On the counting function of irregular primes

Florian Luca*
School of Mathematics, University of the Witwatersrand
P. O. Box Wits 2050, South Africa
fluca@matmor.unam.mx

Amalia Pizarro-Madariaga
Departamento de Matemáticas
Universidad de Valparaiso, Chile
amalia.pizarro@uv.cl

Carl Pomerance
Department of Mathematics
Dartmouth College
Hanover, NH 03755-3551, USA
carl.pomerance@dartmouth.edu

Abstract

It is well-known that there are infinitely many irregular primes. We prove a quantitative version of this statement, namely the number of such primes $p \le x$ is at least $(1+o(1))\log\log x/\log\log\log x$ as $x \to \infty$. We show that the same conclusion holds for the irregular primes corresponding to the Euler numbers. Under some conditional results from diophantine approximation, the above lower bounds can be improved to $\gg \log x/(\log\log x)^2$.

2010 Mathematics Subject Classification: Primary 11B68

1 Introduction

The Bernoulli numbers $\{B_m\}_{m\ge 0}$ are defined via their exponential generating function

$$\frac{t}{e^t-1} = \sum_{m\ge 0} B_m \frac{t^m}{m!}. \tag{1}$$

*Also affiliated with the Mathematical Institute, UNAM Juriquilla 76230 Santiago de Querétaro, México

The first few values are

$$B_0 = 1,\ B_1 = -1/2,\ B_2 = 1/6,\ B_3 = 0,\ B_4 = -1/30,\ B_5 = 0,\ B_6 = 1/42.$$

Evaluating relation (1) at $-t$ and subtracting the resulting formula from (1), one gets that $B_m = 0$ for all odd $m \geq 3$. There are several explicit formulas for computing B_m as well as the recurrences among them such as

$$\sum_{k=0}^{m-1} \binom{m}{k} B_k = 0 \qquad \text{for all} \qquad m \geq 2,$$

which can be used to compute B_{m-1} in terms of B_j for $j \in \{0, \ldots, m-2\}$.

A prime $p > 2$ is called regular if it divides the class number of the cyclotomic field $\mathbb{Q}(\zeta_p)$, where $\zeta_p = e^{2\pi i/p}$ is a nontrivial pth root of unity. In 1850, Kummer [9] showed that p is regular if and only if it does not divide the numerator of any of the numbers

$$B_2, B_4, \ldots, B_{p-3}.$$

The first few regular primes are

$$3,\ 5,\ 7,\ 11,\ 13,\ 17,\ 19,\ 23,\ 29,\ 31,\ 41,\ \ldots.$$

This is sequence A007703 in [18]. In 1964, Siegel [17] conjectured that the regular primes have relative density $1/\sqrt{e}$ as a subset of all the primes. However, it is not even known that there are infinitely many regular primes. An odd prime which is not regular is called irregular. The first few irregular primes are

$$37,\ 59,\ 67,\ 101,\ 103,\ 131,\ 149,\ \ldots. \tag{2}$$

This is sequence A000928 in [18]. Unlike with the regular primes, it is known that there are infinitely many irregular primes. The first proof of this fact was given in 1915 by Jensen in [8], who in fact showed that there are infinitely many irregular primes congruent to 3 modulo 4. Almost 40 years later, in 1954, Carlitz [3], gave a simple proof of the weaker result that there are infinitely many irregular primes. Jensen's result was extended to the existence of irregular primes in other congruence classes by Montgomery [15] and Metsänkylä [14].

Let

$$\mathcal{I}_B = \{p : p \text{ irregular}\}$$

and let $\mathcal{I}_B(x) = \mathcal{I}_B \cap [1, x]$. Our main result is the following.

Theorem 1. *The inequality*

$$\#\mathcal{I}_B(x) \geq (1+o(1))\frac{\log\log x}{\log\log\log x}$$

holds as $x \to \infty$.

Let $\{E_m\}_{m\geq 0}$ be the sequence of Euler numbers whose exponential generating function is given by

$$\sec t = \sum_{m\geq 0} E_m \frac{t^m}{m!}$$

This is sequence A122045 in [18]. There are some similarities with the Bernoulli numbers. For example, $E_m = 0$ for all odd m including $m = 1$. However, unlike the Bernoulli numbers, the Euler numbers are integers. Long before Wiles' proof of Fermat Last Theorem, Vandiver [20] proved that if

$$x^p + y^p = z^p$$

holds for some positive integers $x,\ y,\ z$ and an odd prime p such that $p \nmid xyz$, then $E_{p-3} \equiv 0 \pmod{p}$. Gut [6], proved that if

$$x^{2p} + y^{2p} = z^{2p}$$

is satisfied for some positive integers $x,\ y,\ z$ and an odd prime $p \geq 11$ with $p \nmid xyz$, then

$$E_{p-3} \equiv E_{p-5} \equiv E_{p-7} \equiv E_{p-9} \equiv E_{p-11} \pmod{p}.$$

Inspired by the above results, Carlitz [3], called an odd prime p to be irregular with respect to the Euler numbers if it divides one of the numbers

$$E_2,\ E_4,\ \ldots, E_{p-3}.$$

He proved that the number of such primes is infinite. Accordingly, we put

$$\mathcal{I}_E = \{p : p \text{ irregular for the Euler numbers}\}$$

and let $\mathcal{I}_E(x) = \mathcal{I} \cap [1, x]$. The following result is the analog of Theorem 1 for the counting function of the irregular primes with respect to the Euler numbers.

Theorem 2. *The inequality*

$$\#\mathcal{I}_E(x) \geq (1+o(1))\frac{\log\log x}{\log\log\log x}$$

holds as $x \to \infty$.

The rest of the paper is organized as follows. In Section 2, we present Carlitz's proof of the infinitude of the irregular primes. As we shall see, this proof does not allow the conclusion that the function $\#\mathcal{I}_B(x)$ grows faster than some finite iterate of the natural logarithm function. In Section 3, we show that if $u > 0$ and v are fixed integers, then most primes p have the property that the linear form $up + v$ does not have any large divisor of the form $q - 1$ with $q \neq p$ a prime. Since this result is of independent interest and might have other applications, we state it here. Let $\pi(x)$ denote the number of primes in $[1, x]$.

