Asymptotic bit frequency in Fibonacci words

Jean-Luc Baril, Sergey Kirgizov, and Vincent Vajnovszki

LIB, Université de Bourgogne Franche-Comté
B.P. 47 870, 21078 Dijon Cedex France
E-mails: {barjl, sergey.kirgizov, vvajnov}@u-bourgogne.fr

May 6, 2022

Abstract

It is known that binary words containing no k consecutive 1s are enumerated by k-step Fibonacci numbers. In this note we discuss the expected value of a random bit in a random word of length n having this property.

1 Introduction

For $n \geqslant 0$ and $k \geqslant 2$, we denote by $\mathcal{B}_n(1^k)$ the set of length n binary words avoiding k consecutive 1s. For example, we have

$$\begin{aligned}\mathcal{B}_4(11) &= \{0000, 0001, 0010, 0100, 0101, 1000, 1001, 1010\}, \text{ and}\\ \mathcal{B}_4(111) &= \{0000, 0001, 0010, 0011, 0100, 0101, 0110, 1000, 1001, 1010, 1011, 1100, 1101\}.\end{aligned}$$

It is well known, see Knuth [12, p. 286], that $\mathcal{B}_n(1^k)$ is enumerated by the k-step Fibonacci numbers, precisely $|\mathcal{B}_n(1^k)| = f_{n+k,k}$, where $f_{n,k}$ is defined, following Miles [14] as

$$f_{n,k} = \begin{cases} 0 & \text{if } 0 \leqslant n \leqslant k-2,\\ 1 & \text{if } n = k-1,\\ \sum_{i=1}^{k} f_{n-i,k} & \text{otherwise.}\end{cases}$$

Denote by $v_{n,k}$ the *frequency* (also called *popularity*) of 1s in $\mathcal{B}_n(1^k)$, i.e. the total number of 1s in all words of $\mathcal{B}_n(1^k)$. For instance, $v_{4,2} = 10$ and $v_{4,3} = 22$. The ratio of frequency of 1s to the overall number of bits in words of $\mathcal{B}_n(1^k)$ is

$$\alpha_{n,k} = \frac{v_{n,k}}{n \cdot |\mathcal{B}_n(1^k)|},$$

and it equals the expected value of a random bit in a random word from $\mathcal{B}_n(1^k)$. In [2], the authors left without proof the fact that, for any $k \geqslant 2$, $\lim_{n\to\infty} \alpha_{n,k}$ converges to a non-zero value as n grows. This note is devoted to proving this fact, which apart from its interest *en soi* has practical counterparts. Indeed, words in $\mathcal{B}_n(1^k)$ play a critical role in some telecommunication frame synchronization protocols, see for example [1, 3, 5], or in particular Fibonacci-like interconnection networks [8].

Our discussion is based on the bivariate generating function

$$F_k(x,y) = \sum_{n=0}^{\infty} \sum_{m=0}^{n-\lfloor \frac{n}{k} \rfloor} a_{n,m} x^n y^m$$

whose coefficient $a_{n,m}$ equals the number of words from $\mathcal{B}_n(1^k)$ containing exactly m 1s. For $k=2$ and $k=3$, Table 1 presents some values of $a_{n,m}$ for small n and m.

$m\backslash n$	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
0	1	1	1	1	1	1	1	1	1
1	1	2	3	4	5	6	7	8	9
2			1	3	6	10	15	21	28
3					1	4	10	20	35
4							1	5	15
5									1

$m\backslash n$	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
0	1	1	1	1	1	1	1	1	1
1	1	2	3	4	5	6	7	8	9
2		1	3	6	10	15	21	28	36
3				2	7	16	30	50	77
4					1	6	19	45	90
5							3	16	51

Table 1: First few values of $a_{n,m}$ for $k=2$ (left) and $k=3$.

