ACTA ARITHMETICA
161.1 (2013)

On the sumset of the primes and a linear recurrence

by

Christian Ballot (Caen) and Florian Luca (México)

1. Introduction. In 1934, Romanoff [6] showed that the set $\mathcal{R}$ of integers that are the sum of a prime number and a power of 2 has positive lower asymptotic density in the positive integers. That is,

$$\liminf_{x\to\infty} \frac{\#\mathcal{R}(x)}{x} > 0, \tag{1}$$

where if $\mathcal{A}$ is a subset of the natural numbers, then $\mathcal{A}(x)$ is the set of integers less than or equal to x which are in $\mathcal{A}$, and $\#\mathcal{A}(x)$ is the cardinality of $\mathcal{A}(x)$. One key element of the proof of Romanoff was to show the convergence of the series

$$\sum_{n\,\text{odd}} \frac{\mu(n)^2}{ne_n(2)}, \tag{2}$$

where μ is the Möbius function and $e_n(2)$ is the order of 2 modulo n.

In 2010, Lee [4] took the interesting step of replacing powers of 2 by Fibonacci numbers and was able to adapt Romanoff's proof to show that integers that are the sum of a prime and a Fibonacci number also have a positive lower asymptotic density. Instead of the series (2), Lee had to show the convergence of the sum

$$\sum \frac{\mu(n)^2}{ne_n(F)}, \tag{3}$$

where $e_n(F)$ is the period of maximal length of the Fibonacci sequence modulo p as p varies through the prime divisors of n.

In Lee's work properties that may seem specific to the Fibonacci sequence were used (for instance, properties of the Zeckendorf representation, and the fact that any given residue can occur at most four times within a period of the Fibonacci sequence modulo a prime, for all primes), so it is natural to ask how far both Romanoff's and Lee's results may be extended.

2010 *Mathematics Subject Classification*: 11P32, 11B37.
Key words and phrases: linear recurrence, prime, sumset, asymptotic density, S-units.

DOI: 10.4064/aa161-1-3 © Instytut Matematyczny PAN, 2013

In the present paper, we fix an integral nondegenerate linear recurring sequence $u = (u_n)_{n\geq 0}$ whose (minimal) characteristic polynomial $P(x)$ in $\mathbb{Z}[x]$ is monic and separable. That is, its roots $\alpha_1, \ldots, \alpha_s$ in $\mathbb{C}$ are nonzero and distinct, where $s \geq 1$. We assume further that $|\alpha_1| \geq 2$ in case $s = 1$. Note that our hypothesis implies that there exist nonzero complex numbers c_i, $i = 1, \ldots, s$, such that for all $n \geq 0$ we have

$$u_n = c_1\alpha_1^n + \cdots + c_s\alpha_s^n. \tag{4}$$

We recall that the linear recurrence u is said to be *nondegenerate* if no quotient of any two roots of $P(x)$ is a root of unity. Note that $u_n = 2^n$ and $u_n = F_n$, the nth Fibonacci number, are particular instances of sequences satisfying all our hypotheses.

Thus, the upshot of the paper is to prove the following theorem.

THEOREM 1. *Let $(u_n)_{n\geq 0}$ be a nondegenerate integral linear recurring sequence whose characteristic polynomial $P(x)$ in $\mathbb{Z}[x]$ is monic and has distinct complex roots. We assume the absolute value of the root of $P(x)$ to be at least 2 in case the degree of P is one. Then, if $\mathcal{S}$ is the sumset $\mathcal{P} + \mathcal{U}$, where $\mathcal{P}$ is the set of prime numbers and $\mathcal{U}$ is the set of terms u_n of the sequence $(u_n)_{n\geq 0}$, we have*

$$\liminf_{x\to\infty} \frac{\#\mathcal{S}(x)}{x} > 0, \tag{5}$$

that is, the set $\mathcal{S}$ has positive lower asymptotic density.

Even if Theorem 1 extends considerably the theorems of Romanoff and Lee, the general structure of our proof remains that of Romanoff, which proceeds in three steps. Denoting by $r(n)$ the number of representations of an integer n as a sum $p+u_k$, the first step is to show that the average number of representations $\sum_{n=1}^{N} r(n)$ is either asymptotic to c_1N, or asymptotically greater than c_1N for some $c_1 > 0$, as N tends to infinity. The second step is to estimate the sum $\sum_{n=1}^{N} r(n)^2$ and show it is bounded above by c_2N for some $c_2 > 0$ and for N large enough. The final step uses Cauchy–Schwarz's inequality, that is,

$$\Big(\sum_{n=1}^{N} r(n)\Big)^2 \leq \Big(\sum_{\substack{n=1\\ n\in\mathcal{S}}}^{N} 1^2\Big) \cdot \Big(\sum_{n=1}^{N} r(n)^2\Big), \tag{6}$$

which implies that $\#\mathcal{S}(N) \gg N^2/N = N$. The proof of these steps will be carried out in Section 3. In particular, in Lemma 16, a general series is proven to converge which, in some way, plays the role of the series (2) and (3). Section 2 is mainly allotted to stating and establishing various propositions and lemmas that depend in an essential way on the fundamental theorem on S-units. Those results will be used in Section 3, particularly in proving Lemma 15, which is instrumental in the proof of Lemma 16.

In what follows, the letters p and q invariably represent prime numbers and $\pi(x)$ stands, as usual, for the number of primes up to x.

