#### ON SPARSELY TOTIENT NUMBERS

### D. W. MASSER AND P. SHIU

Let 
$$\varphi(n)$$
 denote Euler's totient function, defined for  $n > 1$  by 
$$\varphi(n) = n \prod_{p \mid n} (1 - p^{-1}).$$

Let F be the set of integers n>1 with the property that  $\varphi(m)>\varphi(n)$  whenever m>n. The purpose of this paper is to establish a number of results about the set F. For example, we shall prove that each prime divides all sufficiently large elements of F, each positive integer divides some element of F, and that the ratio of successive elements of F approaches 1.

1. Introduction. Similar studies have been carried out in the past, initially by Ramanujan [7] for the divisor function d(n), and then by Alaoglu and Erdös [1] for d(n) and the divisor sum function  $\sigma(n)$ , and by Erdös and Nicolas [2] for the prime divisor function  $\omega(n) = \sum_{p|n} 1$  (see also the last paper for additional references). In particular, Ramanujan considered the set of integers n such that d(m) < d(n) whenever 1 < m < n. He called such integers highly composite, and by analogy it seems appropriate to refer to the elements of our set F as sparsely totient numbers.

Since  $\varphi(n) \to \infty$  as  $n \to \infty$ , it is obvious that F is infinite. Our first result shows how to construct many elements of F explicitly. Let  $p_1 = 2$ ,  $p_2 = 3, \ldots$  denote the primes in ascending order of magnitude.

Theorem 1. Suppose  $k \geq 2$ ,  $d \geq 1$ ,  $l \geq 0$  and

(a) 
$$d < p_{k+1} - 1$$

(b) 
$$d(p_{k+1}-1) < (d+1)(p_k-1)$$
.

Then  $dp_1 \cdots p_{k-1} p_{k+1}$  is in F.

COROLLARY. Let n, n' be consecutive elements of F. Then  $n'/n \to 1$  as  $n \to \infty$ .

For n > 1 denote by P(n) the greatest prime factor of n and by Q(n) the smallest prime not dividing n. Already Theorem 1 above provides some information about large values of P(n) and Q(n) for n in F, as well as showing that there are elements of F divisible by any given integer d. Also, the statement that each prime divides all sufficiently large elements of F is equivalent to  $Q(n) \to \infty$  as  $n \to \infty$  in F. We shall prove this in much more precise form in our next result.

Now we observe that since  $\varphi(2m) = \varphi(m)$  for m odd, it follows that every element of F is even. Also, since  $\varphi(2^{k-1} \cdot 3) = \varphi(2^k)$  for  $k \geq 2$ , we see that the only power of 2 in F is 2 itself. Hence every n > 2 in F has a well-defined second greatest prime factor, which we denote by P'(n). As this function turns out to be of special significance in the study of F, we also give some of its properties in the result below.

THEOREM 2. For n in F we have

(a) 
$$\liminf_{n\to\infty} \frac{P(n)}{\log n} = 1, \limsup_{n\to\infty} \frac{P(n)}{\log n} \ge 2,$$

(b) 
$$\liminf_{n\to\infty} \frac{Q(n)}{\log n} \ge \sqrt{2} - 1$$
,  $\limsup_{n\to\infty} \frac{Q(n)}{\log n} = 1$ ,

(c) 
$$\liminf_{n \to \infty} \frac{P'(n)}{\log n} = 1, \qquad \limsup_{n \to \infty} \frac{P'(n)}{\log n} \le \sqrt{2} + 1,$$

and

(d) 
$$\limsup_{n \to \infty} \frac{P(n)}{\log^2 n} \le 1.$$

Many of the problems concerning sparsely totient numbers are related to the distribution of primes. For example, we shall see that it follows easily from Bertrand's Postulate that  $(P(n))^4$  never divides n in F. Using a deeper result on primes in short intervals we sharpen this as follows.

THEOREM 3. For all sufficiently large n in F, the power  $(P(n))^3$  does not divide n.

We see by taking  $d = p_k$ , l = 0 in Theorem 1 that  $p_1 \cdots p_{k-1} p_k^2$  is sparsely totient for all  $k \ge 2$ , and consequently the exponent in Theorem 3 is best possible.

Finally let F(x) denote the counting function of F; that is, the number of sparsely totient numbers n with  $1 \le n \le x$ . It is not difficult to verify that the explicit constructions in Theorem 1 give the lower bound

$$F(x) \gg \log^2 x / \log \log x$$

for  $x \ge 2$ . In our last result we give the following somewhat larger upper bound.

THEOREM 4. We have

$$\log F(x) \ll \log^{1/2} x$$

for all  $x \geq 2$ .

We also include in this paper an Appendix which contains a brief account of further work on the set F. Glyn Harman very kindly showed us a method of improving (d) in Theorem 2 to  $P(n) \ll \log^{2-\delta} n$  for some  $\delta > 0$ . In addition we describe a plausible gap hypothesis which enables us to obtain best possible versions of all the statements of Theorem 2, thereby considerably illuminating the structure of sparsely totient numbers. Finally we include a table of the 150 elements of F not exceeding  $10^6$ , together with their factorizations.

We end this introduction with a word about the related set  $F^*$  of highly totient numbers n > 1 with the property that  $\varphi(m) < \varphi(n)$  whenever 1 < m < n. Clearly  $F^*$  contains all primes, and it is very probable that there are no other elements in  $F^*$ ; furthermore this can in fact be established with the help of a suitable gap hypothesis (see also (3) of [1], p. 465). So the set  $F^*$  seems comparatively uninteresting.