2 Main result

Proposition 1 gives the expression of the generating function $F_k(x,y)$. Even though this result is already obtained in [2], in order to make the paper self-contained we give an alternative proof of it. Then we calculate the generating functions for the frequency of 1s and for the overall number of bits in $\mathcal{B}_n(1^k)$ by means of classic generating functions manipulations (Propositions 2). Applying Theorem 4.1 from [16], after ensuring that its conditions are satisfied, we obtain the main result of this note, Theorem 1. The evolution of the random bit expectation for $k=2$ and $k=3$ is presented on Figure 1 for small values of n. And numerical estimations for the limit value ($n \to \infty$) of the random bit expectation, for small values of k are given in Table 2.

Proposition 1 ([2])**.**

$$F_k(x,y) = \frac{y\left(1-(xy)^k\right)}{y - xy^2 - xy + (xy)^{k+1}}.$$

Proof. The set $\mathcal{B}(1^k) = \bigcup_{n=0}^{\infty} \mathcal{B}_n(1^k)$ respects the following recursive decomposition

$$\mathcal{B}(1^k) = \mathbb{1}_{k-1} \cup \left(\bigcup_{i=0}^{k-1} \left(1^i 0 \cdot \mathcal{B}(1^k) \right) \right)$$

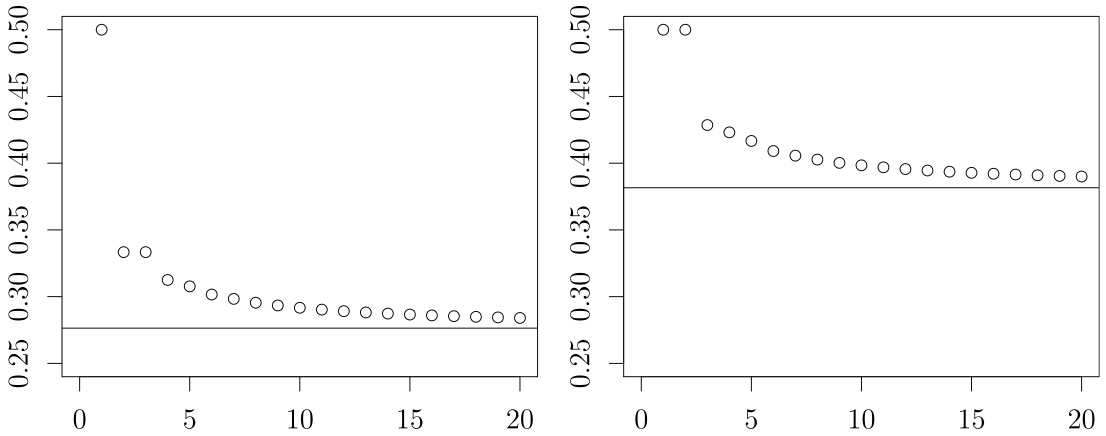


Figure 1: Expected value of a random bit in a random word from $\mathcal{B}_n(1^2)$ (left) and $\mathcal{B}_n(1^3)$ for small values of n.

where $\mathbb{1}_{k-1} = \bigcup_{i=0}^{k-1}\{1^i\}$ is the set of words in $\mathcal{B}(1^k)$ containing no 0s, and $\cdot$ denotes the concatenation. Note that the empty word also lies in $\mathbb{1}_{k-1}$. The claimed generating function is the solution of the following functional equation

$$F_k(x, y) = \sum_{i=0}^{k-1} x^i y^i + F_k(x, y) \sum_{i=0}^{k-1} x^{i+1} y^i.$$

□

In the proof of Theorem 1 we need the following easy to derive results.

Proposition 2.