2. Tool box and lemmas. One of the main tools behind several of the results we state in this section is the fundamental theorem on S-units (see [1, pp. 19–22] for history and references).

THEOREM 2. *Let G be a finitely generated subgroup of $\mathbb{C}^*$ and $\tau \geq 1$ be an integer. The number of solutions of the equation*

$$X_1 + \cdots + X_\tau = 1,$$

under the assumption that no subsum of the left-hand side vanishes and the X_i's are in G, is finite.

EXAMPLE. Consider the subgroup of $\mathbb{C}^*$ generated by ± 2 and the equation $X_1 + X_2 + X_3 = 1$. The triplets $(1, 2^n, (-2)^n)$, n odd, are all solutions, but for all of them the subsum $X_2 + X_3$ is zero.

We prove a pair of corollaries of Theorem 2, before stating much stronger related results. We do this so as not to act as total thieves, or, more seriously, so as to get a feel of where essential ingredients of our arguments come from.

COROLLARY 3. *Let c_i and α_i, $i = 1, \ldots, s$, be $2s$ nonzero complex numbers such that α_i/α_j is not a root of unity if $i \neq j$. Then the equation*

$$c_1\alpha_1^n + \cdots + c_s\alpha_s^n = 0 \tag{7}$$

has at most finitely many solutions in integers n.

Proof. We proceed by strong induction on s. Since c_1 and α_1 are not zero, $c_1\alpha_1^n$ is nonzero for all n. Thus, the lemma holds for $s = 1$. Now suppose that $s \geq 2$ and that $\sum_{i=1}^{s} c_i\alpha_i^n = 0$. Then, dividing by $c_s\alpha_s^n$ yields

$$-\sum_{i=1}^{s-1} \frac{c_i}{c_s}\left(\frac{\alpha_i}{\alpha_s}\right)^n = 1. \tag{8}$$

By the inductive hypothesis, each subsum of the left-hand side may take the value 0 for only finitely many n's. Let $\mathcal{I}$ be this finite set of integers n. If (8) holds for some integer n outside $\mathcal{I}$, then (8) yields up to $(s-1)!$ solutions $(X_1, \ldots, X_{s-1})$ to the equation $X_1 + \cdots + X_{s-1} = 1$, where the X_i's lie in the group G generated by the c_i's and the α_i's. Indeed, $(X_1, \ldots, X_{s-1})$ may be any permutation of the $s-1$-tuplet

$$(-(c_1/c_s)(\alpha_1/\alpha_s)^n, \ldots, -(c_{s-1}/c_s)(\alpha_{s-1}/\alpha_s)^n).$$

By Theorem 2 there can be only finitely many such n's. For suppose $-(c_1/c_s)(\alpha_1/\alpha_s)^n = X_1$ for two distinct values of n, say n_1 and n_2; then we would have $(\alpha_1/\alpha_s)^{n_1-n_2} = 1$, which would contradict the fact that the ratio α_1/α_s is not a root of unity. ∎

COROLLARY 4. *Let $u = (u_n)_{n\geq 0}$ be a nonconstant nondegenerate linear recurring sequence whose characteristic polynomial in $\mathbb{C}[x]$ has distinct roots, none of which is a primitive root of order ≥ 2. Let c be a complex number. Then the equation*

$$u_n = c$$

has finitely many integer solutions n.

Proof. We may apply Corollary 3 to $u_n - c$, since $u_n - c$ is not identically zero and may be written as an expression of the form (7). ∎

In fact, much more is known. Let $\mathcal{R}$ be a subring of $\mathbb{C}$. Put $u = (u_n)_{n\geq 0}$ for a nondegenerate linear recurring sequence u of order s over $\mathcal{R}$ and, for $c \in \mathcal{R}$, put $m(u, c)$ for the number of solutions n of $u_n = c$. Finally, if $\mu(s, \mathcal{R})$ is the supremum over all nondegenerate linear recurring sequences u of order s over $\mathcal{R}$ and all c's in $\mathcal{R}$ of all $m(u, c)$'s, then $\mu(s, \mathcal{R})$ is finite. As explained on pp. 26–27 of [1], this comes as a consequence of work of Schlickewei and Schmidt. In particular, we deduce the two propositions below, the second one being stronger than the first.

PROPOSITION 5. *If $u = (u_n)_{n\geq 0}$ is a nondegenerate integral linear recurring sequence, then there is a positive integer M such that for all $c \in \mathbb{Z}$,*

$$m(u, c) \leq M.$$

PROPOSITION 6. *Let $\mathcal{O}$ be the ring of integers of a number field. Then there is a positive integer M_s such that for all nondegenerate linear recurring sequences $u = (u_n)_{n\geq 0}$ over $\mathcal{O}$ of order less than or equal to s, the equation $u_n = c$, where c is any element of $\mathcal{O}$, has at most M_s solutions.*

LEMMA 7. *Assume the hypotheses of Theorem 1. Then, in proving that theorem, we may suppose that the roots of $P(x)$ are not roots of unity.*

Proof. We use the notation of Theorem 1. No root α_i, $i = 1, \ldots, s$, is a primitive root of unity of order $h \geq 3$. Otherwise, as $P(x)$ has integral coefficients, all primitive roots of order h would be zeros of $P(x)$ and, thus, their quotient would be a root of unity, which contradicts the hypothesis.

