THE GREATEST PRIME FACTOR AND RECURRENT SEQUENCES

GREG BACK AND MIHAI CARAGIU

Abstract. The recurrent sequences considered in the present paper are prime sequences of the form $p_j = \text{gpf}(a_1 p_{j-1} + a_2 p_{j-2} + \cdots + a_d p_{j-d} + a_0)$, where for any integer $x \geq 2$, we denote by $\text{gpf}(x)$ the greatest prime factor of x. In the simple case of the 'GPF-Fibonacci' sequences corresponding to $d = 2$, $a_0 = 0$, and $a_1 = a_2 = 1$, we find that regardless of the initial conditions p_0 and p_1, all such sequences ultimately enter the cycle 7, 3, 5, 2. A computational exploration of the 'GPF-Tribonacci' analogue $d = 3$, $a_0 = 0$, and $a_1 = a_2 = a_3 = 1$ reveals four cycles of lengths, listed in the decreasing order of frequencies, 100, 212, 28 and 6, with the two larger cycles collecting more than 98% of the sequences as defined by the initial conditions p_0, p_1, and p_2. The paper concludes with a general ultimate periodicity conjecture and discusses its plausibility.

1. Introduction

For any integer $x \geq 2$, let $\text{gpf}(x)$ be the greatest prime factor of x. The idea of combining the (generalized) Fibonacci recursion

$$G_0 = u,\ G_1 = v,\ G_n = G_{n-1} + G_{n-2} \text{ for } n \geq 2$$

with the greatest prime factor function naturally leads to prime sequences $\{p_n\}$ of the following type:

$$p_0 = a,\ p_1 = b,\ p_n = \text{gpf}\,(p_{n-1} + p_{n-2}) \quad \text{for } n \geq 2. \tag{1}$$

Let us agree to call such sequences, "GPF-Fibonacci" sequences. The behavior of the greatest prime factor function has already been studied in various contexts, including polynomials [4, 5], arithmetic progressions [7], and integers in an interval [6]. In the present paper we show that all GPF-Fibonacci sequences are ultimately periodic, eventually entering the cycle 7,3,5,2. For example, if $a = 19$ and $b = 13$, the sequence is

$$19, 13, 2, 5, 7, 3, 5, 2, 7, 3, 5, 2, \ldots.$$

This is a special case of a general class of recurrent sequences that bring together the idea of a linear recursion and the greatest prime factor function, so that they satisfy a recurrence relation of the form

$$p_j = \text{gpf}(a_1 p_{j-1} + a_2 p_{j-2} + \cdots + a_d p_{j-d} + a_0). \tag{2}$$

We will present data suggesting very interesting facts about the Tribonacci analogues of sequences of type (1). A computational exploration of the 'GPF-Tribonacci' sequences obtained by setting $d = 3$, $a_0 = 0$, and $a_1 = a_2 = a_3 = 1$ in (2) reveals four cycles of lengths, listed in the decreasing order of frequencies, 100, 212, 28 and 6, with the two cycles of largest length collecting more than 98% of the sequences as defined by the initial conditions p_0, p_1, and p_2. The paper concludes with a brief discussion regarding a general conjecture on sequences given by (2).

2. The Main Result

The proof that all GPF-Fibonacci sequences (1) are ultimately periodic rests on the following simple fact.

Lemma 1. *Let p be a prime such that $p+2$ is composite. Then the set*

$$K_p := \{r | r \text{ prime}, \ r \leq p\}$$

of all primes up to and including p is closed under the binary operation

$$(r, s) \mapsto \operatorname{gpf}(r+s). \tag{3}$$

Proof. Since the operation (3) is idempotent, we only need to consider $r, s \in K_p$, with $r < s$. If both r, s are odd, then $r+s$ is even, so

$$\operatorname{gpf}(r+s) \leq (r+s)/2 < s \leq p,$$

and thus $\operatorname{gpf}(r+s) \in K_p$. On the other hand, if $r = 2$, then the fact that $\operatorname{gpf}(2+s) \leq p$ for all $s \in K_p$ follows from the fact that $p+2$ is composite. □

For a more detailed discussion on the properties of the commutative and non-associative magma [1] operation defined by (3) on the set of all primes, see [2].

