JOURNAL OF NUMBER THEORY 43, 220–227 (1993)

# Primes in a Sparse Sequence

K. SOUNDARARAJAN

*Department of Mathematics, University of Michigan, Ann Arbor, Michigan 48109*

*Communicated by H. Zassenhaus*

Received July 6, 1990; revised September 3, 1991

In this paper, we consider the sequence $1^1$, $+1^1+2^2$, $1^1+2^2+3^3$, ... and prove some of its congruence properties. Surprisingly, this sequence is uniformly distributed in the residue classes (mod $m$) where $m \not\equiv 0 \pmod 4$. Using these results and Selberg's sieve, we obtain an upper bound for the number of primes in the sequence which are $\leqslant x$. © 1993 Academic Press, Inc.

## 1. INTRODUCTION

In [1], Hampel studied the sequence $n^n$ and proved some of its congruence properties. In this paper, we study the related sequence $1^1, 1^1+2^2, 1^1+2^2+3^3, \ldots$ with particular reference to the number of primes in this sequence.

The distribution of primes in some other sparse sequences has been dealt with in Chapter 7 of [2].

Our main result is

THEOREM. *Let $a_1 = 1^1$ and $a_n = a_{n-1} + n^n$ for $n \geqslant 2$. Let*

$$A = \{a_n : n \in \mathbb{N}\} \qquad \textit{and} \qquad \pi_A(x) = \sum_{\substack{p \leqslant x \\ p \in A}} 1.$$

*Then, we have*

$$\pi_A(x) \ll \frac{\log x}{(\log \log x)^2}.$$

Our proof of the theorem is based on an application of Selberg's upper bound sieve (see, for example, [3]). For this, it suffices to estimate

$$\sum_{\substack{n \in A \\ n \leqslant x \\ n \equiv 0 \ (\mathrm{mod}\ d)}} 1$$

for square free $d$.



0022-314X/93 $5.00
Copyright © 1993 by Academic Press, Inc.
All rights of reproduction in any form reserved.

Our methods are, in fact, capable of yielding estimates for

$$\sum_{\substack{n \in A \\ n \leqslant x \\ n \equiv a \pmod{d}}} 1$$

for all $a$ and $d$.

## 2. Periodicities of $n^n$ and $A$

The following lemma was proved in a slightly more general form by Hampel [1].

Lemma 1. *The sequence $n^n$ is ultimately periodic* $(\bmod\, m)$ *for all $m \geqslant 2$. Further if $g(m)$ denotes the fundamental period of the sequence* $(\bmod\, m)$, *then we have*

$$g(m) = \begin{cases} \text{l.c.m.}\{p_1^{\alpha_1}(p_1-1), ..., p_s^{\alpha_s}(p_s-1)\} \\ \qquad \textit{if} \quad m = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_s^{\alpha_s} \textit{ is odd} \\ \text{l.c.m.}\{2, g(m/2)\} \qquad \textit{if} \quad m \equiv 2 \pmod 4)\}. \end{cases}$$

It is quite easy to see that the phrase "ultimate periodicity of $n^n$ (mod m)" may be replaced by the congruence

$$l^l \equiv (l+g(m))^{l+g(m)} \pmod m \qquad \text{for all} \quad l > m.$$

We now consider the sequence $A$ to the modulus $p^\alpha$, where $p$ is an odd prime and $\alpha \in \mathbb{N}$. We begin with

Lemma 2. *Let $p$ be an odd prime, and $\alpha \geqslant 1$ be a natural number. Then, we have*

$$a_{2p^\alpha(p-1)} - a_{p^\alpha(p-1)} \equiv p^{\alpha-1}(p-1) \qquad (\bmod\, p^\alpha).$$