If one of the roots, say α_s, is equal to 1, then, by hypothesis, $s \geq 2$ and -1 is not a root of $P(x)$. Assuming u_n is expressed as in (4), then the c_i's lie in the root field of $P(x)$ and c_s is a rational number. Say $c_s = a/b$, where a is an integer and b a positive integer. Then, if the set $b\mathcal{U} + \mathcal{P}$ has positive lower asymptotic density, so does $\mathcal{S} = \mathcal{U} + \mathcal{P}$ (consider the map $bu_k + p \mapsto u_k + p$). Thus, we may work with $v = (bu_n - a)_{n\geq 0}$ instead of $u = (u_n)_{n\geq 0}$. Now, if say $\alpha_s = -1$, then 1 is not a root of $P(x)$ and we may work with the sequence $v = (bu_{2n} - a)_{n\geq 0}$, where again we have put $c_s = a/b$. Neither 1 nor -1 is a root of the characteristic polynomial of v, and v satisfies the

hypotheses of Theorem 1. Moreover, the set $\mathcal{P} + \{u_{2n};\ n \geq 0\}$ is a subset of $\mathcal{S}$. Hence, if it has positive lower density, then so does the set $\mathcal{S}$. ∎

The assumption that none of the α_i's is a root of unity will be made from now on throughout the rest of the paper. In particular, the hypothesis that $|\alpha_1| \geq 2$ in case $s = 1$ no longer needs to be stated.

LEMMA 8. *Let $(u_n)_{n\geq 0}$ be a nonzero nondegenerate integral linear recurring sequence whose monic characteristic polynomial in $\mathbb{Z}[x]$ has distinct roots $\alpha_1, \ldots, \alpha_s$. Then the equation*

$$u_m = u_n \quad (m \neq n)$$

has finitely many solutions (m, n) in nonnegative integers ([1]).

Proof. The lemma clearly holds if $s = 1$ for then $u_m = u_n$ implies that $\alpha_1^m = \alpha_1^n$, forcing the integer α_1 to be ± 1, a case ruled out after the proof of Lemma 7. Assume now $s \geq 2$. Dividing the equation $u_m - u_n = 0$ through by $c_s\alpha_s^n$, we get

$$\sum_{i=1}^{s} \frac{c_i}{c_s}\,\frac{\alpha_i^m}{\alpha_s^n} - \sum_{i=1}^{s-1} \frac{c_i}{c_s}\left(\frac{\alpha_i}{\alpha_s}\right)^n = 1. \tag{9}$$

If the left-hand side contains no zero subsum then, by Theorem 2, there are finitely many solutions of the equation $X_1 + \cdots + X_{2s-1} = 1$ in the subgroup G of $\mathbb{C}^*$ generated by the c_i's and the α_i's. But there can be at most one value of n for which $(c_i/c_s)(\alpha_i/\alpha_s)^n$ is equal to a given X in G, otherwise α_i/α_s would be a root of unity. Once n is determined, then α_1^m is equal to some fixed element of G and so m may take at most one value.

We now prove the lemma in complete generality. Note that the left-hand side of $u_m - u_n = 0$ can be written as the sum of $2s$ terms, since it is

$$\sum_{i=1}^{s} c_i\alpha_i^m - \sum_{i=1}^{s} c_i\alpha_i^n. \tag{10}$$

Thus, there are $2^{2s} - (2s + 1)$ subsums of the sum (10) that contain at least two terms, and only those subsums may potentially be 0. Suppose S_0 is one such 0-subsum which is *minimal* in that it contains no further smaller 0-subsums. If all its terms are of the type $c_i\alpha_i^k$ for the same k, say $k = m$, then, by Corollary 3, that can occur for only finitely many values of m. Once m is fixed, then the equation $u_n = u_m$ has finitely many solutions in n by Corollary 4. If, on the contrary, S_0 contains a mixture of terms corresponding to the exponent m and the exponent n, then,

([1]) Nonnegativity of m and n is essential here since, as our proof will show, there can be infinitely many solutions in integers for some specific recurrences. For instance, $F_{2n+1} = F_{-2n-1}$ for all integers n, where $(F_n)_{n\in\mathbb{Z}}$ is the Fibonacci sequence.

assuming S_0 contains at least three terms, one exponent, say m, occurs at least in two terms. So, let us suppose that the 0-subsum S_0 contains the terms $c_i\alpha_i^m$, $c_j\alpha_j^m$ and $-c_k\alpha_k^n$ for some $i \neq j$. We may reiterate the reasoning made earlier, that is, divide the subsum through by $c_i\alpha_i^m$ and obtain an equation like (9) with the left-hand side containing the term $-(c_j/c_i)(\alpha_j/\alpha_i)^m$ as well as the term $(c_k/c_i)\alpha_k^n/\alpha_i^m$. By the minimality of our 0-subsum, we may again apply the argument from the beginning of this proof and see that S_0 may equal 0 for at most finitely many pairs of integers (m, n).

