ON COMPUTING THE DENSITY OF INTEGERS OF THE FORM $2^n + p$

GIANNA M. DEL CORSO, ILARIA DEL CORSO, ROBERTO DVORNICICH, AND FRANCESCO ROMANI

Abstract. The problem of finding the density of odd integers which can be expressed as the sum of a prime and a power of two is a classical one. In this paper we tackle the problem both with a direct approach and with a theoretical approach, suggested by Bombieri. These approaches were already introduced by Romani in [16], but here the methods are extended and enriched with statistical and numerical methodologies. Moreover, we give a proof, under standard heuristic hypotheses, of the formulas claimed by Bombieri, on which the theoretical approach is based. The different techniques produce estimates of the densities which coincide up to the first three digits.

1. Introduction

In 1849 de Polignac [6] conjectured that any odd integer can be expressed as 2^n+p for suitable n and prime p. Actually Euler, long before, had found explicit counterexamples, and this was recognised later by de Polignac himself. In the following years, the naive de Polignac conjecture turned into the difficult question of deciding which portion of integers has this kind of representation. The first important contribution to this problem was given in 1934 by Romanov [18], who proved that there is a positive proportion of odd numbers that can be indeed represented in such a way, but the question of the existence of infinitely many counterexamples remained unsolved until 1950, when van der Corput [7] and Erdős [8], independently, proved that there is also a positive proportion of odd numbers that are not of the form $2^n + p$. Recently other authors examined related problems such as the extension of Romanov's theorem in number fields [13], or the proof that the sumset of integers of the form $2^n + q$, where q is a prime or the product of two primes, has a positive density [12]. In [14] it is proved that the sumset $\{p^2 + b^2 + 2^n : p \text{ is prime and } b, n \in \mathbb{N}\}$ has a positive lower density. Very interesting is the paper by Pintz [15] where the lower asymptotic density is estimated and connections with classical problems in number theory such as the Goldbach-Linnik problem are shown.

Let $\mathbb{N}^+$ and $\mathcal{P}$ denote the set of positive integers and the set of odd primes, let

$$(1.1)\qquad \begin{aligned} a(n) &= \#\{(p, v) \mid n = p + 2^v; v \in \mathbb{N}^+, p \in \mathcal{P}\},\\ A(x) &= \#\{n \leq x \mid a(n) > 0\}. \end{aligned}$$

The result of Romanov essentially states that there is a constant $d_1 > 0$ such that

$$d_1 = \liminf_{x\to\infty} \frac{A(x)}{x}$$

2010 *Mathematics Subject Classification.* Primary 11P32; Secondary 11-04, 65C60.
Key words and phrases. Sums of primes and powers; odd integers.

even if Romanov did not provide any explicit value for d_1. The results of van der Corput [7] and Erdös [8] gave an upper bound on the distribution, mainly that there exists a positive constant $d_2 < 1/2$ such that

$$d_2 = \limsup_{x\to\infty} \frac{A(x)}{x}.$$

In [5] it is mentioned that the argument of Erdős can be a little generalized to reduce the bound on d_2 to $0.5 - 9 \cdot 10^{-8}$. The bound on d_1 has been refined in the recent years [5, 15, 9] but the current value of $d_1 \geq 0.09368$ is still very distant from the theoretical upper bound on $d_2 \leq 0.49095$ given in [9] as well.

In [16] the problem has been tackled in a computational way, assuming that $d_1 = d_2 = \delta$ and proposing two different computational approaches to estimate the density

$$\delta = \lim_{x\to\infty} \frac{A(x)}{x}. \tag{1.2}$$

In this paper we treat this problem both theoretically and computationally making a major progress to the initial work in [16]. In particular, we first give a proof, under very natural assumptions, of some formulas in [17] proposed by Bombieri about 40 years ago. Assuming the existence of the density δ, Bombieri described the way to construct a sequence, which he conjectured to converge to δ, and claimed a formula to compute it. More precisely, for each $k \geq 1$, he considered a sequence of squarefree numbers $\{Q_k\}_{k\in\mathbb{N}}$ with $Q_k \mid Q_{k+1}$ and for each of them a set $\mathcal{T}_{Q_k}$ of integers coprime to Q_k and with the same asymptotic distribution of the primes. Bombieri conjectured that the set of integers representable as $2^n + t$, where t belongs to $\mathcal{T}_{Q_k}$, has a density δ_{Q_k}, which depends only on Q_k and claimed an explicit formula for δ_{Q_k}. In Section 2 we illustrate this approach and, using the circle method, we prove Bombieri's formulas for δ_{Q_k} when $\mathcal{T}_{Q_k}$ is a *generic* sequence with the above property, which, by probabilistic and heuristic arguments, can be assumed to be "good" for our purposes. Finally, again by Bombieri's conjectures, the sequence of the densities δ_{Q_k} is decreasing and converges to δ, so by computing the δ_{Q_k}'s we obtain approximations to δ.

The computational goal is to estimate the density (upper density / lower density) both via Bombieri's formulas and with a direct approach extrapolating the data obtained directly generating an odd number as a sum of a prime and a power of two.

In Section 3, following the theoretical argument introduced in Section 2, we analyze how to approximate the sequence δ_{Q_k}, for Q_k the product of the first k-primes. The formulas proved in Section 2 express δ_{Q_k} as the average over Q_k terms, a number of terms that increases tremendously with k, so that the exact computation of δ_{Q_k} is impractical for $k > 12$. We approximate these quantities for larger k using statistical techniques to compute the sample mean and to validate the estimate of the error. The estimate of the density δ is then obtained extrapolating the data (both exact and approximated).

The direct approach is presented in Section 4, where, after generating large odd numbers as a prime plus a power of 2, we use extrapolating techniques to estimate the lower and upper densities. In particular, we were able to measure $A(x)/x$ directly for x as large as 2^{48}, and up to 2^{61} on selected intervals. Note that a similar approach used in [16] allowed to estimate the density only for $x < 2^{31}$.

With these methods we produce different estimates of d_1, d_2 and δ that coincide for the first three digits. The fact that we produced similar results using different techniques is certainly interesting even if we are not able to give a conclusive answer about the coincidence of upper and lower densities.

2. Theoretical approach

In this section we present the approach to the computation of the density δ proposed by Bombieri (see [17] and [16]) whose main idea is to define a decreasing sequence $\{\delta_Q\}$ of values that conjecturally converges to δ. Bombieri claimed explicit formulas to compute δ_Q. Here we give a proof of Bombieri's formulas **(B1)** and **(B2)** under a very natural heuristic assumption.

We will assume throughout the existence of the limit in (1.2), namely $d_1 = d_2 = \delta$. Recall that the function $a(n)$ introduced in (1.1) counts the number of ways one can represent a number n as the sum of a prime and a power of 2. Then $A(x) = \sum_{\ell=1}^{\infty} \#\{n \le x \,|\, a(n) = \ell\}$ is the number of representable integers up to x. We set

$$\delta_\infty(\ell) = \lim_{x\to\infty} \frac{1}{x}\#\{n \le x \,|\, a(n) = \ell\},$$

and

$$\delta_\infty = \sum_{\ell=1}^{\infty} \delta_\infty(\ell). \tag{2.1}$$

Bombieri suggested to proceed as follows to estimate δ.

Let $\mathcal{T}$ denote an infinite sequence of natural numbers. Now define

$$\begin{aligned}
a_{\mathcal{T}}(n) &= \#\{(t,v) \,|\, n = t + 2^v; v \in \mathbb{N}^+, t \in \mathcal{T}\},\\
\delta_{\mathcal{T}}(\ell) &= \lim_{x\to\infty} \frac{1}{x}\#\{n \le x \,|\, a_{\mathcal{T}}(n) = \ell\},\\
\delta_{\mathcal{T}} &= \sum_{\ell=1}^{\infty} \delta_{\mathcal{T}}(\ell).
\end{aligned}$$

Clearly, the quantities $a(n)$, $A(x)$ and $\delta_\infty(\ell)$ previously defined coincide with $a_{\mathcal{P}}(n)$, $A_{\mathcal{P}}(x)$ and $\delta_{\mathcal{P}}(\ell)$, where $\mathcal{P}$ is the set of odd prime numbers.

Using the definition of $a_{\mathcal{T}}(n)$ and rearranging the summation we get, for each integer s,

$$\sum_{n\le N} a_{\mathcal{T}}(n)^s \sim N \sum_{\ell=1}^{\infty} \ell^s \delta_{\mathcal{T}}(\ell). \tag{2.2}$$

The theory formulated so far holds for any integer sequence $\mathcal{T}$, in particular for the sequence of the prime numbers or for any sequence $\mathcal{T}_Q$, where Q is a squarefree integer, with the following properties

- **(i):** for any $t \in \mathcal{T}_Q$ then $\gcd(t, Q) = 1$,
- **(ii):** $\sum_{t\in\mathcal{T}_Q, t<N} \log t \sim N$.

