On simultaneous binary expansions of $n$ and $n^2$

Giuseppe Melfi

*Université de Neuchâtel*
*Groupe de Statistique*
*Espace de l’Europe 4, CH-2002 Neuchâtel, Switzerland*
E-mail: Giuseppe.Melfi@unine.ch

A new family of sequences is proposed. An example of sequence of this family is more accurately studied. This sequence is composed by the integers $n$ for which the sum of binary digits is equal to the sum of binary digits of $n^2$. Some structure and asymptotic properties are proved and a conjecture about its counting function is discussed.

**Key Words:** Sequences and sets. Binary expansion. Asymptotic behaviour.

1. INTRODUCTION

This paper arises from some natural questions about digits of numbers. We are interested in numbers $n$ such that $n$ and $n^m$ have a certain relation involving sums of digits. Let $B(n)$ be the sum of digits of the positive integer $n$ written on base 2. This function represents the numbers of ones in the binary expansion of $n$, or from the code theory point of view, the number of nonzero digits in the bits string representing $n$, i.e., the so-called ‘Hamming weight’ of $n$. It is obvious that $B(n) = B(2n)$, but for a prime $p > 2$ a relation between $B(n)$ and $B(pm)$ is less trivial [2, 4, 5, 6, 9].

In this paper we are interested in comparing $B(n)$ with $B(n^m)$. Stolarsky [11] proved some inequalities for the functions $r_m(n) = B(n^m)/B(n)$. Lindström [8] proved that $\limsup_{n \to \infty} B(n^m)/\log_2 n = m$, for $m \geq 2$.

One can naturally define a family of positive integer sequences.

**Definition 1.1.** Let $k \geq 2$, $l \geq 1$, $m \geq 2$ be positive integers. We say that a positive integer $n$ satisfies the property $(k, l, m)$ if the sum of its digits in its expansion in base $k$ is $l$ times the sum of the digits of the expansion in base $k$ of $n^m$.
For every triplet \((k, l, m)\) we have a sequence made up of the positive integers of the type \((k, l, m)\).

The simplest case is \((k, l, m) = (2, 1, 2)\), which corresponds to the positive integers \(n\) for which the numbers of ones in their binary expansion is equal to the number of ones in \(n^2\).

This paper is consacrated to the study of the \((2,1,2)\)-numbers. The list shows several interesting facts (see Table 1). The distribution is not regular. A huge amount of questions, most of which of elementary nature, can be raised.

In despite of its elementary definition, this sequence surprisingly does not appear in literature. Recently, proposed by the author, it appeared on \([10]\).

Several questions, concerning both the structure properties and asymptotic behaviour, can be raised. Is there a necessary and sufficient condition to assure that a number is of type \((2,1,2)\)? What is the asymptotic behaviour of the counting function of \((2,1,2)\)-numbers?

The irregularity of distribution does not suggest a clear answer to these questions.

In Section 3 we provide some sufficient conditions in order that a number is of type \((2,1,2)\). In particular we explicitely provide several infinite sets of \((2,1,2)\)-numbers, but all these sets are quite thin, and the problem of an exhaustive answer does not appear easy at first sight.

Let \(p(n)\) be the number of \((2,1,2)\)-numbers which does not exceed \(n\).

**Conjecture 1.1.** Let \(p(n) = \sum_{m \leq n \mod \text{type } (2,1,2)} 1\) be the counting function of isosquare numbers. There exists a continuous function \(F(x)\), periodic of period 1, such that

\[
p(n) = \frac{n^\alpha}{\log n} F\left(\frac{\log n}{\log 2}\right) + R(n),
\]

where \(\alpha = \log 1.6875/\log 2 \approx 0.7548875\), and \(R(n) = o(n^{\alpha}/\log n)\).

Further \(F(n)\) is nowhere differentiable.

Sequences constructed in Section 3 yield \(p(n) \gg \log n\). In Section 4 we prove a lower bound \(p(n) \gg n^{0.025}\). Conjecture 1.1 is suggested by a more detailed observation of the list, as well as some similarities with certain functions studied by Boyd et al. \([1]\).