Thus, if there are to be infinitely many solutions (m, n) to the equation $u_m = u_n$ with m and n distinct and nonnegative, then infinitely often each term $c_i\alpha_i^m$ of (10) must be paired with a term $c_j\alpha_j^n$, and their difference be zero. Clearly, $i \neq j$, for $i = j$ would lead to α_i being a root of unity, a case we no longer need to consider by Lemma 7. That is, we have a permutation σ of $\{1, \dots, s\}$ with no fixed point such that

$$c_i\alpha_i^m - c_{\sigma(i)}\alpha_{\sigma(i)}^n = 0 \quad \text{for each } i = 1, \dots, s. \tag{11}$$

Note that there are finitely many such permutations of $\{1, \dots, s\}$. We show that there can be at most finitely many pairs of nonnegative integers (m, n) satisfying (11) for a given σ. Suppose (m_1, n_1) and (m_2, n_2) are two solutions in nonnegative integers of (11) for the same permutation σ. Then, for each $i \in \{1, \dots, s\}$, we have

$$c_i\alpha_i^{m_1} = c_{\sigma(i)}\alpha_{\sigma(i)}^{n_1} \quad \text{and} \quad c_i\alpha_i^{m_2} = c_{\sigma(i)}\alpha_{\sigma(i)}^{n_2}.$$

Hence, $\alpha_i^{m_1-m_2} = \alpha_{\sigma(i)}^{n_1-n_2}$. Put $m = m_1 - m_2$ and $n = n_1 - n_2$, and, after possibly relabelling the α_i's and assuming the orbit of $i = 1$ under the action of σ has length h, we have

$$\alpha_1^m = \alpha_2^n, \quad \alpha_2^m = \alpha_3^n, \quad \alpha_3^m = \alpha_4^n, \quad \dots, \quad \alpha_h^m = \alpha_1^n.$$

Therefore, $\alpha_1^{m^h} = \alpha_2^{nm^{h-1}} = \alpha_3^{n^2m^{h-2}} = \cdots = \alpha_1^{n^h}$. Thus, $\alpha_1^{m^h-n^h} = 1$. Hence, $m^h = n^h$ and either $m = n$, or $m = -n$. The case $m = n$ implies $(\alpha_1/\alpha_2)^m = 1$ and contradicts the nondegeneracy of (u_n).

If $m = -n$, then $(\alpha_1\alpha_2)^m = 1$. Thus, $\alpha_2 = \alpha_1^{-1}\zeta$, where $\zeta^m = 1$. But $c_1\alpha_1^{m_1} = c_2\alpha_2^{n_1}$. So $\alpha_1^{m_1+n_1} = (c_2/c_1)\zeta^{n_1}$. Therefore, $|\alpha_1|^{m_1+n_1} = |c_2/c_1|$. Not all roots α_i's can be of absolute value ≤ 1 or, by Kronecker's theorem [3], they would all be roots of unity. So choose α_1 of absolute value > 1. Now, if $|\alpha_1|^\ell = |c_2/c_1|$ for some integer ℓ, then m_1+n_1 must be equal to ℓ. However, only $\ell + 1$ pairs of *nonnegative* integers (m_1, n_1) satisfy $m_1 + n_1 = \ell$. ∎

Definition 9. Suppose $\alpha_1, \dots, \alpha_s$ are s fixed algebraic integers. Then given s positive integers $m_1, \dots, m_s$, we define $\Delta_{m_1,\dots,m_s}$ as the $(s+1)\times(s+1)$ determinant

$$\Delta_{m_1,\dots,m_s} := \begin{vmatrix} 1 & 1 & \dots & 1 & 1 \\ \alpha_1^{m_1} & \alpha_2^{m_1} & \dots & \alpha_s^{m_1} & 1 \\ \vdots & \dots & \dots & \dots & \vdots \\ \alpha_1^{m_s} & \alpha_2^{m_s} & \dots & \alpha_s^{m_s} & 1 \end{vmatrix}.$$

LEMMA 10. *Let $\alpha_1, \dots, \alpha_s$ be s nonzero algebraic integers such that neither of them, nor the ratio of any two of them, is a root of unity. Let $\mathcal{M}$ be a set of $M_s + s$ positive integers, where M_s is the constant defined in Proposition* 6 *and $\mathcal{O}$ is the ring of integers of the number field generated by the α_i's. Then there are s integers $m_1, \dots, m_s$ in $\mathcal{M}$ such that $\Delta_{m_1,\dots,m_s}$ is not zero.*

Proof. We proceed by induction on s. For $s = 1$, note that $\Delta_m = 1-\alpha_1^m \neq 0$ for any m in $\mathcal{M}$. For $s = 2$, choose any m_1 in $\mathcal{M}$ and note that expanding the determinant $\Delta_{m_1,m}$ along its third line yields $c_1\alpha_1^m + c_2\alpha_2^m + c_3$, where c_1, c_2 and c_3 are, up to sign, the corresponding minors. Thus if $\Delta_{m_1,m} = 0$, then $u_m := c_1\alpha_1^m + c_2\alpha_2^m = -c_3$. Since $c_2 = -\Delta_{m_1} \neq 0$, the sequence $u = (u_m)_{m\geq 0}$ is an honest nonzero linear recurrence, which is nondegenerate of order at most 2 over $\mathcal{O}$. Thus, by Proposition 6, there are at most M_2 values of m that satisfy $u_m = -c_3$. But $\mathcal{M} \setminus \{m_1\}$ contains $M_2 + 1$ integers, so there is an m in $\mathcal{M} \setminus \{m_1\}$ for which $\Delta_{m_1,m} \neq 0$. This argument may be repeated. Suppose the lemma holds up to $s - 1 \geq 2$. Let $\mathcal{M}$ contain $M_s + s$ positive integers. Since $M_s \geq M_{s-1}$, $\mathcal{M}$ contains strictly more than $M_{s-1} + (s-1)$ positive integers. By the inductive hypothesis, there are integers $m_1, \dots, m_{s-1}$ in $\mathcal{M}$ such that the determinant $\Delta_{m_1,\dots,m_{s-1}}$ is nonzero. Expanding the determinant $\Delta_{m_1,\dots,m_{s-1},m}$ along its last line and assuming it is 0, we get an equation of the type $u_m := \sum_{i=1}^{s} c_i\alpha_i^m = -c_{s+1}$. Again the sequence $u = (u_m)_{m\geq 0}$ is not identically 0 since $c_s = \pm\Delta_{m_1,\dots,m_{s-1}} \neq 0$ and u is a nondegenerate linear recurring sequence of order at most s over $\mathcal{O}$. Thus, by Proposition 6, at most M_s integer values of m may annihilate $u_m + c_{s+1}$. But $\mathcal{M} \setminus \{m_1, \dots, m_{s-1}\}$ contains $M_s + 1$ integers. One of them, say m_s, is such that $\Delta_{m_1,\dots,m_{s-1},m_s} \neq 0$. ∎