# 2. Proof of Theorem 1. We start with the following lemma.

LEMMA 1. For  $r \ge 1$  let  $x_1, \ldots, x_r, y_1, \ldots, y_r, X, Y$  be real numbers satisfying

$$\max(x_1,\ldots,x_r) \le Y, \qquad 1 < x_i \le y_i (1 \le i \le r).$$

Then if also

$$(2.1) y_1 \cdots y_r Y > x_1 \cdots x_r X$$

we have

$$(2.2) \quad (y_1-1)\cdots(y_r-1)(Y-1) > (x_1-1)\cdots(x_r-1)(X-1).$$

*Proof.* We note first that (2.2) is trivial if X < 1, since the left-hand side is positive. Similarly if  $1 \le X < y$ , then

$$(y_r - 1)(Y - 1) > (X - 1)(Y - 1) \ge (x_r - 1)(X - 1)$$

and again (2.2) follows immediately. Next, if  $X \ge y_r$  we have

$$(X-y_r)(y_r-x_r)\geq 0,$$

and on adding  $y_r x_r X + y_r$  to both sides, rearranging, and dividing by  $y_r$  we get

$$(2.3) (y_r - 1)(X' - 1) \ge (x_r - 1)(X - 1),$$

where  $X' = Xx_r/y_r$ . Note that all this proves the lemma for r = 1, as X' < Y.

We can now argue by induction on r. Suppose the lemma has been proved with r replaced by r-1 for some  $r \ge 2$ . From the above, it suffices to establish (2.2) when  $X \ge y_r$ , so that (2.3) holds. We can write (2.1) as

$$y_1 \cdot \cdot \cdot y_{r-1}Y > x_1 \cdot \cdot \cdot x_{r-1}X'$$

with  $X' = Xx_r/y_r$  as before, and now the inductive hypothesis shows that

$$(y_1-1)\cdots(y_{r-1}-1)(Y-1)>(x_1-1)\cdots(x_{r-1}-1)(X'-1).$$

Multiplying by  $y_r - 1$  and using (2.3) completes the inductive step. This proves the lemma.

Now we start on the proof of Theorem 1. Let

$$(2.4) n = dp_1 \cdots p_{k-1} p_{k+1}$$

satisfy the conditions (a) and (b) of the theorem. Then

$$(2.5) \varphi(n) \le d(p_1 - 1) \cdots (p_{k-1} - 1)(p_{k+1} - 1)$$

and so by (b)

(2.6) 
$$\varphi(n) < (d+1)(p_1-1)\cdots(p_k-1).$$

To prove that n is in F we pick any m > n and we eventually show that

$$\varphi(m) > \varphi(n).$$

There is a unique integer  $t \ge 1$  such that

$$p_1 \cdots p_t \leq m < p_1 \cdots p_{t+1}$$
.

Then the number of distinct prime divisors  $\omega(m)$  of m satisfies  $\omega(m) \le t$ . We deduce that

$$\varphi(m)/m \geq \left(1-p_1^{-1}\right)\cdots\left(1-p_t^{-1}\right)$$

and so

$$\varphi(m) \geq (p_1 - 1) \cdots (p_t - 1).$$

If now  $t \ge k + 1$ , then

$$\varphi(m) \ge (p_1 - 1) \cdots (p_{k+1} - 1) \ge (d+1)(p_1 - 1) \cdots (p_k - 1)$$

by (a), and therefore (2.7) holds because of (2.6).

Hence we may assume  $t \le k$ . Thus  $\omega(m) \le k$ . If now  $\omega(m) \le k - 1$  then

$$\varphi(m)/m \geq (1-p_1^{-1})\cdots(1-p_{k-1}^{-1}).$$

But (2.4), (2.5) give

$$\varphi(n)/n \leq (1-p_1^{-1})\cdots(1-p_{k-1}^{-1})(1-p_{k+1}^{-1}),$$

so that

$$\varphi(m) > (m/n)\varphi(n) > \varphi(n)$$

and again (2.7) holds.

Thus we may henceforth assume that  $\omega(m) = k$ , so that

$$m = eq_1 \cdots q_k$$

for primes  $q_1, \ldots, q_k$  with  $q_1 < \cdots < q_k$  and an integer  $e \ge 1$  composed only of primes from  $q_1, \ldots, q_k$ . So

$$(2.8) q_1 \ge p_1, \dots, q_k \ge p_k$$

and

(2.9) 
$$\varphi(m) = e(q_1 - 1) \cdots (q_k - 1).$$

Suppose now that  $e \ge d + 1$ . Then (2.6) gives

$$\varphi(n) < e(p_1 - 1) \cdot \cdot \cdot (p_k - 1),$$

whence (2.7) follows from (2.8) and (2.9). Finally if  $e \le d$  then we write m > n in the form

$$q_1 \cdots q_{k-1}Y > p_1 \cdots p_{k-1}X$$

with  $Y = q_k$ ,  $X = dp_{k+1}/e$ . Using (2.8), we apply Lemma 1 with r = k - 1 to deduce that

$$(q_1-1)\cdots(q_{k-1}-1)(Y-1)>(p_1-1)\cdots(p_{k-1}-1)(X-1).$$

Multiplying by e and recalling (2.9) we get

$$\varphi(m) > (p_1 - 1) \cdots (p_{k-1} - 1)(dp_{k+l} - e),$$

and since  $e \le d$  this gives (2.7) by virtue of (2.5). This completes the proof of Theorem 1.

We pause here to note that Theorem 1 would become false if either of the strict inequalities (a) or (b) were relaxed. In fact if

$$d = p_{k+1} - 1$$

the number  $n = dp_1 \cdots p_{k-1} p_{k+1}$  is never in F if condition (b) is satisfied. For (b) implies in this case

$$p_{k+l} - 1 < (1 + p_{k+1}^{-1})(p_k - 1) < p_k$$

so l=0; also for  $k \ge 2$  the number  $p_{k+1}-1$  is divisible only by primes from  $p_1, \ldots, p_k$ . Thus

$$\varphi(n) = (p_1 - 1) \cdots (p_k - 1)(p_{k+1} - 1).$$

But now

$$m = p_1 \cdot \cdot \cdot p_k p_{k+1} > n$$

and  $\varphi(m) = \varphi(n)$ . Hence condition (a) is sharp.