Observe that the second condition requires that the sequence $\mathcal{T}_Q$ has the same asymptotic distribution as the primes. Since the set $\mathcal{T}_Q$ is a sieved set, in some sense, and for suitable choices, it behaves like the set of primes bigger than the prime factors of Q.

Bombieri (see [17]) proposes to compute the $\delta_{\mathcal{T}_Q}$'s to approximate δ. More precisely, he claims the formulas **(B1)** and **(B2)** which show that $\delta_{\mathcal{T}_Q}$ does not depend on the particular choice of the sequence $\mathcal{T}_Q$ but only on the squarefree number Q, so we will denote it simply by δ_Q. With this notation, he conjectured that for any increasing sequence $\{Q_k\}_{k\geq 1}$ with $Q_k \mid Q_{k+1}$

$$\lim_{k\to\infty} \delta_{Q_k}(\ell) = \delta_\infty(\ell)$$

and

$$\delta_{Q_{k+1}} \leq \delta_{Q_k}$$

from which clearly

$$\delta_\infty = \lim_{k\to\infty} \delta_{Q_k}.$$

Bombieri also conjectured that $\delta = \delta_\infty$. Following this idea, evaluating δ_Q leads to an approximation of δ.

2.1. **Bombieri's formulas.** Let $Q > 0$ be a squarefree integer and let $\mathcal{T}_Q$ be a sequence satisfying **(i)** and **(ii)**. Let

$$\gamma_Q(m) = \frac{Q}{\phi^2(Q)} \sum_{\substack{u=1 \\ \gcd(m+2^u,Q)=1}}^{\phi(Q)} 1$$

where ϕ is Euler's totient function. Then

$$\textbf{(B1)} \qquad \sum_{n\leq N} a_{\mathcal{T}_Q}(n)^s \sim \frac{N}{Q}\sum_{m=1}^{Q}\sum_{r=1}^{\infty} \sum_{\substack{i_1+i_2+\cdots+i_r=s \\ i_j>0}} \frac{s!}{i_1!i_2!\dots i_r!}\frac{1}{r!}\left(\frac{\gamma_Q(m)}{\log 2}\right)^r$$

and

$$\textbf{(B2)} \qquad \delta_Q = \sum_{l=1}^{\infty}\delta_Q(l) = \frac{1}{Q}\sum_{m=1}^{Q}\left(1-\exp(-\frac{\gamma_Q(m)}{\log 2}\right).$$

To prove his formulas, Bombieri [3] suggests to use the circle method. In the following, we prove **(B1)** and **(B2)** by a heuristic argument based on some widely accepted assumptions connected with the good distribution of a generic sequence $\mathcal{T}_Q$ satisfying **(i)** and **(ii)** into the arithmetic progressions *admissible* under **(i)** and similar phenomena. When using the circle method, these assumptions allow us to disregard the contribution of minor arcs and to limit ourselves to the main terms which occur in the estimate of the major arcs. A similar approach has been used many times (see for instance [1], [2], [19] and [20]).

2.2. **Auxiliary lemmas.** The following immediate arithmetic lemma describes the arithmetic progressions which are admissible for $\mathcal{T}_Q$ (i.e., those which may contain elements of $\mathcal{T}_Q$).

Lemma 2.1. *Let Q be a squarefree positive integer. For any positive integer q let $d = d(q)$ be the largest factor of q coprime to Q and set $q = q_0 d$. Then the classes modulo q admissible for T_Q are those coprime to q_0.*

We now state and prove some simple lemmas that we will need in the course of the proof.

Lemma 2.2. *For any real z with $z > 0$*

$$\sum_{s=1}^{\infty} \frac{\log^s(z)}{s!} \sum_{n\le N} a_{\mathcal{T}}(n)^s \sim N \sum_{\ell=1}^{\infty} (z^\ell - 1)\delta_{\mathcal{T}}(\ell).$$

Proof. Multiplying both sides of (2.2) by $\log^s(z)/s!$ and summing over s we have

$$\sum_{s=1}^{\infty} \frac{\log^s(z)}{s!} \sum_{n\le N} a_{\mathcal{T}}(n) \sim N \sum_{\ell=1}^{\infty} \delta_{\mathcal{T}}(\ell) \sum_{s=1}^{\infty} \ell^s \frac{\log^s(z)}{s!};$$

we get the thesis observing that, from the exponential generating function, we have

$$\sum_{s=1}^{\infty} \frac{\log^s(z)}{s!} \ell^s = z^\ell - 1.$$

□

Lemma 2.3. *For any real z, with $z > 0$, and for each integer $r \ge 1$ we have*

$$(z-1)^r = \sum_{s=1}^{\infty} \frac{\log^s(z)}{s!} \sum_{\substack{i_1+i_2+\cdots+i_r=s \\ i_j>0}} \frac{s!}{i_1! i_2! \dots i_r!}.$$

Proof. Using the binomial theorem and the Taylor expansion of $z^h = e^{h\log(z)}$, we get

$$(z-1)^r = \sum_{h=0}^{r} \binom{r}{h} z^h (-1)^{r-h} = \sum_{s\ge 0} \frac{\log^s(z)}{s!} \sum_{h=0}^{r} \binom{r}{h} h^s (-1)^{r-h}.$$

We note that $\sum_{h=0}^{r} \binom{r}{h}(-1)^{r-h} = 0$ for each $r \ge 1$, hence the term corresponding to $s = 0$ vanishes. Moreover, by the inclusion-exclusion principle we have the equality

$$\sum_{h=0}^{r} \binom{r}{h} h^s (-1)^{r-h} = \sum_{\substack{i_1+i_2+\cdots+i_r=s \\ i_j>0}} \frac{s!}{i_1! i_2! \dots i_r!}$$

and the lemma follows. □

2.3. **Proof of (B1) and (B2).** In this proof we decided to sketch the details when applying the circle method, avoiding the most technical facts, which are anyway standard and can be found for instance in [21].

We rewrite the sum $\sum_{n\le N} a_{\mathcal{T}_Q}(n)^s$ as

$$\sum_{n\le N} \Big(\sum_{\substack{(p_1,v_1) \\ n=p_1+2^{v_1}}} 1 \Big) \dots \Big(\sum_{\substack{(p_s,v_s) \\ n=p_s+2^{v_s}}} 1 \Big) = \sum_{v_1,\dots,v_s} \sum_{\substack{p_1,\dots,p_s \\ p_i+2^{v_i}\le N}} 1, \tag{2.3}$$

where $p_1, \dots, p_s$ vary among the elements of the sequence $\mathcal{T}_Q$ which in the following will be denoted simply by $\mathcal{T}$, since Q is fixed throughout the proof.

We can rearrange the previous sum by grouping together all the v_j's with the same value. Assuming that $i_1, \dots, i_r$ among the v_j have values $u_1, \dots, u_r$, respectively, where $u_1 < u_2 < \cdots < u_r$, we obtain

$$\sum_{r=1}^{\infty} \sum_{\substack{i_1+\cdots+i_r=s \\ i_k\ge 1}} \frac{s!}{i_1! \dots i_r!} \sum_{u_1<\cdots<u_r\le L=\log_2 N} \sum_{p_1+2^{u_1}=\cdots=p_r+2^{u_r}\le N} 1\,. \tag{2.4}$$

To use the classical circle method, we introduce the trigonometric sums

$$S(\alpha) = \sum_{\substack{p\le N\\ p\in\mathcal{T}}} e(\alpha p)\,.$$

where we are using the standard notation $e(x) = e^{2\pi i x}$. In this notation we have

$$\begin{aligned}(2.5)\qquad &\sum_{p_1+2^{u_1}=\dots=p_r+2^{u_r}\le N} 1\\ =\ &\int_0^1\cdots\int_0^1 S(\alpha_1)\dots S(\alpha_r)\sum_{n\le N} e(-((n-2^{u_1})\alpha_1+\dots+(n-2^{u_r})\alpha_r))\mathrm{d}\alpha_1\dots\mathrm{d}\alpha_r\\ =\ &\int_0^1\cdots\int_0^1 S(\alpha_1)\dots S(\alpha_r)K(\alpha_1,...,\alpha_r)\mathrm{d}\alpha_1\dots\mathrm{d}\alpha_r\,,\end{aligned}$$

where

$$K(\alpha_1,...,\alpha_r) = \sum_{n\le N} e(-((n-2^{u_1})\alpha_1+\dots+(n-2^{u_r})\alpha_r))\,.$$

The next step is again classical. We split the interval $(0,1)$ into two parts: the so-called major arcs $\mathfrak{M}$ and minor arcs $\mathfrak{m}$.