We will use a particular case of Theorem 2.3 of [1], which was proved by Evertse [2] and also by van der Poorten and Schlickewei [5].

PROPOSITION 11. *For any algebraic nondegenerate linear recurring sequence $u = (u_n)_{n\geq 0}$ and any $\epsilon > 0$, there exists a constant n_ϵ for which*

$$|u_n| \geq |\alpha_1|^{(1-\epsilon)n} \quad \text{if } n \geq n_\epsilon,$$

where α_1 is a root of maximal absolute value of the characteristic polynomial of u.

3. Proof of the main theorem. We address the first step of Romanoff's line of proof by noting that it suffices to prove a lower bound for the average number of representations.

Recall here that $r(n)$ is the cardinality of the set of pairs of nonnegative integers (j,k) such that $n = p_j + u_k$, p_j being the jth prime.

LEMMA 12. *Let $u = (u_n)_{n\geq 0}$ be a nondegenerate integral linear recurring sequence whose characteristic polynomial has distinct roots. Then*

$$\sum_{n=1}^{N} r(n) \gg N.$$

Proof. If s is the number of characteristic roots of $(u_n)_{n\geq 0}$ and c is s times the maximum of the $|c_i|$'s, where we assume u_n to have the representation (4), then we immediately have $|u_n| \leq c|\alpha_1|^n$, where $|\alpha_1| \geq |\alpha_i|$, $i = 2, \ldots, s$. Thus, the number of terms of u in an interval $[-N, N]$ is at least equal to k_N, where k_N is the largest integer n satisfying $c|\alpha_1|^n \leq N$. But k_N is asymptotically equal to $\log N/\log|\alpha_1|$ as N tends to infinity. Therefore,

$$\begin{aligned}\sum_{n=1}^{N} r(n) &\geq (\pi(2N/3) - \pi(N/3)) \cdot \#\{n;\, u_n \in [-N/3, N/3]\} \\ &\geq \frac{N}{3\log N} \cdot \log\left(\frac{N}{3}\right)((\log|\alpha_1|)^{-1} + o(1)) \quad (N \to \infty) \\ &\gg N,\end{aligned}$$

where in the second inequality above we used the prime number theorem. ∎

The second step of Romanoff's method is, as usual, the most involved step in proving the positive lower density of the sumset. We decompose it into several lemmas.

DEFINITION 13. Given an integral linear recurring sequence $u = (u_n)_{n\geq 0}$ and a prime p, we denote by $k(p)$ the period of u modulo p, that is, the (minimal) positive integer k such that $u_{n+k} \equiv u_n \pmod p$ for all integers $n \geq 0$. If y is an integer, then $\nu(y,p)$ denotes the number of appearances of $y \pmod p$ in an interval of length $k(p)$, i.e., the number of n's in $[1, k(p)]$ such that $u_n \equiv y \pmod p$. Then ν_p is the maximum over all y's of the $\nu(y,p)$'s. Finally, $z(p)$ denotes the ratio $k(p)/\nu_p$. Note that if u is of order s, then $k(p) \leq p^s$.

NOTATION. Let $u = (u_n)_{n\geq 0}$ be an integral nondegenerate linear recurring sequence with s distinct characteristic roots $\alpha_1, \ldots, \alpha_s$, where α_1 will always stand for a root of maximal absolute value. We will denote by n_0 a positive integer $\geq M_s$, where M_s was defined in Proposition 6, with the additional property that for all $n > n_0$, we have $u_n \neq u_m$ for all $m \neq n$.

Note that the existence of n_0 is guaranteed by Lemma 8 and Proposition 6. We then define $\kappa := 2(n_0 + s + 1)$.

DEFINITION 14. Let $u = (u_n)_{n\geq 0}$ be an integral nondegenerate linear recurring sequence with s distinct characteristic roots, so that the terms u_n admit the representation (4). We will say that a prime p is *u-irregular*, or just irregular, if it belongs to one of the three sets $\mathcal{B}_i$, $i = 1, 2, 3$, where

$$\begin{aligned}
&p \in \mathcal{B}_1 \quad \text{iff} \quad p < p_0 \text{ or } p \text{ divides } \prod_{i=1}^{s} c_i\alpha_i,\\
&p \in \mathcal{B}_2 \quad \text{iff} \quad p \notin \mathcal{B}_1 \text{ and } k(p) < p^{1/(s+2)},\\
&p \in \mathcal{B}_3 \quad \text{iff} \quad p \notin \mathcal{B}_1 \cup \mathcal{B}_2 \text{ and } \nu_p > \kappa\frac{k(p)}{p^{\beta}}, \text{ where } \beta = \frac{1}{s+3}.
\end{aligned}$$

The prime p_0 is the smallest prime p such that both the inequality

$$(p^s)^{\frac{s+4}{s+2}} > \kappa p^s$$

holds and p does not divide the discriminant of $P(x)$, the characteristic polynomial of u. Primes not in $\mathcal{B}_1 \cup \mathcal{B}_2 \cup \mathcal{B}_3$ are called *u-regular*, or just regular.

