Journal of Prime Research in Mathematics Vol. **8** *(2012), 28-35*

GAPS IN BINARY EXPANSIONS OF SOME ARITHMETIC FUNCTIONS, AND THE IRRATIONALITY OF THE EULER CONSTANT

JORGE JIMÉNEZ URROZ[1], FLORIAN LUCA[2], MICHEL WALDSCHMIDT[3]

Abstract. We show that if $F_n = 2^{2^n} + 1$ is the nth Fermat number, then the binary digit sum of $\pi(F_n)$ tends to infinity with n, where $\pi(x)$ is the counting function of the primes $p \le x$. We also show that if F_n is not prime, then the binary expansion of $\phi(F_n)$ starts with a long string of 1's, where ϕ is the Euler function. We also consider the binary expansion of the counting function of irreducible monic polynomials of degree a given power of 2 over the field $\mathbb{F}_2$. Finally, we relate the problem of the irrationality of Euler constant with the binary expansion of the sum of the divisor function.

Key words : Binary expansions, Prime Number Theorem, Rational approximations to log 2, Fermat numbers, Euler constant, Irreducible polynomials over a finite field.
AMS SUBJECT : 11A63, 11D75, 11N05.

1. Fermat numbers

1.1. **The prime counting function.** Let $F_n = 2^{2^n} + 1$ be the nth Fermat number. In 1650, Fermat conjectured that all the numbers F_n are prime. However, to date it is known that F_n is prime for $n \in \{0, 1, 2, 3, 4\}$ and for no other n in the set $\{5, 6, \ldots, 32\}$. It is believed that F_n is composite for all $n \ge 5$. For more information on Fermat numbers, see [3].

[1] Departamento de Matemática Aplicada IV, Universidad Politecnica de Catalunya, Barcelona, 08034, España. Email: jjimenez@ma4.upc.edu
[2] Instituto de Matemáticas, Universidad Nacional Autonoma de México, C.P. 58089, Morelia, Michoacán, México. Email: fluca@matmor.unam.mx
[3] Université Pierre et Marie Curie Paris 6, Institut de Mathématiques de Jussieu, 4 Place Jussieu, 75252 Paris, Cedex 05, France. Email: miw@math.jussieu.fr.

For a positive real number x we put $\pi(x) = \#\{p \le x\}$ for the counting function of the primes $p \le x$. Consider the sequence $P_n = \pi(F_n)$ for all $n \ge 0$. Observe that F_n is prime if and only if $\pi(F_n) > \pi(F_n - 1)$.

Here, we look at the binary expansion of P_n. In particular, we prove that P_n cannot have few bits in its binary expansion. To quantify our result, let $s_2(P_n)$ be the binary sum of digits of P_n.

Theorem 1. *There exists a constant c_0 such that the inequality*

$$s_2(P_n) > \frac{\log n}{2\log 2} - c_0$$

holds for all $n \ge 0$.

Before proving the theorem we need a preliminary lemma. Let

$$\begin{aligned} (\log 2)^{-1} &= 2^0 + 2^{-2} + 2^{-3} + 2^{-4} + 2^{-8} + \cdots + 2^{-a_k} + \cdots \\ 2 - (\log 2)^{-1} &= 2^{-1} + 2^{-5} + 2^{-6} + 2^{-7} + 2^{-9} + \cdots + 2^{-b_k} + \cdots \end{aligned}$$

where $\mathcal{A} = \{a_k\}_{k\ge 0}$ is the sequence $(0, 2, 3, 4, 8, \ldots)$ giving the position of the kth bit in the binary expansion of $(\log 2)^{-1}$, and $\mathcal{B} = \{b_k\}_{k\ge 0}$ is the sequence $(1, 5, 6, 7, 9, \ldots)$ giving the position of the kth zero coefficient in the binary expansion of $2 - (\log 2)^{-1}$. These sequences are disjoint and their union is the sequence of nonnegative integers.