Let R be a small power of N and let $V = N/R$. We consider a Farey dissection of the interval $[0,1[$ of total level V and level of the major arcs equal to R. More precisely, for $1\le a\le q\le R$ and $(a,q)=1$, a major arc $M(q,a)$ will be an interval with center in the rational point a/q and radius approximately $1/qV$. We define $\mathfrak{M}$ as the union of the major arcs $M(q,a)$ and $\mathfrak{m}$ just as $(0,1)\setminus\mathfrak{M}$.

As mentioned above, we shall only consider the integral over $\mathfrak{M}$. By performing the change of variable $\alpha_j = a_j/q_j+\beta_j$ each arc $M(q_j,a_j)$ is translated into the arc $\xi(q_j)$ with center in 0. The multiple integral over $\mathfrak{M}$ becomes

$$(2.6)\qquad \sum_{\substack{1\le a_1\le q_1\le R\\(a_1,q_1)=1}}\int_{\xi(q_1)} S(a_1/q_1+\beta_1)\dots\sum_{\substack{1\le a_r\le q_r\le R\\(a_r,q_r)=1}}\int_{\xi(q_r)} S(a_r/q_r+\beta_r)\cdot$$
$$K(a_1/q_1+\beta_1,...,a_r/q_r+\beta_r)\,\mathrm{d}\beta_1\dots\mathrm{d}\beta_r\,.$$

We now make precise our assumption that $\mathcal{T}$ is well distributed in the arithmetic progressions. As in Lemma 2.1, set $q=q_0d$. Recalling that the classes modulo q admissible for $\mathcal{T}=\mathcal{T}_Q$ are those coprime to q_0, the required distribution of $\mathcal{T}$ is of type of the distribution of primes in the arithmetic progressions, namely,

$$\sum_{\substack{p\in\mathcal{T}\\ p\le N\\ p\equiv l (\mathrm{mod}\ q)}} 1 \sim \frac{1}{\phi(q_0)d}\frac{N}{\log N}\qquad \text{if } (l,q_0)=1\,,\ \text{for each}\qquad q\le R.$$

We point out that this is the reason for not choosing $\mathcal{T}$ just as the sequence of primes not dividing Q; in fact, in this case this kind of distribution is known only when q is small with respect to N, which is not sufficient for our needs.

Let us now estimate $S(a/q)$ We have

$$S(a/q) = \sum_{\substack{p\le N\\ p\in\mathcal{T}}} e(ap/q) = \sum_{1\le l\le q} \sum_{\substack{p\in\mathcal{T}\\ p\le N\\ p\equiv l\ (q)}} e(ap/q) \sim \sum_{\substack{1\le l\le q\\ (l,q_0)=1}} e(al/q)\cdot\left(\frac{1}{\phi(q_0)d}\frac{N}{\log N}\right)$$

$$= \frac{1}{\phi(q_0)d}\frac{N}{\log N}\sum_{\substack{1\le l\le q\\ (l,q_0)=1}} e(al/q) = \frac{1}{\phi(q_0)d}\frac{N}{\log N}\left(\sum_{\substack{1\le i\le q_0\\ (i,q_0)=1}} e(ai/q_0)\right)\left(\sum_{1\le j\le d} e(aj/d)\right).$$

Now, if $d>1$, $\sum_{1\le j\le d} e(aj/d) = 0$, so we need to consider only the terms $S(a/q)$ where $d=1$ and $q=q_0$. On the other hand,

$$\sum_{\substack{1\le k\le q\\ (k,q)=1}} e(k/q) = \mu(q),$$

so we can further restrict to the terms for which $q \mid Q$. In this case we have

$$S(a/q) \quad\sim\quad \frac{1}{\phi(q)}\frac{N}{\log N}\sum_{\substack{1\le l\le q\\ (l,q)=1}} e(al/q) = \frac{\mu(q)}{\phi(q)}\frac{N}{\log N}.$$

In conclusion, we get

$$\sum_{\substack{q\le R\\ (a,q)=1}} S(a/q) \sim \sum_{\substack{q\mid Q\\ (a,q)=1}} \frac{\mu(q)}{\phi(q)}\frac{N}{\log N}. \tag{2.7}$$

Moreover, we can rewrite the function $K(a_1/q_1+\beta_1,\dots,a_r/q_r+\beta_r)$ of (2.6) as

$$\prod_{j=1}^{r} e(2^{u_j}(a_j/q_j+\beta_j)) \sum_{n\le N} e\big(-n\big((a_1/q_1+\cdots+a_r/q_r)+(\beta_1+\cdots+\beta_r)\big)\big) = P(c+\boldsymbol{\beta})B(c+\boldsymbol{\beta}),$$

where $\boldsymbol{\beta} = \sum_{j=1}^{r}\beta_j$, $c = c(a_1,q_1,\dots,a_r,q_r) = \sum_{j=1}^{r} a_j/q_j$, $P(c+\boldsymbol{\beta})$ is the arithmetic product over j, and $B(c+\boldsymbol{\beta})$ denotes the sum over n.

The function $B(x)$ is the classical bell-shaped function centered in 0 and periodic of period 1. So, denoting by $||\rho||$ the distance of a real number ρ from its nearest integer, we have $B(c+\boldsymbol{\beta}) = B(\|c+\boldsymbol{\beta}\|)$. Moreover, $B(0)=N$ and, for N sufficiently large, it decreases very rapidly outside 0; in our asymptotic estimate it can be approximated with the 2-piece linear function through the points $(-1/N,0)$, $(0,N)$ and $(1/N,0)$ in $[-1/N,1/N]$ and with 0 in $1/N<|x|\le 1/2$.

It follows that in the integrals in (2.6) we may neglect the contributions over the points for which

$$\|(a_1/q_1+\cdots+a_r/q_r)+(\beta_1+\cdots+\beta_r)\| > \frac{1}{N}, \tag{2.8}$$

while if

$$\|(a_1/q_1+\cdots+a_r/q_r)+(\beta_1+\cdots+\beta_r)\| \le \frac{1}{N}, \tag{2.9}$$

we can estimate the integral by approximating $B(c+\boldsymbol{\beta})$ with the 2-piece linear function above.

Moreover, recalling that all the relevant q are divisors of Q, we need only consider the pairs (a_j, q_j) for which

$$a_1/q_1 + \cdots + a_r/q_r \in \mathbb{N}\,. \tag{2.10}$$

In fact, if $a_1/q_1 + \cdots + a_r/q_r$ is not a integer, then the distance from the nearest integer is at least $1/Q$ and therefore the distance of $(a_1/q_1 + \cdots + a_r/q_r) + (\beta_1 + \cdots + \beta_r)$ from the nearest integer is greater than $1/Q - \sum_j 1/q_j V \geq 1/Q - rR/N$. If N is sufficiently large, this distance is always $> 1/N$ on the domain of integration, so (2.8) holds and the integral is negligible.

Hence condition (2.9) becomes

$$|\beta_1 + ... + \beta_r| \leq \frac{1}{N}. \tag{2.11}$$

Moreover, (2.11) can be refined by considering just the r-tuples $(\beta_1, \ldots, \beta_r)$ for which any single β_j is smaller than $1/N$. In fact, (2.11) individuates an r-dimensional strip out of which the integral is negligible. By symmetry, we can change the signs of the β_i's in all possible ways, obtaining 2^{r-1} strips such that the integral is negligible outside each of them. These strips divide the hypercube $[-1/2, 1/2]^r$ into a finite number of regions and intersect in a region contained in $[-1/N, 1/N]^r$. Every region other than the intersection is outside some strip, and therefore its contribution to the integral is negligible.

We are then left to study the integrals in (2.6) over the domain

$$D = \{(\beta_1, \ldots, \beta_r)\colon |\pm\beta_1 \pm \cdots \pm \beta_r| < 1/N\}.$$

The argument above shows also that enlarging the domain of integration to a $D' \supseteq D$ does not change the asymptotic value of the integral, since its value over $D' \setminus D$ is negligible.