Thus, primes in $\mathcal{B}_2$ may be thought of as having a small period and those in $\mathcal{B}_3$ as having a high frequency residue.

LEMMA 15. *The sum of the reciprocals of irregular primes converges.*

Proof. Irregular primes in $\mathcal{B}_1$ are finitely many so clearly the lemma holds for such primes. Primes $p \leq T$ in $\mathcal{B}_2$ satisfy $k(p) < T^{1/(s+2)}$. We estimate their number by observing that

$$\begin{aligned}
2^{\#\{p\leq T;\, k(p)<T^{1/(s+2)}\}} &\leq \prod_{\substack{p\leq T\\ k(p)\leq T^{1/(s+2)}}} p \leq \prod_{n<T^{1/(s+2)}} |u_{n+n_0} - u_{n_0}|\\
&\leq \exp\Big(O\Big(\sum_{n<T^{1/(s+2)}} n\Big)\Big)\\
&= \exp(O(T^{2/(s+2)})),
\end{aligned}$$

so the counting function of such primes $p \leq T$ is $O(T^{2/(s+2)})$. In particular, as $2/(s+2) < 1$, the sum of their reciprocals converges by Abel summation. Here we used $0 < |u_{n+n_0} - u_{n_0}| \ll |\alpha_1|^{n+n_0} \ll e^{\lambda n}$ for some $\lambda > 0$.

It remains to see that the sum of the inverses of the primes in $\mathcal{B}_3$ also converges. Suppose first that $s = 1$. Then, as $p \nmid c_1\alpha_1$, $k(p)$ is the order of α_1 modulo p and $\nu_p = 1$. Thus, p being in $\mathcal{B}_3$ says that $k(p) < \kappa^{-1}p^{1/4} < p^{1/3}$ for p large enough. Hence, we may reconduct the argument used for primes in $\mathcal{B}_2$.

So we assume $s \geq 2$. Let p be in $\mathcal{B}_3$. Then slicing the interval $[0, k(p)-1]$ into subintervals $[0, p^\beta), [p^\beta, 2p^\beta)$, etc., of length not exceeding p^β, we get at most $\lfloor k(p)/p^\beta \rfloor + 1 < 2k(p)/p^\beta$ such subintervals. Since $\nu(y,p) > 2(n_0+s+1)k(p)/p^\beta$ for some y, it follows that there is one subinterval that catches at least n_0+s+1 solutions n to $u_n \equiv y \pmod p$. Thus, there are integers $n_1, n_1+\ell_1, \ldots, n_1+\ell_{n_0+s}$ such that $0 < \ell_1 < \cdots < \ell_{n_0+s} \in [1, p^\beta)$ and $u_{n_1} \equiv u_{n_1+\ell_i} \equiv y \pmod p$ for $i = 1, \ldots, n_0+s$. Because $n_0+s \geq M_s+s$, there is a subset of s integers $m_1 < \cdots < m_s$ among the ℓ_i's such that $\Delta_{m_1,\ldots,m_s}$ is not zero, by Lemma 10. Exploiting relation (4) and the fact that $u_{n_1}, u_{n_1+m_1}, \ldots, u_{n_1+m_s}$ are all congruent to y modulo p, we deduce that the linear system

$$\begin{cases} X_1 + X_2 + \cdots + X_s + X_{s+1} = 0, \\ \alpha_1^{m_1} X_1 + \alpha_2^{m_1} X_2 + \cdots + \alpha_s^{m_1} X_s + X_{s+1} = 0, \\ \quad\vdots \qquad\quad \vdots \qquad\qquad \cdots \qquad\quad \vdots \\ \alpha_1^{m_s} X_1 + \alpha_2^{m_s} X_2 + \cdots + \alpha_s^{m_s} X_s + X_{s+1} = 0 \end{cases}$$

has the solution

$$(X_1, X_2, \ldots, X_s, X_{s+1}) := (c_1\alpha_1^{n_1}, c_2\alpha_2^{n_1}, \ldots, c_s\alpha_s^{n_1}, -y)$$

modulo p, and this solution is nontrivial modulo p because $p > p_0$ implies that each c_i and each α_i, $i = 1, \ldots, s$, is invertible modulo p. Hence, letting π be any prime ideal above p in $\mathbb{K} = \mathbb{Q}(\alpha_1, \ldots, \alpha_s)$, the determinant of the system is 0 modulo π.

