NUMBERS WHICH ARE ORDERS ONLY OF CYCLIC GROUPS

PAUL POLLACK

Abstract. We call n a cyclic number if every group of order n is cyclic. It is implicit in work of Dickson, and explicit in work of Szele, that n is cyclic precisely when $\gcd(n, \phi(n)) = 1$. With $C(x)$ denoting the count of cyclic $n \le x$, Erdős proved that

$$C(x) \sim e^{-\gamma} x / \log\log\log x, \quad \text{as } x \to \infty.$$

We show that $C(x)$ has an asymptotic series expansion, in the sense of Poincaré, in descending powers of $\log\log\log x$, namely

$$\frac{e^{-\gamma} x}{\log\log\log x}\left(1 - \frac{\gamma}{\log\log\log x} + \frac{\gamma^2 + \frac{1}{12}\pi^2}{(\log\log\log x)^2} - \frac{\gamma^3 + \frac{1}{4}\gamma\pi^2 + \frac{2}{3}\zeta(3)}{(\log\log\log x)^3} + \dots\right).$$

1. Introduction

Call the positive integer n *cyclic* if the cyclic group of order n is the unique group of order n. For instance, all primes are cyclic numbers. It is implicit in work of Dickson [Dic05], and explicit in work of Szele [Sze47], that n is cyclic precisely when $\gcd(n, \phi(n)) = 1$, where $\phi(n)$ is Euler's totient. (In fact, this criterion had been stated as "evident" already by Miller in 1899 [Mil99, p. 235].) If $C(x)$ denotes the count of cyclic numbers $n \le x$, Erdős proved in [Erd48] that

$$C(x) \sim e^{-\gamma} x / \log\log\log x, \tag{1}$$

as $x \to \infty$, where γ is the Euler–Mascheroni constant. Thus, the relative frequency of cyclic numbers decays to 0 but "with great dignity" (Shanks).

Several authors have investigated analogues of (1) for related counting functions from enumerative group theory. See, for example, [May79, MM84, War85, Sri87, EMM87, EM88, NS88, Sri91, NP18]. Our purpose in this note is somewhat different; we aim to refine the formula (1). Begunts [Beg01], optimizing the method of [Erd48], showed that $C(x)$ is given by $e^{-\gamma} x / \log\log\log x$ up to a multiplicative error of size $1 + O(\log\log\log\log x / \log\log\log x)$ (the same result appears as Exercise 2 on p. 390 of [MV07]). We improve this as follows.

Theorem 1.1. *The function $C(x)$ admits an asymptotic series expansion, in the sense of Poincaré (see* [dB81, §1.5]*), in descending powers of* $\log\log\log x$. Precisely: *There is a sequence of real numbers $c_1, c_2, c_3, \dots$ such that, for each fixed positive integer N and all*

2010 *Mathematics Subject Classification.* Primary 11N37; Secondary 20D60.

The author is supported by the National Science Foundation (NSF) under award DMS-2001581.

large x,

$$C(x) = \frac{e^{-\gamma}x}{\log\log\log x}\left(1 + \frac{c_1}{\log\log\log x} + \frac{c_2}{(\log\log\log x)^2} + \dots + \frac{c_N}{(\log\log\log x)^N}\right) + O_N\left(\frac{x}{(\log\log\log x)^{N+2}}\right).$$

Our proof of Theorem 1.1 yields the following explicit determination of the constants c_k. Write the Taylor series for the Γ-function, centered at 1, in the form $\Gamma(1+z) = 1 + C_1 z + C_2 z^2 + \dots$. Then the coefficients $c_1, c_2, \dots$ are determined by the formal relation

$$1 + c_1 z + c_2 z^2 + c_3 z^3 + \dots = \exp(0! C_1 z + 1! C_2 z^2 + 2! C_3 z^3 + \dots).$$

For computations of the C_k and c_k, it is useful to recall that

$$\Gamma(1+z) = \exp\left(-\gamma z + \sum_{k=2}^{\infty} \frac{(-1)^k}{k}\zeta(k) z^k\right). \tag{2}$$

