Ruth-Aaron numbers revisited

in memory of my mentor and friend, Paul Erdős

Carl Pomerance, Bell Labs, Murray Hill, NJ

Abstract. Let $S(n)$ denote the sum of the prime factors of n taken with multiplicity. We say that n is a Ruth-Aaron number if $S(n) = S(n+1)$ in honor of the famous American baseball players Babe Ruth and Hank Aaron. Ruth's lifetime homerun record was 714, which Aaron broke on April 8, 1974, by hitting number 715 towards his own eventual record of 755. Note that $S(714) = S(715)$. Erdős and I proved in 1978, in our first joint paper, that the number of Ruth-Aaron numbers up to x is $O(x \log\log x \log\log\log x / \log x)$. We mentioned that we could improve this to $O(x/\log x)$, but that we did not know how to prove $o(x/\log x)$. So, we could prove the Ruth-Aaron numbers have density 0, but we could not prove the sum of their reciprocals is bounded. In this note we are able to improve on this old result by a factor of nearly $\log x$, which is enough to show that the sum of the reciprocals is indeed bounded. Along the way, we prove a lemma, of perhaps independent interest, concerning the average number of divisors of a quadratic polynomial over a fairly short interval.

Introduction.

Paul Erdős is famous not only for his theorems, but also for his discovery and development of young mathematicians, with stories of his "epsilons" being legend. I am fortunate beyond words that Erdős played a pivotal role in my career, helping me to develop my talents in combinatorial number theory, eventually leading to my involvement in the analysis of number-theoretical algorithms. However, Paul did not discover me when I was an epsilon. Rather, I had already done my doctorate at Harvard University under John Tate, and was several years into an assistant professorship at the University of Georgia. Unlike most Harvard-trained number theorists, I had not become an expert in algebraic geometry; my true love was elementary number theory. Among the first papers I studied at Georgia was Erdős's 1956 paper "On pseudoprimes and Carmichael numbers" in Publicationes Mathematicae Debrecen. Here he found a good upper bound for the distribution of Carmichael numbers, and gave his famous heuristic argument on why they should be relatively numerous. More than twenty years after I read this paper, Red Alford, Andrew Granville and I would add some new elements and make this heuristic argument into a proof. We dedicated the paper, which appeared in Annals of Mathematics, to Erdős on his eightieth birthday.

But I get ahead of the story. In the spring of 1974, I had still not met Erdős, I was still trying to find my true character as a mathematician, and I happened to be watching a baseball game on television. Many European mathematicians are not so knowledgeable about baseball, so let me explain that a "homerun" is a special event for a player. It is very difficult to do, and the best homerun hitters are strong, and with very quick reflexes. One of the greatest homerun hitters of all time was George Herman "Babe" Ruth, who, when he retired in 1935, had hit a total of 714 homeruns. This feat seemed impregnable, being about one-third higher than anyone else of his or prior eras.

After the Second World War, American baseball became integrated with African-American players and also players from Latin America. Among the first American blacks to play for the major teams were Jackie Robinson and Henry "Hank" Aaron. They are considered as American heroes since not only were they excellent athletes, but, with great personal dignity, they put up with constant degrading insults and even death threats. It became especially bad for Aaron, when in the spring of 1974 it appeared that he might actually surpass Babe Ruth's supposedly unbeatable record. On April 8 of that year, he succeeded, hitting number 715. By his retirement several years later, he had hit a total of 755 homeruns, and the address of the baseball field in Atlanta is given with this in mind: it is 755 Hank Aaron Boulevard.

At the actual moment when the 714 record was broken, being more of a mathematician than a baseball fan I guess, I started looking at interesting properties of the numbers 714 and 715. The first thing that I noticed was that they factor very easily, and in fact, their product is the product of the first 7 primes. (It is conjectured now that this is the last pair of consecutive integers whose product is the product of the first k primes for some k.) The next day, I challenged my colleague, David Penney, to find an interesting property of 714 and 715. He discovered the same property, but he also challenged a class he was teaching that morning, and one of the students, Jeremy Jordan, discovered that $S(714) = S(715)$, where $S(n)$ is the sum of the prime factors of n taken with multiplicity. (Since 714 and 715 are both squarefree, one might also have taken a sum of distinct prime divisors, but it is a little simpler to consider the completely additive function $S(n)$.)