The principal ideal generated by p inside $\mathcal{O}_\mathbb{K}$, the ring of integers of $\mathbb{K}$, is the product of all π's in $\mathcal{O}_\mathbb{K}$ that lie above p, that is, its prime ideal decomposition is squarefree. Indeed, p does not divide the discriminant of $P(x)$. Therefore,

$$p = N_{\mathbb{K}/\mathbb{Q}}(\pi) \text{ divides the nonzero rational integer } N_{\mathbb{K}/\mathbb{Q}}(\Delta_{m_1,\ldots,m_s}).$$

Note that $\Delta_{m_1,\ldots,m_s}$ is the sum of $(s+1)!$ products each containing factors of the type $\alpha_i^{m_j}$. Since $m_j < p^\beta$, we see, assuming $p \leq T$, that $m_j < T^\beta$. Thus,

$$|\Delta_{m_1,\ldots,m_s}| = O((s+1)!|\alpha_1|^{sT^\beta}) = \exp(O(T^\beta)).$$

Further, since Galois conjugation of $\mathbb{K}$ over $\mathbb{Q}$ permutes $\alpha_1, \ldots, \alpha_s$, it follows that all conjugates of $|\Delta_{m_1,\ldots,m_s}|$ in $\mathbb{K}$ are also of size at most $\exp(O(T^\beta))$. Thus, the integer $|N_{\mathbb{K}/\mathbb{Q}}(\Delta_{m_1,\ldots,m_s})|$ is itself $\exp(O(T^\beta))$.

Since $m_1, \ldots, m_s$ may vary with each p in $\mathcal{B}_3$, we denote the dependence of $\Delta_{m_1,\ldots,m_s}$ on p by writing it as $\Delta_{m_1,\ldots,m_s}(p)$.

Hence, the product of the primes $p \leq T$ in $\mathcal{B}_3$ may be estimated as follows:

$$2^{\#\mathcal{B}_3(T)} \le \prod_{p\in\mathcal{B}_3(T)} p \le \prod_{p\in\mathcal{B}_3(T)} |N_{\mathbb{K}/\mathbb{Q}}(\Delta_{m_1,\dots,m_s}(p))|$$
$$\le \prod_{m_1<\cdots<m_s<T^\beta} \exp(O(T^\beta)) = \exp\Big(O\Big(T^\beta \cdot \sum_{m_1<\cdots<m_s<T^\beta} 1\Big)\Big)$$
$$= \exp(O(T^\beta)^{s+1}) = \exp(O(T^{\frac{s+1}{s+2}})).$$

So, the number of primes $\le T$ in $\mathcal{B}_3$ is $O(T^{\frac{s+1}{s+2}})$, therefore the sum of their reciprocals also converges. ∎

Lemma 16. *The series*

$$\sum_{n\ge1} \frac{\mu(n)^2}{nz(n)}$$

converges, where $z(n)$ is the maximum, over all prime factors p of n, of the $z(p)$'s.

Proof. Numbers n built entirely of u-irregular primes satisfy

$$\sum_{\substack{n\ge1\\ p|n\Rightarrow p\,\text{irregular}}} \frac{\mu(n)^2}{nz(n)} \le \sum_{\substack{n\ge1\\ p|n\Rightarrow p\,\text{irregular}}} \frac{\mu(n)^2}{n} \le \prod_{p\,\text{irregular}} \left(1+\frac{1}{p}\right) = O(1),$$

since the product $\prod_{p\,\text{irregular}}(1+1/p)$ converges iff the sum of the reciprocals of irregular primes converges, a convergence shown in Lemma 15.

We now consider integers n that have at least one regular prime factor. Then

$$\sum_{\substack{n\\ \exists p|n,\,p\,\text{regular}}} \frac{\mu(n)^2}{nz(n)} \le \sum_{p\,\text{regular}} \sum_{z(n)=z(p)} \frac{\mu(n)^2}{nz(n)} \le \sum_{p\,\text{regular}} \frac{1}{pz(p)} \sum_{z(n)\le z(p)} \frac{\mu(n)^2}{n}$$
$$\le \sum_{p\,\text{regular}} \frac{\kappa}{p^{\frac{3+s}{2+s}}} \Bigg(\prod_{q\,\text{irregular}} \left(1+\frac{1}{q}\right) \prod_{\substack{q\,\text{regular}\\ q<p^{s(s+4)}}} \left(1+\frac{1}{q}\right)\Bigg).$$

Indeed, if p is regular, then $p\notin\mathcal{B}_3$. So $\nu_p \le \kappa k(p)/p^\beta$, i.e., $z(p)\ge \kappa^{-1}p^\beta$, or

$$\frac{1}{pz(p)} \le \frac{\kappa}{p^{1+\beta}} = \frac{\kappa}{p^{\frac{s+3}{s+2}}}.$$

The explanation of the second product above is that if $q\ge p^{s(s+4)}$ is regular and appears in the factorization of an n with $z(n)\le z(p)$, then

$$z(q)\ge \kappa^{-1}q^\beta \ge \kappa^{-1}p^{\frac{s(s+4)}{s+2}} \ge \kappa^{-1}(p^s)^{\frac{s+4}{s+2}} > p^s \ge z(p),$$

since $p^s\ge k(p)\ge z(p)$. We have also used the hypothesis that, p being regular, $p\ge p_0$. But $z(q)>z(p)$ contradicts the hypothesis $z(n)\le z(p)$. Thus, regular prime factors of an n with $z(n)\le z(p)$ must be less than $p^{s(s+4)}$.

