Decompostion of natural numbers into $weight \times level + jump$ and application to a new classification of prime numbers

Rémi Eismann
reismann@free.fr

Abstract

In this article we introduce a decomposition of elements a_n of an increasing sequence of natural numbers $(a_n)_{n\in\mathbb{N}^*}$ into *weight* $\times$ *level* + *jump* which we use to classify the numbers a_n either by *weight* or by *level*. We then show that this decomposition can be seen as a generalization of the Eratosthenes sieve (which is the particular case of the whole sequence of natural numbers). Finally, we apply this decomposition to prime numbers in order to obtain a new classification of primes, and we analyze a few properties of this classification to make a series of conjectures based on numerical data.

1 Introduction

We start this paper by introducing a decomposition algorithm of the elements a_n of an increasing sequence of natural numbers $(a_n)_{n\in\mathbb{N}^*}$ into *weight* $\times$ *level* + *jump*. We then show that a necessary and sufficient condition for this decomposition to hold is

$$a_{n+1} < \tfrac{3}{2}a_n.$$

We use this decomposition to classify the numbers a_n either by *weight* or by *level* by introducing a classification principle.

By applying this algorithm to the whole sequence of natural numbers we find that it reduces to the Eratosthenes sieve (shifted by one unit).

In section 4 we consider this decomposition algorithm on the sequence of prime numbers and prove that such a decomposition is also in that case possible, except for $p_1 = 2$, $p_2 = 3$ and $p_4 = 7$. Moreover we show that the smallest member of a twin prime pair (except 3) always have a *weight* equal to 3.

We then use this decomposition to establish a new classification of prime numbers by *weight* and by *level* and we define *levels* of type $(1; i)$. Furthermore, we provide some related results as well as a series of precise and non-trivial conjectures on prime numbers.

Finally we show how composite numbers and 2−almost primes behave under the decomposition.

2 Decomposition algorithm of numbers into *weight* $\times$ *level* + *jump* and application to a classification scheme

We introduce an algorithm whose input is an increasing sequence of positive integers $(a_n)_{n\in\mathbb{N}^*}$ and whose output is a sequence of unique triplets of positive integers $(k_n, L_n, d_n)_{n\in\mathbb{N}^*}$.

We define the *jump* of a_n by

$$d_n := a_{n+1} - a_n.$$

Then let l_n be defined by

$$l_n := \begin{cases} a_n - d_n \text{ if } a_n - d_n > d_n; \\ 0 \text{ otherwise.} \end{cases}$$

The *weight* of a_n is defined to be

$$k_n := \begin{cases} \min\{k \in \mathbb{N}^* \text{ s.t. } k > d_n, k|l_n\} \text{ if } l_n \neq 0; \\ 0 \text{ otherwise.} \end{cases}$$

Finally we define the *level* of a_n by

$$L_n := \begin{cases} \frac{l_n}{k_n} \text{ if } k_n \neq 0; \\ 0 \text{ otherwise.} \end{cases}$$

We then have a decomposition of a_n into *weight* $\times$ *level* + *jump*: $a_n = l_n + d_n = k_n \times L_n + d_n$ when $l_n \neq 0$.

In the Euclidian division of a_n by its *weight* k_n, the quotient is the *level* L_n, and the remainder is the *jump* d_n.

Lemma 2.1. *A necessary and sufficient condition for the decomposition of a number a_n belonging to an increasing sequence of positive integers $(a_n)_{n\in\mathbb{N}^*}$ into weight $\times$ level + jump to hold is that*

$$a_{n+1} < \tfrac{3}{2} a_n.$$

.

Proof. The decomposition is possible if $l_n \neq 0$, that is if $a_n - d_n > d_n$, which can be rewritten as $a_{n+1} < \frac{3}{2} a_n$. □

In order to use this algorithm to classify the numbers a_n we introduce the following rule (whose meaning will become clearer in the next section): if for a_n we have $k_n > L_n$ then a_n is said to be classified by *level*, if not then a_n is said to be classified by *weight*.