Penney and I, together with another student, Carol Nelson, wrote a short, humorous paper [10] on our observations. On the issue of $S(n) = S(n+1)$, we had actually managed to give a proof that there are infinitely many solutions if one is prepared to assume Schinzel's Hypothesis H. Namely if n is an integer such that

$$s = 2n + 1, \quad p = 8n + 5, \quad q = 48n^2 + 24n - 1, \quad r = 48n^2 + 30n - 1$$

are all prime, then $pq + 1 = 4sr$ and $S(pq) = S(4sr)$. Schinzel's Hypothesis H implies that a collection of polynomials such as these will indeed assume simultaneously prime values infinitely often.

Say an integer n with $S(n) = S(n+1)$ is a "Ruth-Aaron number." (We had originally called n an "Aaron number," but in retrospect it seems fairer to honor both baseball greats.) In our paper [10] we wrote: "The numerical data suggest that Aaron numers are rare. We suspect they have density 0, but we cannot prove this." These words started my life over as a mathematician in the Erdős school. Paul had read this article, which was published within a few months of the actual baseball event, and wrote to me that he knew how to prove density 0, and would like to visit me at Georgia to discuss it. This then became the subject of our first joint paper [6], in which we also discussed the joint distribution of the largest prime factors of n and $n + 1$.

We could prove that the number of Ruth-Aaron numbers up to x is $O(x/\log x)$, but only gave the details of a slightly weaker result. We wrote "We expect the estimate $O(x/(\log x)^k)$ is true for every k, but we cannot prove this for any $k > 1$. In fact, we cannot even get $o(x/\log x)$." So, in particular, we were not able to prove that the sum of the reciprocals of the Ruth-Aaron numbers is bounded. In this note, I am able to give an improvement by nearly a factor $\log x$, and thus show that the sum of the reciprocals is indeed bounded. Along the way, I prove a lemma, of perhaps independent interest, that is based on an old result of Erdős [5] concerning the average number of divisors of consecutive polynomial values. The result established here is for quadratic polynomials, where the length of the interval for the consecutive arguments is permitted to be fairly short in comparison to the size of the coefficients.

Theorem. *The number of integers $n \le x$ with $S(n) = S(n+1)$ is*

$$O\left(\frac{x(\log\log x)^4}{(\log x)^2}\right).$$

In particular, the sum of the reciprocals of the Ruth-Aaron numbers is bounded.

An amusing footnote to this story concerns the awarding of an honorary degree by Emory University to Paul Erdős in 1995. Paul invited me and my wife to attend a reception the evening before for the honorees and their guests. Completely unknown to me until just before entering the room, one of the others to receive an honorary degree was Hank Aaron! I introduced myself to him and tried to tell him how his athletic feat had such important consequences to my career as a mathematician. He smiled diplomatically and said he was happy for this, though I believe he thought he had just met a very strange person. I introduced him to Erdős and the two chatted for awhile. A photo of them exists in one of the recent biographies of Paul's life. Ron Gould, a professor at Emory who was one of the people instrumental in arranging the honorary degree for Erdős, knew that Aaron was to be there, but of course he had no idea of the connection to Erdős. He and his wife had come supplied with some new baseballs for Aaron to sign as souvenirs. They graciously let me have one of these, and I had both Aaron and Erdős sign the same baseball. Though the writing is unfortunately fading with time, it is a prized possession. I joke that Aaron should have Erdős-number 1 since, though he does not have a joint paper with Erdős, he does have a joint baseball.

Proof of the theorem.

Let $P(n)$ denote the largest prime factor of n. Say $n \le x$ and $S(n) = S(n+1)$. Write $n = pk$, $n+1 = qm$ where $p = P(n)$, $q = P(n+1)$.

We first note that we may assume that

$$p > x^{1/\log\log x}, \quad q > x^{1/\log\log x} \tag{1}$$

since the number of integers $n \le x$ for which (1) does not hold is $O(x/(\log x)^2)$, see [3].