A discussion on Conjecture 1.1 will be developed in Section 5. Our arguments appear to be confirmed by experimental results done for \(n < 10^8\).
2. NOTATIONS

If the binary expansion of \( n \) is \( c_1c_2\ldots c_k \), \((c_i \in \{0,1\})\) we will write \( n = (c_1c_2\ldots c_k) \). The first digit may be 0 so, for example, we allow the notation \( 3 = (011) \). We will also denote \( k \) times \( 1\ldots 1 \) as \( (1^{(k)}) \), so for example,

\[
(11\ldots 11\ 00\ldots 00\ 11\ldots 11\ 010) = (1^{(k)})0^{(h)}1^{(l)}010.
\]

This allows to perform arithmetical operations in a more compact manner. For example, if \( k > h + h' \), one has \( (1^{(k)}) - (1^{(h)})0^{(h')} = (1^{(k-h-h')}0^{(h)}) \).

Let \( B(n) \) be the number of 1’s in the binary expansion of \( n \). This function, known also as the Hamming weight of \( n \), is used mainly in the theory of algorithms, namely for the study of computational aspects and complexity. So, \( n \) is of type \((2,1,2)\) if \( B(n) = B(n^2) \).

Let \( c \in \{0,1\} \) a binary digit. We will use the notation \( c' \) to indicate the other digit: \( c' = 1 - c \). This notation will be very useful in computing the Hamming weight of a number. We will use the property that if \( n = (c_1c_2\ldots c_k) \), then \( 2^k - n - 1 = (c'_1c'_2\ldots c'_k) \), and \( B(n) = k - B(2^k - n - 1) \).

3. ARITHMETICS AND STRUCTURE PROPERTIES

There are infinitely many \((2,1,2)\)-numbers. Note that if \( n \) is an \((2,1,2)\)-number, then \( 2n \) is an \((2,1,2)\)-number, and that if \( n \) is an even \((2,1,2)\)-number, then \( n/2 \) is a \((2,1,2)\)-number. There is no direct dependence between \( 2n+1 \) or \( 2n-1 \) and \( n \); classical arguments for the study of asymptotic properties of the counting function (see [12]) cannot be applied. In the following remarks we show some sets of \((2,1,2)\)-numbers.

Remark 3. 1. For every \( k > 1 \), the number \( n_k = 2^k - 1 \) is of type \((2,1,2)\).

In fact, \( B(n_k) = k \), and

\[
(2^k - 1)^2 = 2^{2k} - 2^{k+1} + 1 = (11\ldots 11\ 00\ldots 00\ 1),
\]

so \( n_k^2 \) also contains \( k \) times the digit 1.

Table 1 contains 4-tuples of consecutive integers of \((2,1,2)\)-numbers. After the 4-tuple \((1,2,3,4)\) the second one is \((316,317,318,319)\).

Remark 3. 2. There are infinitely many 4-tuples of consecutive integers composed by \((2,1,2)\)-numbers.
In fact it is an exercise to show that for every \( k \geq 9 \) and \( n = 2^k - 2^{k-2} - 2^{k-3} - 4 \), the numbers \( n, n+1, n+2 \) and \( n+3 \) are all of type \((2,1,2)\).

The following proposition shows that the set of \((2,1,2)\)-numbers presents a certain gap structure.

**Proposition 3.1.** There are infinitely many \( n \) such that the interval \([n, n + n^2] \) does not contain any \((2,1,2)\)-number.

**Proof.** Let \( n = 2^{2k} = (10_{(2k)}) \). Each \( m \in [n, n + n^2] \) is of the form \( n + r \) with \( r < n^2 \). In its binary expansion \( m \) is of the kind \((10_{(2k)})c_1c_2\ldots c_k\).

Here \( c_i \in \{0, 1\} \) are binary digits and \( B((c_1c_2\ldots c_k)) = B(r) \geq 1 \). Let \( r^2 = (q_1q_2\ldots q_{2k}) \). We have again \( B((q_1q_2\ldots q_{2k})) \geq 1 \). Hence

\[
m^2 = (2^{2k} + (c_1c_2\ldots c_k))^2
\]

\[
= 2^{4k} + 2^{2k+1}(c_1c_2\ldots c_k) + (c_1c_2\ldots c_k)^2
\]

\[
= (10_{(k-1)}c_1c_2\ldots c_k00q_1q_2\ldots q_{2k})
\]

Therefore \( B(m^2) = 1 + B(r) + B(r^2) \geq 1 + B(r) = B(m) \). ☐

The preceding construction cannot be improved. One can easily prove that if \( n > 5 \) is odd, then there exists an \((2,1,2)\)-number between \( 2^n \) and \( 2^n + 4 \cdot 2^n \). Namely if \( n = 2m + 1 \) such a number is \( 2^{2m+1} + 2^{m+2} - 1 \). The proof by check digits in column operations is straightforward.