(This is one version of a well-known expansion for the digamma function; see, e.g., entries 5.7.3 and 5.7.4 in [OLBC10].) The first few c_k are given by

$$c_1 = -\gamma, \quad c_2 = \gamma^2 + \frac{1}{2}\zeta(2) = \gamma^2 + \frac{\pi^2}{12}, \quad c_3 = -\left(\gamma^3 + \frac{1}{4}\gamma\pi^2 + \frac{2}{3}\zeta(3)\right).$$

Owing to (2), each c_k belongs to the ring $\mathbb{Q}[\gamma, \zeta(2), \zeta(3), \dots, \zeta(k)]$. From the fact that the coefficients of $\log\Gamma(1+z)$ are alternating in sign, one deduces that both the C_k and the c_k are alternating as well. Moreover,

$$|c_k| \ge (k-1)!|C_k| \ge (k-1)!\zeta(k)/k \ge (k-1)!/k$$

for each $k \ge 2$. It follows that the series $1 + c_1/\log\log\log x + c_2/(\log\log\log x)^2 + \dots$ is purely an asymptotic series, in that it diverges for all values of x.

The proof of Theorem 1.1 has many ingredients in common with the related work cited above (see also [PP, Pol]). But we must be more careful about error terms than in earlier papers, and somewhat delicate bookkeeping is required to wind up with a clean result.

Notation. The letters p and q, with subscripts or other decorations, are reserved for primes. We use K_0, K_1, K_2, etc. for absolute positive constants. To save space, we write $\log_k$ for the kth iterate of the natural logarithm.

2. Lemmata

We will use Mertens' theorem in the following form, which is a consequence of the prime number theorem with the classical $x\exp(-K_0\sqrt{\log x})$ error estimate of de la Vallée Poussin.

Lemma 2.1. *There is an absolute constant c such that, for all $X \ge 3$,*

$$\sum_{p \le X} \frac{1}{p} = \log_2 X + c + O(\exp(-K_1\sqrt{\log X})).$$

Moreover, for all $X \ge 3$,

$$\prod_{p \le X}\left(1 - \frac{1}{p}\right) = \frac{e^{-\gamma}}{\log X}\left(1 + O(\exp(-K_2\sqrt{\log X}))\right).$$

The following sieve result is a special case of [HR74, Theorem 7.2].

Lemma 2.2. *Suppose that $X \geq Z \geq 3$. Let $\mathcal{P}$ be a set of primes not exceeding Z. The number of $n \leq X$ coprime to all elements of $\mathcal{P}$ is*

$$X \prod_{p \in \mathcal{P}} \left(1 - \frac{1}{p}\right) \left(1 + O\left(\exp\left(-\frac{1}{2} \frac{\log X}{\log Z}\right)\right)\right).$$

The final estimate of this section was proved independently by Pomerance (see Remark 1 of [Pom77]) and Norton (see the Lemma on p. 699 of [Nor76]).

Lemma 2.3. *For every positive integer m and every $X \geq 3$,*

$$\sum_{\substack{p \leq X \\ p \equiv 1z \pmod{m}}} \frac{1}{p} = \frac{\log_2 X}{\phi(m)} + O\left(\frac{\log(2m)}{\phi(m)}\right).$$

3. Proof of Theorem 1.1

3.1. **Outline.** We summarize the strategy of the proof, deferring the more intricate calculations to later sections. Put

$$y = \frac{\log_2 x}{2 \log_3 x} \quad \text{and} \quad z = (\log_2 x) \cdot \exp(\sqrt{\log_3 x}).$$

Let us call a prime p dividing n a standard divisor of $\gcd(n, \phi(n))$ if there is a prime $q \leq x^{1/\log_2 x}$ dividing n with $q \equiv 1 \pmod{p}$. Clearly, each standard divisor of $(n, \phi(n))$ is a divisor of $\gcd(n, \phi(n))$.