The following easy result was established in [6]: for $P(N) \ge 5$,

$$P(N) \le S(N) \le P(N)\log N/\log P(N). \tag{2}$$

In light of (1), we may assume $P(n), P(n+1) \ge 5$, so that (2) holds for n and $n+1$.

We next note that the numbers k, m determine the primes p, q. Indeed, from the two equations

$$pk+1 = qm, \quad p + S(k) = q + S(m)$$

we get that

$$p = \frac{(S(k)-S(m))m-1}{k-m}, \quad q = \frac{(S(k)-S(m))k-1}{k-m}. \tag{3}$$

Thus, the number of choices for n corresponding to choices of k, m with $k, m < x^{1/2}/\log x$ is at most $x/(\log x)^2$. We hence may assume that

$$p \le x^{1/2}\log x \ \text{ or } \ q \le x^{1/2}\log x. \tag{4}$$

Suppose $p > x^{1/2}\log x$. Then (2) and (4) imply that

$$p \le S(n) = S(n+1) \le q\frac{\log(n+1)}{\log q} \le \frac{x^{1/2}\log x\ \log(n+1)}{\log(x^{1/2}\log x)} < 2x^{1/2}\log x.$$

A similar inequality holds if $q > x^{1/2}\log x$. We conclude that

$$p < 2x^{1/2}\log x \ \text{ and } \ q < 2x^{1/2}\log x. \tag{5}$$

Suppose that

$$S(k) < p/(\log x)^2, \quad S(m) < q/(\log x)^2. \tag{6}$$

Then, since $p + S(k) = q + S(m)$, we have

$$|p-q| < \frac{p+q}{(\log x)^2}. \tag{7}$$

For p satisfying (1), the number of primes q such that (7) holds is $O(p\log\log x/(\log x)^3)$ and the sum of $1/q$ for such primes q is $O(\log\log x/(\log x)^3)$. Now, for a given choice of p, q the number of $n \le x$ with $p|n$ and $q|n+1$ is at most $1 + x/(pq)$. Thus, if (6) holds, the number of n that we are counting is at most

$$\sum_{p,q \text{ subject to } (1),(5),(7)} \left(1 + \frac{x}{pq}\right) \ll \sum_{p<2x^{1/2}\log x} \left(\frac{p\log\log x}{(\log x)^3} + \frac{x\log\log x}{p(\log x)^3}\right) \ll \frac{x\log\log x}{(\log x)^2}.$$

We thus may assume that (6) does not hold. The arguments for the cases $S(k) \ge p/(\log x)^2$ and $S(m) \ge q/(\log x)^2$ are parallel, so we shall only give the details for the first case. That is, we shall assume that

$$S(k) \ge p/(\log x)^2. \tag{8}$$

Write $k = rl$ where $r = P(k)$. As in the proof of (5), the inequality (2) gives us

$$q \le p\frac{\log x}{\log p}, \quad p \le q\frac{\log(x+1)}{\log q}.$$

The second inequality implies that $\log q \geq \frac{2}{3}\log p$, so that $q \geq \frac{1}{2} p \log p / \log x$. That is, we have

$$p\frac{\log p}{2\log x} \leq q \leq p\frac{\log x}{\log p}. \tag{9}$$

Similarly, (8) gives us

$$p\frac{\log p}{2(\log x)^3} \leq r \leq p. \tag{10}$$

For a given choice of p, r, q the number of $n \leq x$ with $pr|n$ and $q|n+1$ is at most $1 + x/(prq)$. Suppose $p \leq x^{1/3}$. Then the number of n in this case is at most

$$\sum_{\substack{p,q,r \text{ subject to } (1),(9),(10) \\ p \leq x^{1/3}}} \left(1 + \frac{x}{prq}\right) \ll \frac{x}{(\log x)^3} + \sum_{p > x^{1/\log\log x}} \frac{x}{p} \cdot \frac{\log\log x}{\log p} \cdot \frac{\log\log x}{\log p}$$
$$\ll \frac{x(\log\log x)^4}{(\log x)^2}.$$

Thus, we may assume that $p > x^{1/3}$. It follows from (9) that $p/6 < q < 3p$, and it follows from (10) that $p/(6(\log x)^2) \leq r \leq p$.