### 4. A LOWER BOUND FOR THE COUNTING FUNCTION

We begin this section with some preliminary lemmata.

**Lemma 4.1.** If \( n < 2^\nu \), then \( B(n(2^\nu - 1)) = \nu \).

**Proof.** We assume that \( n \) is odd. Let \( n = (c_1c_2\ldots c_k) \), with \( c_1 = c_k = 1 \). Since \( n < 2^\nu \) we have that \( k \leq \nu \). We have that \( n(2^\nu - 1) = (c_1c_2\ldots c_k0(\nu)) - (c_1c_2\ldots c_k) \).

Hence \( n(2^\nu - 1) = (c_1c_2\ldots c_{k-1}c'_k1_{(\nu-k)}c'_1c'_2\ldots c'_{k-1}c_k) \). Therefore

\[
B(n(2^\nu - 1)) = B((c_1c_2\ldots c_{k-1}c'_k1_{(\nu-k)}c'_1c'_2\ldots c'_{k-1}c_k))
\]

\[
= B((c_1c_2\ldots c_{k-1}c_k1_{(\nu-k)}c'_1c'_2\ldots c'_{k-1}c_k))
\]

\[
= B(n) + (\nu - k) + (k - B(n))
\]

\[
= \nu.
\]
If \( n \) is even, \( n' = n/2^h \) is an odd integer for a certain \( h \), and \( n' < 2^\nu \). Hence \( B(n'(2^\nu - 1)) = \nu \) and \( B(2^{h}n'(2^\nu - 1)) = B(n'(2^\nu - 1)) = \nu \), so \( B(n(2^\nu - 1)) = \nu \).

**Lemma 4.2.** Let \( n \in \mathbb{N} \) and let \( \nu \) be such that \( n < 2^{\nu-1} \). Then
\[
B(2^\nu n + 1) = B(n) + 1 \quad \text{and} \quad B((2^\nu n + 1)^2) = B(n^2) + B(n) + 1.
\]

**Proof.** The proof is straightforward.

**Lemma 4.3.** Let \( n = (c_1c_2\ldots c_k), m = (d_1d_2\ldots d_h) \) odd positive integers. If \( \nu \geq \max\{2h - 1, h + k + 1\} \), we have
\[
B(n2^\nu - m) = B(n) - B(m) + \nu
\]
and
\[
B((n2^\nu - m)^2) = B(n^2) + B(m^2) - B(mn) + \nu - 1.
\]

**Proof.** Let \( n = (c_1c_2\ldots c_k) \) and \( m = (d_1d_2\ldots d_h) \). If \( h \leq \nu \) we have \( n2^\nu - m = (c_1c_2\ldots c_k-1c_k'1_{\nu-h}d_1'd_2'\ldots d_{h-1}'d_h) \). Since \( m \) and \( n \) are odd, \( c_k' + d_h = c_k + d_h' \), so \( B(n2^\nu - m) = B((c_1c_2\ldots c_k'1_{\nu-h}d_1'd_2'\ldots d_h')) = B(n) + \nu - B(m) \).

Let \( n^2 = (q_1q_2\ldots q_{2k}) \) and \( m^2 = (r_1r_2\ldots r_{2h}) \). Let \( mn = (s_1s_2\ldots s_{k+h}) \). We have \( (n2^\nu - m)^2 = (n^22^{2\nu} + m^2) - nm2^{\nu+1} \), and if \( 2h \leq \nu + 1 \) and \( k + h \leq \nu - 1 \), we have
\[
(n2^\nu - m)^2 = (q_1\ldots q_{2k-1}q_{2k}'1_{\nu-k-h-1}s_1'\ldots s_{k+h-1}'s_{k+h}0_{\nu+1-2k}r_1'\ldots r_{2h}) \]
hence \( B((n2^\nu - m)^2) = B(n^2) + B(m^2) - B(mn) + \nu - 1 \).

**Corollary 4.1.** If \( B(n^2) = 2B(n) - 1 \), and \( \nu \geq k + 2 \), then \( 2^\nu n - 1 \) is of type \((2,1,2)\).

**Corollary 4.2.** Let \( n \) an odd positive integer. Let \( m = 2^h - 1 \) with \( n < m \). If \( \nu \geq 2h + 1 \) then
\[
B(n2^\nu - m) = B(n) + \nu - h
\]
and
\[
B((n2^\nu - m)^2) = B(n^2) + \nu - 1.
\]
Proof. These statements are an easy consequence of Lemma 4.3 of Lemma 4.1 and of Remark 3.1.