Let $\mathcal{S}_0$ be the set of $n \leq x$ with no prime factor in $[2, y]$. For each positive integer k, let $\mathcal{S}_k$ be the set of $n \in \mathcal{S}_0$ having exactly k distinct prime factors from the interval $(y, z]$, all of which divide n to the first power only, and at least one of which is a standard divisor of $\gcd(n, \phi(n))$. We will estimate $C(x)$ by

$$\#\left(\mathcal{S}_0 \setminus \bigcup_{1 \leq k \leq \log_3 x} \mathcal{S}_k\right) = \#\mathcal{S}_0 - \sum_{1 \leq k \leq \log_3 x} \#\mathcal{S}_k. \tag{3}$$

Suppose n is counted by $C(x)$ but not by (3). Then n has a prime factor $p \leq y$. Since n is counted by $C(x)$, it must be that $p \nmid \phi(n)$, so that n is not divisible by any $q \equiv 1 \pmod{p}$. By Lemma 2.2, for a given p the number of those $n \leq x$ is $\ll x \prod_{q \leq x,\ q \equiv 1 \pmod{p}} (1 - 1/q) \leq x \exp(-\sum_{q \leq x, q \equiv 1 \pmod{p}} 1/q)$. And by Lemma 2.3,

$$\sum_{\substack{q \leq x \\ q \equiv 1 \pmod{p}}} \frac{1}{q} = \frac{1}{p-1} \log_2 x + O(1) \geq 2 \log_3 x + O(1).$$

Thus, the number of n corresponding to a given p is $\ll x \exp(-2 \log_3 x) = x/(\log_2 x)^2$. Summing on $p \leq y$, we deduce that the total number of n counted by $C(x)$ but not (3) is $O(x/\log_2 x)$.

Working from the opposite side, suppose that n is counted by (3) but not by $C(x)$. Then at least one of the following holds:

(i) there is a prime $p > y$ for which $p^2 \mid n$,
(ii) there is a prime $p > z$ that divides n and $\phi(n)$,
(iii) there is a prime p in $(y, z]$ dividing n and a prime $q \equiv 1 \pmod p$ dividing n with $q > x^{1/\log_2 x}$,
(iv) n has more than $\log_3 x$ different prime factors in $(y, z]$.

The number of $n \le x$ for which (i) holds is $\ll x \sum_{p>y} 1/p^2 \ll x/y \log y \ll x/\log_2 x$. In order for (ii) to hold but (i) to fail, there must be a prime $q \equiv 1 \pmod p$ dividing n. Clearly, there are most x/pq such n corresponding to a given p, q. Thus, the number of n that arise this way is

$$\ll x \sum_{p>z} \frac{1}{p} \sum_{\substack{q \le x \\ q \equiv 1 \pmod p}} \frac{1}{q} \ll x \sum_{p>z} \frac{\log_2 x + \log p}{p^2} \ll \frac{x \log_2 x}{z} = \frac{x}{\exp(\sqrt{\log_3 x})}.$$

For similar reasons, the number of $n \le x$ for which (iii) holds is

$$\ll x \sum_{y<p\le z} \frac{1}{p} \sum_{\substack{x^{1/\log_2 x} < q \le x \\ q \equiv 1 \pmod p}} \frac{1}{q} \ll x \sum_{p>y} \frac{\log_3 x}{p^2} \ll x \frac{\log_3 x}{\log_2 x}.$$

To handle (iv), observe that $\sum_{y<p\le z} 1/p \le K_3/\sqrt{\log_3 x} < 1/2$ for large values of x. Thus, the number of $n \le x$ for which (iv) holds is (crudely) at most

$$x \sum_{k > \log_3 x} \left(\sum_{y<p\le z} 1/p \right)^k \le 2x(K_3/\sqrt{\log_3 x})^{\log_3 x} \le x/\log_2 x.$$

Collecting estimates, we conclude that

$$C(x) = \# \left(\mathcal{S}_0 \setminus \bigcup_{1 \le k \le \log_3 x} \mathcal{S}_k \right) + O(x/\exp(\sqrt{\log_3 x})).$$

Since the error term is $O_N(x/(\log_3 x)^{N+2})$ for any fixed N, for the sake of proving Theorem 1.1 we may replace $C(x)$ by $\#(\mathcal{S}_0 \setminus \bigcup_{1\le k\le \log_3 x} \mathcal{S}_k)$.

In §3.2 we prove suitable estimates for the numbers $\#\mathcal{S}_k$ and in §3.3 we tie everything together and complete the proof of Theorem 1.1.