Hence, as the first product is $O(1)$ and the second, by Mertens' theorem, is $O(\log p)$, we have

$$\sum_{\substack{n \\ \exists p|n,\, p \text{ regular}}} \frac{\mu(n)^2}{nz(n)} \ll \sum_{p \text{ regular}} \frac{\log p}{p^{\frac{3+s}{2+s}}} \ll 1. \quad \blacksquare$$

LEMMA 17. *We have*

$$\sum_{n=1}^{N} r(n)^2 \ll N.$$

Proof. Following the proof of the comparable Lemma 3 of [4], we have

$$\sum_{n=1}^{N} r(n)^2 = \sum_{n=1}^{N} \sum_{\substack{p+u=n \\ p'+u'=n}} 1 \le \sum_{h\in[-N,N]} \Big(\sum_{\substack{p-p'=h \\ p,p'\in[2,N]}} 1\Big)\delta_N(h),$$

where

$$\delta_N(h) = \#\{(m,n) \in \mathbb{N}^2;\ h = u_m - u_n, u_m \text{ and } u_n \text{ belonging to } [-N,N]\}.$$

Put

$$\alpha_N(h) := \delta_N(h) \sum_{\substack{p-p'=h \\ p,p'\in[2,N]}} 1.$$

Note that, by Proposition 11, say with $\epsilon = 1/2$, $|u_n| \in [0,N]$ implies that $|\alpha_1|^{n/2} \in [0,N]$, if $n \ge n_\epsilon$. Thus, $n \le (2/\log|\alpha_1|)\log N$. Hence, $\delta_N(0)$ is of order $O(\log N)$ and

$$\alpha_N(0) \ll (\log N) \sum_{p\le N} 1 \ll N,$$

by the prime number theorem. Clearly, $\alpha_N(h) = \alpha_N(-h)$, so it suffices to show $\sum_{h\ge1} \alpha_N(h) \ll N$. If h is odd and positive and $p - p' = h$, then $p' = 2$. Since the number of n's such that $|u_n| \in [0,N]$ is $\ll \log N$ and, by Proposition 5, there are at most M values of m such that $u_m = u_n + h$, it follows that for h odd, we have $\delta_N(h) \ll \log N$. Hence,

$$\sum_{1\le h \text{ odd} \le N} \alpha_n(h) \ll (\log N) \sum_{\substack{1\le h\le N \\ 2+h=p}} 1 \ll N,$$

by the prime number theorem.

By Brun's combinatorial sieve, we know that for h even, the number of primes $p \le x$ such that $p + h$ is prime is

$$\ll \prod_{p|h} \left(1 + \frac{1}{p}\right) \cdot \frac{x}{\log^2 x}.$$

Thus,

$$\sum_{1\le h\,\text{even}\le N}\alpha_n(h) \ll \frac{N}{\log^2 N}\sum_{h>0}\sum_{d\mid h}\frac{\mu(d)^2}{d}\delta_N(h)$$
$$\le \frac{N}{\log^2 N}\sum_{d\ge 1}\frac{\mu(d)^2}{d}\sum_{\substack{d\mid h\\ h\le N}}\delta_N(h).$$

Now for $d>1$ choose a prime factor p of d such that $z(p)=z(d)$. Then

$$\sum_{\substack{d\mid h\\ h\le N}}\delta_N(h) = \#\{(u_m,u_n)\in[-N,N]^2;\, u_m\equiv u_n \pmod d\}$$
$$\le \#\{(u_m,u_n)\in[-N,N]^2;\, u_m\equiv u_n \pmod p\}.$$

The number of choices for u_n is $O(\log N)$ and, once n is chosen, u_m is equal to u_n modulo p at most ν_p times per period $k(p)$. So the number of choices for m is $\ll (\nu_p\log N)/k(p)$. That is,

$$\sum_{\substack{d\mid h\\ h\le N}}\delta_N(h) \ll \frac{\log^2 N}{z(d)}.$$

Therefore,

$$\sum_{1\le h\,\text{even}\le N}\alpha_n(h)\ll N \quad\text{if and only if}\quad \sum_{d\ge 1}\frac{\mu(d)^2}{dz(d)}\ \text{converges},$$

a convergence that was shown in Lemma 16. ∎

Our main theorem, Theorem 1, follows from Cauchy–Schwarz's inequality (6), Lemma 12 and Lemma 17.

References

[1] G. Everest, A. van der Poorten, I. Shparlinski and T. Ward, *Recurrence Sequences*, Amer. Math. Soc., Providence, RI, 2003.

[2] J.-H. Evertse, *On sums of S-units and linear recurrences*, Compos. Math. 53 (1984), 225–244.

[3] L. Kronecker, *Zwei Sätze über Gleichungen mit ganzzahligen Coefficienten*, J. Reine Angew. Math. 53 (1857), 173–175; Werke, Chelsea, New York, 1968, Vol. I, 105–108.

[4] K. S. E. Lee, *On the sum of a prime and a Fibonacci number*, Int. J. Number Theory 6 (2010), 1669–1676.

[5] A. J. van der Poorten and H. P. Schlickewei, *Additive relations in fields*, J. Austral. Math. Soc. Ser. A 51 (1991), 154–170.

[6] N. P. Romanoff, *Über einige Sätze der additiven Zalhentheorie*, Math. Ann. 109 (1934), 668–678.

Christian Ballot
L.M.N.O., CNRS UMR 6139
Université de Caen
F-14032 Caen Cedex, France
E-mail: christian.ballot@unicaen.fr

Florian Luca
Fundación Marcos Moshinsky
UNAM
Circuito Exterior, C.U., Apdo. Postal 70-543
México, D.F. 04510, México
E-mail: fluca@matmor.unam.mx

Received on 23.12.2012
and in revised form on 22.4.2013 (7299)