ON DUCCI SEQUENCES WITH PRIMES

MIHAI CARAGIU, ALEXANDRU ZAHARESCU, AND MOHAMMAD ZAKI

ABSTRACT. We introduce an analogue of the Ducci game that involves \(d\)-tuples of prime numbers subjected to the iteration \(G\) sending such a \(d\)-tuple \((p_1, p_2, \ldots, p_d)\) into \((\text{gpf}(p_1 + p_2), \text{gpf}(p_2 + p_3), \ldots, \text{gpf}(p_d + p_1))\), where for any \(x \geq 1\), \(\text{gpf}(x)\) represents the greatest prime factor of the integer \(x\). We show that the iteration of \(G\) always leads into a limit cycle \(C\). Moreover, if \(C\) has length greater than one, then not only every vector in \(C\) has all components \(p_i\) for any investigation into the lengths of the limit cycles of \(G\) would necessarily involve \((x_0, x_1, \ldots, x_{d-1}) = (|x_0 - x_1|, |x_1 - x_2|, \ldots, |x_{d-1} - x_0|)\).

For example, if \(d\) is a power of 2, iterating the above map always leads to the null vector. For an arbitrary \(d\), it is known that \(\phi\) leads into limit cycles in which the vectors are essentially binary (the components of each \(d\)-tuple in a cycle being either 0 or some constant \(c\)), so that any investigation into the lengths of the limit cycles of \(\phi\) would necessarily involve its binary reduction \(\overline{\phi}_d : \mathbb{F}_2^d \to \mathbb{F}_2^d\) given by \(\overline{\phi}(u_0, u_1, \ldots, u_{d-1}) = (u_0 + u_1, u_1 + u_2, \ldots, u_{d-1} + u_0)\).

Many outstanding problems involve the number of iterations until the null vector is reached, the lengths of the Ducci cycles and the asymptotic growth of the number of cycles with distinct lengths, and generalized Ducci-type mappings incorporating various weights. Analogues of the integer/binary Ducci problem have been suggested and investigated in various contexts such as real numbers, matrices and multi-dimensional arrays, abelian groups, algebraic numbers, \(p\)-adic integers, etc.

In the present paper we will introduce a new analogue of the Ducci problem that involves prime numbers and the greatest prime factor function. The reason for considering this analogue lies in a series of intriguing recent results involving the “greatest prime factor sequences” \([1, 12]\). These are prime sequences \(\{x_n\}\) satisfying a recurrence relation of the form

\[x_n = \text{gpf}(a_1 x_{n-1} + a_2 x_{n-2} + \cdots + a_k x_{n-k} + b),\] (1)

where \(a_i > 0\) and \(b \geq 0\) are integers, while \(\text{gpf}(x)\) represents the greatest prime factor of the positive integer \(x\) (with the convention \(\text{gpf}(1) = 1\)). It was conjectured [1] that all prime sequences satisfying a recurrence of the form (1) are ultimately periodic. Computational evidence supports this ultimate periodicity conjecture and special cases have been proved, such as the case of prime sequences satisfying \(x_n = \text{gpf}(ax_{n-1} + b)\) where \(a\) divides \(b\) \([11, 12]\), and the case of the “GPF-Fibonacci” sequences [1] satisfying \(x_n = \text{gpf}(x_{n-1} + x_{n-2})\). All
ON DUCCI SEQUENCES WITH PRIMES

GPF-Fibonacci sequences eventually enter the unique limit cycle $(3,5,2,7)$. However, a computer search [1] revealed multiple limit cycles of lengths 100, 212, 28 and 6 for the “GPF-Tribonacci” sequences, satisfying

\[ x_n = \text{gpf}(x_{n-1} + x_{n-2} + x_{n-3}). \]