Next, Schinzel's Hypothesis H (see for example [3] p. 2) applied to the polynomials

$$F_1(t) = 1 + dt, \qquad F_2(t) = 1 + (d+1)t$$

shows that for each  $d \ge 1$  there are infinitely many  $k \ge 2$  such that  $((d+1)p_k-1)/d$  is integral and prime (moreover for any fixed d such as d=1 we can find plenty of examples in practice). Denoting this prime by  $p_{k+l}$  we see that  $l \ge 0$  and  $d(p_{k+l}-1)=(d+1)(p_k-1)$ . Clearly also  $p_k$  does not divide d. But then the number  $n=dp_1\cdots p_{k-1}p_{k+1}$  cannot be in F if condition (a) is satisfied. For (a) implies that d is divisible only by primes from  $p_1, \ldots, p_{k-1}$ , so

$$\varphi(n) = d(p_1 - 1) \cdots (p_{k-1} - 1)(p_{k+l} - 1).$$

But now

$$m = (d+1) p_1 \cdots p_k > n$$

and

$$\varphi(m) \leq (d+1)(p_1-1)\cdots(p_k-1) = \varphi(n).$$

Hence condition (b) is also sharp.

3. Proof of Corollary. The idea of this can be explained very easily. We observe that the elements of F given by Theorem 1 form blocks that neatly fit together. For, putting d = 1, we see that the numbers

$$p_1 \cdots p_{k-1} p_{k+l} \qquad (l \ge 0)$$

lie in F as long as  $p_{k+l} - 1 < 2(p_k - 1)$ ; so this takes us from  $p_1 \cdots p_k$  to roughly  $2p_1 \cdots p_k$ . Then, putting d = 2, we see that the numbers

$$2p_1\cdots p_{k-1}p_{k+l} \qquad (l\geq 0)$$

lie in F as long as  $p_{k+l} - 1 < 3(p_k - 1)$ ; so these take us roughly up to  $3p_1 \cdots p_k$ . Then we put  $d = 3, 4, \ldots$  and so on, up to  $d = p_{k+1} - 2$ . By then the elements of F have reached roughly  $p_1 \cdots p_k p_{k+1}$ , and so we can begin again with d = 1.

The details are as follows. Let  $0 < \varepsilon < 1$ . We have to show that for every sufficiently large n in F there exists n' in F with

$$(3.1) n < n' \le (1 + \varepsilon)n.$$

Let k be the integer satisfying

$$(3.2) p_1 \cdots p_k \le n < p_1 \cdots p_{k+1}$$

Since n is large, k is also large, and in particular we may assume  $k \ge 2$  and

$$(3.3) p_k - 2 \ge 2/\varepsilon$$

as well as

$$(3.4) p_{k+l} \le (1 + \frac{1}{2}\varepsilon) p_{k+l-1}$$

for all  $l \ge 1$ .

Next define the integer m by

$$(3.5) mp_1 \cdots p_k \le n < (m+1) p_1 \cdots p_k,$$

so that

$$1 \le m < p_{k+1}.$$

Our construction of n' depends on the size of m, and we consider four cases in turn:

- (i)  $p_{k+1} 2 \le m < p_{k+1}$
- (ii)  $\varepsilon^{-1} \leq m < p_{k+1} 2$
- (iii)  $1 \leq m < \varepsilon^{-1}$ ;  $(1 + \varepsilon)n \geq (m + 1)p_1 \cdots p_k$
- (iv)  $1 \leq m < \varepsilon^{-1}$ ;  $(1 + \varepsilon)n < (m + 1)p_1 \cdots p_k$ .

In case (i) we choose  $n' = p_1 \cdots p_{k+1}$ . By Theorem 1 this lies in F, and n' > n from (3.5). Also

$$n'/n \le p_{k+1}/m \le p_{k+1}/(p_{k+1}-2) \le 1+\varepsilon$$

by (3.3). Thus (3.1) holds.

In cases (ii) and (iii) we choose  $n' = (m+1)p_1 \cdots p_k$ . In both cases we have  $m < p_{k+1} - 2$  by (3.3), and so by Theorem 1 with d = m + 1, l = 0 we see that n' lies in F. Again from (3.5) we have n' > n. And in case (ii)

$$n'/n \le (m+1)/m \le 1+\varepsilon$$

while in case (iii) this inequality is immediate. Thus (3.1) holds once more. Finally, in case (iv) let p' be the least prime satisfying

$$(3.6) p' > n/(mp_1 \cdots p_{k-1}).$$

By (3.5) we see that  $p' > p_k$ , and so  $p' = p_{k+l}$  for some  $l \ge 1$ . Also we have

$$(3.7) p_{k+l-1} \le n/(mp_1 \cdots p_{k-1}) < (m+1)p_k/(m(1+\varepsilon)).$$

Then from (3.4)

$$p_{k+l} \le \left( (m+1)/m \right) \left( p_k \left( 1 + \frac{1}{2}\varepsilon \right) / (1+\varepsilon) \right)$$

which does not exceed  $(m+1)(p_k-1)/m$ , by (3.3). It follows now from Theorem 1 with d=m that the number  $n'=mp_1\cdots p_{k-1}p_{k+l}$  lies in F. By (3.6) we have n'>n, and from (3.7)

$$n \geq mp_1 \cdots p_{k-1}p_{k+l-1}$$

which gives using (3.4)

$$n'/n \leq 1 + \frac{1}{2}\varepsilon$$
.

Thus (3.1) holds, and this completes the proof of the Corollary.