Putting together all the previous considerations, and noticing that

$$D \subseteq \{(\beta_1, \ldots, \beta_r)\colon |\beta_j| < 1/N \ \forall j\},$$

we get that (2.6) is asymptotic to

$$\sum_{q_j|Q} \sum_{\substack{a_1/q_1+\cdots+a_r/q_r\in\mathbb{Z}\\(a_j,q_j)=1}} \int_{|\beta_1|<1/N} S(\frac{a_1}{q_1}+\beta_1)\cdots \int_{|\beta_r|<1/N} S(\frac{a_r}{q_r}+\beta_r)P(c+\boldsymbol{\beta})B(\boldsymbol{\beta})\mathrm{d}\beta_1 \ldots \mathrm{d}\beta_r. \tag{2.12}$$

Moreover, to obtain an asymptotic estimate of (2.12) we can replace $P(c+\boldsymbol{\beta})$ with $P(c)$ and $S(a_j/q_j + \beta_j)$ with $\frac{\mu(q_j)}{\phi(q_j)}\frac{N}{\log N}$ and equation (2.12) is asymptotic to

$$\sum_{\substack{q_j|Q\\1\leq j\leq r}} \prod_{h=1}^{r} \frac{\mu(q_h)}{\phi(q_h)} \left(\sum_{\substack{a_1/q_1+\cdots+a_r/q_r\in\mathbb{Z}\\(a_j,q_j)=1}} \prod_{j=1}^{r} e(2^{u_j}(a_j/q_j)) \right) \frac{N^r}{(\log N)^r} \int_D \cdots \int B(\boldsymbol{\beta})\mathrm{d}\beta_1 \ldots \mathrm{d}\beta_r$$

as $N \to \infty$. We now choose

$$D' = \{(\beta_1, \ldots, \beta_r)\colon |\sum_{j=1}^{r} \beta_j| < 1/N\,,\ |\sum_{j=1}^{r} \beta_j - 2\beta_i| < 1/N \text{ for } 1 < i \leq r\}$$

and perform the change of variables $y = \sum_{j=1}^{r} \beta_j \quad x_i = y - 2\beta_i$ for $1 < i \le r$. We have

$$\int_{D'} \cdots \int B(\boldsymbol{\beta}) \mathrm{d}\beta_1 \ldots \mathrm{d}\beta_r = \frac{1}{2^{r-1}} \int_{-1/N}^{1/N} \cdots \int_{-1/N}^{1/N} B(y) \mathrm{d}y \mathrm{d}x_2 \ldots \mathrm{d}x_r$$

$$= \frac{1}{N^{r-1}} \int_{-1/N}^{1/N} B(y) \mathrm{d}y \sim \frac{1}{N^{r-1}}.$$

Putting everything together, we obtain that (2.5) is asymptotic to

$$\sum_{\substack{q_j \mid Q \\ 1 \le j \le r}} \prod_{h=1}^{r} \frac{\mu(q_h)}{\phi(q_h)} \left(\sum_{\substack{a_1/q_1 + \cdots + a_r/q_r \in \mathbb{Z} \\ (a_j, q_j) = 1}} \prod_{j=1}^{r} e(2^{u_j}(a_j/q_j)) \right) \frac{N}{(\log N)^r} \tag{2.13}$$

as $N \to \infty$.

Summing over the u's, we obtain

$$\sum_{u_1 < \cdots < u_r \le L} \; \sum_{p_1 + 2^{u_1} = \cdots = p_r + 2^{u_r} \le N} 1 \sim \tag{2.14}$$

$$\sim \sum_{\substack{q_j \mid Q \\ 1 \le j \le r}} \sum_{\substack{a_1/q_1 + \cdots + a_r/q_r \in \mathbb{Z} \\ (a_j, q_j) = 1}} \prod_{h=1}^{r} \frac{\mu(q_h)}{\phi(q_h)} \left(\sum_{u_1 < \cdots < u_r \le L} \prod_{j=1}^{r} e(2^{u_j}(a_j/q_j)) \right) \frac{N}{(\log N)^r}.$$

Taking into account that $2^{u+\phi(q)} \equiv 2^u \pmod{q}$, the expression in the last parentheses is asymptotic to

$$\frac{L^r}{r!} \prod_{j=1}^{r} \frac{1}{\phi(q_j)} \left(\sum_{u_j=1}^{\phi(q_j)} e(2^{u_j}(a_j/q_j)) \right)$$

and, after substituting in (2.14) we get the asymptotic estimate

$$\frac{N(\log 2)^{-r}}{r!} \sum_{\substack{q_j \mid Q \\ 1 \le j \le r}} \sum_{\substack{a_1/q_1 + \cdots + a_r/q_r \in \mathbb{Z} \\ (a_j, q_j) = 1}} \prod_{h=1}^{r} \frac{\mu(q_h)}{\phi(q_h)^2} \sum_{u=1}^{\phi(q_h)} e(2^u(a_h/q_h)). \tag{2.15}$$

Then, substituting in (2.4), we get

$$\sum_{n \le N} a_{\mathcal{T}_Q}(n)^s \sim \tag{2.16}$$

$$\sim N \sum_{r=1}^{\infty} \sum_{\substack{i_1 + \cdots + i_r = s \\ i_k \ge 1}} \frac{s!}{i_1! \cdots i_r!} \frac{(\log 2)^{-r}}{r!} \sum_{\substack{q_j \mid Q \\ 1 \le j \le r}} \sum_{\substack{a_1/q_1 + \cdots + a_r/q_r \in \mathbb{Z} \\ (a_j, q_j) = 1}} \prod_{h} \frac{\mu(q_h)}{\phi(q_h)^2} \sum_{u=1}^{\phi(q_h)} e(2^u(a_h/q_h)).$$

For each sequence $\{\omega(m)\}_{m \in \mathbb{Z}}$, denote for simplicity

$$\operatorname*{mean}_{m \in \mathbb{Z}} \omega(m) := \lim_{M \to \infty} \frac{1}{2M+1} \sum_{m=-M}^{M} \omega(m).$$

With this notation we have

$$\operatorname*{mean}_{m\in\mathbb{Z}} e(m\sum_j a_j/q_j) = \begin{cases} 0 & \text{if } \sum_j a_j/q_j \notin \mathbb{Z}; \\ 1 & \text{if } \sum_j a_j/q_j \in \mathbb{Z}, \end{cases}$$

and hence

$$\sum_{n\le N} a_{\mathcal{T}_Q}(n)^s \sim \tag{2.17}$$

$$\sim \operatorname*{mean}_{m\in\mathbb{Z}} N \left(\sum_{r=1}^{\infty} \sum_{\substack{i_1+\cdots+i_r=s\\ i_k\ge 1}} \frac{s!}{i_1!\cdots i_r!} \frac{(\log 2)^{-r}}{r!} \sum_{\substack{q_j|Q\\ 1\le j\le r}} \sum_{(a_j,q_j)=1} \prod_h \frac{\mu(q_h)}{\phi(q_h)^2} \sum_{u=1}^{\phi(q_h)} e((2^u+m)a_h/q_h) \right)$$

$$\sim N\cdot \operatorname*{mean}_{m\in\mathbb{Z}} \sum_{r=1}^{\infty} \frac{1}{r!} \left(\frac{\Gamma_Q(m)}{\log 2}\right)^r \sum_{\substack{i_1+\cdots+i_r=s\\ i_k\ge 1}} \frac{s!}{i_1!\cdots i_r!},$$

where

$$\Gamma_Q(m) = \sum_{q|Q} \sum_{(a,q)=1} \frac{\mu(q)}{\phi(q)^2} \sum_{u=1}^{\phi(q)} e((2^u+m)a/q)\,. \tag{2.18}$$

Since the function $\Gamma_Q(m)$ is periodic of period Q, the mean over $\mathbb{Z}$ in (2.17) is equal to the mean over one period, hence equation (2.17) can be rewritten as

$$\sum_{n\le N} a_{\mathcal{T}_Q}(n)^s \sim \frac{N}{Q} \sum_{m=1}^{Q} \sum_{r=1}^{\infty} \frac{1}{r!} \left(\frac{\Gamma_Q(m)}{\log 2}\right)^r \sum_{\substack{i_1+\cdots+i_r=s\\ i_k\ge 1}} \frac{s!}{i_1!\cdots i_r!}. \tag{2.19}$$

We now multiply (2.19) by $(\log z)^s/s!$ and sum over s. Applying Lemma 2.2 to the left-hand side and Lemma 2.3 to the right-hand side, we obtain

$$\sum_{l=1}^{\infty} (z^l-1)\delta_{\mathcal{T}_Q}(l) = \frac{1}{Q} \sum_{m=1}^{Q} \sum_{r=1}^{\infty} \frac{1}{r!} \left(\frac{\Gamma_Q(m)}{\log 2}\right)^r (z-1)^r, \tag{2.20}$$

and setting $z=0$

$$\delta_{\mathcal{T}_Q} = \sum_{l=1}^{\infty} \delta_{\mathcal{T}_Q}(l) = \frac{1}{Q} \sum_{m=1}^{Q} \left(1 - \exp\left(-\frac{\Gamma_Q(m)}{\log 2}\right)\right).$$

Note that $\delta_{\mathcal{T}_Q}(l)$ is the coefficient of z^l in (2.20): looking at the right-hand side, we have that the coefficient of z^l in the inner sum is

$$\sum_{r\ge l} \binom{r}{l} (-1)^{r-l} \frac{1}{r!} \left(\frac{\Gamma_Q(m)}{\log 2}\right)^r = \frac{1}{l!} \left(\frac{\Gamma_Q(m)}{\log 2}\right)^l \sum_{r\ge 0} \frac{1}{r!} \left(-\frac{\Gamma_Q(m)}{\log 2}\right)^r,$$

hence

$$\delta_{\mathcal{T}_Q}(l) = \frac{1}{Q} \sum_{m=1}^{Q} \left(\frac{1}{l!} \left(\frac{\Gamma_Q(l)}{\log 2}\right)^l \exp\left(-\frac{\Gamma_Q(m)}{\log 2}\right)\right).$$