Theorem 3. *Let* $u > 0$ *and* v *be integers. Let* $2 \leq z \leq x$. *There is a positive absolute constant* c *such that the number of primes* $p \leq x$ *such that* $up + v$ *has a divisor* $q - 1$ *with* $q > z,\ q \neq p$ *and* q *prime is* $O(\pi(x)/(\log z)^c)$. *The implied constant might depend on* u *and* v.

The proof of the above result uses sieve methods and follows some ideas from [5]. In Section 4, we recall Siegel's zero-lemma, and Baker's lower bound on a non-zero linear linear form in logarithms of algebraic numbers in a recent formulation due to Matveev. For us, only the rational case is of interest. We also recall a stronger conjectural form of this result due to Lang and Waldschmidt (see [11], [21]). Sections 5 and 6 are devoted to the proofs of Theorems 1 and 2. These proofs combine Theorem 3 with Matveev's bound. In Section 7, we show that the the lower-bounds can be strengthened to $\gg \log x/(\log\log x)^2$ if instead of Matveev's bound we use the Lang-Waldschmidt conjecture in our arguments.

We use $\omega(n),\ \Omega(n),\ \tau(n)$ for the number of distinct prime factors of n, the total number of prime factors of n (counting multiplicities), the number of divisors of n, respectively. We also let $P^+(n)$ and $P^-(n)$ denote the largest and smallest prime factor of n, respectively, with the conventions that $P^+(1) = 1$ and $P^-(1) = +\infty$. We use $p,\ q,\ r$ with or without subscripts to denote primes. We use the Landau symbols $O,\ o$ and the Vinogradov symbols $\gg$ and $\ll$ with their usual meaning. We recall that $A = O(B),\ A \ll B$ and $B \gg A$ are all equivalent to the fact that $|A| < cB$ for some constant c, whereas $A = o(B)$ means that $A/B \to 0$. The constants implied by such symbols are absolute except for Section 3 where they depend on u and v.

2 Carlitz's proof revisited

Carlitz's proof is based on the following three congruences valid for all even positive integers m:

$$B_m \equiv 0 \pmod{p^r} \qquad (p^r \mid m, \quad p-1 \nmid m), \tag{3}$$

$$pB_m \equiv -1 \pmod{p} \qquad (p-1 \mid m), \tag{4}$$

$$\frac{B_{m'}}{m'} \equiv \frac{B_m}{m} \pmod{p} \qquad (m' \equiv m \not\equiv 0 \pmod{p-1}) \tag{5}$$

(see formulas (2.1), (2.2) and (2.3) on page 329 of [3]). The second congruence (4) follows from the the theorem of Staudt and von Clausen which asserts that

$$B_m \equiv -\sum_{p-1|m} \frac{1}{p} \pmod{1} \qquad \text{for all even } m.$$

Lemma 1. *An odd prime p is irregular if and only if p divides the numerator of some B_m/m, where $m > 0$ is even.*

Proof. If p is irregular, it divides the numerator of some B_m for some even m with $2 \le m \le p-3$, and so it divides the numerator of B_m/m. Conversely, if p divides the numerator of some B_m/m with $m > 0$ even, then (4) implies that $p-1 \nmid m$, so that (5) implies that p divides the numerator of $B_{m'}/m'$, where $m' = m \mod (p-1)$. Note that m' is even and that $m' \neq 0$. This completes the proof. □

Let $k \ge 1$ and $p_1, \ldots, p_k$ be the first k irregular primes. We put

$$M = \operatorname{lcm}[p_1 - 1, \ldots, p_k - 1].$$

Write $|B_M|/M = C_M/D_M$, where C_M and D_M are coprime. It follows from Lemma 1 that every prime factor of C_M is irregular. Since $p_i - 1 \mid M$ for all $i = 1, \ldots, k$, it follows from (4), that p_i divides the denominator of B_M for all $i = 1, \ldots, k$. Hence, $\gcd(C_M, M) = 1$, so either $C_M = 1$, or C_M is divisible by some irregular prime p_{k+1} not among $p_1, \ldots, p_k$, and in particular, $C_M \ge p_{k+1}$. We now exploit the relation

$$\frac{C_M}{D_M} = \frac{|B_M|}{M} = \frac{2(M-1)!}{(2\pi)^M}\zeta(M)$$

due to Euler. Since $n! > (n/e)^n$ for all $n \ge 1$ and $\zeta(M) > 1$, it follows that

$$\frac{C_M}{D_M} \ge \frac{2e}{M-1}\left(\frac{M-1}{2\pi e}\right)^M > \frac{2e}{M-1} \cdot 2^M > M^2 \tag{6}$$

where the above inequalities hold because $M \geq p-1 \geq 36$. This shows that the case $C_M = 1$ is not possible, therefore

$$p_{k+1} \leq C_M \leq 2D_M \frac{(M-1)!}{(2\pi)^M} \zeta(M).$$

Using the Staudt–von Clausen theorem, which implies that the denominator of B_M is $\prod_{p-1|M} p$, and putting $\tau(n)$ for the number of divisors of n, we have

$$D_M \leq M \prod_{p-1|M} p \leq M(M+1)^{\tau(M)} \leq M(M+1)^M,$$

because $\tau(n) \leq n$ for all $n \geq 1$. Since $\zeta(M) < 2$ and $(M-1)! < (M-1)^{M-1}$, we get

$$p_{k+1} \leq C_M < \frac{4M(M+1)}{(2\pi)^M} (M+1)^{M-1} (M-1)^{M-1} < (M^2-1)^{M-1} < M^{2M}.$$

The above upper bound can be somewhat improved but not by much due to the presence of the factor $(M-1)!/(2\pi)^{M-1} > ((M-1)/(2e\pi))^{M-1}$. Note also that $M \geq p_k - 1$. Hence,

$$p_{k+1} < M^{2M} < (p_1 \cdots p_k)^{2p_1 \cdots p_k}$$

The above argument reveals that the Carlitz argument only produces a bound on p_{k+1} which is exponential in p_k, so as a consequence, it cannot produce a lower bound on $\mathcal{I}_B(x)$ which is of order a finite number of iterates of the natural logarithm.