Let us note that standard results on gaps between primes enable the Corollary to be strengthened to

$$n'/n = 1 + O(\log^{-\delta} n)$$

for some  $\delta > 0$ . But the conditional results of the Appendix show that n, n' probably have a very large common factor, and in particular the relation

$$n'/n = 1 + O(n^{-\varepsilon})$$

is probably false for every  $\varepsilon > 0$ . In practice the convergence does seem rather slow; for example when n = 810810 we get  $n' \ge 870870$  so  $n'/n \ge 1.074...$ 

Finally we remark that the result of this Corollary is mentioned by Alaoglu and Erdös in [1] (p. 465). However, the simple proof they give of the corresponding property of highly abundant numbers (p. 463) does not immediately seem to generalize to sparsely totient numbers, because it could happen (and indeed probably will) that  $\varphi(n(p-1)/p) = \varphi(n)$ . Even so, it does lead to a quick proof of the corresponding property for the larger set  $\overline{F}$  of numbers n such that  $\varphi(m) \ge \varphi(n)$  whenever  $m \ge n$ .

## 4. Proof of Theorem 2. For positive integers h and k we write

$$f(k,h) = h([k/h] + 1)$$

for the unique integer x satisfying  $k < x \le k + h$  which is a multiple of h. Given n in F, our basic strategy is to replace a suitable divisor k of n by f(k, h) for some h, and thereby obtain the number

$$m = nf(k, h)/k > n.$$

Thus  $\varphi(m) > \varphi(n)$ . But if h is small and prime to n, for example, then m will have acquired all the prime factors of h in exchange for those of k. So on the other hand  $\varphi(m)/m$  might be expected to be quite small compared with  $\varphi(n)/n$ . We note that

$$1 < m/n = f(k, h)/k = 1 + (h/k)(1 - \{k/h\}) \le 1 + h/k$$

where  $\{x\} = x - [x]$  denotes the fractional part of x. Here the presence of the term  $\{k/h\}$  sometimes leads to interesting problems of diophantine approximation.

We shall need the following lemmas.

LEMMA 2. Suppose n is in F and

$$p_1 \cdots p_k \leq n < p_1 \cdots p_{k+1}$$

for some  $k \geq 1$ . Then  $\omega(n)$  is either k-1 or k.

*Proof.* It is clear from the upper bound for n that  $\omega(n) \le k$ . This proves the lemma if  $k \le 2$ , so we may assume  $k \ge 3$ , and, if possible,  $\omega(n) \le k - 2$ . Then

$$(4.1) \varphi(n)/n \ge (1-p_1^{-1})\cdots(1-p_{k-2}^{-1}).$$

Now put  $m = f(n, p_1 \cdots p_{k-1})$  so that

$$(4.2) 1 < m/n < 1 + p_1 \cdots p_{k-1}/n \le 1 + p_k^{-1}.$$

Moreover, since  $p_1 \cdots p_{k-1}$  divides m we have

$$\varphi(m)/m \le (1-p_1^{-1})\cdots(1-p_{k-1}^{-1}) \le (1-p_{k-2}^{-1})\varphi(n)/n$$

using (4.1). This together with (4.2) yields

$$\varphi(m) \leq (1 - p_{k-1}^{-1})(1 + p_k^{-1})\varphi(n) < \varphi(n),$$

and so contradicts the fact that n is in F, proving the lemma.

We remark that the conclusion of this lemma cannot be strengthened. For by Theorem 1 the number  $n = p_1 \cdots p_k$  lies in F for all k, and it can be shown that  $n = p_1 \cdots p_{k-2} p_k^2$  lies in F for infinitely many k.

LEMMA 3. For n in F we have

$$P(n) < (Q(n))^2.$$

*Proof.* Suppose not. Then P = P(n) and Q = Q(n) satisfy  $Q^2 \le P$ . Put m = nf(P, Q)/P, so that

$$(4.3) 1 < m/n \le 1 + Q/P \le 1 + Q^{-1}.$$

Since m has acquired the factor Q but possibly lost the factor P, we have

$$\varphi(m)/m \le (1 - Q^{-1})(1 - P^{-1})^{-1}\varphi(n)/n$$
  
 $\le (1 - Q^{-1})(1 - Q^{-2})^{-1}\varphi(n)/n.$ 

Hence (4.3) gives

$$\varphi(m) \leq (1 + Q^{-1})(1 - Q^{-1})(1 - Q^{-2})^{-1}\varphi(n) = \varphi(n),$$

again a contradiction. This proves the lemma.

LEMMA 4. For n > 2 in F we have

$$P'(n) < (\sqrt{2} + 1)Q(n).$$

*Proof.* Suppose not, and put  $\lambda = \sqrt{2} - 1$ . Then with P' = P'(n) we have

$$Q \leq \lambda P' < \lambda P$$
.

Put m = nf(P'P,Q)/P'P, so that

$$(4.4) 1 < m/n < 1 + Q/P'P < 1 + \lambda^2 Q^{-1}.$$

Also

(4.5) 
$$\varphi(m)/m \le (1 - Q^{-1})(1 - P'^{-1})^{-1}(1 - P^{-1})^{-1}\varphi(n)/n$$
$$< (1 - Q^{-1})(1 - \lambda Q^{-1})^{-2}\varphi(n)/n.$$

But because  $1 - \lambda^2 = 2\lambda$  we see that

$$(1-Q^{-1})(1+\lambda^2Q^{-1})<1-(1-\lambda^2)Q^{-1}=1-2\lambda Q^{-1}<(1-\lambda Q^{-1})^2,$$

and so (4.4), (4.5) lead to  $\varphi(m) < \varphi(n)$ , again a contradiction. This proves the lemma.

Now let us establish Theorem 2 by examining each of the limits in turn. First, to prove (a) and (c) we start by observing that

$$(4.6) P(n) > P'(n) \ge (1 + o(1)) \log n$$

as  $n \to \infty$  in F. For let k be the integer defined by

$$(4.7) p_1 \cdots p_k \le n < p_1 \cdots p_{k+1}.$$

By the Prime Number Theorem  $p_k = (1 + o(1)) \log n$ , and by Lemma 2 we see that  $h = \omega(n)$  satisfies  $h \ge k - 1$ . Hence

$$P(n) > P'(n) \ge p_{h-1} \ge p_{k-2} = (1 + o(1)) \log n.$$

Next we know from Theorem 1 that  $n = p_1 \cdots p_k$  is in F for all  $k \ge 2$ . Hence for these n we have

$$(4.8) P'(n) < P(n) = p_{k} = (1 + o(1)) \log n.$$

Comparing (4.6) and (4.8), we obtain the first limits in (a) and (c).