Using (3), we have that

$$p(rl - m) = p(k - m) = (S(l) + r - S(m))m - 1.$$

Hence

$$pl(rl - m) - rml = (S(l) - S(m))ml - l$$

and we conclude that

$$(pl - m)(rl - m) = (S(l) - S(m))ml - l + m^2. \tag{11}$$

Thus, given l, m the number of choices of r, and hence for n, is at most

$$\tau((S(l) - S(m))ml - l + m^2) \leq x^{o(1)},$$

where τ denotes the divisor function. Suppose that $p \geq x^{2/5}$. Since $l \leq x/(pr) \ll x(\log x)^2/p^2$ and $m \ll x/p$, we conclude that the number of choices for n is at most $x^{4/5+o(1)}$. Hence we may assume that

$$x^{1/3} < p < x^{2/5}. \tag{12}$$

Suppose that

$$P(l) < x^{1/6}, \quad P(m) < x^{1/6}. \tag{13}$$

Then $p + r = q + O(x^{1/6})$. Given p, r it follows that the number of choices for q is $O(x^{1/6})$. But the number of choices for p, r with $r \leq p$ and (12) holding is $O(x^{4/5})$. Thus the number of triples p, q, r is $O(x^{29/30})$. But $prq \gg x/(\log x)^2$, so the number of choices for

n given p, r, q is $O((\log x)^2)$. It follows that but for $O(x^{29/30}(\log x)^2)$ choices for $n \le x$ we have that (13) does not hold.

We first consider the case that $P(l) \ge x^{1/6}$. Write $l = sj$ where $s = P(l)$. We rewrite (11) as

$$(psj - m)(rsj - m) = mjs^2 + ((S(j) - S(m))mj - j)s + m^2. \tag{14}$$

We shall fix a choice for j, m and sum over choices for s. The following lemma is useful.

Lemma. *Suppose A, B, C are integers with $\gcd(A, B, C) = 1$, $D := B^2 - 4AC \ne 0$, $A \ne 0$. Suppose the maximum value of $|At^2 + Bt + C|$ on the interval $[1, x]$ is M_0. Let $M = \max\{M_0, |D|, x\}$, let $\mu = \lceil \log M / \log x \rceil$ and assume that $\mu \le \frac{1}{7} \log\log x$. Then*

$$\sum_{n \le x} \tau(|An^2 + Bn + C|) \le x(\log x)^{2^{3\mu+1}+4}$$

holds uniformly for $x \ge x_0$. (We interpret $\tau(0)$ as 0 should it occur in the sum. The number x_0 is an absolute constant independent of the choice of A, B, C.)

Proof. For a positive integer m, let $N(m)$ denote the number of solutions to the congruence $An^2+Bn+C \equiv 0 \pmod{m}$. By the Chinese remainder theorem, $N(m)$ is a multiplicative function. If p is a prime that does not divide the discriminant D, then $N(p^a) \le 2$ for all positive integers a. Suppose $p|D$ is an odd prime. Note that we may assume that p does not divide A. For if $p|A$, then it would also divide B, since it divides $B^2 - 4AC$, and so it does not divide C, since A, B, C are coprime. So if $p|A$ and $p|D$, we have $N(p^a) = 0$ for all integers a. So assume the odd prime p divides D and not A. Making the change of variables that sends $2Ax + B$ to x, we see that $N(p^a)$ for our polynomial is the same as for the polynomial $x^2 - D$. Say $D = \delta p^j$, where p does not divide δ. Then $N(p^a) = p^{\lfloor a/2 \rfloor}$ for $a \le j$. If j is odd, then $N(p^a) = 0$ for $a > j$. If j is even, $N(p^a) = 2p^{j/2}$ for $a > j$ or $N(p^a) = 0$ depending on whether $(\delta/p) = 1$ or -1, respectively. Now assume $2|D$. Thus, B is even. If A is even, then C is odd and $N(2^a) = 0$ for all positive integers a. So, we may assume that A is odd. Replacing $Ax + B/2$ with x, we may assume our polynomial is $x^2 - D/4$. Say $D/4 = 2^j\delta$, where δ is odd. Then $N(2^a) = 2^{\lfloor a/2 \rfloor}$ for $a \le j$. If j is odd, then $N(2^a) = 0$ for $a > j$. Assume now that that j is even. Then $N(2^{j+1}) = 2^{j/2}$. If $\delta \not\equiv 1 \pmod{4}$, then $N(2^a) = 0$ for $a > j+1$. If $\delta \equiv 1 \pmod{4}$, then $N(2^{j+2}) = 2^{1+j/2}$. If $\delta \not\equiv 1 \pmod{8}$, then $N(2^a) = 0$ for $a > j + 2$, and if $\delta \equiv 1 \pmod{8}$, then $N(2^a) = 2^{2+j/2}$ for $a > j + 2$. We conclude that if we factor m as $m_0m_1m_2$ where the prime factors of m_0 divide D and m_0 to at least the second power, the prime factors of m_1 are different from the prime factors of m_0 and they divide D, and the prime factors of m_1 do not divide D, then $N(m) \ll \tau(m_2)\sqrt{m_0}$.