*Proof.* The proof proceeds by induction on $\alpha$. Now,

$$a_{2p(p-1)} - a_{p(p-1)} = \sum_{p(p-1) < r \leqslant 2p(p-1)} r^r.$$

We write $r = q(p-1)+s$, where $1 \leqslant s \leqslant (p-1)$ and $p \leqslant q \leqslant 2p-1$. Then, clearly

$$a_{2p(p-1)} - a_{p(p-1)} = \sum_{p(p-1) < r \leqslant 2p(p-1)} r^r \equiv \sum_{\substack{p \leqslant q \leqslant 2p-1 \\ 1 \leqslant s \leqslant p-1}} (s-q)^s$$

$$\equiv \sum_{0 \leqslant t \leqslant p-1} \sum_{1 \leqslant s \leqslant p-1} t^s$$

$$\equiv 0 + (p-1) + \sum_{2 \leqslant t \leqslant p-1} \sum_{1 \leqslant s \leqslant p-1} t^s$$

$$\equiv (p-1) + \sum_{2 \leqslant t \leqslant p-1} \frac{t(t^{p-1}-1)}{(t-1)} \equiv (p-1) \pmod p$$

by Fermat's little theorem. Hence the lemma is true for the case $\alpha = 1$.

Suppose it is true for the case $\alpha = k \geqslant 1$. Then,

$$a_{2p^{k+1}(p-1)} - a_{p^{k+1}(p-1)} = \sum_{p^{k+1}(p-1) < r \leqslant 2p^{k+1}(p-1)} r^r.$$

As before, we write $r = qp^k(p-1) + s$ with $p^k(p-1) + 1 \leqslant s \leqslant 2p^k(p-1)$ and $2p-2 \geqslant q \geqslant p-1$. Then, clearly

$$a_{2p^{k+1}(p-1)} - a_{p^{k+1}(p-1)} \equiv \sum_{\substack{p-1 \leqslant q \leqslant 2p-2 \\ p^k(p-1)+1 \leqslant s \leqslant 2p^k(p-1)}} (qp^k(p-1)+s)^s$$

$$\equiv \sum_{\substack{p-1 \leqslant q \leqslant 2p-2 \\ p^k(p-1)+1 \leqslant s \leqslant 2p^k(p-1)}} (s^s + s^s \cdot qp^k(p-1))$$

(by the binomial theorem)

$$\equiv \sum_{p^k(p-1)+1 \leqslant s \leqslant 2p^k(p-1)} s^s \left(p + \frac{p(p-1)}{2} p^k(p-1)\right)$$

$$\equiv p \sum_{p^k(p-1)+1 \leqslant s \leqslant 2p^k(p-1)} s^s$$

$$\equiv p^k(p-1) \pmod{p^{k+1}}$$

by the induction hypothesis.

This completes the proof.

An easy extension of Lemma 2 is

LEMMA 3. *Let $l$ be any natural number. Then*

$$a_{l+2p^\alpha(p-1)} - a_{l+p^\alpha(p-1)} \equiv a_{2p^\alpha(p-1)} - a_{p^\alpha(p-1)}$$

$$\equiv p^{\alpha-1}(p-1) \mod (p^\alpha),$$

*where, as before, $p$ is an odd prime and $\alpha \geqslant 1$ is an integer.*

Our next lemma determines the fundamental period of $A \pmod{p^\alpha}$.

LEMMA 4. *Let $p$ be an odd prime and $\alpha \geqslant 1$ be an integer. Then, for all $l > p^\alpha(p-1)$, we have*

$$a_{l+p^{\alpha+1}(p-1)} \equiv a_l \pmod{p^\alpha}.$$

*Further $p^{\alpha+1}(p-1)$ is the fundamental period.*

*Proof.* It is easily seen, from Lemma 3, that

$$a_{l+p^{\alpha+1}(p-1)} \equiv a_l \pmod{p^\alpha}.$$

Let $t$ be the fundamental period. Since $t$ is a period, clearly,

$$a_{l+t+1} - a_{l+t} \equiv a_{l+1} - a_l \pmod{p^\alpha} \qquad \text{for all} \quad l > p^\alpha(p-1),$$