Now the second limit in (a) also follows from Theorem 1, which shows that  $n = p_1 \cdots p_{k-1} p_{k+1}$  is in F whenever  $k \ge 2$  and

$$(4.9) p_k \le p_{k+1} < 2(p_k - 1) + 1.$$

This *n* satisfies (4.7) and so  $p_k = (1 + o(1)) \log n$ . But also the largest *l* satisfying (4.9) is such that  $p_{k+l} = (2 + o(1)) \log n$ . Hence for these *n* we have

$$P(n) = (2 + o(1)) \log n.$$

This proves the second limit in (a).

Next, the first limit in (b) follows immediately from Lemma 4 and the first limit in (c). Also for n satisfying (4.7) we have  $Q(n) \le p_{k+1}$  and therefore

$$Q(n) \le (1 + o(1)) \log n$$

for any n whatsoever; and for the numbers  $n = p_1 \cdots p_k$  in F we see on the other hand that  $Q(n) = p_{k+1}$ . These together establish the second limit in (b).

Now the second limit in (c) is a consequence of Lemma 4 and the second limit in (b). Finally the limit in (d) follows from Lemma 3 together with the second limit in (b). This completes the proof of Theorem 2.

2. Proof of Theorem 3. We first record the following simple result about the repeated factors

$$R(n) = n \prod_{p|n} p^{-1}$$

of a sparsely totient number n.

LEMMA 5. Let n in F let r be any factor of R(n), and let q be any prime not dividing n. Then

$$r<\big(1-\big\{r/q\big\}\big)q^2.$$

*Proof.* Suppose not. We put m = nf(r, q)/r, so that

$$1 < m/n = 1 + (q/r)(1 - \{r/q\}) \le 1 + q^{-1}$$
.

Also since m has all the prime factors of n together with q, we have

$$\varphi(m)/m \leq (1-q^{-1})\varphi(n)/n,$$

and therefore

$$\varphi(m) \le (1 - q^{-1})(1 + q^{-1})\varphi(n) < \varphi(n)$$

a contradiction. This proves the lemma.

COROLLARY. For n in F we have

$$R(n) < (Q(n))^2.$$

Using this corollary, we see quite quickly that  $P^4 = (P(n))^4$  never divides n in F. For otherwise  $R = R(n) \ge P^3$  and we would deduce

$$(5.1) P3 < Q2$$

for Q = Q(n). Now Q can be at most the smallest prime exceeding P, and so Bertrand's Postulate implies  $Q \le 2P - 1$ . Hence by (5.1) we see that  $P^3 < (2P - 1)^2$ , which forces P = 2, Q = 3. But we have already noted that the only power of 2 in F is n = 2, and this is certainly not divisible by  $2^4$ .

To make further progress we have to take into account the curly brackets in Lemma 5. The solution of the corresponding diophantine approximation problem is given in the next lemma.

LEMMA 6. For all sufficiently large integers m there exists a prime q > m with

$$\{m^2/q\} > 1 - m^2/q^2.$$

*Proof.* Let  $\varepsilon = 1/40$ . It suffices to prove that there exists q with

(5.2) 
$$\{m^2/q\} > \varepsilon, \quad m < q < m + m^{2/3};$$

this is because when m is large

$$1 - m^2/q^2 = (q + m)(q - m)/q^2 < 2m^{2/3}/q < \varepsilon.$$

For any interval I = (x, x + y] we write  $\pi(I) = \pi(x + y) - \pi(x)$  for the number of primes in I. It is well-known ([5]) that for any  $\vartheta$  with  $7/12 < \vartheta \le 1$ , the number  $\pi(I)$  is asymptotic to  $y/\log x$  provided  $x^{\vartheta} \le y \le x$ ; hence there exists  $x_0 = x_0(\varepsilon)$  such that

$$(5.3) \pi(I) \ge \frac{1}{2} y / \log x$$

whenever  $x \ge x_0$  and  $x^{\vartheta} \le y \le x$ . Our proof actually requires  $\vartheta < 2/3$ . It is also known (see for example [4] p. 523) that  $\pi(I)$  is asymptotically at most  $2y/\log y$  (and indeed according to [6] the inequality  $\pi(I) \le 2y/\log y$  holds for all  $x \ge 1, y > 1$ ); at any rate

$$(5.4) \pi(I) \le 3y/\log y$$

for all  $x \ge 1$  and all sufficiently large y.

We take  $\vartheta = 5/8$  and  $R = [m^{1/4}]$ . The interval

$$J = (m, m + (R + 1)^{1/2} m^{1/2})$$

has length exceeding  $m^{5/8}$ , so that (5.3) gives

(5.5) 
$$\pi(J) \ge \frac{1}{2} m^{5/8} \log m.$$

We next claim that at least one of the intervals

$$I_r = (m + (r + 2\varepsilon)^{1/2} m^{1/2}, m + (r + 1)^{1/2} m^{1/2})$$
  $(0 \le r \le R)$ 

contains a prime q. For if this were not so, then all the primes in J would lie in the complementary intervals

$$J_r = (m + r^{1/2}m^{1/2}, m + (r + 2\varepsilon)^{1/2}m^{1/2}), \qquad (0 \le r \le R)$$

and we would then have

(5.6) 
$$\pi(J) \leq \sum_{r=0}^{R} \pi(J_r).$$

But the length  $L_r$  of  $J_r$  satisfies

$$m^{1/3} < \frac{1}{2} \varepsilon m^{1/2} r^{-1/2} < L_r < \varepsilon m^{1/2} r^{-1/2}$$
  $(1 \le r \le R)$ 

and

$$m^{1/3} < L_0 = (2\varepsilon)^{1/2} m^{1/2} < m^{1/2}$$
.