Lemma 2. *There exist k_0 such that for any $k \ge k_0$ we have $a_{k+1} < 4a_k$ and $b_{k+1} < 4b_k$.*

Proof. By definition

$$(\log 2)^{-1} = \sum_{i=0}^{k} 2^{-a_i} + M,$$

where $M < 2^{-a_{k+1}+1}$. We use the fact (see below) that there exists a constant K such that

$$\left|\frac{1}{\log 2} - \frac{p}{q}\right| > \frac{1}{q^K} \tag{1}$$

holds for all positive rational numbers p/q. We take $q = 2^{a_k}$ and $p = \sum_{i=0}^{k} 2^{a_k - a_i}$, to get

$$\frac{1}{2^{a_{k+1}-1}} = \sum_{m \ge a_{k+1}} \frac{1}{2^m} > \frac{1}{\log 2} - \frac{p}{q} > \frac{1}{2^{Ka_k}},$$

so

$$a_{k+1} < Ka_k + 1 \quad \text{forall} \quad k \ge 0.$$

It is known that we can take $K = 3.58$ for k sufficiently large, say $k \ge k_0$ (see [4]; see also [8] for the fact that we can take $K = 3.9$ for $k \ge k_0$). The result for a_k now follows trivially. The exact same reasoning, substituting $1/\log 2$ by $2 - 1/\log 2$ everywhere, gives the result for b_k. □

Corollary 3. *For each integer n let $\kappa_0 := \kappa_0(n)$ be the largest positive integer k such that $b_k < n-3$ and $\kappa_1 := \kappa_1(n)$ be the largest positive integer k such that $a_k < b_{\kappa_0}$. Then the inequalities $b_{\kappa_0} \geq (n-3)/4$ and $a_{\kappa_1} \geq (n-3)/16$ hold for all sufficiently large n.*

Proof. We have

$$a_{\kappa_1} < b_{\kappa_0} < n-3 \leq b_{\kappa_0+1} < 4b_{\kappa_0}$$

and

$$b_{\kappa_0} < a_{\kappa_1+1} < 4a_{\kappa_1}.$$

□

We now start the proof of theorem 1.1.

Proof. Assume now that $n \geq 4$. By Theorem 1 in [7], we have

$$\frac{x}{\log x}\left(1+\frac{1}{2\log x}\right) < \pi(x) < \frac{x}{\log x}\left(1+\frac{3}{2\log x}\right) \qquad \text{forall} \qquad x \geq 59. \quad (2)$$

Since $F_n \geq F_4 > 59$, we may apply inequality (2) with $x = F_n$. Since both functions $x \mapsto x/\log x$ and $x \mapsto x/(\log x)^2$ are increasing for $x > e^2$, and $F_n > 2^{2^n} > e^2$ for $n \geq 4$, we have that

$$\pi(F_n) \geq \frac{2^{2^n-n}}{\log 2}\left(1+\frac{1}{2^{n+1}\log 2}\right) > 2^{2^n-n}\left(\frac{1}{\log 2}+\frac{1}{2^{n+1}}\right),$$

where we used the fact that $1/2 < \log 2 < 1$. Further,

$$\begin{aligned}
\pi(F_n) &\leq \frac{F_n}{\log F_n}\left(1+\frac{3}{2\log F_n}\right)\\
&< \frac{2^{2^n-n}}{\log 2}\left(1+\frac{3}{2^{n+1}\log 2}\right)\left(1+\frac{1}{2^{2^n}}\right)\\
&< \frac{2^{2^n-n}}{\log 2}\left(1+\frac{3}{2^n}+\frac{1}{2^{2^n}}+\frac{3}{2^{2^n+n}}\right)\\
&< \frac{2^{2^n-n}}{\log 2}\left(1+\frac{1}{2^{n-2}}\right)\\
&< 2^{2^n-n}\left(\frac{1}{\log 2}+\frac{1}{2^{n-3}}\right).
\end{aligned}$$

Hence,

$$P_n = 2^{2^n-n}\left(\frac{1}{\log 2}+\theta_n\right), \qquad \text{where} \qquad \theta_n \in \left(\frac{1}{2^{n+1}}, \frac{1}{2^{n-3}}\right).$$

Thus, the binary digits of P_n are the same as the binary digits of the number $(\log 2)^{-1} + \theta_n$ and, in fact, the first binary digits of P_n are exactly the a_k for all $k \le \kappa_1$ since $a_{\kappa_1} < b_{\kappa_0}$ and, hence, θ_n does not induce a carry over a_{κ_1}. Applying Lemma 2, iteratively, we get that $a_k \le 4^{k-k_0} a_{k_0}$ for all $k \ge k_0$. Hence,

$$s_2(P_n) > \sum_{j \le \kappa_1, a_j \in \mathcal{A}} 1 = \kappa_1 \ge \frac{\log a_{\kappa_1}/a_{k_0}}{\log 4} + k_0 > \frac{\log n}{\log 4} - c_0,$$

by Corollary 3. □

1.2. **The Euler function.** Let $\phi(n)$ be the Euler function of n. If F_n is prime, then $\phi(F_n) = 2^{2^n}$. We show that if F_n is not prime, then the binary expansion of $\phi(F_n)$ starts with a long string of 1's. More precisely, we have the following result.