To conclude the proof of **(B1)** and **(B2)**, it remains to show that $\Gamma_Q(m) = \gamma_Q(m)$, namely,

$$\Gamma_Q(m) = \frac{Q}{\phi(Q)^2} \sum_{\substack{v=1 \\ (m+2^v,Q)=1}}^{\phi(Q)} 1\,.$$

Now, from equation (2.18) we have

$$\Gamma_Q(m) = \sum_{q|Q} \frac{\mu(q)}{\phi(q)^2} \sum_{u=1}^{\phi(q)} c_q(2^u + m),$$

where $c_q(n) = \sum_{(a,q)=1} e(na/q)$ is the Ramanujan sum, and it is well known that (see for example [21, formula (3.14)])

$$c_q(n) = \mu\left(\frac{q}{(q,n)}\right) \frac{\phi(q)}{\phi(q/(q,n))}.$$

Since $q|Q$, then q is squarefree and hence

$$c_q(2^u + m) = \mu(q)\mu((q, 2^u + m))\phi((q, 2^u + m)).$$

Using again that $2^{u+\phi(q)} \equiv 2^u \pmod q$, we have

$$\begin{aligned}
\Gamma_Q(m) &= \sum_{q|Q} \sum_{u=1}^{\phi(q)} \left(\frac{\mu(q)}{\phi(q)}\right)^2 \mu((q, m+2^u))\phi((q, m+2^u)) \\
&= \frac{1}{\phi(Q)} \sum_{u=1}^{\phi(Q)} \sum_{q|Q} \frac{\mu(q)^2}{\phi(q)} \mu((q, m+2^u))\phi((q, m+2^u)).
\end{aligned}$$

We note that $F(q) = \frac{\mu(q)^2}{\phi(q)}\mu((q, m+2^u))\phi((q, m+2^u))$ is a multiplicative function and since Q is squarefree we have

$$\begin{aligned}
\Gamma_Q(m) &= \frac{1}{\phi(Q)} \sum_{u=1}^{\phi(Q)} \sum_{q|Q} F(q) = \frac{1}{\phi(Q)} \sum_{u=1}^{\phi(Q)} \prod_{\substack{p \text{ prime} \\ p|Q}} (1 + F(p)) \\
&= \frac{1}{\phi(Q)} \sum_{u=1}^{\phi(Q)} \prod_{\substack{p \text{ prime} \\ p|Q}} \left\{ 1 + \frac{1}{p-1} \left(\begin{cases} 1 & \text{if } p \nmid m + 2^u \\ -(p-1) & \text{if } p \mid m + 2^u \end{cases} \right) \right\} \\
&= \frac{1}{\phi(Q)} \sum_{\substack{u=1 \\ (m+2^u,Q)=1}}^{\phi(Q)} \prod_{\substack{p \text{ prime} \\ p|Q}} \frac{p}{\phi(p)} = \frac{1}{\phi(Q)} \sum_{\substack{u=1 \\ (m+2^u,Q)=1}}^{\phi(Q)} \frac{Q}{\phi(Q)} \\
&= \frac{Q}{\phi(Q)^2} \sum_{\substack{u=1 \\ (m+2^u,Q)=1}}^{\phi(Q)} 1 = \gamma_Q(m)\,.
\end{aligned}$$

This concludes the proof.

3. Computation and approximation of δ_{Q_k}

We apply the results of the previous section to the sequence $\{Q_k\}_{k\geq 1}$ where $Q_k = p_1 \cdots p_k$ is the product of the first k primes. This sequence satisfies the hypothesis requested by Bombieri's conjecture, i.e. that Q_k is a squarefree number and that $Q_k \mid Q_{k+1}$. Moreover, the fast increase of Q_k should provide a sequence of values δ_{Q_k} rapidly converging to a limit, so that the values for small k can give a good insight into the limiting value. The quantities $\delta_k := \delta_{Q_k}$ are evaluated using formula **(B2)**. In order to compute the quantities δ_k for increasing values of k we need to compute the values of $\gamma_{Q_k}(m)$ for any integer $m \in [1, Q_k]$. The computation of δ_k for a given k requires a number of operations as large as $k\, Q_k\, \phi(Q_k)$, and can be time consuming for a large k.

With direct calculation and careful programming, we were able to compute the exact values of δ_k for $k = 1, 2, \ldots, 11$ producing the values reported in Table 1. The values of δ_k for $k \leq 9$ can be computed with any precision using Mathematica®. Two other values δ_{10} and δ_{11}, can be computed using double precision arithmetic. The direct calculation for $k > 11$ seems unfeasible with the current computational resources.

k	p_k	Q_k	δ_k
1	2	2	0.4720834970750688
2	3	6	0.4595145846080572
3	5	30	0.4496940203924133
4	7	210	0.4456422156631219
5	11	2 310	0.4446427853024636
6	13	30 030	0.4429385971709090
7	17	510 510	0.4409762694311436
8	19	$\approx 9.69 \cdot 10^6$	0.4405196746923989
9	23	$\approx 2.2 \cdot 10^8$	0.4402994162717804
10	29	$\approx 6.4 \cdot 10^9$	0.4400964391954299
11	31	$\approx 2.00 \cdot 10^{11}$	0.4393679993121766

Table 1. The first 11 values of δ_k.

Note that δ_k is obtained as the mean of Q_k values, and for $k = 12$, the values to be summed are approximately $7.4 \cdot 10^{12}$. For larger values of k, Q_k grows incredibly fast so that, from a computational point of view, we can assume Q_k as infinite. The idea presented in this section is to approximate the average over Q_k values in **(B2)** with the sample mean taken over a set of N_k values, with $N_k \ll Q_k$ but as large as possible.

3.1. Statistical estimate.

Let $\mathcal{I}_k = \{1, 2, \ldots, Q_k\}$ and define the function $\beta : \mathcal{I}_k \to \mathbb{R}$, as

$$\beta(m) = 1 - \exp\left(-\frac{\gamma_{Q_k}(m)}{\log 2}\right).$$

Note that the even values of m do not contribute to the average since $\beta(2\nu) = 0$, hence $\delta_k = \frac{1}{Q_k} \sum_{\nu=0}^{Q_k/2-1} \beta(2\nu + 1)$.

Let $\mathcal{M}$ be a sample set of $N_k/2$ values in $\mathcal{I}_k$. Let us consider the sample mean defined as

$$\Delta_k = \frac{1}{N_k} \sum_{m \in \mathcal{M}} \beta(m), \tag{3.1}$$

which is an approximation of δ_k. To measure how good is the approximation of δ_k with Δ_k, we need to estimate the variance of the quantities involved in the sum. Given a set of n samples X_i, the *unbiased sample variance* of the population, and the *corrected sample standard deviation* are defined as

(3.2)

$$\mathcal{V}^2(X) = \frac{1}{n-1} \sum_{i=1}^{n} \left(X_i - \frac{1}{n} \sum_{j=1}^{n} X_j \right)^2, \quad \mathcal{V}(X) = \frac{1}{\sqrt{n-1}} \sqrt{\sum_{i=1}^{n} \left(X_i - \frac{1}{n} \sum_{j=1}^{n} X_j \right)^2},$$

and represent an estimate of the population variance and standard deviation.

From the central limit theorem we know that the distribution of Δ_k is approximately $N(\mu, \frac{\sigma}{\sqrt{N_k}})$, where σ is the standard deviation of the whole population, meaning that with probability at least $1 - \alpha$ the actual mean $\mu = \delta_k$ lies in the interval $[\Delta_k - z_{\alpha/2} \frac{\sigma}{\sqrt{N_k}}, \Delta_k + z_{\alpha/2} \frac{\sigma}{\sqrt{N_k}}]$ where $z_{\alpha/2}$ is the critical value of the standard normal distribution which can be read from the tables of normal distribution. For example the critical value to have $1 - \alpha = 0.95$ is $z_{0.025} = 1.96$. The application of the central limit theorem requires the knowledge of the standard deviation, but for a sufficiently large N_k the standard deviation can be substituted by the corrected sample standard deviation $\mathcal{V}(X)$.

Of course this result holds when the sample set is large enough, or for a population already distributed normally. To have a population whose distribution is closer to the normal one we chose to compute the sample mean as the average over partial sample means. As we will see in Section 3.2 this has also the effect of partially avoiding instability due to floating point arithmetic.