and so

$$(l+t+1)^{l+t+1} \equiv (l+1)^{l+1} \pmod{p^\alpha}.$$

Hence by Lemma 1, $p^\alpha(p-1) \mid t$ (as $g(p^\alpha) = p^\alpha(p-1)$). Hence either $t = p^\alpha(p-1)$ or $t = p^{\alpha+1}(p-1)$. However,

$$a_{l+p^\alpha(p-1)} - a_l \equiv p^{\alpha-1}(p-1) \not\equiv 0 \pmod{p^\alpha},$$

where $l > p^\alpha(p-1)$, by Lemma 3. Hence, $p^{a+1}(p-1)$ is the fundamental period which proves the result.

It is not hard to see that, if $\alpha \leqslant p$, then the sequence is periodic mod $p^\alpha$, from the beginning.

From Lemma 4, we deduce

LEMMA 5. *Let $m$ be an integer $\geqslant 1$, which is not divisible by 4. Then, the sequence $A$ is ultimately periodic* (mod $m$). *Further, if $f(m)$ denotes the fundamental period of $A$* (mod $m$), *then we have*

$$f(m) = \begin{cases} \text{l.c.m.}\{p_1^{\alpha_1+1}(p_1-1), ..., p_s^{\alpha_s+1}(p_s-1)\} \\ \qquad \textit{if} \quad m = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_s^{\alpha_s} \textit{ is odd} \\ \text{l.c.m.}\{4, f(m/2)\} \qquad \textit{if} \quad m \equiv 2 \pmod 4. \end{cases}$$

*Proof.* The case in which $m$ is odd is easily seen by Lemma 4. Further, as $A$ is periodic (mod 2) with fundamental period 4, we get the result for $m \equiv 2 \pmod 4$.

We note that if $\alpha_i \leqslant p_i$ for all $s \geqslant i \geqslant 1$, then the sequence is periodic (mod $m$) from the beginning.

## 3. Uniform Distribution

In this section we show that the sequence $A$ is uniformly distributed in the residue classes mod $m$ for all $m \not\equiv 0 \pmod 4$. More precisely, we show

LEMMA 6. *Let* $m = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_s^{\alpha_s}$ *be an odd number. Then among the numbers* $a_{g(m)+1}, a_{g(m)+2}, \ldots, a_{g(m)+f(m)}$ *exactly* $f(m)/m$ *numbers belong to each residue class* mod $m$.

*Proof.* The lemma is easily verified for the case $m = 3$. Suppose the lemma is true for all odd numbers $t$ satisfying $m > t \geqslant 1$. We now show it for $m$. Let

$$a_{2g(m)+l} - a_{g(m)+l} \equiv a_{2g(m)} - a_{g(m)} \equiv UV \qquad (\text{mod } m),$$

where $1 \leqslant VU \leqslant m$ and $(V, m) = 1$.

Let $2g(m) \geqslant i \geqslant g(m)+1$ be an integer. Let $r$ run over all integers in the range $0 \leqslant r \leqslant (m/(m, U)) - 1$.

Then clearly

$$a_{i+rg(m)} \equiv a_i + rUV \equiv a_i + (m, U)(rUV/(m, U)) \equiv a_i + s(m, U) \qquad (\text{mod } m),$$

where $s$ runs over all integers in the range $0 \leqslant s \leqslant (m/(m, U)) - 1$. (This is because $(U/(m, U), m) = 1$ and $(V, m) = 1$ and so $(UV/(m, U), m) = 1$.) Let there be $\lambda_b$ solutions to the congruence $a_x \equiv b \ (\text{mod}\,(m, U))$ with $g(m)+1 \leqslant x \leqslant 2g(m)$. Then by the previous remark there are $\lambda_b$ solutions to each of the congruences $a_x \equiv b + s(m, U) \pmod m$, where $0 \leqslant s \leqslant (m/(m, U)) - 1$ with $g(m)+1 \leqslant x \leqslant g(m) + mg(m)/(m, U)$.