3 Application of the algorithm to the sequence of natural numbers

In this situation we have $a_n = n$ et $d_n = 1$. The decomposition is impossible for $n = 1$ and $n = 2$ ($l_1 = l_2 = 0$). Apart from those two cases, we have the decomposition of n into *weight* $\times$ *level* + *jump*: $n = k_n \times L_n + 1$ when $n > 2$ and we also have the following relations

$$\begin{aligned} & L_n = 1 \\ \Leftrightarrow \quad & k_n = l_n = n - 1 \\ \Leftrightarrow \quad & l_n = n - 1 \text{ is prime,} \end{aligned}$$

$$\begin{aligned} & L_n \neq 1 \\ \Leftrightarrow \quad & k_n \times L_n = l_n = n - 1 \\ \Leftrightarrow \quad & l_n = n - 1 \text{ is composite.} \end{aligned}$$

Furthermore, we remark that they do not exist natural numbers except the numbers $(p_n + 1)$ for which $k_n > L_n$. Indeed, since we have $n - 1 = l_n = k_n \times L_n$ then according to the definitions of k_n and L_n if $n - 1$ is not prime we necessarily have $k_n \leq L_n$. We can thus characterize the fact that a number $l_n = n - 1$ is prime by the fact that n is classified by *level*, (or equivalently here by the fact that n is of *level* 1).

Since there is an infinity of prime numbers, there is an infinity of natural numbers of *level* 1. Similarly there is an infinity of natural numbers with a *weight* equal to k with k prime.

The algorithm allows to separate prime numbers (l_n or *weights* of natural numbers of *level* 1) from composite numbers, and is then indeed a reformulation of the Eratosthenes sieve. Thus applying this algorithm to any other increasing sequence of positive integers, for example to the sequence of prime numbers itself, can be seen as an extension of that sieve.

n A000027	k_n A020639$(n-1)$	L_n A032742$(n-1)$	d_n	l_n
1	0	0	1	0
2	0	0	1	0
3	2	1	1	2
4	3	1	1	3
5	2	2	1	4
6	5	1	1	5
7	2	3	1	6
8	7	1	1	7
9	2	4	1	8
10	3	3	1	9
11	2	5	1	10
12	11	1	1	11
13	2	6	1	12

Table 1: The 13 first terms of the sequences of *weights*, *levels*, *jumps*, and l_n in the case of the sequence of natural numbers.

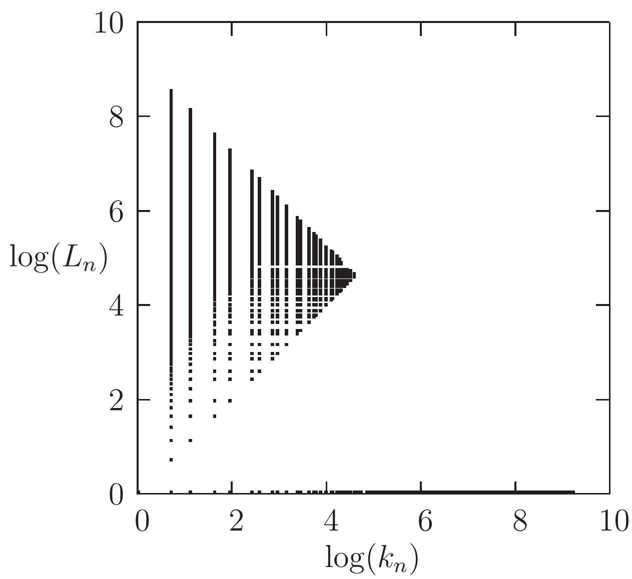

Figure 1: Plot of natural numbers in $\log(k_n)$ vs. $\log(L_n)$ coordinates (with $n \leq 10000$). The Eratosthenes sieve.

4 Application of the algorithm to the sequence of primes

We can wonder what happens if we try to apply the decomposition to the sequence of primes itself: for any $n \in \mathbb{N}^*$ we have $a_n = p_n$ and $d_n = g_n$ (the prime gap). The algorithm of section 1 can then be rewritten with these new notations as follows.