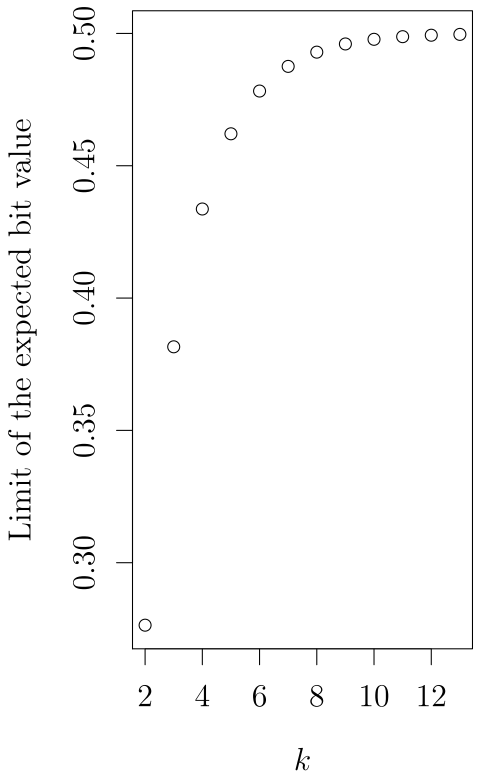
- $P_k(x) = \frac{\partial F_k(x,y)}{\partial y}|_{y=1}$ *is the generating function where the coefficient of* x^n *is the frequency of 1s in* $\mathcal{B}_n(1^k)$*. We have*

$$P_k(x) = \frac{x \cdot \sum_{i=0}^{k-2}(i+1)x^i}{(x^k + x^{k-1} + \cdots + x^2 + x - 1)^2}.$$

- $T_k(x) = x\frac{\partial F_k(x,1)}{\partial x}$ *is the generating function where the coefficient of* x^n *equals the total number of all bits in* $\mathcal{B}_n(1^k)$*. We have*

$$T_k(x) = \frac{x\left(\sum_{i=0}^{k-2}(2i+2)x^i + \sum_{i=k-1}^{2k-2}(2k-i-1)x^i\right)}{(x^k + x^{k-1} + \cdots + x^2 + x - 1)^2}.$$

Every root r of a polynomial $h(x)$ of degree n with a non-zero constant term corresponds to the root $1/r$ of its negative reciprocal $-x^n h(1/x)$. The denominator of both $P_k(x)$ and $T_k(x)$ involves $x^k + x^{k-1} + \cdots + x^2 + x - 1$ and its negative reciprocal is $x^k - x^{k-1} - \cdots - x^2 - x - 1$ which is known in the literature as Fibonacci polynomial, see for instance [6, 7, 9, 10, 11, 13, 14, 15, 19] and references therein. In particular, Dubeau



k	Limit of the expected bit value
2	0.276393202250021
3	0.381580077680607
4	0.433657112297348
5	0.462073883180840
6	0.478227505713290
7	0.487545982771861
8	0.492928265543398
9	0.496019724266083
10	0.497779940783496
11	0.498772398758879
12	0.499326557312936
13	0.499633184444604

Table 2: Numerical estimations for the limit of the expected value of a random bit in a random word from $\mathcal{B}_n(1^k), n \to \infty$.

proved [7, Theorem 1] that its root of the largest modulus is $\varphi_k = \lim_{n\to\infty} f_{n+1,k}/f_{n,k}$, the generalized golden ratio, and φ_k approaches 2 when $k \to \infty$ [7, Theorem 2]. Wolfram [19, Lemma 3.6] showed that any other root r of the Fibonacci polynomial satisfies $3^{-1/k} < |r| < 1$. See Figure 2 for an illustration of this fact. Moreover, Corollary 3.8 in [19] proves that Fibonacci polynomial is irreducible over $\mathbb{Q}$. In order to refer later to them we summarize these results in the next proposition.

Proposition 3. *The polynomial $g_k(x) = x^k + x^{k-1} + \cdots + x^2 + x - 1$ is irreducible over $\mathbb{Q}$, its root of the smallest modulus is unique and equal to $1/\varphi_k$.*

The next lemma says that both fractions representing $P_k(x)$ and $T_k(x)$ are irreducible.