Theorem 4. *If F_n is not prime, then the binary expansion of $\phi(F_n)$ starts with a string of* 1*'s of length at least* $n - \lfloor \log n / \log 2 \rfloor - 1$.

We need the following well known lemma (see Proposition 3.2 and Theorem 6.1 in [3]).

Lemma 5. *Any two Fermat numbers are coprime. Further, for $n \ge 2$, each prime factor of F_n is congruent to* 1 *modulo* 2^{n+2}.

Proof. Let $n \ge 0$ and let p be a prime factor of F_n. Since 2^{2^n} is congruent to -1 modulo p, the order of the class of 2 in the multiplicative group $(\mathbb{Z}/p\mathbb{Z})^\times$ is 2^{n+1}. This shows that two Fermat numbers have no common prime divisor.

Assume now $n \ge 2$. Then p is congruent to 1 modulo 8, hence 2 is a square modulo p. Let a satisfy $a^2 \equiv 2 \pmod p$. Then the order of the class of a in the multiplicative group $(\mathbb{Z}/p\mathbb{Z})^\times$ is 2^{n+2} and therefore 2^{n+2} divides $p-1$. □

Proof of Theorem 4. Assume that F_n is not prime. Then $n \ge 5$. Write $F_n = \prod_{i=1}^k p_i^{\alpha_i}$, where $p_1 < \cdots < p_k$ are distinct primes and $\alpha_1, \ldots, \alpha_k$ are positive integer exponents. Using Lemma 5, we can write $p_i = 2^{n+2} m_i + 1$ for each $i = 1, \ldots, k$. Further, no m_i is a power of 2, for otherwise p_i itself would be a Fermat prime, which is false because any two Fermat numbers are coprime by Lemma 5. Let $\mathcal{P}$ be the set of primes $p \equiv 1 \pmod{2^{n+2}}$ which are not Fermat primes and for any positive real number x let $\mathcal{P}(x) = \mathcal{P} \cap [1, x]$. Then $p_i \in \mathcal{P}(2^{2^n})$. Thus,

$$\sum_{i=1}^k \frac{1}{p_i} \le \sum_{\substack{2^{n+2}\cdot 3 \le p \le 2^{2^n} \\ p \in \mathcal{P}}} \frac{1}{p} = \frac{\#\mathcal{P}(t)}{t}\Big|_{t=2^{n+2}\cdot 3}^{t=2^{2^n}} + \int_{2^{n+2}\cdot 3}^{2^{2^n}} \frac{\#\mathcal{P}(t)}{t^2} dt,$$

where the above equality follows from the Abel summation formula. In order to estimate the first term and the integral, we use the fact that

$$\#\mathcal{P}(t) \le \pi(t,1,2^{n+2}) \le \frac{2t}{\phi(2^{n+2})\log(t/2^{n+2})} \quad \text{forall} \quad t \ge 2^{n+2}\cdot 3,$$

where $\pi(t;a,b)$ is the number of primes $p \le t$ in the arithmetic progression $a \pmod b$. The right–most inequality is due to Montgomery and Vaughan [5]. Thus,

$$\begin{aligned}
\sum_{i=1}^{k}\frac{1}{p_i} &\le \frac{\#\mathcal{P}(2^{2^n})}{2^{2^n}} + \int_{2^{n+2}\cdot 3}^{2^{2^n}} \frac{\#\mathcal{P}(t)}{t^2}dt\\
&\le \frac{1}{2^n\log(2^{2^n-n-2})} + \frac{1}{2^n}\int_{2^{n+2}\cdot 3}^{2^{2^n}} \frac{dt}{t\log(t/2^{n+2})}\\
&< \frac{1}{2^n} + \frac{1}{2^n}\int_{3}^{2^{2^n-n-2}} \frac{du}{u\log u} \qquad (u := t/2^{n+2})\\
&= \frac{1}{2^n} + \frac{\log\log u}{2^n}\Big|_{u=3}^{u=2^{2^n-n-2}} < \frac{1}{2^n} + \frac{\log((2^n-n-2)\log 2)}{2^n} < \frac{n}{2^n}.
\end{aligned}$$