The following is a direct consequence of Lemma 1.

Proposition 2. *All GPF-Fibonacci sequences are ultimately periodic.*

Proof. Pick up a prime $p \equiv 1 \pmod 3$ (this will ensure that $p+2$ is composite) such that $p \geq \max(a, b)$, where $a = p_0$ and $b = p_1$ are the initial two terms of (1). Then $\{a, b\} \subset K_p$. But by Lemma 1, K_p is gpf-closed, so that by applying it inductively we find that $p_n \in K_p$ for all $n \geq 0$. To conclude the proof, note that a recurrent sequence of order 2 with terms in a finite set is eventually periodic. □

Remark. *Note that all terms p_j of (1) satisfy $p_j \leq \max(a, b) + 4$. In fact (since 3 is the only prime p such that p, $p+2$, and $p+4$ are all primes), with the single exception of $\{a, b\} = \{2, 3\}$ they satisfy $p_j \leq \max(a, b) + 2$.*

Next we show that not only the GPF-Fibonacci sequences are ultimately periodic but that all of them eventually enter the same 4-cycle.

Theorem 3. *All GPF-Fibonacci sequences (1) in which $a \neq b$ eventually enter the same 4-cycle 7,3,5,2.*

Proof. From Proposition 2, such a sequence is ultimately periodic. The cycle length cannot be 1 (this happens if and only if $a = b$). The cycle length cannot be 2 either, since if the cycle consists of just p, q with $p \neq q$, then $\operatorname{gpf}(p+q)$ cannot possibly be in the set $\{p, q\}$. Thus, the limit cycle is of the form $p_1, p_2, \ldots, p_k$ with $k \geq 3$ and $\operatorname{gpf}(p_j + p_{j+1}) = p_{j+2}$ for all $j = 1, 2, \ldots, k$, where $p_{k+1} = p_1$, etc. In what follows, each of the subscripts j in $\{p_j\}$ will be considered large enough so that all of their shifts considered in the proof will keep us in the cycle $p_1, p_2, \ldots, p_k$. Let $M = p_i$ be the largest element in the cycle. The cycle elements p_{i-1} and p_{i-2} are both distinct (otherwise the cycle would have length 1) and strictly smaller than M, otherwise, if exactly one out of p_{i-2}, p_{i-1} is M, then $\operatorname{gpf}(p_{i-2} + p_{i-1})$ cannot be equal to $M = p_i$. Moreover, one out of p_{i-2}, p_{i-1} must be equal to 2. Indeed, if both of them would be odd (and, as we already know, distinct and smaller than M), then $M = p_i = \operatorname{gpf}(p_{i-2} + p_{i-1}) \leq (p_{i-2} + p_{i-1})/2 < M$, a contradiction. At this moment we distinguish two cases: either $p_{i-2} = p$ and $p_{i-1} = 2$ (Case I), or $p_{i-2} = 2$ and $p_{i-1} = p$ (Case II), where $2 < p < M$ and $M = p + 2$ (which easily follows in both cases).

Case I. $p_{i-2} = 2$ and $p_{i-1} = p$ (odd), with $p + 2 = M$. Let $p_{i-3} = q$. Since $q \leq M = p + 2$ and $\mathrm{gpf}(q + 2) = \mathrm{gpf}(p_{i-3} + p_{i-2}) = p_{i-1} = p$, it follows that $p - 2 \leq q \leq p + 2$. But clearly $q \notin \{2, p\}$, so $q \in \{p - 2, p + 2\}$. The case $q = p + 2$ can be ruled out, since then $p = \mathrm{gpf}(q + 2) = \mathrm{gpf}(p + 4)$ which cannot happen. Therefore $p_{i-3} = p - 2$ and thus at this moment we get $p_{i-3} = p - 2$, $p_{i-2} = p$ and $p_i = p + 2$ in an arithmetic progresion of primes with difference 2, which necessarily implies $p_{i-3} = p - 2 = 3$, $p_{i-1} = p = 5$, and $M = p_i = 7$. Iterating the sequence we immediately enter the cycle 7,3,5,2, which indicates that p_{i-3}, p_{i-2}, p_{i-1} are not yet in the limit cycle. In fact, Case I never happens, and the problem reduces to Case II, which will be considered next.