Thus if we show that the numbers $a_{g(m)+1}, \ldots, a_{2g(m)}$ are equally distributed in the residue classes (mod $(m, U)$), it would follow that the numbers $a_{g(m)+1}, \ldots, a_{g(m)+mg(m)/(m, U)}$ are equally distributed in the residue classes (mod $m$).

Now as $a_{i+mg(m)/(m, U)} \equiv a_i + mUV/(m, U) \equiv a_i \pmod m$ for all $i \geqslant g(m)$ it follows that $f(m) \mid mg(m)/(m, U)$. Also, as

$$a_{i+\mu g(m)} \not\equiv a_i \ (\text{mod } m) \qquad \text{for any} \quad 0 < \mu < m/(m, U),$$

it follows that $f(m) = mg(m)/(m, U)$. This together with the earlier remark shows that (to prove the lemma) it suffices to prove that the numbers $a_{g(m)+1}, \ldots, a_{2g(m)}$ are equally distributed in the residue classes (mod$(m, U)$).

As $g((m, U)) \mid g(m)$, it suffices to show that among the numbers $a_{g((m,U))+1}, \ldots, a_{g((m,U))+g(m)}$ exactly $g(m)/(m, U)$ numbers belong to each residue class mod $(m, U)$. Now as $f(m) \neq g(m)$ clearly $(m, U) \neq m$. Hence by the induction hypothesis it suffices to show $f((m, U)) \mid g(m)$.

Let $p_i^{\beta_i} \parallel (m, U)$, where $s \geqslant i \geqslant 1$. Then $p_i^{\beta_i + 1} \mid g(m)$. (For, suppose $p_i^{\gamma_i} \parallel g(m)$ with $1 \leqslant \gamma_i < \beta_i + 1$. Also, clearly $p_i - 1 \mid g(m)$. Hence

$$\begin{aligned} a_{2g(m)} - a_{g(m)} &= a_{2rp_i^{\gamma_i}(p_i - 1)} - a_{rp_i^{\gamma_i}(p_i - 1)} \\ &\equiv rp_i^{\gamma_i - 1}(p_i - 1) \pmod{p_i^{\gamma_i}}, \qquad \text{where} \quad (r, p_i) = 1. \end{aligned}$$

Hence $p_i^{\gamma_i} X(m, U)$, which is false.) Hence $f((m, U)) \mid g(m)$, which, as noted before, yields the result.

We observe that if $\alpha_i \leqslant p_i$ for all $s \geqslant i \geqslant 1$ then the result is true from the beginning. We next state

LEMMA 7. *Let $m = 2p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_s^{\alpha_s} \equiv 2 \pmod 4$. Then among the numbers $a_{g(m)+1}, a_{g(m)+2}, \ldots, a_{g(m)+f(m)}$ exactly $f(m)/m$ numbers lie in each residue class* mod $m$. *Further if $\alpha_i \leqslant p_i$ for all $s \geqslant i \geqslant 1$, the result is true from the beginning.*

The proof of this runs along the same lines as that of Lemma 6 and is accordingly omitted.

## 4. PROOF OF THE MAIN THEOREM

Let $a$ and $b$ be any two positive integers such that $b \not\equiv 0 \pmod 4$. Let $N$ be a natural number. Then thanks to Lemmas 6 and 7, we have

$$\sum_{\substack{n \in A \\ n \leqslant a_N \\ n \equiv a \pmod b}} 1 = \frac{N}{b} + O\left(\frac{f(b)}{b}\right) = \frac{N}{b} + O(b^2).$$

In particular this result holds if $a = 0$ and $b = d$ is square free.

We now state a well-known version of Selberg's sieve.