Let $m_0(n)$ denote the largest divisor of $An^2 + Bn + C$ consisting of primes $p \le x^{1/3}$ for which p^2 divides An^2+Bn+C and D. If $m_0(n) > x^{1/3}$, then An^2+Bn+C is divisible by an integer m such that the prime factors of m all divide D, and $x^{1/3} < m \le x^{2/3}$. The number of $n \le x$ with $An^2 + Bn + C$ divisible by such a number m is

$$\ll \sum_m \frac{xN(m)}{m} \ll x \sum_m \frac{1}{\sqrt{m}} < x^{5/6} \sum_m 1.$$

Now the number of integers $m \le x^{2/3}$ whose prime factors all divide D is $\le x^{o(1)}$, see [3]. Also, the maximal contribution for a term in our divisor sum is $\le 2^{(1+o(1))\log M/\log\log M}$, which is $\le x^{1/7+o(1)}$. Thus, the contribution to the divisor sum from integers n with $m_0(n) > x^{1/3}$ is $\le x^{41/42+o(1)}$, which is negligible. Hence we may consider only those numbers n with $m_0(n) \le x^{1/3}$.

Let $m_1(n)$ denote the largest divisor of An^2+Bn+C consisting of primes $\le x^{1/3}$ which do not divide $m_0(n)$, and let $m_2(n)$ denote the largest divisor of An^2+Bn+C consisting of primes $> x^{1/3}$. Then

$$\tau(|An^2+Bn+C|) = \tau(m_0(n))\tau(m_1(n))\tau(m_2(n)) \le 2^{3\mu}\tau(m_0(n))\tau(m_1(n)).$$

Indeed, the number of prime factors of $m_2(n)$, counted with multiplicity, is at most 3μ. Also, there is a divisor $m(n)$ of $m_1(n)$ which is $\le x^{2/3}$ and such that

$$\tau(m_1(n)) \le \tau(m(n))^{3\mu}.$$

Indeed, since the prime factors of $m_1(n)$ are all $\le x^{1/3}$, there is a factorization of $m_1(n)$ as $a_1a_2\dots a_t$ where each $a_i \in (x^{1/3}, x^{2/3}]$, so that $t \le 3\mu$. But

$$\tau(m_1(n)) \le \tau(a_1)\tau(a_2)\dots\tau(a_t) \le \left(\max\{\tau(a_i)\}\right)^t.$$

We conclude that

$$\tau(|An^2+Bn+C|) \le 2^{3\mu}\tau(m_0(n))\tau(m(n))^{3\mu}.$$

Using our reduction to the case $m_0(n) \le x^{1/3}$, we have $m_0(n)m(n) \le x$. Let m_0 run over squarefull integers composed of primes p with $p^2|D$, and let m run over integers composed of primes p such that p^2 does not divide D. Then,

$$\begin{aligned}\sum_{n\le x, m_0(n)\le x^{1/3}} & \tau(|An^2+Bn+C|) \\ & \le 2^{3\mu} \sum_{m_0\le x^{1/3}, m\le x^{2/3}} \tau(m_0)\tau(m)^{3\mu} \sum_{n\le x, m_0 m | An^2+Bn+C} 1 \\ & \ll 2^{3\mu} x \sum_{m_0\le x^{1/3}, m\le x^{2/3}} \frac{\tau(m_0)N(m_0)}{m_0}\frac{\tau(m)^{3\mu}N(m)}{m} \\ & \ll 2^{3\mu} x \sum_{m_0\le x^{1/3}} \frac{\tau(m_0)}{\sqrt{m_0}} \sum_{m\le x^{2/3}} \frac{\tau(m)^{3\mu+1}}{m}.\end{aligned}$$