3.2. **How to compute the sample mean.** In order to sum the $N = N_k$ different samples, we may proceed by levels. We first form N/L_1 groups of size L_1, obtaining N/L_1 partial sums, and then we group them again in groups of size L_2 and so on. In particular if $N = L_1 L_2 \cdots L_r$ we can proceed computing the partial means at level h obtained from the partial means at level $h-1$, treating the means at level $h-1$ as observations. More rigorously, denoting by $\mathcal{D}^{(0)}(m) = \beta(m)$ we compute the quantities $\mathcal{D}^{(h)}(g)$ as follows

$$\mathcal{D}^{(h)}(g) = \frac{1}{L_h} \sum_{i=1}^{L_h} \mathcal{D}^{(h-1)}(i + (g-1)L_h), \quad g = 1, 2, \ldots, \frac{N}{L_1 \cdots L_h},$$

and finally we have $\Delta_k = \mathcal{D}^{(r)}(1) = \frac{1}{L_r} \sum_{i=1}^{L_r} \mathcal{D}^{(r-1)}(i)$, where $L_r = N/(L_1 \cdots L_{r-1})$.

As already observed, computing the final sum proceeding by layers has an interesting side effect since the partial sums $\mathcal{D}^{(h)}(g)$ tend to be more normally distributed than the sums at the previous level. Each $\mathcal{D}^{(h)}(g)$ can be interpreted as an observation for the sums at level $h+1$. Moreover we note that the variance of the values $\mathcal{D}^{(h)}(g)$ decreases with h since at each level the partial sums concentrate better and better around the mean.

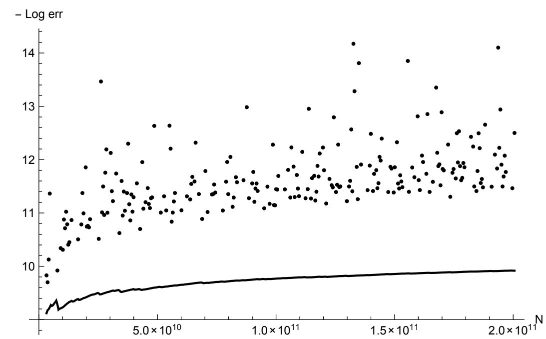

FIGURE 1. The solid line represent the value $\mathrm{est}_{11}^{L_2}(i)$, for $i = 2, 3, \ldots, 437$. The dots are the values of $\mathrm{err}_{11}^{L_2}(i)$ for $i = 1, 2, \ldots, L_3$, with $L_2 = 899$ amd $L_3 = 437$.

3.3. **Validation of the Error Estimate.** As observed in Section 3.1 the central limit theorem allows us to estimate the error of approximating δ_k with the sample average Δ_k. The theoretical bound is expressed in terms of the corrected standard deviation and of the critical value $z_{\alpha/2}$.

To understand how conservative this estimate is, and to use this result for large values of k, we consider the value $k = 11$ for which we are able to compute directly (see Table 1) the values of δ_k as the mean over all the Q_k values. The value δ_{11} obtained and reported in Table 1 has been computed in double precision. We compute Q_k/L_1 partial means of $L_1 = Q_7 = 510\,510$, values and we form $L_3 = \frac{Q_k}{L_1 L_2}$ groups of $L_2 = 899 = 29 \cdot 31$ partial means. We compute the $L_3 = 437 = 19 \cdot 23$ partial means obtained averaging over i consecutive values of $\mathcal{D}^{(2)}(g)$, i.e.

$$\mu(n) = \frac{1}{n} \sum_{g=1}^{n} \mathcal{D}^{(2)}(g), \quad n = 1, 2, \ldots, L_3.$$

We know that $\mu(n)$ approaches δ_{11} and that the last value $\mu(L_3) = \delta_{11}$. By fixing $L_2 = 899$ independently of the values of L_1 and Q_k, we consider the following two error measures, the *actual error*

$$\mathrm{err}_k^{L_2}(n) = |\delta_k - \mu(n)|, \quad n = 1, 2, \ldots, L_3, \tag{3.3}$$

and the *error estimate*

$$\mathrm{est}_k^{L_2}(n) = \frac{2}{\sqrt{n}} \mathcal{V}(\mathcal{D}^{(2)}) = \frac{2}{\sqrt{n}\sqrt{n-1}} \sqrt{\sum_{g=1}^{n} \left(\mathcal{D}^{(2)}(g) - \mu(n)\right)^2} \qquad n = 2, 3, \ldots, L_3, \tag{3.4}$$

which should approach with high probability the theoretical value $z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$, where σ is the standard deviation of the population, n is the sample size, and $z_{\alpha/2} \approx 2$ if we require an accuracy of 95%. The behavior of these two error estimates is shown in Figure 1 and we see that the theoretical bound given by (3.4) is always better than expected since no value of the vector $\mathrm{est}_{11}^{L_2}$ is below the actual error even if this can happen with a probability up to 5%.

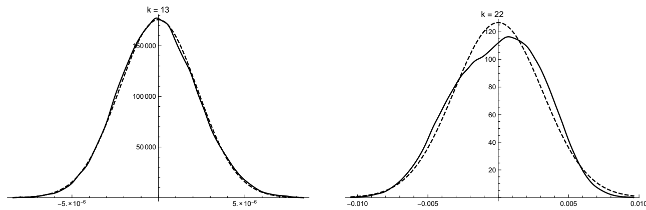

FIGURE 2. We see that for $k = 13, L_1 = 510\,510$ the histogram almost overlaps with the theoretical Gaussian distribution, while for $k = 22$, where we are taking only $L_1 = 210$ samples for each mean value there is a discrepancy between the histogram and the theoretical distribution.

3.4. **Numerical stability.** When summing N_k positive numbers particular attention should be placed on the error produced by floating point arithmetic. When using the customary algorithms, denoting by ε the machine precision, the error is proportional to $\varepsilon\, N_k$ or $\varepsilon \log_2 N_k$ depending on the algorithm used. With a little computational overhead we can compute the sum with a floating point error of the order of machine precision, independently of the number of terms to be added (as long as $N_k\, \varepsilon < 1$). This technique, due to Kahan [11] (see also [10] and references therein for more details) is the one adopted in this paper for

accumulating all the sums in our computation, which can therefore be considered correct up to machine precision. The only error produced is then due to the truncation of the original sum in **(B2)** to the first N_k terms, which can be estimated by $\mathrm{err}_k^{L_2}$ as defined in (3.4).

The problem of computing the sample variance of N data points may be difficult [4] in particular when N is large and the variance of our data is small. The algorithm given by (3.2), requires two passes through the data and it is a stable algorithm [10] because it is minimally affected by cancellation.

3.5. **Results of the statistical approximation.** The choice of the sample set is crucial since the samples selected should be representative of the whole population. Selecting consecutive values of m ensures that we have no repeated samples and we are exploring all the residual classes.

Of course as k grows we can only select a small portion of values to average because the cost of computing $\beta(m)$ increases with k. The central limit theorem, which is the main tool we are using here to estimate the error, requires the samples to be independent and sufficiently large in number. We are not able to prove independence of the observations – and, indeed, we doubt that the initial observations are independent – however, computing the means of the first layer of data for $k = 13$ and $k = 22$ we get the two histograms[1] in Figure 2 where they are compared with the theoretical Gaussian distributions. A four months computation produced the results reported in Table 2. In the table we report the values of L_1 and L_3, while we omit L_2 which is set to $899 = 29 \cdot 31$ for all the k, the sample size N_k is

[1]Each histogram is been obtained from $L_2 L_3$ points and smoothed with the corresponding command in Mathematica.

k	$\approx Q_k$	L_1	L_3	Δ_k	$\mathrm{est}^{(L_2)}(L_3)$
12	7.42e+12	510 510	65	0.4391667926	6.16e-10
13	3.04e+14	510 510	65	0.4388427016	9.76e-10
14	1.31e+16	510 510	65	0.4385677997	1.03e-09
15	6.15e+17	510 510	43	0.4385202254	1.27e-09
16	3.26e+19	30 030	71	0.4384639848	1.60e-08
17	1.92e+21	210	371	0.4384402240	7.07e-07
18	1.17e+23	210	318	0.4383560374	8.42e-07
19	7.86e+24	210	265	0.4383219193	9.30e-07
20	5.58e+26	210	265	0.4382922878	9.23e-07
21	4.07e+38	210	238	0.4380775133	1.07e-06
22	3.28e+30	210	212	0.4380501081	1.14e-06

TABLE 2. The table reports the estimates of δ_k for values of k between 12 and 22. The value of L_2 is set to 899. From the last column, we note that the accuracy is roughly of 6 digits for the higher values of k. Due to the increase in the complexity of computing the values $\beta(m)$ for large k, the values of $N_k = L_1L_2L_3$ are decreasing as k increases.

the product of L_1, L_2 and L_3. The values $\Delta_k \approx \delta_k$ are obtained using (3.1) and the error estimate $\mathrm{est}_k^{(L_2)}(L_3)$ is computed accordingly with the definition in (3.4).