Hence by (5.4)

$$\pi(J_r) \le 9\varepsilon m^{1/2} r^{-1/2} / \log m \qquad (1 \le r \le R)$$

and

$$\pi(J_0) \le 9m^{1/2}/\log m < m^{1/2}.$$

Therefore

$$\sum_{r=0}^{R} \pi(J_r) \le m^{1/2} + 9\varepsilon m^{1/2} (\log m)^{-1} \sum_{r=1}^{R} r^{-1/2}.$$

Since  $\sum_{r=1}^{R} r^{-1/2} < 2R^{1/2}$  we conclude that

$$\sum_{r=0}^{R} \pi(J_r) < 20 \varepsilon m^{5/8} / \log m = \frac{1}{2} m^{5/8} / \log m.$$

From (5.5) and (5.6) we see that this is impossible. Hence indeed there exists r with  $0 \le r \le R$  such that  $I_r$  contains a prime q, and we can write q = m + d with d > 0 and

$$(5.7) (r+2\varepsilon)m < d^2 \le (r+1)m.$$

Thus the integer  $N = d^2 - rq$  satisfies

$$(5.8) N < q.$$

On the other hand

$$N > (r + 2\varepsilon)m - rq = 2\varepsilon q - (r + 2\varepsilon)d$$

and since

$$(r+2\varepsilon)d < (R+1)^{3/2}m^{1/2} < 3m^{7/8} < 3q^{7/8}$$

we see that

$$(5.9) N > \varepsilon q.$$

So by (5.8) and (5.9) we have

$$\left\{m^2/q\right\} = \left\{d^2/q\right\} = \left\{N/q\right\} > \varepsilon$$

as required by (5.2). Also d > 0 and (5.7) gives  $d < 2m^{5/8}$ , so the other inequalities of (5.2) for q = m + d are obvious. This proves the lemma.

Now Theorem 3 is immediate. Suppose  $P^3 = (P(n))^3$  divides n for some sufficiently large n in F. By Lemma 6 with m = P, there exists a prime q > P with

$${P^2/q} > 1 - P^2/q^2$$
.

Since q > P, the prime q does not divide n; on the other hand  $r = P^2$  does divide R(n), so Lemma 5 gives the contradictory

$$P^2 < (1 - \{P^2/q\})q^2$$
.

This establishes Theorem 3.

**6.** Proof of Theorem 4. For any positive integers n, k we may define  $Q_k(n)$  as the kth smallest prime not dividing n. If further  $k \le \omega(n)$  we may define  $P_k(n)$  as the kth greatest prime factor of n. We shall need to consider the equation

(6.1) 
$$x^k + kx = k - 1;$$

it is easily seen that this has a unique positive root  $\lambda_k$  satisfying

$$(6.2) 1 - 2/k < \lambda_k < 1.$$

We have already noted that  $\omega(n) \ge 2$  for all n > 2 in F.

LEMMA 7. For n > 2 in F and  $2 \le k \le \omega(n)$  we have

$$Q_{k-1}(n) > \lambda_k (P_k(n) - 1).$$

*Proof*. We write  $P_i = P_i(n)$ ,  $Q_i = Q_i(n)$  for  $1 \le i \le k$ . Put  $r = P_1 \cdots P_k$ ,  $s = Q_1 \cdots Q_{k-1}$ .

and m = nf(r, s)/r, so that

$$1 < m/n \le 1 + s/r \le 1 + Q_{k-1}^{k-1}/P_k^k$$
.

We also have

$$\varphi(m)/m \le (1 - Q_1^{-1}) \cdots (1 - Q_{k-1}^{-1})(1 - P_1^{-1})^{-1}$$
$$\cdots (1 - P_k^{-1})^{-1} \varphi(n)/n$$

which does not exceed

$$(1-Q_{k-1}^{-1})^{k-1}(1-P_k^{-1})^{-k}\varphi(n)/n.$$

Using  $\varphi(m) > \varphi(n)$  we deduce that

$$(6.3) (1 + Q_{k-1}^{k-1}/P_k^k)(1 - Q_{k-1}^{-1})^{k-1}(1 - P_k^{-1})^{-k} > 1.$$

We now note the inequalities

$$1 + t < e^t$$
,  $1 - t < e^{-t}$   $(t \ge 0)$ 

and

$$(1-t)^{-1} < e^{t/(1-t)}$$
  $(0 \le t < 1).$ 

These transform (6.3) into

$$Q_{k-1}^k/P_k^k + kQ_{k-1}/(P_k-1) > k-1.$$

Thus  $x = Q_{k-1}/(P_k - 1)$  satisfies

$$x^k + kx > k - 1,$$

which implies from (6.1) that  $x > \lambda_k$ . This proves the lemma.

COROLLARY. For n in F and fixed  $k \ge 2$  we have

(a) 
$$\liminf_{n \to \infty} \frac{P_k(n)}{\log n} = 1$$
,  $\limsup_{n \to \infty} \frac{P_k(n)}{\log n} \le \lambda_k^{-1}$ 

(b) 
$$\liminf_{n\to\infty} \frac{Q_{k-1}(n)}{\log n} \ge \lambda_k, \qquad \limsup_{n\to\infty} \frac{Q_{k-1}(n)}{\log n} = 1.$$

*Proof.* The equalities in (a) and (b) are established exactly as in the proof of Theorem 2. We omit the details. The inequalities then both follow immediately from Lemma 7.

