n^{\frac{1}{2}}$, the corresponding function $f_r(n)$ will have an exponent that increases with r. Finally suppose we only want that for almost all m in $[n+1, n+F(n)]$ there should be proper factors d_m of m, distinct for different values of m. Then from the inequality $\delta(c)/(c-1) < 1$ of Theorem 5.2 it follows that $F(n) = o(n)$. Put another way, if $F(n)$ is any function such that for each n, $[n+1, n+F(n)]$ contains a subset of size $(1+o(1))\,F(n)$ with the distinct divisor property, then Theorem 5.2 implies $F(n) = o(n)$.

From the proof of Theorem 5.2, it follows that the function $c/2 - \delta(c)$ is positive and non-increasing. It therefore tends to a limit. Is this limit 0? That is, do we have $\delta(c) = c/2 + o(1)$ as $c \to \infty$?

Is the function $\delta(c)/(c-1)$ monotone? Is it strictly monotone? If the latter is so, the following corollary holds. For each number α, $1/2 < \alpha < 1$, let $F(n;\alpha)$ denote the largest integer so that in $[n, n+F(n;\alpha)]$ there is a subset of cardinality at least $\alpha F(n;\alpha)$ with the distinct divisor property. Then there is a number $\gamma > 0$ such that $F(n;\alpha) \sim \gamma n$ as $n \to \infty$. In fact, if γ exists, then clearly $\delta(\gamma+1)/\gamma = \alpha$. If we knew that $\delta(c)/(c-1)$ were strictly monotone and if β denotes the inverse function, then $\gamma = \beta(\alpha) - 1$. Can γ be proven to exist without using $\delta(c)/(c-1)$ strictly monotone?

In [2] we consider a problem that is in a sense "dual" to the considerations with $f(n)$. With $f(n)$ we map all composite numbers just above n to distinct divisors. In [2] we map the first n integers to distinct multiples just above m. Specifically, we let $f(n,m)$ denote the least integer so that in $[m+1, m+f(n,m)]$ we can find $a_1, \ldots, a_n$ with $i \mid a_i$ for $i = 1, \ldots, n$. We establish some results on the average order and maximal order of $f(n,m)$ (considered as a function of m) and we also obtain estimates for $f(n,n)$.

REFERENCES

[1] R. J. Cook, *An upper bound for the sum of large differences between prime numbers*, Proc. Amer. Math. Soc. *81* (1981), 33-40.

[2] P. Erdős and C. Pomerance, *Matching the natural numbers up to* n *with distinct multiples in another interval*, Nederl. Akad. Wetensch. Proc. Ser. A *83* (1980), 147-161.

[3] P. Erdős and J. L. Selfridge, *Some problems on the prime factors of consecutive integers. II*, Proc. Washington State U. Conf. on Number Th. at Pullman (1971), 13-21.

[4] S. W. Graham, *The greatest prime factor of the integers in an interval*, J. London Math. Soc., to appear.

[5] C. A. Grimm, *A conjecture on consecutive composite numbers*, Amer. Math. Monthly *76* (1969), 1126-1128.

[6] M. N. Huxley, *On the difference between consecutive primes*, Invent. Math. *15* (1972), 164-170.

[7] C. Pomerance, *Some number theoretic matching problems*, Queen's Papers in Pure and Applied Math. *54* (1980), 237-247.

[8] K. Ramachandra, *A note on numbers with a large prime factor*, J. London Math. Soc. (2) *1* (1969), 303-306.

[9] K. T. Ramachandra, N. Shorey, and R. Tijdeman, *On Grimm's problem relating to factorization of a block of consecutive integers*, J. reine angew. Math. *273* (1975), 109-124.

[10] J. W. M. Turk, *Multiplicative properties of neighbouring integers*, Doctoral Dissertation at the University of Leiden, The Netherlands, September, 1979.

[11] R. Warlimont, *Über die Haufigkeit grosser Differenzen konsekutiver Primzahlen,* Monatsh. Math. *83* (1977), 59-63.

Mathematical Institute
Hungarian Academy of Sciences
Budapest, Hungary

Department of Mathematics
University of Georgia
Athens, Georgia 30602 U.S.A.

Received September 5, 1980; Revised November 19, 1982.