Lemma 1. *The polynomials $\sum_{i=0}^{k-2}(i+1)x^i$ and $x^k + x^{k-1} + \cdots + x^2 + x - 1$ are relatively prime; and so are $\sum_{i=0}^{k-2}(2i+2)x^i + \sum_{i=k-1}^{2k-2}(2k-i-1)x^i$ and $x^k + x^{k-1} + \cdots + x^2 + x - 1$.*

Proof. The polynomial $x^k + x^{k-1} + \cdots + x^2 + x - 1$ is irreducible due to Proposition 3. It does not divide $\sum_{i=0}^{k-2}(i+1)x^i$ as it has a greater degree. And it also cannot divide

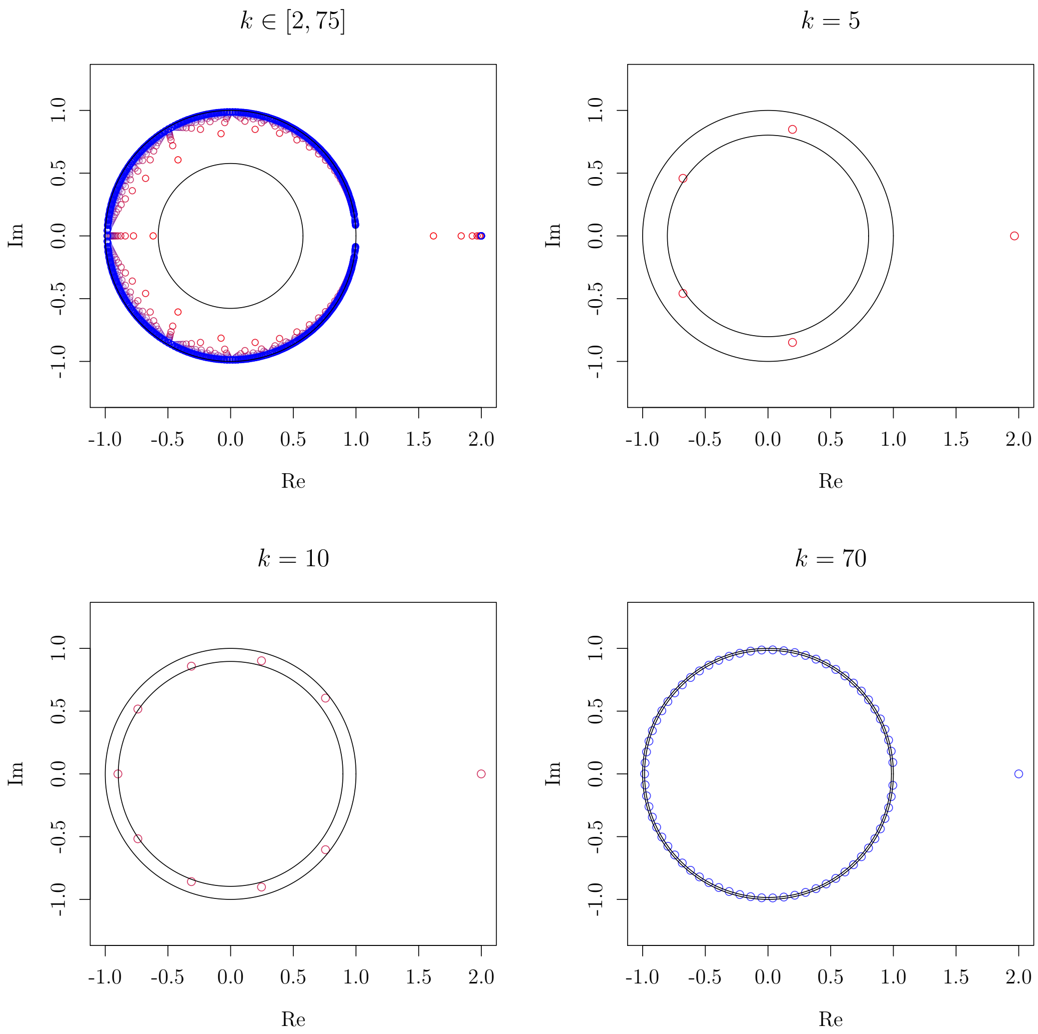


Figure 2: Roots of the polynomial $x^k - x^{k-1} - \cdots - x^2 - x - 1$ (the negative reciprocal of $g_k(x)$) for certain values of k.