3 Linear forms in primes with large shifted prime divisors

Here, we prove Theorem 3. We follow some of the ideas in the proof of [5, Theorem 2]. Our proof uses results on the distribution of y–smooth numbers $n \leq x$, that is numbers $n \leq x$ with $P^+(n) \leq y$, due to de Bruijn [1], the distribution of primes $q \leq x$ having $\Omega(q-1)/\log\log x$ away from 1 due to Erdős [4] and Timofeev [19], and Brun's sieve. For most of our sieving applications, Theorem 2.3 on page 70 in [7] will suffice. Throughout this section, the letters $c_1, c_2, \ldots$ denote absolute positive constants and the symbols O, $\ll$ and $\gg$ depend on u and v.

We start with the case $v = 0$. In this case, up is divisible by $q-1$, for some $q > z$. Taking $z > u$, we conclude that $q - 1 = u_1 p$ for some divisor

u_1 of u. For fixed $u_1 \mid u$, the number of primes $p \le x$ such that $u_1 p + 1$ is also prime is, by Brun's sieve, $O(\pi(x)/\log x)$. Summing this bound up over all divisors u_1 of u, we get the desired conclusion of Theorem 3 with $c = 1$.

From now on, we assume that $v \neq 0$.

Lemma 2. *The number of primes $p \le x$ with $up + v$ divisible by some $q - 1$ with $q > z$ prime, $q \neq p$, and $\Omega(q-1) \le \frac{2}{3}\log\log q$ is $O(\pi(x)/(\log z)^{c_1})$.*

Proof. First assume $q \le \sqrt{x}$. For each choice of q having $\Omega(q-1) \le \frac{2}{3}\log\log q$, we count integers $a \le (ux+v)/(q-1)$ such that $a(q-1) = up+v$ for some prime $p \le x$. This puts $p \le x$ into a certain arithmetic progression p^* modulo $(q-1)/\gcd(q-1, u)$. By Brun's method (in this case, the Brun-Titchmash theorem), this count is

$$O\left(\frac{x}{\varphi(q-1)\log(x/q)}\right) = O\left(\pi(x)\frac{\log\log q}{q}\right), \tag{7}$$

where φ is Euler's function. We know from [4], or [19], that the counting function of the primes $q \le t$ with $\Omega(q-1) \le \frac{2}{3}\log\log q$ is $O(\pi(t)/(\log t)^{c_2})$. Applying this result and partial summation, the above estimate (7) summed for $q > z$ gives the count $O(\pi(x)\log\log z/(\log z)^{c_2})$. This is consistent with the conclusion of the lemma for any choice of $c_1 < c_2$.

It remains to consider the case of $q > z \ge \sqrt{x}$. By de Bruijn's standard result [1], the number of integers $n \le t$ for which $P^+(n) \le t^{2/\log\log t}$ is $O(t/(\log t)^{10})$ (actually any fixed number may be used here instead of "10"). Thus, putting $y = x^{1/\log\log x}$, we may assume that $q - 1 = br$ for r prime and $r > y$. For a prime $p \neq q$ with $q - 1 = br \mid up + v$, let $a = (up+v)/br$. We now fix a, b with $\Omega(b) \le \frac{2}{3}\log\log x$ and we count primes $r \le (ux+v)/ab$ with $br + 1 = q$ prime and abr of the form $up + v$ with p prime, $p \neq q$. That is, for large x, the three integers r, $br + 1$ and $abr - v$ are all free of prime factors from the interval $(u, y^{1/3}]$. We put

$$E_1 = abv(v + a).$$

Note that $E_1 \neq 0$, for if $E_1 = 0$, then either $v = 0$, which has been treated above, or $v = -a$. Should the last instance occur, we would then have $up - a = up + v = abr = a(q-1) = aq - a$, therefore $up = aq$. Taking $z > u$, we conclude that $q = p$, which is excluded. Thus, $E_1 \neq 0$. By Theorem 2.3 on page 70 in [7], the count of such r is

$$O\left(\frac{x}{ab(\log(y^{1/3}))^3}\left(\frac{E_1}{\varphi(E_1)}\right)^2\right) = O\left(\frac{x(\log\log x)^5}{ab(\log x)^3}\right). \tag{8}$$

It is known, see [10, Corollary 2.5] for example, that sum of $1/b$ over all numbers with $P^+(b) \le x$ and $\Omega(b) \le \frac{2}{3}\log\log x$ is $O((\log x)^{1-c_3})$. The sum on a contributes a factor $\log x$, so we have a final count of $O(\pi(x)/(\log x)^{c_4})$ for any fixed $c_4 < c_3$. This completes the proof of the lemma. □

We let $\Omega_y(n)$ denote the number of prime factors $\ell \le y$ of n counted with multiplicity.

Lemma 3. *The number of primes $p \le x$ with $up+v$ of the form $a(q-1)$ with $q > z$ prime, $q \ne p$, and $\Omega_q(a) \le \frac{2}{3}\log\log q$ is $O(\pi(x)/(\log z)^{c_5})$.*

Proof. It is convenient to consider three ranges for q.

Case 1. $q \le e^{\sqrt{\log x}}$.