Now,

$$\sum_{m_0\le x^{1/3}} \frac{\tau(m_0)}{\sqrt{m_0}} \le \prod_{p^2|D}\left(1+\frac{3}{p}+\frac{4}{p^{3/2}}+\dots\right) \ll (\log|D|)^3 = o((\log x)^4).$$

Using the inequalities

$$\tau^j(n) \le \tau_{2^j}(n), \quad \sum_{n\le x} \frac{\tau_k(n)}{n} \le \frac{1}{k!}(\log x + k)^k \quad (\text{for } x \ge 1),$$

where $\tau_k(n)$ denotes the number of ordered factorizations of n into k positive integral factors, we have

$$\sum_{m\le x^{2/3}} \frac{\tau(m)^{3\mu+1}}{m} \le \frac{1}{(2^{3\mu+1})!}\left(\frac{2}{3}\log x + 2^{3\mu+1}\right)^{2^{3\mu+1}}.$$

Now $2^{3\mu+1} \le \sqrt{\log x}$ for $x \ge x_0$, so that

$$\frac{2}{3}\log x + 2^{3\mu+1} \le \log x,$$

for $x \ge x_0$. Assembling these estimates, we have for $x \ge x_0$ that

$$\sum_{n\le x} \tau(|An^2 + Bn + C|) \le x(\log x)^{2^{3\mu+1}+4},$$

which proves the lemma.

We now apply the lemma with $A = mj$, $B = (S(j) - S(m))mj - j$, $C = m^2$. Since j comprises the small prime factors of n and m comprises the small prime factors of $n+1$, we have that $\gcd(j,m) = 1$ and so $\gcd(A,B,C) = 1$. Note that $4AC = 4m^3j$, that $j^2|B^2$ and that $B^2 \equiv j^2 \pmod m$. Thus, if $D = 0$, then $j|4$ and $m = 1$. Then $j = 2$ or $j = 4$. This gives the triples $2,2,1$ or $4,6,1$ for A,B,C, and neither choice has $D = 0$. Thus $D \ne 0$. Further, assuming that $j < 6x^{1/6}(\log x)^2$, $m \ll x^{2/3}$, and $s \le 6x^{1/3}(\log x)^2/j$, we have that the maximum of $|As^2 + Bs + C|$ for the range of s is $\ll x^{4/3}(\log x)^2$. It follows from the lemma that

$$\sum_{s\le 6x^{1/3}(\log x)^2/j} \tau(|As^2 + Bs + C|) \le (1/j)x^{1/3}(\log x)^c \tag{15}$$

for some positive constant c. (We have ignored the condition that s is prime.)

If $x^{1/3} < p \le x^{1/3}(\log x)^{c+5}$, the number of n in this case is at most

$$\sum_{p\asymp q}\left(1 + \frac{x}{pq}\right) \ll x^{2/3}(\log x)^{2c+10} + \frac{x}{\log x}\sum \frac{1}{p} \ll \frac{x\log\log x}{(\log x)^2}.$$

Thus, we may assume that $p > x^{1/3}(\log x)^{c+5}$. Then $m \ll x^{2/3}/(\log x)^{c+5}$, so that summing (15) over all choices for m, j we get a quantity that is $\ll x/(\log x)^2$, which is negligible.

Finally, we consider the remaining case in (13) when $P(m) \ge x^{1/6}$. Let $m = tu$ where $t = P(m)$. Then, from (11), we obtain

$$(pl - tu)(rl - tu) = t^2(u^2 - ul) + t(ulS(l) - ulS(u)) - l. \tag{16}$$

We apply the Lemma to the quadratic polynomial with $A = u^2 - ul$, $B = ul(S(l) - S(u))$, $C = -l$. As before, we may assume that $p \geq x^{1/3}(\log x)^{c+5}$ so that $l \leq 6x^{1/3}/(\log x)^{2c+3}$. We have $u \ll x^{1/2}$, and $t \ll (1/u)x^{2/3}$. Summing the number of divisors of the right side of (16) for t, u, l ranging as stated, we get an estimate that is $\ll x/(\log x)^{2c+2}$, which is negligible. This completes the proof.