Using the inequality

$$1 - \prod_{i=1}^{k}(1-x_i) < \sum_{i=1}^{k} x_i$$

valid for all $k \ge 1$ and $x_1,\ldots,x_k \in (0,1)$ with $x_i = 1/p_i$ for $i = 1,\ldots,k$, we get

$$1 - \frac{\phi(F_n)}{F_n} = 1 - \prod_{i=1}^{k}\left(1 - \frac{1}{p_i}\right) < \sum_{i=1}^{k}\frac{1}{p_i} < \frac{n}{2^n},$$

therefore

$$\phi(F_n) > F_n\left(1 - \frac{n}{2^n}\right) > 2^{2^n}\left(1 - \frac{n}{2^n}\right),$$

which together with the fact that $\phi(F_n) < F_n - 1 = 2^{2^n}$ (since F_n is composite) implies the desired conclusion. □

2. Digits of the number of irreducible polynomials of a given degree over a finite field

Let $\mathbb{F}_q$ be a finite field with q elements. For any positive integer m, denote by $N_q(m)$ the number of monic irreducible polynomials over $\mathbb{F}_q$ of degree m. Then (see for instance §14.3 of [1]) for each $m \ge 1$, we have

$$q^m = \sum_{d|m} dN_q(d) \quad \text{and} \quad N_q(m) = \frac{1}{m}\sum_{d|m}\mu(d)q^{m/d}.$$

The two formulae are equivalent by Möbius inversion formula. From the first one, given the fact that all the elements in the sum are positive, we deduce

$$N_q(m) < \frac{q^m}{m} \quad \text{for } m \geq 2.$$

A consequence of the second one is

$$q^m - mN_q(m) = -\sum_{d|m, d<m} \mu(d)q^{d/m} \leq q^{m/2} + \sum_{d \leq m/3} q^d < 2q^{m/2} \quad \text{for } m \geq 2.$$

Hence, we have

$$\frac{q^m}{m} - \frac{2q^{m/2}}{m} < N_q(m) < \frac{q^m}{m}\cdot$$

For $q = 2$ and $m = 2^n$ we deduce that the number $\tilde{P}_n := N_2(2^n)$ satisfies

$$2^{2^n - n} - 2^{2^{n-1} - n + 1} < \tilde{P}_n < 2^{2^n - n}.$$

It follows that the binary expansion of $\tilde{P}_n$ starts with a number of 1's at least $2^{n-1} - 1$.

3. Irrationality of the Euler constant

Let $T_k = \sum_{n \leq 2^k} \tau(n)$ with $\tau(n) = \sum_{d|n} 1$ and let $T_k = \sum_{i=0}^{v_k} a_i 2^i$ be its binary expansion. If we have $a_{\ell+i} = 0$, for any $0 \leq i \leq L-1$, we say that T_k has a *gap of length at least L starting at ℓ.*

We introduce the following condition depending on a parameter $\kappa > 0$ and involving Euler's constant γ.

Assumption (A_κ): *There exists a positive constant B_0 with the following property. For any $(b_0, b_1, b_2) \in \mathbb{Z}^3$ with $b_1 \neq 0$, we have*

$$|b_0 + b_1 \log 2 + b_2\gamma| \geq B^{-\kappa}$$

with

$$B = \max\{B_0,\ |b_0|,\ |b_1|,\ |b_2|\}.$$

From Dirichlet's box principle (see [9]), it follows that if condition (A_κ) is satisfied, then $\kappa \geq 2$. According to (1), condition (A_κ) is satisfied with $\kappa = 3$ if Euler's constant γ is rational. It is likely that it is also satisfied if γ is irrational, but this is an open problem. A folklore conjecture is that $1, \log 2$ and γ are linearly independent over $\mathbb{Q}$. If this is true, then (A_κ) can be seen as a measure of linear independence of these three numbers. It is known (see [9]) that for almost all tuples $(x_1, \ldots, x_m)$ in $\mathbb{R}^m$ in the sense of Lebesgue's measure, the following measure of linear independence holds:

For any $\kappa > m$, there exists a positive constant B_0 such that, for any $(b_0, b_1, \ldots, b_m) \in \mathbb{Z}^{m+1} \setminus \{0\}$, we have

$$|b_0 + b_1x_1 + \cdots + b_mx_m| \geq B^{-\kappa}$$

with

$$B = \max\{B_0,\ |b_0|,\ |b_1|, \ldots,\ |b_m|\}.$$

It is also expected that *most* constants from analysis, like $\log 2$ and γ, behave, from the above point of view, as almost all numbers. Hence, one should expect condition (A_κ) to be satisfied for any $\kappa > 2$.

Theorem 6. *Assume κ is a positive number such that the condition (A_κ) is satisfied. Then, for any sufficiently large k, any ℓ and L satisfying*

$$2 + \kappa \frac{\log k}{\log 2} \leq k - \ell \leq L,$$

T_k does not have a gap of length at least L starting at ℓ.