LEMMA 8. *Let $B$ be a finite set of integers. Let $\wp$ be a set of primes and*

$$p(z) = \prod_{\substack{p < z \\ p \in \wp}} p,$$

*where $z \geqslant 2$ is a real number.* Let $B_d = \{b : b \in B$ and $b \equiv 0 \pmod d\}$. *Let $X$ be an approximation to $|B| = |B_1|$ and $w$ be a multiplicative function such that $w(p)X/p$ is an approximation to $|B_p|$. Let*

$$R_d = |B_d| - \frac{w(d)}{d} X.$$

*Suppose* $0 \leqslant w(p)/p \leqslant 1 - 1/A_1$ *for some constant* $A_1 \geqslant 1$, $w(p) = 0$ *for primes* $p$ *not in* $\wp$. *If*

$$g(d) = \prod_{p \mid d} \frac{w(p)}{p - w(p)}, \qquad G(z) = \sum_{d < z} \mu^2(d)\, g(d),$$

*then we have*

$$\begin{aligned} S(B, \wp, z) &= |\{b : b \in B, (b, p(z)) = 1\}| \\ &\leqslant \frac{X}{G(z)} + \sum_{\substack{d_1, d_2 \leqslant z \\ d_1 \mid P(z) \\ d_2 \mid P(z)}} |R_{\mathrm{l.c.m.}(d_1, d_2)}|. \end{aligned}$$

This is Theorem 9.1 of [3].

By our earlier remark, we can choose $\wp$ to be all primes $< z$, $w(p) = 1$ for these primes, and

$$R_d = \sum_{\substack{n \in A \\ n \leqslant a_N \\ n \equiv 0 \pmod{d}}} 1 - \frac{N}{d} = O(d^2).$$

Then $g(d) = 1/\varphi(d)$ for square free $d$ and by standard arguments

$$G(z) = \sum_{\substack{d \mid p(z) \\ d \leqslant z}} \frac{1}{\varphi(d)} \gg \log z.$$

Hence, applying Lemma 8 to our sequence gives

$$\begin{aligned} \pi_A(a_N) &\ll \frac{N}{\log z} + \Big( \sum_{\substack{d_1, d_2 \leqslant z \\ d_1 \mid p(z) \\ d_2 \mid p(z)}} d_1^2\, d_2^2 \Big) \\ &\ll \frac{N}{\log z} + O(z^{10}) \\ &\ll \frac{N}{\log N} \qquad \text{by choosing } z = N^{1/11}, \text{ say.} \end{aligned}$$

We note that if

$$N = \left[ \frac{\log x}{\log \log x} + \frac{2 \log x \log \log \log x}{(\log \log x)^2} \right]$$

then $a_N \geqslant x$ and so

$$\pi_A(x) \leqslant \pi_A(a_N) \ll \frac{\log x}{(\log\log x)^2},$$

which completes the proof of our main theorem.

## 5. Concluding Remarks

We have only obtained an upper bound for the number of primes in $A$ less than $x$. We have not been able to show that there exist infinitely many primes in $A$. We believe, however, that this is indeed the case. We also believe that

$$\pi_A(x) \sim \frac{C \log x}{(\log\log x)^2} \qquad \text{for large } x, \text{ where } C > 0 \text{ is a constant.}$$

We have been able to derive results similar to Lemma 6, for the modulus $m \equiv 0 \pmod 4$. The details will appear elsewhere.

## Acknowledgments

The author is indebted to Professor R. Balasubramanian for help and encouragement and also thanks Professor M. S. Rangachari for encouragement. A part or the paper was written while the author was at the Research Science Institute. He thanks RSI for hospitality and Professors W. W. Adams and L. C. Washington for encouragement during this stay there. The author is grateful to the referee for his valuable suggestions.

## References

1. R. Hampel, The length of the shortest period of rests of numbers $n^n$, *Ann. Polon. Math.* **1** (1955), 360–366.
2. C. Hooley, "Applications of Sieve Methods to the Theory of numbers," Cambridge Univ. Press, Cambridge, 1976.
3. H. E. Richert, "Lectures on Sieve Methods," Lectures on Mathematics and Physics, No. 55, Tata Institute of Fundamental Research, 1976.