$\sum_{i=0}^{k-2}(2i+2)x^i + \sum_{i=k-1}^{2k-2}(2k-i-1)x^i$ as the latter does not have any positive real roots. $\square$

From Propositions 2, 3, Dubeau's results [7], and Lemma 1 we have:

Lemma 2. *Both generating functions $P_k(x)$ and of $T_k(x)$ have the same and unique pole of the smallest modulus with multiplicity 2. The pole equals $1/\varphi_k$, where φ_k is the generalized golden ratio.*

For our main result of this note we need the Theorem 4.1 from [16]:

Theorem 4.1 from [16]. *Assume that a rational generating function $\frac{f(x)}{g(x)}$, with $f(x)$ and $g(x)$ relatively prime and $g(0) \neq 0$, has a unique pole $1/\beta$ of the smallest modulus. Then, if the multiplicity of $1/\beta$ is ν, we have*

$$[x^n]\frac{f(x)}{g(x)} \sim \nu \frac{(-\beta)^\nu f(1/\beta)}{g^{(\nu)}(1/\beta)} \beta^n n^{\nu-1}.$$

Both $P_k(x)$ and $T_k(x)$ are rational generating functions, and by Lemmas 1 and 2 they fulfill the conditions in the above theorem, so

$$[x^n]P_k(x) \sim 2n\varphi_k^{n+2} \cdot \left. \frac{x\left(\sum_{i=0}^{k-2}(i+1)x^i\right)}{\left((x^k+x^{k-1}+\cdots+x^2+x-1)^2\right)''} \right|_{x=1/\varphi_k}$$

$$[x^n]T_k(x) \sim 2n\varphi_k^{n+2} \cdot \left. \frac{x\left(\sum_{i=0}^{k-2}(2i+2)x^i + \sum_{i=k-1}^{2k-2}(2k-i-1)x^i\right)}{\left((x^k+x^{k-1}+\cdots+x^2+x-1)^2\right)''} \right|_{x=1/\varphi_k}.$$

The expected value of a random bit in a random word from $\mathcal{B}_n(1^k)$ is $\frac{[x^n]P_k(x)}{[x^n]T_k(x)}$. Taking the limit, we obtain:

Theorem 1. *The expected value of a random bit in a random word from $\mathcal{B}_n(1^k)$ tends to*

$$\left. \frac{kx^k - kx^{k-1} - x^k + 1}{kx^k - kx^{k-1} + x^{2k} - 3x^k + 2} \right|_{x=1/\varphi_k} \quad \text{when } n \to \infty,$$

where $\varphi_k = \lim_{n\to\infty} f_{n+1,k}/f_{n,k}$ is the generalized golden ratio, in particular φ_2 is the golden ratio.

See Table 2 for some numerical estimations of the result obtained in the previous theorem. This result involves the generalized golden ratio. More than 20 years ago it was conjectured by Wolfram [19] that the Galois group of the polynomial $x^k - x^{k-1} - \cdots - x^2 - x - 1$ is the symmetric group S_k, and so there is no algebraic expression for φ_k (the root of the largest modulus of this polynomial) when $k \geqslant 5$. In case of even or prime k the conjecture was settled by Martin [13]. Cipu and Luca [6] showed that φ_k cannot be constructed by ruler and compass for $k \geqslant 3$. Nevertheless, good approximations are available, for instance Hare, Prodinger and Shallit [11] expressed φ_k and $1/\varphi_k$ in terms of rapidly converging series.

The generalized golden ratio φ_k tends to 2 as k grows, and we deduce the following.

Corollary 1. *The limit of the expected bit value of binary words avoiding k consecutive 1s, whose length tends to infinity, approaches $1/2$ as k grows:*

$$\lim_{k\to\infty} \lim_{n\to\infty} \frac{v_{n,k}}{n \cdot |\mathcal{B}_n(1^k)|} = \frac{1}{2}.$$

Finally, note that other sets of restricted binary words are counted by the generalized Fibonacci numbers, for instance q-decreasing words [2] for $q \geqslant 1$. In this case every length maximal factor of the form 0^a1^b satisfies $a = 0$ or $q \cdot a > b$. Theorem 1 and Corollary 1 apply to this case (with the same limit, see [2, Corollary 5]) by setting $k = q + 1$.