Write $a = a_1a_2$, where $P^+(a_1) \le q$ and $P^-(a_2) > q$. Again, by [1] and an easy argument, we may assume that $a_1 \le \sqrt{x}$. For fixed a_1, q we count integers $a_2 \le (ux+v)/(a_1(q-1))$ with $P^-(a_2) > q$ and $a_1a_2(q-1)$ of the form $up+v$ with p prime. By Brun's method again, this count is

$$O\left(\frac{x}{\varphi(a_1(q-1))\log q\log x}\right) = O\left(\pi(x)\frac{\log\log q}{a_1q\log(a_1q)}\right).$$

By [10, Corollary 2.5], the sum on a_1 with $\Omega(a_1) \le \frac{2}{3}\log\log q$ then introduces a factor $(\log q)^{1-c_3}$. The sum over $q > z$ now introduces a factor $(\log\log z)/(\log z)^{c_3}$, so we have a final count that is $O(\pi(x)/(\log z)^{c_4})$ as desired. (If $z > e^{\sqrt{\log x}}$, the count is 0.)

Case 2. $e^{\sqrt{\log x}} < q \le \sqrt{x}$.

As in the proof of Lemma 2 we may assume that $P^+(a) > y = x^{1/\log\log x}$. Write a as $a'r$ where r is prime, $r > y$, and $\Omega_q(a') \le \frac{2}{3}\log\log q$. For q, a' fixed, we then have $r \le (ux+v)/a'(q-1)$ is prime and $a'r(q-1)$ is of the form $up+v$ with p prime. Thus, r and $a'(q-1)r - v$ are both free of primes in $(u, y^{1/3}]$. Taking

$$E_2 = a'(q-1)v,$$

we have $E_2 \ne 0$ and the count on the number of such r is

$$O\left(\frac{x}{a'(q-1)(\log(y^{1/3}))^2}\left(\frac{E_2}{\varphi(E_2)}\right)\right) = O\left(\frac{x(\log\log x)^3}{a'(q-1)(\log x)^2}\right).$$

As above, the sum over a' introduces a factor $(\log x)/(\log q)^{c_3}$, and then the sum over $q > \max\{z, e^{\sqrt{\log x}}\}$ of the resulting expression gets us to $O(\pi(x)/(\log z)^{c_4})$.

Case 3. $q > \sqrt{x}$.

As in the proof of Lemma 2, we write $q - 1 = br$ where r is prime and $r > y$. For a given integer a with $\Omega(a) \le \frac{2}{3}\log\log x$ and a given value of b, we count primes $r \le (ux+v)/ab$ with $br + 1 = q$ prime and abr of the form $up + v$ with p prime, $p \ne q$. By Brun's method, this count is given by (8), so the argument that follows in the proof of Lemma 2, with the roles of a and b reversed, finishes the argument here. □

With the following lemma, the proof of Theorem 3 will be complete.

Lemma 4. *The number of primes $p \le x$ for which there is some number $w > z$ with $\Omega_w(up+v) > \frac{4}{3}\log\log w$ is $O(\pi(x)/(\log z)^{c_6})$.*

Proof. This time we consider two cases.

Case 1. $w \le e^{\sqrt{\log x}}$.

Write $up + v = a_1 a_2$, where $P^+(a_1) \le w$ and $P^-(a_2) > w$. Again by [1], we may assume that $a_1 \le \sqrt{x}$. For a_1 fixed with $\Omega(a_1) > \frac{4}{3}\log\log w$, the number of choices for $a_2 \le (ux+v)/a_1$ with $P^-(a_2) > y$ and $a_1 a_2$ of the form $up + v$ with p prime is, via Brun's method, at most

$$O\left(\frac{x}{\varphi(a_1)\log w \log x}\right) = O\left(\frac{x\log\log w}{a_1 \log w \log x}\right).$$

Summing the above inequality on a_1 and using [10, Corollary 2.5], we have a count of

$$O\left(\pi(x)\frac{\log\log w}{(\log w)^{c_7}}\right).$$

Case 2. $w > e^{\sqrt{\log x}}$.

Write $up + v = a_1 a_2 r$, where $r = P^+(up+v)$, $P^+(a_1) \le w$, $P^-(a_2) > w$. As before, we may assume that $r > y$, where we recall that $y = x^{1/\log\log x}$. We fix a_1, a_2 and count primes $r \le (ux+v)/a_1 a_2$ with $a_1 a_2 r$ of the form $up + v$ with p prime. By Brun's method, it is

$$O\left(\frac{x(\log\log x)^2}{\varphi(a_1 a_2)(\log x)^2}\right) = O\left(\frac{x(\log\log x)^3}{a_1 a_2 (\log x)^2}\right).$$

We then sum over values of a_2 with $P^-(a_2) > w$, which by Brun's method introduces a factor $(\log x)/\log w$, and then sum over a_1 with $P^+(a_1) \le w$ and $\Omega(a_1) > \frac{4}{3}\log\log w - 1$, getting a factor $(\log w)^{1-c_8}$.

Thus, the final count for a given value of w is $O(\pi(x)/(\log w)^{c_6})$, where $c_6 < \min\{c_7, c_8\}$. To complete the proof let w run over numbers z^{2^j} where $j = 0, 1, 2, \ldots$ and sum the resulting estimates to obtain $O(\pi(x)/(\log z)^{c_6})$. □

4 Some results from transcendence theory

We start with Siegel's zero-lemma (see Chapter 1 in [16]).

Lemma 5. *Assume we are given a system of M linear equations in $N > M$ unknowns*

$$\sum_{j=1}^{N} a_{ij} x_j = 0, \qquad i = 1, \ldots, M,$$

whose coefficients are integers not all zero and are bounded by A in absolute value. The system has an integer non-zero solution $\mathbf{x} = (x_1, \ldots, x_N)$ such that

$$\max\{|x_i| : 1 \le i \le N\} \le (NA)^{M/(N-M)}.$$

For a nonzero rational number α written in reduced form as $\alpha = a/b$ with integers $a,\ b \ge 1$ and $\gcd(a, b) = 1$, its *height* is $h(\alpha) = \max\{\log |a|, \log b\}$. The following result is a particular case of a theorem of Matveev [13].