The *jump* (gap) of p_n is

$$g_n := p_{n+1} - p_n.$$

Let l_n be defined by

$$l_n := \begin{cases} p_n - g_n \text{ if } p_n - g_n > g_n; \\ 0 \text{ otherwise.} \end{cases}$$

The *weight* of p_n is then

$$k_n := \begin{cases} \min\{k \in \mathbb{N}^* \text{ s.t. } k > g_n, k|l_n\} \text{ if } l_n \neq 0; \\ 0 \text{ otherwise.} \end{cases}$$

The *level* of p_n is

$$L_n := \begin{cases} \frac{l_n}{k_n} \text{ if } k_n \neq 0; \\ 0 \text{ otherwise.} \end{cases}$$

So the decomposition of p_n into *weight* $\times$ *level* + *jump* reads $p_n = k_n \times L_n + g_n$ when $l_n \neq 0$. So one should investigate for which n we have $l_n \neq 0$, which is provided by the following result.

Theorem 4.1. *This decomposition is always possible except for $p_1 = 2$, $p_2 = 3$ and $p_4 = 7$ (i.e., $p_{n+1} \geq \frac{3}{2}p_n$ holds only for $n = 1$, $n = 2$ and $n = 4$).*

Proof. The decomposition if possible if, and only if, l_n is not equal to zero. But $l_n \neq 0$ if and only if $p_{n+1} < \frac{3}{2}p_n$, that is $p_n - g_n > g_n(*)$ by lemma 2.1. Let us now apply results of Pierre Dusart on the prime counting function π to show that this is always true except for $n = 1$, $n = 2$ and $n = 4$.

Indeed this last equation $(*)$ can be rewritten in terms of π as $\pi(\frac{3}{2}x) - \pi(x) > 1$ (i.e., there is always a number strictly included between x and $\frac{3}{2}x$ for any $x \in \mathbb{R}^+$). But Dusart has shown [1, 2] that on the one hand for $x \geq 599$ we have

$$\pi(x) \geq \frac{x}{\log x}\Big(1 + \frac{1}{\log x}\Big)$$

and on the other hand for $x > 1$ we have

$$\pi(x) \leq \frac{x}{\log x}\Big(1 + \frac{1.2762}{\log x}\Big)$$

So for $x \geq 600$ we have

$$\pi(\frac{3}{2}x) - \pi(x) > \frac{900}{\log 900}\Big(1 + \frac{1}{\log 900}\Big) - \frac{600}{\log 600}\Big(1 + \frac{1.2762}{\log 600}\Big)$$

and since the right hand side of this inequality is approximately equal to 39.2 we indeed have that $\pi(\frac{3}{2}x) - \pi(x) > 1$, so the inequality $(*)$ holds for any prime greater than 600. We check numerically that it also holds in the remaining cases when $x < 600$, except for the aforementioned exceptions $n = 1$, $n = 2$ and $n = 4$ which ends the proof. $\square$

Let us now state a few direct results. For any p_n different from $p_1 = 2$, $p_2 = 3$ and $p_4 = 7$ we have

$$\begin{gathered} gcd(g_n, 2) = 2, \\ gcd(p_n, g_n) = gcd(p_n - g_n, g_n) = gcd(l_n, g_n) = gcd(L_n, g_n) = gcd(k_n, g_n) = 1, \\ 3 \le k_n \le l_n, \\ 1 \le L_n \le \tfrac{l_n}{3}, \\ 2 \le g_n \le k_n - 1, \\ 2 \times g_n + 1 \le p_n. \end{gathered}$$

Lemma 4.1. *p is a prime such that $p > 3$ and $p + 2$ is also prime if and only if p has a weight equal to* 3.

Proof. The primes $p > 3$ such that $p + 2$ is also prime are of the form $6n - 1$, so $p - 2$ is of the form $6n - 3$. The smallest divisor greater than 2 of a number of the form $6n - 3$ is 3. If p_n has a weight equal to 3 then $p_n > 3$ and the jump g_n is equal to 2 since we know that $2 \le g_n \le k_n - 1$ and $2 \times g_n + 1 \le p_n$. $\square$

n	p_n A000040	k_n A117078	L_n A117563	d_n A001223	l_n A118534
1	2	0	0	1	0
2	3	0	0	2	0
3	5	3	1	2	3
4	7	0	0	4	0
5	11	3	3	2	9
6	13	9	1	4	9
7	17	3	5	2	15
8	19	5	3	4	15
9	23	17	1	6	17
10	29	3	9	2	27
11	31	25	1	6	25
12	37	11	3	4	33
13	41	3	13	2	39
14	43	13	3	4	39
15	47	41	1	6	41
16	53	47	1	6	47
17	59	3	19	2	57

Table 2: The 17 first terms of the sequences of *weights*, *levels*, *jumps* and l_n in the case of the sequence of primes.