Acknowledgments

We would like to greatly thank Dietrich Burde, Ted Shifrin, Igor Rivin and Sil from Mathematics Stack Exchange[1] for insightful discussions and useful references about the

[1] See the original discussions here https://math.stackexchange.com/questions/4120185 and here https://math.stackexchange.com/questions/4125568.

irreducibility of Fibonacci polynomial which is directly related to Proposition 3. This work was supported in part by the project ANER ARTICO funded by Bourgogne-Franche-Comté region (France).

References

[1] Bajic, D. "On construction of cross-bifix-free kernel sets". $2nd$ MCM COST 2100, TD(07)237, Lisbon, Portugal, 2007.

[2] Baril, J.-L., Kirgizov, S., Vajnovszki, V. "Gray codes for Fibonacci q-decreasing words". *arXiv:2010.09505 [Math.CO]*, 2020.

[3] Bernini, A., Bilotta, S., Pinzani, R., Vajnovszki, V. "A Gray code for cross-bifix-free sets". *Mathematical Structures in Computer Science*, 27(2), 2017, 184-196.

[4] Brauer, A. "On algebraic equations with all but one root in the interior of the unit circle. To my teacher and former colleague Erhard Schmidt on his 75th birthday". *Mathematische Nachrichten*, 4(1-6), 1950, 250-257.

[5] Chee, Y.M., Kiah, H.M., Purkayastha, P., and Wang, C. "Cross-bifix-free codes within a constant factor of optimality". *IEEE Transactions on Information Theory*, 59(7), 2013, 4668–4674.

[6] Cipu, M., and Luca, F. "On the Galois group of the generalized Fibonacci polynomial". *Ann. Şt. Univ. Ovidius Constanţa*, 9(1), 2001, 27–38.

[7] Dubeau, F. "On r-generalized Fibonacci numbers". *The Fibonacci Quarterly*, 27(3), 1989, 221-229.

[8] Eğecioğlu, Ö., and Iršič, V. "Fibonacci-run graphs I: Basic properties". *Discrete Applied Mathematics*, 295, 2021, 70-84.

[9] Flores, I. "Direct calculation of k-generalized Fibonacci numbers". *Fibonacci Quarterly*, 5(3), 1967, 259-266.

[10] Grossman, G.W., and Narayan, S.K. "On the characteristic polynomial of the j-th order Fibonacci sequence", *Applications of Fibonacci Numbers*, 8, Springer, Dordrecht, 1999, 165-177.

[11] Hare, K., Prodinger, H., and Shallit, J. "Three series for the generalized golden mean". *arXiv:1401.6200 [math.NT]*, 2014.

[12] Knuth, D. "The Art of Computer Programming, Volume 3: Sorting and Searching" 2nd ed., Addison-Wesley, 1998.

[13] Martin, P.A. "The Galois group of $x^n - x^n - 1 - \ldots - x - 1$". *Journal of Pure and Applied Algebra*, 190, 2004, 213–223.

[14] Miles, E. "Generalized Fibonacci numbers and associated matrices". *The American Mathematical Monthly*, 67(8), 1960, 745–752.

[15] Miller, M. D. “On generalized Fibonacci numbers”. *The American Mathematical Monthly*, 78(10), 1971, 1108–1109.

[16] Sedgewick, R., Flajolet, P. “An introduction to the analysis of algorithms” 2nd ed., Addison-Wesley, 2013.

[17] Selmer, Ernst S. “On the irreducibility of certain trinomials”. *Mathematica Scandinavica*, 4(2), 1956, 287–302.

[18] Rouché, E. “Mémoire sur la série de Lagrange”. *Journal de l'École Polytechnique*, 22, 1862, 193–224.

[19] Wolfram, D. A. “Solving generalized Fibonacci recurrences”. *Fibonacci Quarterly*, 36(2), 1998, 129–145.