Proof. Assume k is large enough, in particular $k > B_0$. It is well known (see, for instance, Theorem 320 in [2]), that

$$\sum_{n \leq x} \tau(n) = x \log x + (2\gamma - 1)x + E(x),$$

where $|E(x)| < c_1\sqrt{x}$ for some positive constant c_1. For $x = 2^k$, we get

$$T_k = 2^k k \log 2 + 2^k(2\gamma - 1) + E(2^k). \tag{3}$$

Suppose now that T_k has a gap of length at least L starting at ℓ. Then the binary expansion of T_k is $T_k = \sum_{i=\ell+L}^{v_k} a_i 2^i + \sum_{i=0}^{\ell-1} a_i 2^i$, and, by (3), we get

$$\left| 2^k k \log 2 + 2^{k+1}\gamma - b \right| < 2^\ell + E(2^k),$$

with $b = 2^k + \sum_{i=\ell+L}^{v_k} a_i 2^i$. Now, since $\ell + L \geq k$ and $2^k | b$, we can first divide by 2^k, and then apply Assumption (A_κ) with $b_0 = -2^{-k}b$, $b_1 = k$, $b_2 = 2$ to obtain

$$k^{-\kappa} \leq |k \log 2 + 2\gamma + b_0| < 2^{\ell-k} + 2^{-k/2+2}. \tag{4}$$

Observe that in this case, in Assumption (A_κ) we have $B \leq k$ since for k sufficiently large we have $b \leq 2^k + T_k < k2^k$. We just have to observe that the inequality $2^{-k/2+2} < k^{-\kappa}/2$ holds for all sufficiently large k, to conclude that the last inequality (4) is impossible for any ℓ in the range $k \geq \ell + 2 + \kappa(\log k)/\log 2$. □

As a corollary of Theorem 6 and inequality (1), we give a criterion for the irrationality of the Euler's constant.

Corollary 7. *Assume that for infinitely many positive integers k, there exist ℓ and L satisfying*

$$2 + 3\left(\frac{\log k}{\log 2}\right) \leq k - \ell \leq L$$

and such that T_k has a gap of length at least L starting at ℓ. Then Euler's constant γ is irrational.

4. Further comments

There are other similar games we can play in order to say something about the binary expansion of the average of other arithmetic functions evaluated in powers of 2 or in Fermat numbers, once the average value of such a function involves a constant for which we have a grasp on its irrationality measure. For example, using the fact (see for instance [2] §18.4, Th. 330) that

$$A(x) = \sum_{n \le x} \phi(n) = \frac{1}{2\zeta(2)} x^2 + O(x \log x)$$

together with the fact that the approximation exponent of $\zeta(2) = \pi^2/6$ is smaller than 5.5 (see [6]), then $s_2(A(2^n)) > (\log n)/\log(5.5) - c_2$, where c_2 is some positive constant. We give no further details.

Acknowledgements

This work was done when all authors were in residence at the Abdus Salam School of Mathematical Sciences in Lahore, Pakistan. They thank the institution for its hospitality.

References

[1] D. S. Dummit and R. M. Foote, *Abstract algebra*, John Wiley & Sons Inc., Hoboken, NJ, third ed., 2004.

[2] Hardy-Wright, *An introduction to the theory of numbers.* Sixth edition. Oxford University Press, Oxford, 2008.

[3] M. Křížek, F. Luca and L. Somer, *17 Lectures on Fermat Numbers: From Number Theory to Geometry*, CMS books in mathematics, **10**, Springer New York, 2002.

[4] R. Marcovecchio, *The Rhin-Viola method for* $\log 2$, Acta Arithmetica. **139** (2009), 147–184.

[5] H. L. Montgomery and R. C. Vaughan, *The large sieve*, Mathematika. **20** (1973), 119–134.

[6] G. Rhin and C. Viola, *On a permutation group related to* $\zeta(2)$, Acta Arith. **77** (1996), 23–56.

[7] J. B. Rosser and L. Schoenfeld, *Approximate formulas for some functions of prime numbers*, Illinois J. Math. **6** (1962), 64–94.

[8] E. A. Rukhadze, *A lower bound for the approximation of* $\ln 2$ *by rational numbers*, Vestnik Moskov. Univ. Ser. I Mat. Mekh. **25** (1987) 25–29, 97 (in Russian).

[9] M. Waldschmidt, *Recent advances in Diophantine approximation*, D. Goldfeld et al. (eds.), Number Theory, Analysis and Geometry: In Memory of Serge Lang, Springer Verlag, to appear in 2012.