Remarks. Note that it is not assumed that the polynomial in the Lemma is irreducible. With the coefficients fixed, it is known that the average order of the divisor function on a quadratic polynomial is $\log x$ for the irreducible case, and $(\log x)^2$ for the reducible case, see [2], [5], [9]. It is reasonable to conjecture that the power of $\log x$ in the Lemma can be greatly reduced, and perhaps the methods of [5] will help in this regard. Very little seems to be already known for the case we have dealt with here, which calls for a uniform result as the coefficients vary. There is at least the paper [7] which is a step in this direction, but it is not valid in a wide enough range for our use. Note that the assumption in the Lemma that $AD \neq 0$ is really not necessary; these cases are in fact simpler. In addition, it would not be hard to fashion a version of the Lemma which holds for polynomials of higher degree.

The Lemma is also likely to be of use in other studies, such as the distribution of Carmichael numbers with exactly 3 prime factors. In [1], the authors obtain the number up to x being $\leq x^{5/14+o(1)}$, where the "$o(1)$" comes from the maximal order of the divisor function. Using the Lemma above it should be possible to replace $x^{o(1)}$ with a factor $(\log x)^{O(1)}$, and do so with explicit constants.

One might notice that the equation (11) suggests that we look at the restricted divisors of the right side that are in the residue class m mod l. However, in the range of interest, with the right side of (11) just larger than l^4, the existing work [8], [4], on divisors in residue classes, just fails to be of use. It is for this reason that we considered the Lemma.

It is possible to reduce somewhat the power of $\log\log x$ in the Theorem. As stated in [6], probably the number of Ruth-Aaron numbers up to x is is $\ll x/(\log x)^k$ for any fixed k and $\gg x^{1-\epsilon}$ for any fixed $\epsilon > 0$. It is still unproved if there are infinitely many Ruth-Aaron numbers, but an argument in [10], mentioned above in the Introduction, shows that this would follow on assumption of Schinzel's Hypothesis H.

Finally note that the same result holds by virtually the same proof if the function S is replaced by the sum of the distinct prime factors or is replaced by the sum of the prime powers $p^a \| n$. These functions are not completely additive, but the extra considerations necessary to deal with this are small.

References

[1] R. Balasubramanian and S. V. Nagaraj, *Density of Carmichael numbers with three prime factors*, Math. Comp. **66** (1997), 1705–1708.

[2] R. Bellman and H. N. Shapiro, *On a problem in additive number theory*, Ann. of Math. (2) **49** (1948), 333–340.

[3] N. G. de Bruijn, *On the number of positive integers $\leq x$ and free prime factors $> y$. II*, Nederl. Akad. Wetensch. Proc. Ser. A **69** = Indag. Math. 28 (1966), 239–247.

[4] D. Coppersmith, N. Howgrave-Graham, and S. V. Nagaraj, *Divisors in residue classes, constructively*, preprint 13 May, 2000.

[5] P. Erdős, *On the sum* $\sum_{k=1}^{x} d(f(k))$, J. London Math. Soc. **27** (1952), 7–15.
[6] P. Erdős and C. Pomerance, *On the largest prime factors of* n *and* $n+1$, Aequationes Mathematicae **17** (1978), 311–321.
[7] C. Hooley, *On a new technique and its applications to the theory of numbers*, Proc. London Math. Soc. (3) **38** (1979), 115–151.
[8] H. W. Lenstra, Jr. *Divisors in residue classes*, Math. Comp. **42** (1984), 331–340.
[9] M. Nair, *Multiplicative functions of polynomial values in short intervals*, Acta Arith. **62** (1992), 257–269.
[10] C. Nelson, D. E. Penney, and C. Pomerance, 714 *and* 715, J. Recreational Math. **7** (1974), 87–89.