We considered this was enough to warrant a consideration of a “GPF-Ducci” analogue of the classical Ducci game. The recursion is amazingly simple, to transform a finite sequence of primes we simply take the greatest prime factor of the sums of nearest neighbors. For example, the vector \( v_0 = (5, 103, 7, 23) \) will be transformed into

\[ v_1 = (3, 11, 5, 7), \]

since 3 is the greatest prime factor of 5+103, 11 is the greatest prime factor of 103+7, 5 is the greatest prime factor of 7+23, and finally 7 is the greatest prime factor of 23+5 (note that we assume periodic boundary conditions for the vectors involved in the process). If we continue the iteration we will get \( v_2 = (7, 2, 3, 5), v_3 = (3, 5, 2, 3), v_4 = (2, 7, 5, 3), v_5 = (3, 3, 2, 5), v_6 = (3, 5, 7, 2), v_7 = (2, 3, 3, 5), v_8 = (5, 3, 2, 7), v_9 = (2, 5, 3, 3), \) while \( v_{10} = (7, 2, 3, 5) = v_2 \) signals the entrance into a limit cycle of length 8, with the set of components of the vectors in the limit cycle being precisely the set of the first four primes, \( \{2, 3, 5, 7\} \). In the present paper we will prove that this is not accidental: indeed, for every vector size, the “GPF-Ducci” iteration ultimately enters a limit cycle with all vector components in \( \{2, 3, 5, 7\} \). Conversely, every element of \( \{2, 3, 5, 7\} \) appears in one of the vectors in the limit cycle. This establishes an interesting new analogue of the classical Ducci process. Moreover, a detailed cycle length analysis for vector lengths 3 through 8 will be obtained through exhaustive computer search.

2. THE GPF-DUCCI RECURSION

Let \( P \) be the set of all primes, and let \( d \geq 2 \) be an integer. Our ‘GPF-Ducci’ map

\[ G : P^d \to P^d \]

will be defined by

\[ G(p_1, p_2, \ldots, p_d) = (\text{gpf}(p_1 + p_2), \text{gpf}(p_2 + p_3), \ldots, \text{gpf}(p_d + p_1)) \]

for every \((p_1, \ldots, p_d) \in P_d\). Let \( P_0 := \{2, 3, 5, 7\} \). The main results that will be proved in the present paper are summarized as follows:

- The iteration of \( G \) always leads into a limit cycle \( C \).
- If \( C \) has length greater than 1, every vector in \( C \) has all components in \( P_0 \).
- If \( C \) has length greater than 1, then every element of \( P_0 \) is the component of some vector appearing in \( C \).

The iterations of the GPF-Ducci mapping (2) provide us with a new variation on the Ducci theme, which gives an additional ‘arithmetic touch’ to the Ducci problem. The following elementary result may be thought of as an analogue of the classical result of ultimate periodicity of the Ducci game. It provides a quick argument that all GPF-Ducci iterations ultimately enter a cycle, though without providing any details on the cycle structure.

**Proposition 1.** Every ‘GPF-Ducci’ iteration is ultimately periodic.

**Proof.** Let \((p_1, \ldots, p_d)\) be the initial prime \( d \)-tuple and let \( p \) be a prime such that \( p \equiv 1 \pmod{3} \) and such that \( p \geq \max(p_1, \ldots, p_d) \). One can easily see that the set \( K_p := \{ r \mid r \text{ prime, } r \leq p \} \) is closed under the binary operation \((r, s) \mapsto \text{gpf}(r + s)\). Thus, it follows that the components
of all \( d \)-tuples generated by the GPF-Ducci iteration starting from \((p_1, \ldots, p_d)\) are bounded from above by \( p \), which has ultimate periodicity as an immediate consequence, which proves the proposition. 

For more algebraic properties of the greatest prime factor operation

\[
(r, s) \mapsto \text{gpf}(r + s)
\]

defined on the set of primes, with applications to sequences, we refer to [1] and [10]. Characterizing the cycle structure of the GPF-Ducci iteration is an interesting problem. Note that a cycle of length 1 must consist of a constant \( d \)-tuple \((p, p, \ldots, p)\). If \( d \) is even, for example, every \( d \)-tuple of the form \((p, q, p, q, \ldots, p, q)\) has the property of being mapped into a cycle of length 1 (however, it is possible to find different kinds of \( d \)-tuples with this property, as shown by the \((5, 17, 71, 61)\) 4-tuple). Computational evidence suggested to us a nice result that will be proved in Theorem 1 below: namely, all limit cycles of the GPF-Ducci iteration of lengths greater than 1 contain only vectors with components belonging to \( P_0 \) – which is itself closed under (3).