3.6. **Extrapolating the δ_k.** With our sequence of δ_k we can try to extrapolate a possible limit value for δ. In general, it is not clear which can be a good mathematical model for the convergence of the δ_k's to a possible limit value δ. However, assuming the convergence to the limit of $A(x)/x$ to δ, a good model for extrapolation seems to be the one derived by the prime-counting function $\pi(x) \approx \frac{x}{\log x} + \frac{x}{\log^2 x}$. Accounting also for repeated representations, since we have approximately $\pi(x)\lfloor \log_2 x \rfloor$ values for $A(x)$, then it is reasonable to approximate $A(x)/x$ with the function

$$f_t(x) = \alpha_0 + \sum_{i=1}^{t} \frac{\alpha_i}{\log^i x}. \tag{3.5}$$

The values $\alpha_0, \alpha_1, \dots \alpha_t$ can be computed using the least squares method, and then we can return α_0 as the approximation of δ since for $k \to \infty$ the extrapolating function converges to α_0.

Setting $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \dots, \alpha_t)^T$, the vector of the unknown, and let (Q_k, δ_k) for $k = 2, 3, \dots, m$ be the data points, the least squares problem consists in minimizing the 2-norm

$$\|\boldsymbol{d} - \Phi_m \boldsymbol{\alpha}^*\|_2 = \min_{\boldsymbol{\alpha} \in \mathbb{R}^{t+1}} \|\boldsymbol{d} - \Phi_m \boldsymbol{\alpha}\|_2$$

where $\boldsymbol{d} = (\delta_2, \delta_3, \dots, \delta_{11}, \Delta_{12}, \dots, \Delta_m)^T$, and $\Phi_m \in \mathbb{R}^{(t+1)\times(m-1)}$ defined as

$$\Phi_m = \begin{bmatrix} 1 & \frac{1}{\log Q_2} & \frac{1}{\log^2 Q_2} & \cdots & \frac{1}{\log^t Q_2} \\ 1 & \frac{1}{\log Q_3} & \frac{1}{\log^2 Q_3} & \cdots & \frac{1}{\log^t Q_3} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & \frac{1}{\log Q_m} & \frac{1}{\log^2 Q_m} & \cdots & \frac{1}{\log^t Q_m} \end{bmatrix}.$$

The low dimension of the matrix Φ_m allows us to use one of the customary numerical methods for the normal equations $(\Phi_m^T \Phi_m)\boldsymbol{\alpha}^* = \Phi_m^T \mathbf{d}$. This computation can be performed in Mathematica with high precision in order to keep the round-off negligible. Since we are interested only in the first entry of the solution, and the

	$t = 2$		$t = 3$	
m	α_0^*	ε_{in}	α_0^*	ε_{in}
15	0.437648463	4.81e-10	0.4376394264	7.63e-10
16	0.437619800	2.26e-09	0.4375941534	3.61e-09
17	0.437603036	7.62e-08	0.4375701224	1.20e-07
18	0.437588229	1.51e-07	0.4375500504	2.32e-07
19	0.437578591	2.21e-07	0.4375384734	3.33e-07
20	0.437572717	2.81e-07	0.4375327153	4.15e-07
21	0.437554925	3.43e-07	0.4375091337	4.99e-07
22	0.437540758	4.03e-07	0.4374914180	5.76e-07

TABLE 3. The extrapolated values for different values of m and t and the round-off error obtained in the computation.

solution can be expressed in terms of the pseudo-inverse $F = (\Phi_m^T \Phi_m)^{-1} \Phi_m^T$ of Φ_m as

$$\boldsymbol{\alpha}^* = (\Phi_m^T \Phi_m)^{-1} \Phi_m^T \mathbf{d},$$

we have that $\alpha_0^* = \mathbf{f}_1^T \mathbf{d}$, where $\mathbf{f}_1^T$ is the first row vector of the matrix $(\Phi_m^T \Phi_m)^{-1} \Phi_m^T$. Table 3 reports for different values of m and t in (3.5), the values obtained for α_0^* together with the global error estimate. The global error is given by

$$\varepsilon_{in} = \mathbf{f}_1^T \tilde{\mathbf{d}} - \mathbf{f}_1^T \mathbf{d} = \sum_{k=1}^{m} f_{1k}\, \varepsilon^{(k)}$$

where $\varepsilon^{(k)}$ represent the error in estimating δ_{k+1} with Δ_{k+1} if $k \geq 11$,while is zero for $k \leq 11$ since the values reported in 1 are correct to machine precision. The absolute value of the errors $|\varepsilon^{(k)}|$ can be bounded from above by the values $\text{est}_k^{(L_2)}(L_3)$ reported in the last column of Table 2, hence $|\varepsilon_{in}| \leq \sum_{k=1}^{m-1} |f_{1k}|\, |\text{est}_k^{(L_2)}(L_3)|$.

In Figure 3 we report the fitting of the data on the model with $t = 2$ and $t = 3$.

Accepting as adequate the model proposed in (3.5), we see that the values of α_1^*, $T_2 \approx 0.437540758$ for $t = 2$ and $T_3 \approx 0.437491418$ for $t = 3$ are very close, and then we can propose the value 0.4375 as a possible approximation of the asymptotic distribution δ.

4. Direct Approach

In this section we compute the function $A(x)/x$ for very large values of x. Such a function is obtained with an exhaustive approach by counting all the odd numbers not exceeding x which can be written as the sum of a prime and a power of two. This exhaustive approach is computationally very expensive and requires careful programming. Using a segmented implementation of the sieve of Eratosthenes we computed around 10^4 Billions of primes in the range $[2, 3.3518 \cdot 10^{14}]$ storing them in $n = 187\,805$ chunks of size $c = 2^3 \prod_{i=1}^{9} p_i = 1\,784\,742\,960$. For each chunk

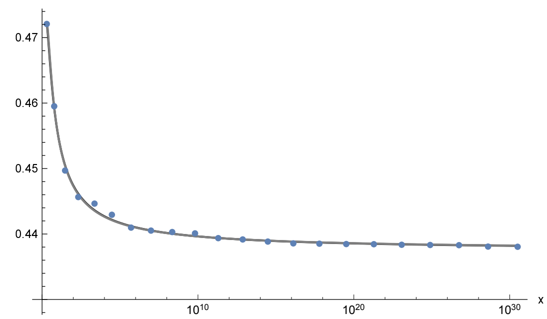

FIGURE 3. Fitting of the data (dots) with the model (3.5) with $t = 2$ and $t = 3$. We see that the two curves are overlapping and so we get a very good approximation of the data.

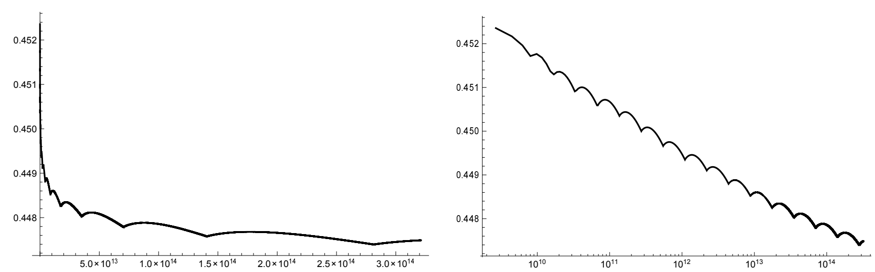

FIGURE 4. Plot of the density function $A(x)/x$ in a normal and in a logarithmic scale. The presence of local minima at powers of two is more evident when we use a logarithmic scale.

we computed the number of integers representable as $p + 2^n$ in the corresponding interval.

Observing that there are few "big" primes and a large number of "small" primes, it is not surprising to find out that the density function $A(x)/x$ is oscillating with local minima corresponding with powers of two as represented in Figure 4. As in the previous section we can use extrapolation techniques to estimate d_1, d_2 and understand if the conjecture $d_1 = d_2 = \delta$ is true, that is if indeed the density function $A(x)/x$ has a limit as x goes to infinity.

As done in Section 3.6, assuming the existence of the limit δ of $A(x)/x$, we can approximate $A(x)/x$ with the function

$$f(x) = \alpha_0 + \alpha_1 \frac{1}{\log x}. \tag{4.1}$$

The values of α_0 and α_1 can be estimated using the least squares method and then we can return α_0 as the approximation of δ since this is the asymptotic behavior of $f(x)$. In Figure 5 we show the fit obtained. The matrix of the least squares problem has size $n \times 2$, with n the number of chunks, that in our case is $n = 187\,805$, but the normal equations give rise to a 2×2 system. Using the computed values of $A(x)$ for x as large as 2^{48} we estimate α_0 with the value of $F \approx 0.437641$.

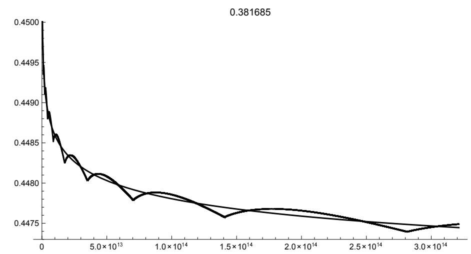

FIGURE 5. Plot of the density function $A(x)/x$ together with the fitting function from (4.1), where the parameters α_0, α_1 have been computed with the least squares method.