Theorem 4. *Let $\alpha_1, \ldots, \alpha_t$ be positive rational numbers which are not 1 and let $b_1, \ldots, b_t$ be integers. Suppose that*

$$B \ge \max\{|b_1|, \ldots, |b_t|\},$$

and put

$$\Lambda := \sum_{i=1}^{t} b_i \log \alpha_i.$$

Then, assuming that $\Lambda \ne 0$, we have

$$2|\Lambda| > \exp\left(-\tfrac{7}{5} t^{9/2} 30^{t+3} (1 + \log B) h(\alpha_1) \cdots h(\alpha_t)\right).$$

The following conjecture of Lang [11] and Waldschmidt [21] strengthens the conclusion of Theorem 4.

Conjecture 1. *For every $\varepsilon > 0$, there exists $C(\varepsilon) > 0$ such that in the hypothesis and notations of Theorem 4, we have*

$$|\Lambda| \ge \frac{C(\varepsilon)^t B}{(|b_1| \cdots |b_t| h(\alpha_1) \cdots h(\alpha_t))^{1+\varepsilon}}. \tag{9}$$

5 The proof of Theorem 1

We follow the proof of Theorem 3 in [12] but introduce some new elements. Recall that

$$B_{2n} = (-1)^{n+1} \frac{2(2n)!}{(2\pi)^{2n}} \zeta(2n). \tag{10}$$

Then

$$\frac{|B_{2n}|}{2n} = \frac{2(2n-1)!}{(2\pi)^{2n}} \left(1 + O\left(\frac{1}{2^{2n}}\right)\right).$$

Write as in Section 2,

$$\frac{|B_{2n}|}{2n} = \frac{C_{2n}}{D_{2n}} \qquad \text{where} \qquad \gcd(C_{2n}, D_{2n}) = 1.$$

Then by Lemma 1, every prime factor of C_{2n} is an irregular prime. Furthermore,

$$D_{2n} = \prod_{p-1|2n} p^{1+\mathrm{ord}_p(2n)}, \tag{11}$$

where for a positive integer m and a prime p we write $\mathrm{ord}_p(m)$ for the exponent with which p appears in the prime factorization of m. Indeed, the above formula (11) follows easily from (3), (4). Let N be large and put

$$K(N) := \omega\left(\prod_{n \le N} C_{2n}\right). \tag{12}$$

Observe that $K(N)$ is a nondecreasing function of N. Our goal is to find a lower bound for $K(N)$ in terms of N. Indeed, assume that we have such an estimate. Then, by (11) and $\tau(n) = n^{o(1)}$ as $n \to \infty$, we have that

$$D_{2n} \le (2n) \prod_{d|2n} (d+1) < (2n+1)^{1+\tau(2n)} \le n^{n^{o(1)}}$$

as $n \to \infty$. Hence, by (10) and $\zeta(2n) < 2$, we get that

$$C_{2n} \le 2\frac{2D_{2n}}{(2\pi)^{2n}}(2n-1)! < N^{2N} \tag{13}$$

for all $n \le N$ once $N > N_0$ is sufficiently large. We now take a large x and put $N(x)$ for the largest positive integer N such that $N^{2N} \le x$. We obtain that if $p \mid C_{2n}$ for some $n \le N(x)$, then certainly $p \in \mathcal{I}_B(x)$. Since certainly

$$N(x) \ge 0.5\frac{\log x}{\log\log x}, \tag{14}$$

we get that

$$\#\mathcal{I}_B(x) \geq K(N(x)). \tag{15}$$

So, any lower bound on $K(N)$ in terms of N, will lead, via inequalities (14) and (15), to a lower bound for $\#\mathcal{I}_B(x)$

To proceed, we write

$$C_{2n} = 2D_{2n}(2n-1)!(2\pi)^{-2n}\left(1 + O\left(\frac{1}{2^{2n}}\right)\right). \tag{16}$$

Taking logarithms, we get

$$\log C_{2n} = \log(2D_{2n}) + \log(2n-1)! - 2n\log(2\pi) + O\left(\frac{1}{2^{2n}}\right).$$

We evaluate the above formula in the numbers $n,\ n+1,\ n+2$ for some $n \leq N-2$, and take the second difference of the resulting relations, getting

$$\log\left(\frac{C_{2n}C_{2(n+2)}}{C_{2(n+1)}^2}\right) - \log\left(\frac{D_{2n}D_{2(n+2)}(2n+2)(2n+3)}{D_{2(n+1)}^2(2n)(2n+1)}\right) = O\left(\frac{1}{2^{2n}}\right). \tag{17}$$

We put

$$F_n = \frac{D_{2n}D_{2(n+2)}(2n+2)(2n+3)}{D_{2(n+1)}^2(2n)(2n+1)},$$

and we evaluate relation (17) at some special $n \leq N-2$. We now describe our values for n. Take $\mathcal{J} = (.99N/6, N/6)$. By the Prime Number Theorem, the number of primes in $\mathcal{J}$ exceeds $1.5 \times 10^{-3}\pi(N)$ when N is large. Let $p \in \mathcal{J}$ and consider the forms $2(6p-8),\ 2(6p-7),\ 2(6p-6)$. Applying Theorem 3 with $(u,v) = (12,-16),\ (12,-14),\ (12,-12)$, we deduce that the subset of primes $p \in \mathcal{J}$ such that one of the above three linear forms has a divisor of the form $q-1$ for some $q \neq p$, $q > z$ has cardinality $< \alpha_1\pi(N)/(\log z)^c$ for some absolute positive constant α_1. Choosing

$$z = \alpha_2 := \exp\left((2000\alpha_1)^{1/c}\right), \tag{18}$$

we conclude that there are $> 10^{-3}\pi(N)$ primes $p \in \mathcal{J}$ such that none of the above linear forms in p has a divisor of the form $q-1$ for $q \neq p$ and $q > \alpha_2$. We may assume that $K := K(N)$ satisfies the inequality

$$2K + 1 < 10^{-3}\pi(N), \tag{19}$$

since otherwise the theorem clearly holds. It follows that we can choose $K+1$ such primes in $\mathcal{J}$, call them $p_1, \ldots, p_{K+1}$, such that additionally none

of them divides $\prod_{1\le n\le N} C_{2n}$. We now take $n_i = 6p_i - 8$ for $i = 1, \dots, K+1$. Fix i and let us first take a closer look at the number F_{n_i}, which is given by:

$$F_{n_i} = \frac{D_{2(6p_i-8)} D_{12(p_i-1)}}{D_{2(6p_i-7)}} \times \frac{(6p_i-7)(12p_i-13)}{6(3p_i-4)(4p_i-5)}.$$