Since every sequence of ‘GPF-Ducci’ iterates is ultimately periodic, for every “seed” \( X = (p_1, \ldots, p_d) \in P^d \), there exist \( L = L_X \) and \( n_0 = n_X \) such that for all \( n \geq n_0 \), \( G^{n+L}(X) = G^n(X) \). Our main result might be thought of as a quaternary analogue of the classical result on the binary character of the vectors in a limit cycle of the classical Ducci iteration.

**Theorem 2.** Let \( X \in P^d \) with \( L_X > 1 \). Then for all \( n \geq n_X \) the components of \( G^n(X) \) belong to \( P_0 \). Moreover, each element of \( P_0 \) appears in \( G^n(X) \) for some \( n \geq n_X \).

A complete computational verification of Theorem 2 for small dimensions \((d \leq 8)\) is detailed in the last section of the paper.

### 3. Proof of the Main Result

For any \( k \) with \( 1 \leq k \leq L = L_X \), let us denote

\[
G^{n_0 + k - 1}(X) := (p_{k,1}, \ldots, p_{k,d}).
\]

Let \( A \) be the \( L \times d \) matrix in which the \( k \)th row is \( G^{n_0 + k - 1}(X) \). Since \( L > 1 \), then \((p_{1,1}, \ldots, p_{1,d})\) is not a constant vector. The cycle \( C \) of the GPF-Ducci iteration thus consists of the rows of the matrix \( A \).

**Proof of Theorem 2.** The proof of our main result will be a direct consequence of the following set of four lemmas.

**Lemma 1.** Let \( q \) be the largest entry in the matrix \( A \). Then \( q \) is odd, and if the primes \( a \) and \( b \) are consecutive entries in a row of \( A \) producing \( q = \text{gpf}(a + b) \) in the immediately following row, one of the following holds true: (i) \( a = b = q \), or (ii) \( q - 2 \) is a prime and either \((a, b) = (2, q - 2)\) or \((a, b) = (q - 2, 2)\).

**Proof.** Indeed \( q \) is odd, since if \( q = 2 \) then \( L_X = 1 \), in contradiction with the assumption of Theorem 2. Assume that (i) is not true, that is either \( a \neq q \), or \( b \neq q \). If \( a \neq q \) and \( b = q \), then \( \text{gpf}(a + b) \neq q \) (otherwise \( q \) will divide \( a \)), which is a contradiction. Similarly we rule out the possibility of \( b \neq q \) and \( a = q \). Also, the case in which both \( a \) and \( b \) are odd primes not equal to \( q \) may be ruled out, since then \( \text{gpf}(a + b) \leq \frac{a+b}{2} < q \). Therefore either \( a = 2 \), or \( b = 2 \). If \( a = 2 \) then \( b \) must be an odd prime. If \( 2 + b \) is not a prime, then \( \text{gpf}(a + b) < 2 + b \leq q \), which contradicts the assumption \( \text{gpf}(a + b) = q \). Therefore \( 2 + b \) is a prime, in which case \( a = 2 \)
and \( b = q - 2 \) is a prime. A similar argument may be made in the case \( b = 2 \), in which case it will follow that \( a = q - 2 \) is prime. Thus we proved that (ii) holds true.

**Lemma 2.** Let \( p := q - 2 \). Then \( p \) must occur in \( A \).

**Proof.** Indeed, from Lemma 1, if \( p \) does not appear in the matrix \( A \) then \( q \) appears in a row exactly when there are consecutive occurrences of \( q \) in the previous row. Since \( q \) does appear in \( A \), and since the cycle matrix \( A \) is subjected to periodic boundary conditions, \( q \) must appear in each row of \( A \). If \( p \) is not an element of \( A \), applying \( G \) to any row of \( A \) produces a strictly smaller number of \( q \)'s in the next row. However, the cyclic structure of \( A \) makes this impossible (the function “number of \( q \)'s in the \( n \)th row” would be periodic and strictly decreasing). \( \Box \)

**Lemma 3.** The largest entry \( q \) of the limit cycle matrix \( A \) satisfies \( q \leq 7 \).

**Proof.** Assume on the contrary that \( q \geq 11 \). Then \( p \) is odd and \( p - 2 = q - 4 \) is not a prime. Let the primes \( a \) and \( b \) be two neighboring entries in a row of \( A \) such that in the next row they will generate \( \text{gpf}(a+b) = p \). We will show that both \( a \) and \( b \) must be equal to \( p \) by showing that all other options will lead to \( \text{gpf}(a+b) \neq p \).