We now prove Theorem 4. It clearly suffices to show that for all x sufficiently large, the number  $F_1(x) = F(x) - F(x/2)$  of elements n of F with  $x/2 < n \le x$  satisfies

$$\log F_1(x) \ll (\log x)^{1/2}.$$

We can suppose that

$$k = \left[ (\log x)^{1/2} / \log \log x \right] \ge 2.$$

Since by Theorem 2 we have  $Q(n) \ge \frac{1}{3} \log n$ , we deduce  $\omega(n) \ge \frac{1}{4} \log n / \log \log n$ , and so  $2 \le k \le \omega(n)$  for all n in F with n > x/2. We now note that each n in F with  $x/2 < n \le x$  is specified uniquely by giving successively the following pieces of information:

- (a) R = R(n)
- (b)  $P_1, ..., P_k$
- (c)  $Q_1, \ldots, Q_{k-1}$
- (d) the prime factors of n in the interval

$$I = (\lambda_k(P_k - 1), P_k).$$

For n/R is squarefree, and it has no prime factors to the right of this interval except  $P_1, \ldots, P_k$ . Further by Lemma 7 it has every prime to the left of this interval as a factor except those of  $Q_1, \ldots, Q_{k-1}$  which do not exceed  $\lambda_k(P_k-1)$ .

Now by Theorem 2 and the Corollary to Lemma 5 there are at most  $2\log^2 x$  possibilities for R in (a). Since  $P_k < P_{k-1} < \cdots < P_1$ , Theorem 2 also shows that there are at most  $(2\log^2 x)^k$  possibilities for  $P_1, \ldots, P_k$  in (b). Next, writing  $P_1 = p_r$  for some  $r \ge 1$ , we see that  $r \le \log^2 x$ , and since

$$Q_1 < \cdots < Q_{k-1} \le p_{r+k-1} < (r+k-1)^2$$
,

we find that there at at most  $(2\log^2 x)^{2(k-1)}$  possibilities for  $Q_1, \ldots, Q_{k-1}$  in (c). Finally, once  $P_k$  has been specified in (b), the length y of the interval I is

$$y = (1 - \lambda_k)P_k + \lambda_k \le 2P_k/k + 1$$

by (6.2). Since  $k \ge 2$ , we have  $P_k \le 3\log x$  from Theorem 2, and using the definition of k we find that

$$y \le 7(\log x)^{1/2} \log \log x.$$

Hence from the inequality (5.4) of the preceding section, the number N of primes in I satisfies

$$N \le 43(\log x)^{1/2}.$$

Since the number of possibilities for (d) is at most  $2^N$ , we conclude from all the estimates above that

$$F_1(x) \le 2\log^2 x (2\log^2 x)^k (2\log^2 x)^{2(k-1)} 2^N \le (\log^3 x)^{3k} 2^N$$

which does not exceed  $\exp(40(\log x)^{1/2})$ . This leads at once to the desired estimate for F(x), and so completes the proof of Theorem 4.

**Appendix.** We discuss here some improvements on our results that can be obtained using deeper methods. The most interesting of these concerns possibly large values of the greatest prime factor P(n) of n. In Theorem 2 we saw that  $P(n) < (1 + \varepsilon) \log^2 n$  for all sufficiently large n in F; and indeed, it appears from the table that occasionally P(n) can be of this order of magnitude. An extreme example occurs for

$$n = 5735730 = 2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdot 13 \cdot 191$$

with P(n) = 191, so

$$P(n)/\log n = 12.273...$$
,  $P(n)/\log^2 n = .78866...$ 

Nevertheless Glyn Harman has substantially improved our upper bound for P(n). He first uses the latest techniques from the theory of exponential sums to prove the following result on diophantine approximation with primes.

THEOREM (HARMAN). There is an absolute constant  $\delta > 0$  with the following property. For any  $N \geq 1$  and  $\varepsilon$ , x with

$$N^{-\delta} < \varepsilon < 1 - N^{-\delta}, \qquad N^{2-\delta} < x < N^{2+\delta}$$

we have

$$\sum_{\substack{\{x/p\}>\varepsilon\\N< p\leq 2N}} 1 = (1-\varepsilon)(\pi(2N)-\pi(N))(1+O(N^{-\delta}))$$

He then observes (compare the proof of Lemma 3) that this leads to

$$(A.1) P(n) \ll \log^{2-\delta} n$$

for all *n* in *F*. In particular, he can prove (A.1) for any  $\delta < 1/10$ .

On the other hand, one can consider products of two primes. The following hypothesis seems plausible.

HYPOTHESIS. For any fixed  $\alpha$ ,  $\beta$  with  $0 < \alpha < \beta$ , there is a function  $\psi(x) = o(x^{1/2})$  such that for every  $x \ge 1$  we can find primes p, q with  $\alpha p < q < \beta p$  and

$$x < pq < x + \psi(x)$$
.

Assuming this hypothesis it can be proved that for n in F we have

$$\limsup_{n\to\infty} \frac{P(n)}{\log n} \le 2,$$

which is best possible in view of Theorem 2. Likewise the Hypothesis implies that for n in F

$$\liminf_{n\to\infty}\frac{Q(n)}{\log n}\geq 1, \qquad \limsup_{n\to\infty}\frac{P'(n)}{\log n}\leq 1,$$

both of which are again best possible by Theorem 2.

We can even refine some of these results to take account of the repeated factors R(n) of n. From the Hypothesis it follows that for n in F

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

and, for fixed  $d \ge 1$ ,

(A.3) 
$$\limsup_{\substack{n \to \infty \\ d \mid R(n)}} \frac{P(n)}{\log n} = 1 + d^{-1}.$$

All this delineates the structure of sparsely totient numbers rather clearly. For any  $\varepsilon > 0$  and sufficiently large n in F, the number n, apart from a repeated factor  $d \le (1 + \varepsilon) \log n$ , is squarefree and divisible by all primes up to  $(1 - \varepsilon) \log n$ . Moreover, it is divisible by no prime larger than  $(1 + \varepsilon) \log n$  except possibly its largest prime factor p. Finally, for fixed d at any rate, the prime p lies between  $(1 - \varepsilon) \log n$  and  $(1 + d^{-1} + \varepsilon) \log n$ . Everything here fits neatly in with the explicit constructions used in Theorem 1, the relations (A.2) and (A.3) corresponding to the inequalities (a) and (b) respectively.