It is clear that for large N, $p_i \| D_{12(p_i-1)}$, and p_i divides neither $D_{2(6p_i-8)}$ nor $D_{2(6p_i-7)}$. Furthermore, p_i does not divide the number

$$(3p_i-4)(4p_i-5)(6p_i-7)(12p_i-13)$$

for large N either. Hence, p_i divides the numerator of F_{n_i} written in reduced form. Let us show that if $i \neq j$ both in $\{1, \dots, K+1\}$, then p_j divides neither the numerator nor the denominator of F_{n_i}. Indeed, assume this is not so. Then p_j either divides one of $D_{2(6p_i-8)},\ D_{2(6p_i-7)},\ D_{12(p_i-1)}$, or one of $3p_i-4,\ 4p_i-5,\ 6p_i-7,\ 12p_i-13$. Say, p_j divides one of the numbers from the first group. Then for large N, $p_j > .99N/6 > \alpha_2$, and $p_j \neq p_i$ has the property that $p_j - 1$ divides one of $2(6p_i-8),\ 2(6p_i-7),\ 12(p_i-1)$, which contradicts the way we choose the prime p_i. Now suppose that p_j divides some number from the second group. Then

$$ap_i - (a+1) = bp_j \qquad \text{for some} \qquad a \in \{3, 4, 6, 12\}. \tag{20}$$

Since $p_j > 0.99p_i$, we get that

$$ap_i > bp_j > (0.99)bp_i \qquad \text{therefore} \qquad (0.99)^{-1}a > b,$$

which implies that $b < a+1$ for all $a \in \{3, 4, 6, 12\}$. Thus, $b \le a \le 12$. Since also

$$a(0.99p_j) < ap_i = bp_j + (a+1) \le bp_j + 13,$$

we get that for large N and $b \le 12$ we must have $a < b+1$. Thus, $a = b$, and now equation (20) shows that $a \mid a+1$, which is false for all $a \in \{3, 4, 6, 12\}$. So indeed, neither the numerator nor the denominator of F_{n_i} is divisible by p_j for any $j \neq i$ in $\{1, \dots, K+1\}$.

We now estimate the size of $h(F_{n_i})$. Since neither one of the numbers $2(6p_i-8),\ 2(6p_i-7),\ 12(p_i-1)$ has any divisor of the form $q-1$ for $q \neq p_i$, $q > \alpha_2$ and q prime, it follows, by (11), that

$$\max\{D_{2n_i},\ D_{2(n_i+1)},\ D_{2(n_i+2)}\} = O(N).$$

Hence,

$$h(F_i) \le \max\{\log(D_{2n_i} D_{2(n_i+2)}(2n_i+3)^2), \log(D_{n_i+1}^2(2n_i+1)^2)\} < 5\log N \tag{21}$$

once N is large enough.

Now let us assume that $\mathcal{Q} = \{q_1, \dots, q_K\}$ is the set of all the prime factors of $\prod_{1 \le n \le N} C_{2n}$. Write

$$\frac{C_{2n_i} C_{2(n_i+2)}}{C_{2(n_i+1)}^2} = \prod_{j=1}^{K} q_j^{a_{i,j}} \qquad (i = 1, \dots, K+1).$$

Then the relation (17) for n_i is

$$\left| \sum_{j=1}^{K} a_{i,j} \log q_j - \log F_{n_i} \right| = O\left(\frac{1}{2^{1.98N}} \right). \tag{22}$$

By (13) and $n \le N-2$, it follows that

$$C_{2n}, C_{2(n+1)}, C_{2(n+2)} < N^{2N}$$

for all sufficiently large N, so that

$$A = \max_{\substack{1 \le i \le K+1 \\ 1 \le j \le K}} |a_{i,j}| < \frac{\log(N^{4N})}{\log 2} < 6N \log N \tag{23}$$

for $N > N_0$. Let $(\Delta_1, \dots, \Delta_{K+1})$ be a nonzero vector of integers in the null-space of the $K \times (K+1)$ matrix

$$\mathcal{A} = \begin{pmatrix} a_{1,1} & a_{2,1} & \cdots & a_{K+1,1} \\ a_{1,2} & a_{2,2} & \cdots & a_{K+1,2} \\ \vdots & \vdots & \cdots & \vdots \\ a_{1,K} & a_{2,K} & \cdots & a_{K+1,K} \end{pmatrix}.$$

Siegel's lemma 5 (with $M = K,\ N = K+1$), tells us that such a vector exists with

$$\max\{|\Delta_i| : i = 1, \dots, K+1\} \le ((K+1)A)^K < (\tfrac{6}{1000} N^2)^K \tag{24}$$

for $N > N_0$ (see (19) and (23)). Then taking the linear combination of relations (22) with coefficients Δ_i for $i = 1, \dots, 2K$, we get

$$\left| \sum_{i=1}^{K+1} \Delta_i \log F_{n_i} \right| = O\left(\frac{(K+1) \max\{|\Delta_i\}|}{2^{1.98N}} \right) = O\left(\frac{1}{2^N} \right). \tag{25}$$