Case II. $p_{i-2} = p$ (odd) and $p_{i-1} = 2$, with $p_i = M = p + 2$. If $p \leq 11$ we can verify by direct calculation that we eventually enter the cycle 7,3,5,2. Let $p \geq 13$, in which case, since p and $p + 2$ are primes, $p + 4$ is composite (divisible by 3), so that

$$p_{i+1} = \mathrm{gpf}(p_{i-1} + p_i) = \mathrm{gpf}(p + 4) \leq (p + 4)/3. \tag{4}$$

With $p_i = p + 2$ and p_{i+1} odd, we have, by using (4),

$$p_{i+2} = \mathrm{gpf}(p_i + p_{i+1}) \leq (p_i + p_{i+1})/2 \leq (4p + 10)/6. \tag{5}$$

From the remark following Proposition 2 in conjunction with (4) and (5), it follows that

$$p_j \leq \max(p_{i+1}, p_{i+2}) + 4 \leq \frac{4p + 34}{6} \tag{6}$$

for all $j > i$. From (6) together with the assumption $p \geq 13$ it follows that $p_j < p + 2 = M$ for all $j > i$, which is a contradiction. As a result, p must be at most 11, and in fact $M = 7$ and $p = 5$. This concludes the proof of Theorem 3. □

Remark. *Note that in* [2] *it was proved that every set of primes with more than 1 element that is closed under the greatest prime factor operation must contain the set* $\{2, 3, 5, 7\}$*, which is itself closed (that is, a submagma which is included in every other submagma that is not a singleton). By applying this we could have strengthened the statement in the beginning of the proof of Theorem 3 from the length of a period being at least 3 to the length of a period being at least 4 (a period of length 3 would produce a nontrivial submagma with 3 elements).*

3. A Tribonacci Surprise and Conjecture

The Tribonacci analogue for sequences of type (1) are sequences defined by recurrence relations of the following form:

$$p_0 = a,\ p_1 = b,\ p_2 = c,\ p_n = \mathrm{gpf}\left(p_{n-1} + p_{n-2} + p_{n-3}\right) \quad \text{for } n \geq 3. \tag{7}$$

Concerning these sequences we found computational evidence revealing intriguing patterns. First of all, a Monte Carlo-like computer experiment supports the following conjecture.

Conjecture 1. All GPF-Tribonacci sequences (7) are ultimately periodic.

This, in conjunction with the distribution of the lengths of the periods, is a surprising fact. Indeed, other than the trivial case of a period 1 (corresponding to the choice $a = b = c$), an analysis of 100,000 randomly generated triples (a, b, c) with components primes less than 1000 suggests a surprising distribution of the spectrum of the periods: in 98.78% of cases, *the periods are 100 and 212.* More precisely, we obtained a period of 100 in 74.73% of cases, a period of 212 in 24.05% of cases, a period of 28 in 0.76% of cases, and a period of 6 in 0.47% of cases. Moreover, in all cases we found the same cycle for each sequence with a period of a given length. As an example of such a sequence with a (most frequent) period of 100, one can take $(a, b, c) = (5, 13, 7)$, in which case the sequence evolves as follows:

[5, 13, 7, 5, 5, 17, 3, 5, 5, 13, 23, 41, 11, 5, 19, 7, 31, 19, 19, 23, 61, 103, 17, 181, 43, 241, 31, 7, 31, 23, 61, 23, 107, 191, 107, 5, 101, 71, 59, 11, 47, 13, 71, 131, 43, 7, 181, 11, 199, 23, 233, 13, 269, 103, 11, 383, 71, 31, 97, 199, 109, 5, 313, 61, 379, 251, 691, 1321, 73, 139, 73, 19, 11, 103, 19, 19, 47, 17, 83, 7, 107, 197, 311, 41, 61, 59, 23, 13, 19, 11, 43, 73, 127, 3, 29, 53, 17, 11, 3, 31], 5, 13, 7, ...