(i) Let \( a \neq p \) and \( b = p \). Then \( \text{gpf}(a+b) \neq p \) (otherwise \( p \) will divide the prime \( a \)). Similarly, \( \text{gpf}(a+b) \neq p \) follows in the case \( a = p \) and \( b \neq p \).

(ii) Let \( a \) and \( b \) be both odd and not equal to \( p \). If \( a < p \) and \( b < p \), then \( a+b \) is even and \( \text{gpf}(a+b) \leq \frac{a+b}{2} < p \), so \( \text{gpf}(a+b) \neq p \). If \( a > p \) and \( b > p \), then \( a = b = q \), and so \( \text{gpf}(a+b) \neq p \). If \( a < p \) and \( b > p \), then \( b = q \) and \( a < p - 2 \) (since \( p - 2 \) is not a prime), in which case \( \text{gpf}(a+b) \leq \frac{a+b}{3} < \frac{p-2}{2} = p \). The symmetric case \( b < p \) and \( a > p \) may be dealt with in a similar way.

(iii) If \( a = b = 2 \), then \( \text{gpf}(a+b) = 2 \neq p \) (\( p \) is odd).

(iv) Let \( a = 2 \) and \( b \) an odd prime. Since \( p - 2 \) is not a prime, \( b \neq p - 2 \) and so \( p \neq b + 2 \). Thus, if \( 2 + b \) is a prime, \( \text{gpf}(a+b) = 2 + b \neq p \), while if \( 2 + b \) is not a prime, then \( \text{gpf}(a+b) \leq \frac{2+b}{3} = \frac{2+q+2}{3} = \frac{4+q}{3} < p \), and hence, \( \text{gpf}(a+b) \neq p \).

We conclude that if \( \text{gpf}(a+b) = p \), then both \( a \) and \( b \) must equal \( p \). That is, \( p \) appears in a row exactly when there are consecutive occurrences of \( p \) in the previous row. Since \( p \) does appear in \( A \) by Lemma 2, and due to the cyclic structure of \( A \), it follows that \( p \) must appear in each row of \( A \). Since \( p - 2 \) does not appear in the matrix \( A \), applying \( G \) to any row of \( A \) produces a strictly smaller number of \( p \)'s in the next row. Thus, the number of occurrences of \( p \) strictly decreases in the subsequent rows. As in the proof of Lemma 2, the cyclic structure of \( A \) makes this impossible. Therefore, \( q \leq 7 \) (so that every entry in the cycle matrix \( A \) is either \( 2 \), \( 3 \), \( 5 \), or \( 7 \)). This completes the proof Lemma 3.

Note that Lemma 3 may be restated as follows: for all \( n \geq n_X \) we have \( G^n(X) \in P^d_0 \).

**Lemma 4.** Each one of the primes \( 2, 3, 5, \) and \( 7 \) is an entry of the matrix \( A \).

**Proof.** First, we show that if \( 3 \) appears in some row of \( A \), then \( 5 \) must appear in \( A \). Indeed, if \( 3 \) appears in a row of \( A \), then that row has entries other than \( 3 \) (recall that \( L > 1 \)). If there is a \( 5 \) in the row, then we are done. If either \( 3 \) and \( 7 \), or \( 3 \) and \( 2 \) appear as nearest neighbors in the row, then \( 5 \) will appear in the next row.

Next, we show that if \( 7 \) appears in some row of \( A \), then \( 3 \) must appear in \( A \). Indeed, if \( 7 \) appears in a row of \( A \), since \( L > 1 \), then that row has entries other than \( 7 \). If there is a \( 3 \) in this row, then we are done. If either \( 7 \) and \( 5 \), or \( 7 \) and \( 2 \) appear as nearest neighbors in this row, then \( 3 \) will appear in the next row. Therefore, if \( 7 \) appears in some row of \( A \), then \( 5 \) and \( 3 \) must appear in \( A \).