But we should emphasize that all these conclusions depend on the above Hypothesis, which, if true, unfortunately seems well beyond the reach of present techniques in analytic number theory.

# Sparsely totient numbers not exceeding 10<sup>6</sup>, with factorizations

	·	
2 2 6 2.3 12 2:3 18 2.3² 30 2.3.5 42 2.3.7 60 2:3.5 66 2.3.11 90 2.3:5 120 2:3.5 126 2:3.7 150 2.3.5² 210 2.3.5.7 240 2:3.5 270 2:3.5 330 2.3.5.11 420 2:3.5.7 462 2.3.7.11 510 2.3.5.17 630 2.3.5.17 630 2.3.5.17 630 2.3.5.7 870 2.3.5.23 840 2:3.5.7 870 2.3.5.29 1050 2.3.5:7	11550 2.3.5.7.11 11970 2.3.5.7.19 12180 2.3.5.7.29 12390 2.3.5.7.59 13860 2.3.5.7.17 14480 2.3.5.7.17 14490 2.3.5.7.23 16170 2.3.5.7.11 16380 2.3.5.7.11 16380 2.3.5.7.11 19110 2.3.5.7.13 19320 2.3.5.7.13 19320 2.3.5.7.13 23100 2.3.5.7.13 23100 2.3.5.7.13 23100 2.3.5.7.11 21840 2.3.5.7.13 23100 2.3.5.7.11 24570 2.3.5.7.11 3030 2.3.5.7.11 32340 2.3.5.7.11 32340 2.3.5.7.11 32340 2.3.5.7.11 32360 2.3.5.7.11 32360 2.3.5.7.11 32370 2.3.5.7.11 32380 2.3.5.7.11 32390 2.3.5.7.11 32390 2.3.5.7.11 35490 2.3.5.7.11.1	278460 2:3:5.7.13.17 300300 2:3.5:7.11.13 307230 2:3.5.7:11.19 314160 2:3.5.7.11.17 330330 2.3.5.7.11:13
1260 2:3:5.7 1320 2:3.5.11 1470 2.3.5.7° 1680 2:3.5.7 1890 2.3:5.7 2310 2.3.5.7.11 2730 2.3.5.7.13 2940 2:3.5.7° 3150 2.3:5:7 3570 2.3:5:7 3570 2.3:5.7.17 3990 2.3.5.7.19 4620 2:3.5.7.11 4830 2.3.5.7.13 5610 2.3.5.7.13 5610 2.3.5.7.13 5610 2.3.5.7.13 5610 2.3:5.7.13 5610 2.3:5.7.17 7350 2.3:5.7.17 7350 2.3:5.7.17 7350 2.3:5.7.17 7350 2.3:5.7.17 8190 2.3:5.7.17 8190 2.3:5.7.11 9660 2:3.5.7.13 9240 2:3.5.7.11	46410 2.3.5.7.13.1 48510 2.3.5.7.13.1 51870 2.3.5.7.11.1 53130 2.3.5.7.11.2 60060 2:3.5.7.11.3 66990 2.3.5.7.11.3 67830 2.3.5.7.11.3 71610 2.3.5.7.11.3 78540 2:3.5.7.11.3 78540 2:3.5.7.11.3 80850 2.3.5.7.11.3 80850 2.3.5.7.11.3 90900 2.3.5.7.11.3 90090 2.3.5.7.11.3 94710 2.3.5.7.11.4 99330 2.3.5.7.11.4 103740 2:3.5.7.11.4 103740 2:3.5.7.11.4 106260 2:3.5.7.11.4 108570 2.3.5.7.11.4 122430 2.3.5.7.11.5 125580 2:3.5.7.11.5 125580 2:3.5.7.11.1	363090 2.3.5.7:13.19 371280 2:3.5.7:13.17 371910 2.3.5.7:11.13 390390 2:3.5.7:11.13 392700 2:3.5.7:11.17 29 395010 2.3.5.7:11.19 420420 2:3.5.7:11.13 431970 2.3.5.7.11.13 510510 2.3.5.7.11.13 510510 2.3.5.7.11.13 510510 2.3.5.7.11.13 13 603630 2.3.5.7.11.13 13 603630 2.3.5.7.11.13 14 690690 2.3.5.7.11.13 15 75070 2.3.5.7.11.13 17 630630 2.3.5.7.11.13 18 690690 2.3.5.7.11.13 19 750750 2.3.5.7.11.13 20 750750 2.3.5.7.11.13 21 780780 2.3.5.7.11.13 22 780780 2.3.5.7.11.13 23 780780 2.3.5.7.11.13 24 795400 2.3.5.7.11.13 25 881790 2.3.5.7.11.13

#### REFERENCES

- [1] L. Alaoglu and P. Erdös, On highly composite and similar numbers, Trans. Amer. Math. Soc., 56 (1944), 448-469.
- [2] P. Erdös and J.-L. Nicolas, Sur la fonction: nombre de facteurs premiers de N, L'Enseignement Math., 27 (1981), 3-27.
- [3] H. Halberstam and H.-E. Richert, Sieve Methods, Academic Press, London 1974.

- [4] L.-K. Hua, *Introduction to Number Theory*, Springer-Verlag, Berlin-Heidelberg-New York 1982.
- [5] M. N. Huxley, On the differences between consecutive primes, Invent. Math., 15 (1972), 164–170.
- [6] H. Montgomery and R. C. Vaughan, The large sieve, Mathematika, 20 (1973), 119–134.
- [7] S. Ramanujan, *Collected papers* (eds G. H. Hardy, P. V. Seshu Aiyar, B. M. Wilson), Chelsea Publishing Company, New York 1962 (pp. 78–128).

Received February 13, 1984.

THE UNIVERSITY OF MICHIGAN ANN ARBOR, MI 48109

AND

LOUGHBOROUGH UNIVERSITY OF TECHNOLOGY LOUGHBOROUGH LEICESTERSHIRE, U. K.