To better represent the oscillatory behaviour of $A(x)/x$ at a power of two, we can design a model where we add to the inverse logarithmic term a periodic oscillating function, where the amplitude of the oscillations is damped as x increases. Setting $t = \log x$, for a fixed integer $s > 0$ we can use as a model for the density function the following

$$g(t) = \alpha_0 + \frac{\alpha_1}{t} - \frac{1}{t}\sum_{k=1}^{s} \beta_k \cos\left(k\frac{2\pi t}{\log 2}\right).$$

Again, since $g(t) \to \alpha_0$ as a $t \to \infty$, the density δ can be estimated with the value α_0. With $s = 1$ we obtain the value $G_1 \approx 0.437641$, and with $s = 6$ the value $G_6 \approx 0.437645$ that are very close, so it is not really interesting to introduce other terms. Note that G_1 and F coincide in the first six digits.

As an alternative, we can estimate separately the lower and upper density d_1 and d_2. In this case, setting as before $t = \log(x)$ we add periodic terms without the dumping terms so that even the model function does not admit a limit. We have, for a fixed integer $s > 0$,

$$h(t) = \alpha_0 + \frac{\alpha_1}{t} - \sum_{k=1}^{s} \beta_k \cos\left(k\frac{2\pi t}{\log 2}\right).$$

In this case, we can use $L_s = \alpha_0 - \sum_{k=1}^{s} |\beta_k|$ as an estimate for d_1 and $U_s = \alpha_0 + \sum_{k=1}^{s} |\beta_k|$ as an estimate for d_2. With different values of s we get different ranges $[L_s, U_s]$. In particular, for $s = 1$ we get $[L_1, U_1]$ with $L_1 \approx 0.437572$ and $U_1 \approx 0.437714$, and with $s = 6$, the range is $[L_6, U_6]$ with $L_6 \approx 0.437541$ and $U_6 \approx 0.437752$.

Another possibility consists in extrapolating only local maxima or local minima. We first identify local maxima between 2^{33} and 2^{48}, and then we extrapolate these values obtaining an approximation of d_2 of $M \approx 0.437863$. We can observe that this value is outside the ranges given before but the values coincide for the first three digits.

When extrapolating the minima, we can take advantage of the fact that they are localized at powers of two hence we can estimate the density on selected "small" intervals $[x - h, x]$ around powers of two. The computational complexity is much lower than the computation of the density up to the value x. This has allowed us to estimate densities on intervals as far as $2^{61} \approx 2.30 \cdot 10^{18}$ while with the first technique we were not able to go further than $3.35 \cdot 10^{14} \approx 2^{48}$.

Since $\lim_{x\to+\infty} \frac{A(x)}{x} = \delta$, we have that, if $A'(x)$ has limit at infinity, then $\lim_{x\to\infty} A'(x) = \delta$ and the derivative can be estimated with the difference quotient

$$A'(x) \approx \frac{A(x+h) - A(x)}{h} \approx \delta, \qquad \text{for } h \ll x, x \to \infty.$$

We can then compute how many odd numbers are representable in the interval $[x_i, x_i + h]$ and divide that number by h to have an estimate of $A(x_i)/x_i$. Applying this reasoning on powers of two, we can estimate the values of the density function at a power of two. The results of the extrapolation with $h = 2^{26}$ gives an estimate for d_1 of $m \approx 0.437588$, which still coincides with the other estimates for the first three digits. The results of the extrapolation are not upper or lower bounds of the quantities d_1, d_2 or δ but only possible estimates. However, the fact that with different models and data points we get results which are so close, is a very encouraging result.

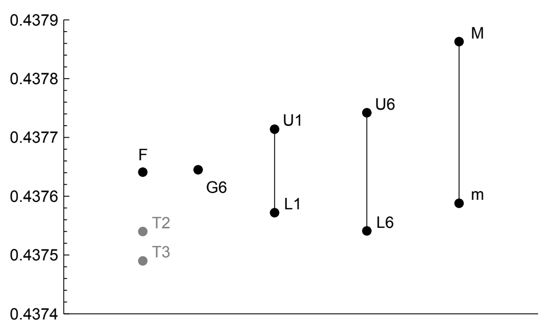

Figure 6. The estimates obtained using the direct and theoretical approach. F is the value obtained using the computed values of $A(x)$ up to 2^{48} and as a model for the density the function $f(x)$ in (4.1), G_6 is the value obtained by incorporating six oscillatory terms, the ranges $[L_1, U_1]$, $[L_6, U_6]$ are obtained as estimate of the lower and upper densities $[d_1, d_2]$ with one or six terms, while $[m, M]$ are obtained by extrapolating minima and maxima of $A(x)$. The gray dots T_2 and T_3 are the values we get pursuing the theoretical approach in Section 2.

5. Conclusions

In this paper we studied the problem of the existence and of the computation of the density of the integers of the form $2^n + p$, where p is a prime. Figure 6 summarizes our results. While we cannot give a conclusive answer about the existence of an asymptotic density for odd integers that can be expressed as the sum of a prime and a power of two, the figures we get are very close and allow us to fill the gap between upper and lower density as produced using only theoretical reasoning. The experimental results moreover give great evidence to the likelihood of Bombieri's conjectures.

Acknowledgments

We would like to thank the Signals and Images Laboratory of the Italian National Research Council (CNR-ISTI) for providing us the computing resources on which we tested our algorithms. Special thanks goes to Marco Righi for his help and assistance with the machines at SILab and to Alberto Perelli for his suggestions on the use of the circle method and for providing us some very helpful references.

References

1. P. T. Bateman and R. A. Horn, *A heuristic asymptotic formula concerning the distribution of prime numbers*, Math. Comp. **16** (1962), 363–367.
2. ———, *Primes Represented by Irreducible Polynomials in One Variable*, Proc. Symp, in Pure Math. **VIII** (1965), 119–132.
3. E. Bombieri, *Private communication*, 1977.
4. Tony F. Chan, Gene H. Golub, and Randall J. LeVeque, *Algorithms for computing the sample variance: analysis and recommendations*, Amer. Statist. **37** (1983), no. 3, 242–247. MR 713834
5. Yong-Gao Chen and Xue-Gong Sun, *On Romanoff's constant*, J. Number Theory **106** (2004), no. 2, 275–284. MR 2059075
6. A. de Polignac, *Réchèrches nouvelles sur les nombres premieres*, C. R. Acad. Sci. Paris Sér. **29** (1849), "397–401".
7. J. G. van der Corput, *On de Polignac's Conjecture*, Simon Stevin **27** (1950), 99–105.
8. P. Erdös, *On integers of the form $2^k + p$ and some related problems*, Summa Brasil. Math. **2** (1950), 113–123. MR 0044558
9. Laurent Habsieger and Xavier-François Roblot, *On integers of the form $p + 2^k$*, Acta Arith. **122** (2006), no. 1, 45–50. MR 2217322
10. Nicholas J. Higham, *Accuracy and stability of numerical algorithms*, second ed., Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2002. MR 1927606
11. W. Kahan, *Further remarks on reducing truncation errors*, Commun. ACM **8** (1965), no. 1, 40.
12. Hongze Li and Hao Pan, *The Romanoff theorem revisited*, Acta Arith. **135** (2008), no. 2, 137–142. MR 2453528
13. M. Madritsch and S. Planitzer, *Romanov's theorem in number fields*, Tech. report, arXiv:1512.04869, 2015.
14. Hao Pan and Wei Zhang, *On the integers of the form $p^2 + b^2 + 2^n$ and $b_1^2 + b_2^2 + 2^{n^2}$*, Math. Comp. **80** (2011), no. 275, 1849–1864. MR 2785483
15. J. Pintz, *A note on Romanov's constant*, Acta Math. Hungar. **112** (2006), no. 1-2, 1–14.
16. F. Romani, *Computations concerning primes and powers of two*, Calcolo **20** (1983), no. 3, 319–336.
17. Francesco Romani, *Computer techniques applied to the study of additive sequences*, Ph.D. thesis, Scuola Normale Superiore di Pisa, 1978, Supervisor: E. Bombieri.
18. N. P. Romanoff, *Über einige Sätze der additiven Zahlentheorie*, Math. Ann. **109** (1934), no. 1, 668–678.
19. A. Schinzel, *Remarks on the paper "sur certains hypotheses concernant les nomnbres premiers"*, Acta Arith. **7** (1961), 1–8.
20. ———, *A remark on a paper of Bateman and Horn*, Math. Comp. **17** (1963), 445–447.
21. R. C. Vaughan, *The Hardy-Littlewood method*, vol. 80, Cambridge University Press, 1981.

Università di Pisa, Italy
E-mail address: gianna.delcorso@unipi.it, ilaria.delcorso@unipi.it, roberto.dvornicich@unipi.it, francesco.romani@unipi.it