The linear form on the left–hand side of (25) above is nonzero (because each F_{n_i} has its numerator divisible by p_i and neither the numerator nor the denominator of F_{n_i} is divisible by p_j for any $j \neq i$ in $\{1, \ldots, K+1\}$). We apply Theorem 4 to get that the expression in the left-hand side of (25) can be bounded below by

$$\frac{1}{2} \exp\left(-\tfrac{7}{5}(K+1)^{9/2} 30^{K+4}(1+\log B) h(F_{n_1}) \cdots h(F_{n_{K+1}})\right),$$

where if we take $B := N^{2K}$, then

$$B \geq \max\{|\Delta_i| : i = 1, \ldots, K+1\},$$

(see (24)). Since for large N, $h(F_{n_i}) < 5 \log N$ for all $i = 1, \ldots, K+1$ (see (21)), inequality (25) gives

$$N \log 2 - \alpha_3 < \tfrac{7}{5}(K+1)^{9/2} 30^{K+4}(1+2K \log N)(5 \log N)^{K+1},$$

with some suitable constant α_3, which implies

$$K(N) \geq (1+o(1)) \frac{\log N}{\log \log N} \qquad (N \to \infty). \tag{26}$$

Theorem 1 follows from estimates (26), (15) and (14).

6 The proof of Theorem 2

This proof is very similar to the proof of Theorem 1. Instead of estimate (10), we use the estimate

$$|E_{2n}| = \frac{4^{2n+1}(2n)!}{\pi^{2n+1}} \left(1 + O\left(\frac{1}{3^{2n}}\right)\right) \tag{27}$$

(see the last display on page 331 of [3]). Write

$$|E_{2n}| = U_{2n} V_{2n}, \qquad \text{where} \qquad V_{2n} = \prod_{p-1 \mid 2n} p^{\mathrm{ord}_p(E_{2n})}.$$

Note that an argument of Carlitz from [3] shows that every prime factor of U_{2n} is an irregular prime for the Euler numbers and that V_{2n} is divisible by every prime $p \equiv 1 \pmod 4$ with $p - 1 \mid 2n$. We put

$$K(N) = \# \left\{ p : p \,\Big|\, \prod_{n \leq N} U_{2n},\ p > \alpha_2 \right\},$$

where α_2 is the constant shown at (18) from the proof of Theorem 1. That is, our $K(N)$ has almost the same definition as in the proof of Theorem 1, except that we only count the distinct prime factors of $\prod_{n\leq N} U_{2n}$ which exceed c_2. Observe that as in the case of the Bernoulli numbers, we have

$$U_{2n} < N^{2N}$$

for all $n \leq N$ and $N > N_0$.

Now the argument proceeds in the same way as in the proof of Theorem 1. Assume that inequality (19) is satisfied for $K = K(N)$. We evaluate relation (27) in $n,\ n+1,\ n+2$, where $n = 6p_i - 8$ and $i = 1, \ldots, K+1$ is one of the primes from the proof of Theorem 1. Except, we now choose the primes p_i from the interval $(.98N/6, N/6)$ and we insist that the primes $p_i \equiv 1 \pmod 4$. We obtain an analog of relation (17) which is

$$\log\left(\frac{U_n U_{n+2}}{U_{n+1}^2}\right) + \log\left(\frac{V_{2n}V_{2(n+2)}(2n+3)(2n+4)}{V_{2(n+1)}^2(2n+1)(2n+2)}\right) = O\left(\frac{1}{3^{1.96N}}\right). \quad (28)$$

From the way we have chosen our primes p_i for $i = 1, \ldots, K+1$, all prime divisors of the rational number

$$\frac{V_{2n_i}V_{2(n_i+2)}}{V_{2(n_i+1)}^2}$$

except for p_i are bounded by the constant α_2, and therefore the rational number

$$\begin{aligned} F_{n_i} &= \frac{V_{2n_i}V_{2(n_i+2)}(2n_i+3)(2n_i+4)}{V_{2(n_i+1)}^2(2n_i+1)(2n_i+2)} \\ &= \frac{2V_{2(6p_i-8)}V_{12(p_i-1)}(6p_i-13)(p_i-1)}{V_{2(6p_i-7)}^2(4p_i-5)(6p_i-7)} \end{aligned}$$

has the prime p_i appearing in its numerator. Further, p_i does not appear in the factorization of anyone of the other rational numbers F_{n_j} corresponding for some $j \neq i$ in $\{1, \ldots, K\}$. Indeed, the only additional fact that we need to check is that $p_j \nmid p_i - 1$, which follows for large N because both p_i and p_j are in $\mathcal{J}$. So, we write again

$$\frac{|E_{2(p_i-2)}E_{2p_i}|}{E_{2(p_i-1)}^2} = \prod_{r\leq\alpha_2} r^{a_{i,r}} \prod_{j=1}^{K} q_j^{a_{i,j}},$$

From here on, the argument continues in exactly the same way and gives the lower bound (26) on $K + \pi(\alpha_2)$ in terms of N, which leads to the desired conclusion.

7 Getting better bounds conditionally

With the notation from the proof of Theorem 1, we put

$$K(N) = \omega\left(\prod_{n\le N} C_{2n}\right).$$

Here we make the following observation.

Theorem 5. *Assume that there exists $\varepsilon_0 > 0$ such that Conjecture 1 holds with $\varepsilon := \varepsilon_0$. Then*

$$K(N) \gg \pi(N). \tag{29}$$

Proof. We follow the proof of Theorem 1 except that instead of (19) we assume that the stronger inequality

$$3K < 10^{-3}\pi(N) \tag{30}$$

holds for $N > N_0$. Note that if this inequality fails, then we are done anyway. So, instead of only $K+1$ primes $p_1, \dots, p_{K+1}$, we can now work with $2K$ primes $p_1, \dots, p_{2K}$. The argument is identical up to choosing the non-zero vector $(\Delta_1, \dots, \Delta_{2K})$, which by Siegel's lemma 5 (with twice as many variables as equations) tells us that we can chose such a vector with

$$\max\{|\Delta_i| : 1 \le i \le 2K\} \le 2KA < N^2, \tag{31}$$

see (24). The analog of (25) is now

$$\left|\sum_{i=1}^{2K} \Delta_i \log F_{n_i}\right| = O\left(\frac{K \max\{|\Delta_i| : 1 \le i \le 2K\}}{2^{1.98N}}\right) = O\left(\frac{1}{2^N}\right), \tag{32}$$

and the left–hand side above is not zero. We now apply the conjectural inequality (9) on the left–hand side above with the obvious choices $t = 2K$, $b_i = \Delta_i$, $\alpha_i = F_{n_i}$ for $i = 1, \dots, 2K$, and $\varepsilon = \varepsilon_0$, getting via (21) and (31), that

$$\left|\sum_{i=1}^{2K} \Delta_i \log F_{n_i}\right| > \exp\left(2K\log(C(\varepsilon_0)) - (1+\varepsilon_0)2K(\log(N^2) + \log(5\log N))\right). \tag{33}$$