FEBRUARY 2014 35
We are now going to show that if 7, 5 and 3 appear in A, then 2 must appear too: for this it will be enough to show that 3 and 5 appear as nearest neighbors in some row of A. Consider a row of A containing a 3, and assume that 2 is not in this row (otherwise there is nothing to prove). If 5 does not appear next to 3, then that row must contain one of the segments 3 3 7, 7 3 3, 7 3 7. From the first two segments we get either 3 5 or 5 3 in the next row. From the segment 7 3 7 we get 5 5 in the next row. Thus we either get 5 5 7 or 7 5 5 if there is no 3 next to 5 in A. But then in the subsequent row we do get 5 3 or 3 5. Therefore if 7 appears in some row of A, then all three numbers 5, 3 and 2 must appear in A. The proof of Lemma 4 (and hence of Theorem 1) will be completed if we prove that 7 appears in A.

Suppose 7 does not appear in A. Then 2 and 5 can’t appear next to each other in a row, since 7 will be immediately generated in the next row. If 2 and 5 appear in a row separated by 3’s, then it is easy to see that a repeated application of G to the particular segment of the form “2 3 3 . . . 3 5” or “5 3 3 . . . 3 2” ultimately produces a 7. Therefore we cannot have 2 and 5 in the same row. If a (necessarily non-constant) row consists of 3’s and 5’s, then 3 and 5 must appear next to each other somewhere in that row. Then, since 7 is not an entry of A, a segment of the form 3 5 or 5 3 may only be produced, via G, from a segment of the form 3 3 2 or 2 3 3 in the previous row, which at its turn may only be a subsegment of either one out of 3 3 2 2, 3 3 2 3, 2 2 3 3, or 3 2 3 3 in the same row. In the first and third case we get 2 and 5 after applying G once. In the second and fourth case we get 2 and 5 after applying G twice. Hence, 3 and 5 can’t appear in the same row. Lastly, the possibility of a non-constant row consisting of 2 and 3 only may be ruled out, since that would imply that the previous row contains both 3 and 5 (only the segments 3 5 and 5 3 may generate a 2 via G).

To summarize, if 7 does not appear in A (i.e., each entry of A is either 2, 3 or 5) and if \( L > 1 \), we proved that 2 and 5 cannot appear in the same row, and yet there is no row of A consisting of 3 and 5 only, and no row of A consisting of 2 and 3 only. This contradiction shows that if \( L > 1 \) then 7 must appear in A. This concludes the proof of Lemma 4, and hence, together with the previous three Lemmas, of Theorem 2.

4. Cycle Length Analysis: Computational Results

As a consequence of Theorem 2, in a computer-assisted search for the lengths of the non-trivial cycles of the GPF-Ducci map \( G \) on \( P_d \), we may restrict the search to the iterates of the \( 4^d \) initial vectors \( X \in P_0^d \) (with the understanding that other trivial cycles, or limit cycles of length 1, may be obtained for other choices of \( X \)). In our analysis of the distribution of the limit cycle lengths corresponding to the \( 4^d \) values of \( X \in P_0^d \) we obtained the following numerical results for \( 3 \leq d \leq 8 \).

- Out of 64 possible seeds \( X \in P_0^3 \), \( G \) leads 4 times into a limit cycle of length 1 and 60 times into a limit cycle of length 12;
- Out of 256 possible seeds \( X \in P_0^4 \), \( G \) leads 232 times into a limit cycle of length 1, 4 times into a limit cycle of length 2, 4 times into a limit cycle of length 4, and 16 times into a limit cycle of length 8;
- Out of 1024 possible seeds \( X \in P_0^5 \), \( G \) leads 4 times into a limit cycle of length 1, 70 times into a limit cycle of length 20, 120 times into a limit cycle of length 30, and 830 times into a limit cycle of length 40;
- Out of 4096 possible seeds \( X \in P_0^6 \), \( G \) leads 1204 times into a limit cycle of length 1, 1428 times into a limit cycle of length 6, 1116 times into a limit cycle of length 12, 174 times into a limit cycle of length 27, and 174 times into a limit cycle of length 54;
ON DUCCI SEQUENCES WITH PRIMES