Lemma 4.4. Let \( n = (c_1 c_2 \ldots c_k) \) be an odd positive integer. Let \( B(n^2) \geq 2B(n) + 1 \). There exist \( \nu \) and \( h \in \mathbb{N} \) such that for \( n' = n2^{\nu} - (2^h - 1) \) we have

\[
B(n^2) = 2B(n') - 1.
\]

Proof. Let \( h = k + 1 \). We have \( n < 2^h - 1 \). Let \( \nu = B(n^2) - 2B(n) + 2h \). We have \( \nu = 2h + a \) with \( a \geq 1 \). Remark that \( \nu < 4k + 2 \). The hypotheses of Corollary 4.2 are satisfied, so \( B(n') = B(n) + \nu - h \) and

\[
B(n^2) = B(n^2) + \nu - 1
= 2B(n) + \nu - 2h + \nu - 1
= 2B(n) + 2\nu - 2h - 1
= 2B(n') - 1.
\]

Lemma 4.5. Let \( n = (c_1 c_2 \ldots c_k) > 1 \) be a positive integer. Let \( n_0 = (c_1 c_2 \ldots c_k 0_{(k)} 10_{(2k+1)} 1) \). Then

\[
B(n_0^2) > 2B(n_0) + 1.
\]

Proof. It suffices to apply twice Lemma 4.2. Remark that \( n_0 \ll n^4 \).

Theorem 4.1. Let \( p(n) \) be the counting function of \((2,1,2)\)-numbers. We have

\[
p(n) \gg n^{0.025}.
\]

Proof. Let \( n = (c_1 c_2 \ldots c_k) \) be an odd positive integer. We will show that for a constant \( A \) not depending on \( n \), it is possible to construct a set of \( n \) distinct \((2,1,2)\)-numbers not exceeding \( An^{10} \).

Let consider the \( n \) integers \( n_i = 2^6 + i \), for \( i = 1, \ldots n \). We have obviously \( n_i < 4n \).
For every $i$, by Lemma 4.5 it exists $n_{0,i} \ll n_i^4$ whose first $k + 1$ digits are the same as those of $n_i$ such that $B(n_{0,i}^2) > 2B(n_{0,i}) + 1$.

By Lemma 4.4 it exists $n_{0,i}' \ll n_{0,i}^5$ such that the first $k$ binary digits of $n_{0,i}'$ are again those of $n_i$ and such that $B(n_{0,i}'^2) = 2B(n_{0,i}') - 1$.

Finally, by Corollary 4.1 it exists $n_{0,i}'' \ll n_{0,i}'^2$, whose the first binary digits are the same as for $n_i$ and such that $B(n_{0,i}''^2) = B(n_{0,i}'')$.

We have $n_{0,i}'' \ll n_{0,i}' \ll n_{0,i}^5 \ll n_{0,i}^4 \ll n_i^4$.

5. A PROBABILISTIC APPROACH

Let $k$ be a sufficiently large positive integer. Let $n$ be such that $2^k \leq n < 2^{k+1}$. In a binary base, these numbers are made up of a ‘1’ digit followed by $k$ binary digits 0 or 1. So $1 \leq B(n) \leq k + 1$. Let us consider $n^2$. We have $2^{2k} \leq n^2 < 2^{2k+2}$, so its binary expansion contains a ‘1’ digit followed by $2k$ or $2k + 1$ binary digits 0 or 1.

In this section we estimate the asymptotic behaviour of $p(n)$ under a suitable assumption.

We consider $B(n)$ and $B(n^2)$ as random variables. Clearly $B(n) - 1$ follow a binomial random distribution of type $b(k, 1/2)$. Schmid [9] studied joint distribution of $B(p_i n)$ for distinct odd integer $p_i$. Here we consider the joint distribution of $B(n)$ and $B(n^2)$. We assume that for sufficiently large $k$, $B(n^2) - 1$ tends to follow a binomial random distribution of type $b(2k, 1/2)$ if $2^{2k} \leq n^2 < 2^{2k+1}$ and a binomial random distribution of type $b(2k + 1, 1/2)$ if $2^{2k+1} \leq n^2 < 2^{2k+2}$. We assume that $B(n)$ and $B(n^2)$ are independent realizations of these random variables.