Comparing (32) with (33), we get that

$$K(N) = K \ge (\alpha_4 + o(1))\frac{N}{\log N}$$

where we can take $\alpha_4 = (\log 2)/(4(1+\varepsilon_0))$, which is what we wanted. □

Via (14) and (29), we get that the lower bound

$$\#\mathcal{I}_B(x) \gg \frac{\log x}{(\log\log x)^2} \tag{34}$$

holds assuming that Conjecture 1 holds for some ε_0.

A similar result as Theorem 5 (with an almost identical proof) holds if $K(N)$ is

$$K(N) = \omega\left(\prod_{n\leq N} U_{2n}\right),$$

with the notation from the proof of Theorem 2. This in turn implies that a conditional lower bound as (34) applies to the counting function of irregular primes with respect to the Euler numbers. However, this method seems to be far away from proving that the irregular primes occupy a positive proportion of all the primes.

Acknowledgements. We thank the referee for comments which improved the quality of the paper. F. L. and A. P. thank Professor Eduardo Friedman for useful advice. F. L. worked on this project while he visited the Mathematics Department of the Universidad de Valparaiso, Chile in June and July of 2013 and during a visit to Dartmouth College in Spring of 2014. He thanks the members of these departments for their hospitality. A. P. was supported in part by project Fondecyt N^{0}11100260.

References

[1] N. G. de Bruijn, "On the number of integers $\leq x$ and free of prime factors $> y$", *Nederl. Akad. Wetensch. Proc. Ser. A* **54** (1951), 50–60.

[2] Y. Bugeaud, M. Mignotte and S. Siksek, "Classical and modular approaches to exponential Diophantine equations I. Fibonacci and Lucas perfect powers", *Ann. of Math.* **163** (2006), 969–1018.

[3] L. Carlitz, "Note on irregular primes", *Proc. Amer. Math. Soc.* **82** (1954), 329–331.

[4] P. Erdős, "On the normal number of prime factors of $p-1$ and some related problems concerning Euler's φ-function", *Quarterly J. Math. (Oxford Ser.)* **6** (1935), 205–213.

[5] P. Erdős and S. S. Wagstaff Jr., "The fractional parts of Bernoulli numbers", *Illinois J. Math.* **24** (1980), 104–112.

[6] M. Gut, "Eulersche Zahlen und grosser Fermat'scher Satz", *Comment. Math. Helv.* **24** (1950), 73–99.

[7] H. Halberstam and H.-E. Richert, *Sieve Methods*, Academic Press, London, 1974. Reprinted by Dover Publications, 2011.

[8] K. L. Jensen, (1915). "Om talteoretiske Egenskaber ved de Bernoulliske Tal", *Nyt Tidsskr. Mat. B* (1915) **26**, 73–83.

[9] E. E. Kummer, "Allgemeiner Beweis des Fermatschen Satzes, dass die Gleichung $x^\lambda+y^\lambda = z^\lambda$ durch ganze Zahlen unlösbar ist, für alle diejenigen Potenz-Exponenten λ, welche ungerade Primzahlen sind und in den Zählern der ersten $(\lambda-3)/2$ Bernoullischen Zahlen als Factoren nicht vorkommen", *J. reine angew. Math.* **40** (1850), 131–138.

[10] P. Kurlberg, J. C. Lagarias, and C. Pomerance, "Product-free sets with high density", *Acta Arith.* **155** (2012), 163–173.

[11] S. Lang, *Eliptic curves: Diophantine analysis*, Springer-Verlag, Berlin–New York, 1978.

[12] F. Luca and A. Pizarro, "Some remarks on the Riemann zeta function and Bernoulli numbers", *Bull. Australian Math. Soc.* **86** (2012), 216–223.

[13] E. M. Matveev, "An explicit lower bound for a homogeneous rational linear form in logarithms of algebraic numbers. II", *Izv. Ross. Akad. Nauk Ser. Mat.* **64** (2000), 125–180; English transl. in *Izv. Math.* **64** (2000), 1217–1269.

[14] T. Metsänkylä, "Note on the distribution of irregular primes", *Ann. Acad. Sci. Fenn. Ser. A I*, **492**, (1971), 7pp.

[15] H. L. Montgomery, "Distribution of irregular primes", *Illinois J. Math.* **9** (1965), 553–558.

[16] W. M. Schmidt, *Diophantine approximations and Diophantine equations*, Lecture Notes in Mathematics, Springer Verlag, 2000.

[17] C. L. Siegel, "Zu zwei Bemerkungen Kummers", *Nachr. Akad. Wiss. Göttingen, Math. Phys. Kl. II* **1964** (1964), 51–62.

[18] N. J. A. Sloane, *The On-line Encyclopedia of Integer Sequences*, https://oeis.org

[19] N. M. Timofeev, "The Hardy–Ramanujan and Halasz inequalities for shifted primes", *Mathematical Notes* **57** (1995), 522–535.

[20] H. S. Vandiver, "Note on Euler number criteria for the first case of Fermat's last theorem", *Amer. J. Math.* **62** (1940) 79–82.

[21] M. Waldschmidt, "Open Diophantine problems", *Mosc. Math. J.* **4** (2004), 245–305, 312.