• Out of 16384 possible seeds $X \in P_0^7$, $G$ leads 4 times into a limit cycle of length 1, 196 times into a limit cycle of length 4, 196 times into a limit cycle of length 28, 3528 times into a limit cycle of length 126, and 12460 times into a limit cycle of length 168;

• Out of 65536 possible seeds $X \in P_0^8$, $G$ leads 19528 times into a limit cycle of length 1, 4 times into a limit cycle of length 2, 4 times into a limit cycle of length 4, 12496 times into a limit cycle of length 8, 14808 times into a limit cycle of length 16, 14816 times into a limit cycle of length 32, 240 times into a limit cycle of length 48, and 3640 times into a limit cycle of length 80.

The next computer-generated supporting data provides explicit examples of choices of the initial vector $X \in P_0^d$ producing a non-trivial cycle of a given length $L_X$:

• For $d = 3$, $L_X = 12$ for $X = (2, 3, 5);
• For $d = 4$, $L_X = 2$ for $X = (2, 7, 3, 5)$, $L_X = 4$ for $X = (5, 3, 7, 2)$, and $L_X = 8$ for $X = (7, 2, 3, 5);
• For $d = 5$, $L_X = 20$ for $X = (5, 3, 3, 7, 7)$, $L_X = 30$ for $X = (5, 7, 2, 5, 5)$, and $L_X = 40$ for $X = (7, 2, 5, 7, 7);
• For $d = 6$, $L_X = 6$ for $X = (2, 5, 3, 5, 3, 3)$, $L_X = 12$ for $X = (5, 2, 3, 5, 2, 2)$, $L_X = 27$ for $X = (5, 5, 3, 2, 7, 3)$, and $L_X = 54$ for $X = (3, 3, 7, 2, 5, 2);
• For $d = 7$, $L_X = 4$ for $X = (5, 7, 3, 3, 7, 7)$, $L_X = 28$ for $X = (7, 7, 7, 5, 7, 5, 5)$, $L_X = 126$ for $X = (3, 5, 7, 7, 5, 7)$, and $L_X = 168$ for $X = (5, 3, 2, 2, 7, 3, 2);
• For $d = 8$, $L_X = 2$ for $X = (2, 7, 3, 5, 2, 7, 3, 5)$, $L_X = 4$ for $X = (2, 5, 3, 7, 2, 5, 3, 7)$, $L_X = 8$ for $X = (3, 5, 2, 7, 7, 3, 7, 5)$, $L_X = 16$ for $X = (7, 5, 5, 7, 3, 5, 3)$, $L_X = 32$ for $X = (5, 7, 7, 3, 7, 2, 3)$, $L_X = 48$ for $X = (3, 7, 3, 7, 3, 2, 5)$, and $L_X = 80$ for $X = (7, 5, 5, 7, 5, 2, 3, 3).

As an interesting fact, the distribution of cycle lengths starting from seeds $X \in P_0^d$ appears to be tilted, for even values of $d$, towards smaller cycle lengths, and for odd values of $d$ towards the larger cycle lengths. The problem of cycle lengths for various values of $d$ (analyzed in the case of the classical Ducci game [8]) presents itself to be a promising combinatorial number theory problem.

5. ACKNOWLEDGMENT

We would like to thank the anonymous reviewer for the helpful and constructive comments.

References

THE FIBONACCI QUARTERLY


MSC2010: 11B37, 11A41, 11B83

DEPARTMENT OF MATHEMATICS AND STATISTICS, OHIO NORTHERN UNIVERSITY, ADA, OHIO 45810
E-mail address: m-caragiu.1@onu.edu

INSTITUTE OF MATHEMATICS OF THE ROMANIAN ACADEMY, PO BOX 1-764, BUCHAREST 014700, ROMANIA, AND DEPARTMENT OF MATHEMATICS, UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN, 1409 W. GREEN STREET, URBANA, IL, 61801
E-mail address: zaharesc@math.uiuc.edu

DEPARTMENT OF MATHEMATICS AND STATISTICS, OHIO NORTHERN UNIVERSITY, ADA, OHIO 45810
E-mail address: m-zaki@onu.edu

38 VOLUME 52, NUMBER 1