It is clear that for very small values of $B(n)$, the numerical value of $B(n^2)$ is also relatively small, since $B(n^2) \leq B(n)^2$, so these variables are not completely independent. But for $B(n) > \sqrt{\log n}$ this phenomenon tends to disappear, and the preceding assumption can be taken in account to an asymptotic behaviour estimate.

Hence

$$\Pr(n \text{ of type } (2,1,2) \text{ and } B(n) = l) = \frac{\binom{2k}{l} + \binom{2k + 1}{l}}{3 \cdot 2^{2k}}$$

This suggests that $p(n) \sim n^\alpha$ with

$$\alpha = -2 + \frac{1}{\log 2} \lim_{k \to \infty} \frac{1}{k} \log \sum_{l=0}^{k} \binom{k}{l} \binom{2k}{l} = \frac{\log 27/16}{\log 2} \simeq 0.75488.$$
The least square method applied to \((2,1,2)\)-numbers not exceeding \(10^8\) gives the value \(\alpha \simeq 0.73\). The difference is due to the fact that there is a small effect of correlation between \(B(n)\) and \(B(n^2)\) for \(n\) in a neighbourhood of a power of 2. Hence we conjecture that \(p(n) \asymp n^\alpha / \log n\).

However, a plot of \(p(n) \log n / n^\alpha\) shows more complex details (see Figure 1). It seems that the limiting function is not derivable. Effectively, as shown by Delange [4] and by Coquet [2], and more recently by Tenenbaum [12], if \(f(n)\) is a function related to the binary expansion of \(n\), often one has that \(F(n) = \sum_{k \leq n} f(k)\) has an expression in which periodic functions, often nowhere differentiable, are involved. These properties probably hold in our case, but a direct approach as shown in [12] does not appear possible. This remark, joint with the observation of Figure 1, justifies Conjecture 1.1.

**ACKNOWLEDGMENTS**

I am grateful to Chu Wen-Chang and to Roberto Avanzi for their useful discussions. I thank Alina Matei for her help in the early phases of the computational approach.

This work has been partly realized during my stay in Cetraro (Italy) with the contribution of C.I.M.E. I am grateful to Carlo Viola for his strong encouragement and support.

**APPENDIX**

![Figure 1](image-url)
Coquet, J. "A summation formula related to the binary digits", Invent. Math.

Larcher, G.; Tichy, R.F. "Some number-theoretical properties of generalized sum-of-\(\pm\) functions", Acta Arith.

Coquet, J. "Power sums of digital sums", J. Number Theory

Algebra Eng. Commun. Comput.

Table 1. The \((2,1,2)\)-numbers not exceeding 13742.

REFERENCES

1. Boyd, D.W.; Cook, J.; Morton, P. “On sequences of \(\pm 1\)’s defined by binary patterns”, Diss. Math. 283, 60 p. (1989).

2. Coquet, J. “A summation formula related to the binary digits”, Invent. Math. 73 (1983), 107-115.

3. Coquet, J. “Power sums of digital sums”, J. Number Theory 22 (1986), 161–176.

4. Delange, H. “Sur la fonction sommatoire de la fonction ‘somme des chiffres’”, Enseignement Math., II. Ser. 21 (1975), 31–47.

5. Dumont, J-M.; Thomas, A. “Systèmes de numération et fonctions fractales relatifs aux substitutions”, Theor. Comput. Sci. 65 (1989), 153–169.

6. Grabner, P.J.; Herendi, T.; Tichy, R.F. “Fractal digital sums and codes”, Appl. Algebra Eng. Commun. Comput. 8 (1997), 33–39.

7. Larcher, G.; Tichy, R.F. “Some number-theoretical properties of generalized sum-of-digit functions”, Acta Arith. 52, 183–196 (1989).

8. Lindström, B. “On the binary digits of a power”, J. Number Theory 65 (1997), 321–324.
9. Schmid, J. “The joint distribution of the binary digits of integer multiples”, Acta Arith. 43 (1984), 391–415.

10. Sloane, N.J.A., “The On Line Encyclopedia of integer sequences”,
http://www.research.att.com/~njas/sequences/index.html

11. Stolarsky, K.B. “The binary digits of a power”, Proc. Am. Math. Soc. 71 (1978), 1–5.

12. Tenenbaum, G. “Sur la non-dérivabilité de fonctions périodiques associées à certaines formules sommatoires”, in The mathematics of Paul Erdős, R.L. Graham and J. Nesetřil eds. (1997), Springer Verlag, 117–128.