If $(a, b, c) = (31, 13, 7)$, the sequence is periodic with the period

[31, 13, 7, 17, 37, 61, 23, 11, 19, 53, 83, 31, 167, 281, 479, 103, 863, 17, 983, 23, 31, 61, 23, 23, 107, 17, 7, 131], 31, 13, 7,. . .,

while if $(a, b, c) = (59, 43, 37)$, the sequence is periodic with the period:

[59, 43, 37, 139, 73, 83], 59, 43, 37,. . ..

It is not always (and most time it is not) the case that the period begins right away. In most sequences of the form (7), the cycle entrance occurs later. As an example, we can provide the case of the initial conditions $(a, b, c) = (17, 31, 41)$. The recursion (7) will generate the above sequence of period 100, but the entrance in the period does not happen until the 535th term. In the collection of sequences considered above, most had an entrance point before the 700th term. There is also an interesting low in the interval (200,300). Note that a GPF-Tribonacci sequence can attain relatively high values even if the initial terms a, b, c are relatively small.

Finally, one might ask whether or not a uniform bound L can be found such that the terms of (7) satisfy

$$p_n \leq L \cdot \max(a, b, c) \quad \text{for all } n. \tag{8}$$

Of course, if (8) is true, then Conjecture 1 would be proved.

4. Epilogue: A General Ultimate Periodicity Conjecture

To conclude, let us go back to the most general class of prime sequences satisfying (2). For the sequences of that type, we will formulate the following conjecture.

Conjecture 2. (The *General Ultimate Periodicity Conjecture*). Every sequence of primes $\{p_j\}$ satisfying a recurrence relation (2) is ultimately periodic.

Concerning the plausibility of Conjecture 2, in addition to the material incorporated to the present paper, we refer to the class of prime sequences ('GPF sequences') introduced in [3], satisfying (2) with $d = 1$, but with $a_0 > 0$ (as was not the case for all sequences in this paper). That is,

$$p_j = \text{gpf}(a_1 p_{j-1} + a_0), \tag{9}$$

where a_0, a_1 are positive integers. The sequences (9) are ultimately periodic if $a_1 = 1$ [3] and, more generally, whenever a_1 divides a_0 [2]. Other computer checks suggested to us that the

General Ultimate Periodicity Conjecture may be true, and we believe it is a project worth being investigated.

References

[1] N. Bourbaki, *Algebra I: Chapters 1-3*, Vol. 1, Springer, 1998.
[2] M. Caragiu and G. Back, *The greatest prime factor and related magmas*, JP J. of Algebra, Number Theory and Appl., **15** (2009), 127–136.
[3] M. Caragiu and L. Scheckelhoff, *The greatest prime factor and related sequences*, JP J. of Algebra, Number Theory and Appl., **6** (2006), 403–409.
[4] C. Hooley, *On the greatest prime factor of a quadratic polynomial*, Acta Math., **117** (1967), 281–299.
[5] C. Hooley, *On the greatest prime factor of a cubic polynomial*, J. Reine Angew. Math., **303/304** (1978), 21-50.
[6] H. Q. Liu, *The greatest prime factor of the integers in an interval*, Acta Arith., **65** (1993), 301–328.
[7] T. N. Shorey and R. Tijdeman, *On the Greatest Prime Factor of an Arithmetical Progression. A Tribute to Paul Erdös*, Cambridge Univ. Press, Cambridge (1990), 385–389.

MSC2010: 11B37, 11A41

Department of Mathematics and Statistics, Ohio Northern University, Ada, OH 45810
E-mail address: gtback@ieee.org

Department of Mathematics and Statistics, Ohio Northern University, Ada, OH 45810
E-mail address: m-caragiu1@onu.edu