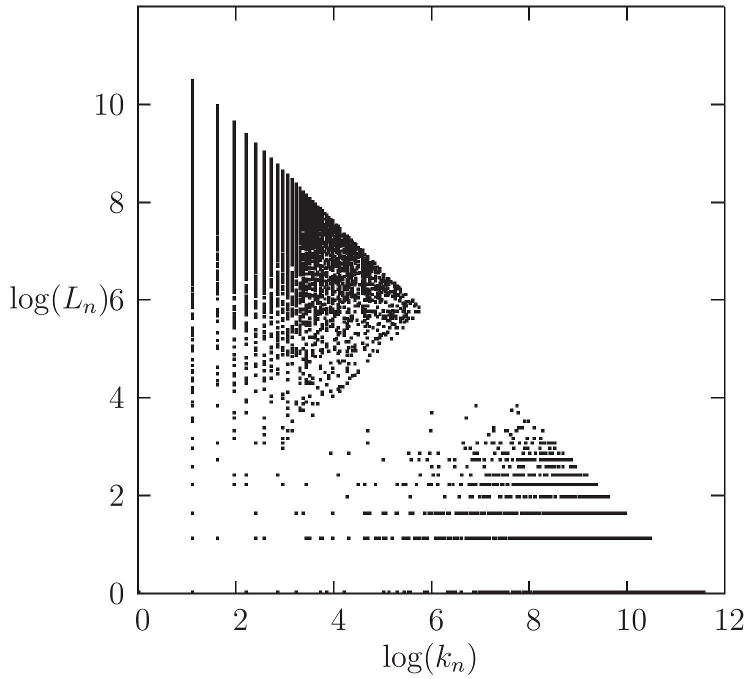

Figure 2: Plot of prime numbers in $\log(k_n)$ vs. $\log(L_n)$ coordinates (with $n \leq 10000$).

5 Classification of prime numbers

We introduce the following classification principle:
- if for p_n we have $k_n > L_n$ then p_n is classified by *level*, if not p_n is classified by *weight*;
- furthermore if for p_n we have that l_n is equal to some prime p_{n-i} then p_n is of *level* $(1; i)$.

For $n \leq 5.10^7$, $17,11\%$ of the primes p_n are classified by *level* and $82,89\%$ are classified by *weight*.

We have the following direct results:

If p_n is classified by *weight* then

$$g_n + 1 \leq k_n \leq \sqrt{l_n} \leq L_n \leq \tfrac{l_n}{3}.$$

If p_n is classified by *level* then

$$L_n + 2 \leq g_n + 1 \leq k_n \leq l_n.$$

Number of primes	which have a *weight* equal to	% / total of primes classified by *weight*	% / total
3370444	3	8.132	6.741
1123714	5	2.711	2.247
1609767	7	3.884	3.219
1483560	9	3.579	2.967
1219514	11	2.942	2.439
1275245	13	3.077	2.550
1260814	15	3.042	2.522
1048725	17	2.530	2.097
1051440	19	2.546	2.103
1402876	21	3.385	2.806
893244	23	2.155	1.786

Table 3: Distribution of primes for the 11 smallest *weights* (with $n \leq 5.10^7$).

Number of primes	which are of *level*	% / total of primes classified by *level*	% / total
2664810	1	31.15	5.330
2271894	3	26.56	4.544
963665	5	11.27	1.927
444506	7	5.197	0.8890
640929	9	7.493	1.282
254686	11	2.978	0.5094
155583	13	1.819	0.3112
351588	15	4.110	0.7032
115961	17	1.356	0.2319
78163	19	0.9138	0.1563
148285	21	1.734	0.297

Table 4: Distribution of primes for the 11 smallest *levels* (with $n \leq 5.10^7$).

If p_n is of *level* $(1;i)$ then

$$L_n = 1 \text{ and } l_n = k_n = p_{n-i},$$
$$p_n = p_{n-i} + g_n \text{ or } p_{n+1} - p_n = p_n - p_{n-i}.$$

If p_n is of *level* $(1;1)$ then

$$L_n = 1 \text{ and } l_n = k_n = p_{n-1},$$
$$g_n = g_{n-1} \text{ or } p_{n+1} - p_n = p_n - p_{n-1},$$
$$p_n = \tfrac{p_{n+1}+p_{n-1}}{2}.$$

Primes of *level* $(1;1)$ are the so-called "balanced primes" (A006562).

Number of primes	which are of *level* $(1;i)$ i
1307356	1
746381	2
345506	3
153537	4
65497	5
27288	6
11313	7

Table 5: Distribution of primes of *level* $(1;i)$ (with $n \leq 5.10^7$, $i \leq 7$).

One can wonder whether the notion of *level* $(1;i)$ can be generalized to primes classified by *level* themselves (*level* $(3;i)$, *level* $(5;i)$ for exemple). We shall not address this issue in this paper.

6 Conjectures on primes

From our numerical data on the decomposition of primes p_n until $n = 5.10^7$ we make the following conjectures.

Since we have shown previously that the smallest number of each twin prime pair (except 3) has a *weight* equal to 3, the well-known conjecture on the existence of an infinity of twin primes can be rewritten as

Conjecture 1. *The number of primes with a weight equal to* 3 *is infinite.*

To extend this conjecture, and by analogy with the decomposition of natural numbers for which we know that for any prime k there exist an infinity of natural numbers with a *weight* equal to k and that there exist an infinity of natural numbers of *level* 1, we make this two conjectures

Conjecture 2. *The number of primes with a weight equal to k is infinite for any $k \geq 3$ which is not a multiple of* 2.

Conjecture 3. *The number of primes of level L is infinite for any $L \geq 1$ which is not a multiple of* 2.

Now, based on our numerical data and again by analogy with the decomposition of natural numbers for which we know that the natural numbers which are classified by *level* have a l_n or a *weight* which is always prime we conjecture

Conjecture 4. *Except for* $p_6 = 13$*,* $p_{11} = 31$*,* $p_{30} = 113$*,* $p_{32} = 131$ *et* $p_{154} = 887$*, primes which are classified by level have a weight which is itself a prime.*

Finally, we make the following conjectures, for which we have no rigorous arguments yet

Conjecture 5. *The number of primes of level* $(1; i)$ *is infinite pour for any* $i \geq 1$*.*

Conjecture 6. *If the jump* g_n *is not a multiple of* 6 *then* l_n *is a multiple of* 3*.*

Conjecture 7. *If the* l_n *is not a multiple of* 3 *then jump* g_n *is a multiple of* 6*.*

Furthermore, we do wonder whether one could generalize the concept of primes in this setting, namely find an $n-$ary composition law $\star$ and a subset of the primes $\mathbb{P}_\star \subset \mathbb{P}$ or a subset of the integers $\mathbb{N}_\star \subset \mathbb{N}$ such that any prime would uniquely decompose into a $\star$-composition of elements of $\mathbb{P}_\star$ or $\mathbb{N}_\star$.

7 Decomposition of composite numbers and of 2−almost primes.

In this section we only provide the plots of the distribution of composite numbers and of 2−almost primes in $\log(k_n)$ vs. $\log(L_n)$ coordinates.

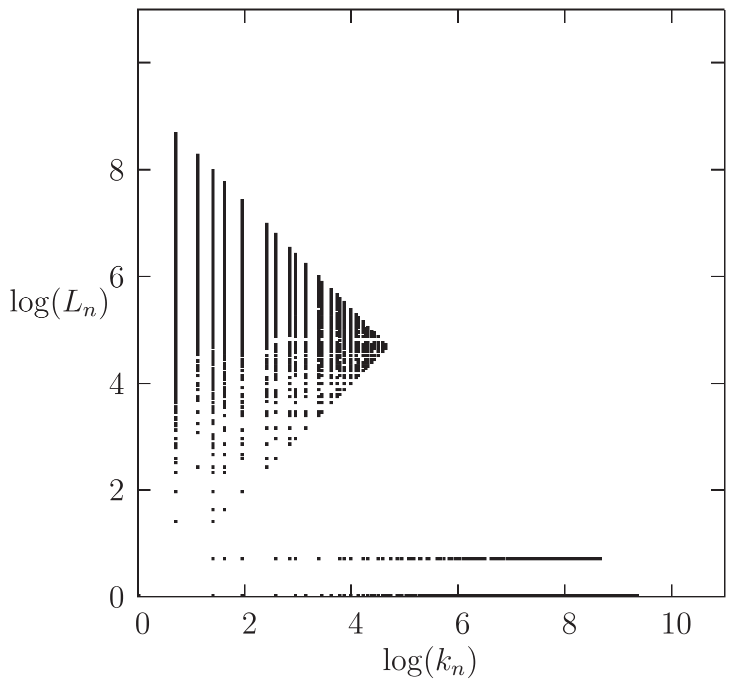

Figure 3: Plot of composite numbers (A002808) in $\log(k_n)$ vs. $\log(L_n)$ coordinates (with $n \leq 9999$).

The sequence of *weights* of composite numbers is A130882.

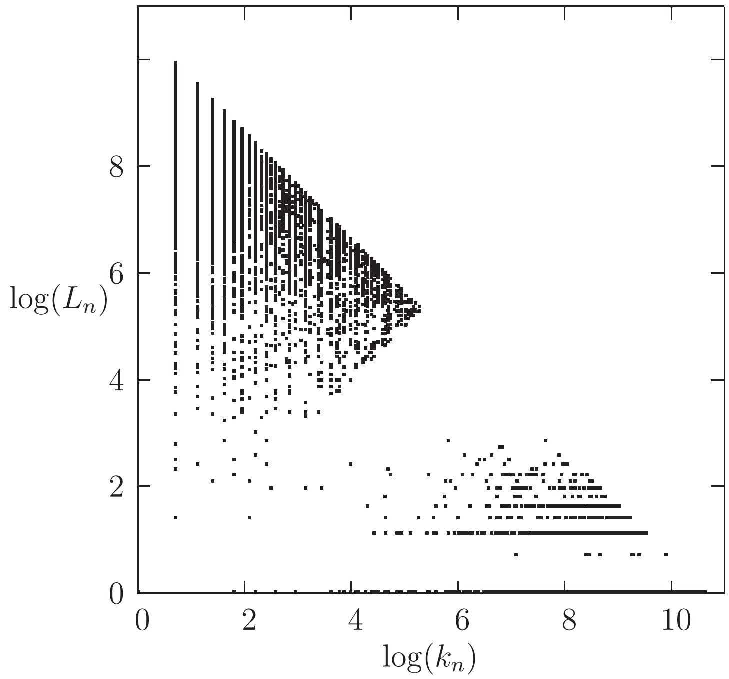

Figure 4: Plot of 2−almost primes (A001358) in log(k_n) vs. log(L_n) coordinates (with $n \leq 9999$).

The sequence of *weights* of 2−almost primes is A130533.

8 Acknowledgements

The author wishes to thank Thomas Sauvaget for the proof of theorem 4.1, his advices and translating the text into English. The author whishes to thank Fabien Sibenaler for developping programs in Java and Assembly implementing the decomposition algorithm and for his encouragements. The author also whishes to thank Jean-Paul Allouche for his advices and for his encouragements and N. J. A. Sloane for its help through the OEIS.

References

[1] Pierre Dusart, Autour de la fonction qui compte le nombre de nombres premiers. Thèse, Université de Limoges, 1998, page 36.

[2] Pierre Dusart, The k^{th} prime is greater than $k(lnk + lnlnk - 1)$ pour $k \geq 2$. *Math. Comp.* 68 (1999), 411 - 415.

[3] N. J. A. Sloane, *The On-Line Encyclopedia of Integer Sequences*, published electronically at www.research.att.com/∼njas/sequences.

2000 *Mathematics Subject Classification*: Primary 11B83; Secondary 11P99.
Keywords: integer sequences, Eratosthenes, sieve, primes, prime gaps, balanced primes.

(Concerned with sequences A000027, A020639, A032742, A000040, A117078, A117563, A001223, A118534, A001359, A006562, A125830, A117876, A074822, A002808, A130882, A001358 and A130533.)

Submitted to *Journal of Integer Sequences*.

Return to Journal of Integer Sequences home page.