3.2. **Estimating $\#\mathcal{S}_k$.** The case $k = 0$ is easy to dispense with. By Lemmas 2.1 and 2.2,

$$\#\mathcal{S}_0 = e^{-\gamma} \frac{x}{\log y} + O(x/\exp(K_4 \sqrt{\log_3 x})). \tag{4}$$

Now suppose that $1 \le k \le \log_3 x$. In order for the integer $n \le x$ to belong to $\mathcal{S}_k$, it is sufficient than $n = p_1 \cdots p_k m$ where

(a) $p_1, \dots, p_k$ are distinct primes belonging to $(y, z]$,
(b) the integer m is free of prime factors in $[2, z]$, and
(c) m has a prime factor $q \le x^{1/\log_2 x}$ with $q \equiv 1 \pmod{p_i}$ for some $i = 1, 2, \dots, k$.

Moreover, these conditions are close to necessary: If n belongs to $\mathcal{S}_k$ but does not satisfy all of (a)–(c), then n is divisible by some product pp' where $p, p' \in (y, z]$ with $p' \equiv 1 \pmod p$. Writing $p' = pt + 1$, where $t < z/y$, we see that the number of such $n \le x$ is at most

$$\sum_{p \in (y,z]} \sum_{t < z/y} \frac{x}{p(pt+1)} < x \sum_{p > y} \frac{1}{p^2} \sum_{t < z/y} \frac{1}{t} \ll \frac{x\sqrt{\log_3 x}}{\log_2 x}.$$

Thus, $\#\mathcal{S}_k$ is given by the count of $n \le x$ satisfying (a)–(c), up to an error term of $O(x\sqrt{\log_3 x}/\log_2 x)$.

Now fix distinct primes $p_1, \ldots, p_k \in (y, z]$. We will count the number of $n \le x$ for which (a)–(c) hold with $p_1, \ldots, p_k$ the prime divisors of n in $(y, z]$. To get at this, we count all $n = p_1 \ldots p_k m \le x$ where condition (b) holds and then subtract the contribution from n for which (b) holds but (c) fails. By Lemma 2.2, this is approximately

$$\frac{x}{p_1 \cdots p_k} \prod_{p \le z} \left(1 - \frac{1}{p}\right) \left(1 - \prod_{\substack{z < q \le x^{1/\log_2 x} \\ q \equiv 1 \ (\mathrm{mod}\ p_i) \text{ for some } i}} \left(1 - \frac{1}{q}\right)\right). \tag{5}$$

In fact, taking $X = x/p_1 \cdots p_k$ (which exceeds $x^{1/2}$) and $Z = x^{1/\log\log x}$ in Lemma 2.2, we see that the error in this approximation is (very crudely) bounded by $O(x/(p_1 \ldots p_k \log_2 x))$.

Now we replace $\prod_{p \le z}(1 - 1/p)$ in (5) with $e^{-\gamma}/\log z$. This introduces another error of size $x/(p_1 \cdots p_k \exp(K_5\sqrt{\log_3 x}))$.

It remains to estimate the product over q in (5). We have that

$$\begin{aligned}
\prod_{\substack{z < q \le x^{1/\log_2 x} \\ q \equiv 1 \ (\mathrm{mod}\ p_i) \text{ for some } i}} \left(1 - \frac{1}{q}\right) &= \exp\left(-\sum_{\substack{z < q \le x^{1/\log_2 x} \\ q \equiv 1 \ (\mathrm{mod}\ p_i) \text{ for some } i}} \frac{1}{q} + O\left(\sum_{q > z} \frac{1}{q^2}\right)\right) \\
&= \exp\left(-\sum_{\substack{z < q \le x^{1/\log_2 x} \\ q \equiv 1 \ (\mathrm{mod}\ p_i) \text{ for some } i}} \frac{1}{q}\right)(1 + O(1/z)).
\end{aligned}$$

Continuing, we observe that

$$\sum_{\substack{z < q \le x^{1/\log_2 x} \\ q \equiv 1 \ (\mathrm{mod}\ p_i) \text{ for some } i}} \frac{1}{q} = \sum_{i=1}^{k} \sum_{\substack{z < q \le x^{1/\log_2 x} \\ q \equiv 1 \ (\mathrm{mod}\ p_i)}} \frac{1}{q} + O\left(\sum_{1 \le i < j \le k} \sum_{\substack{z < q \le x^{1/\log_2 x} \\ q \equiv 1 \ (\mathrm{mod}\ p_i p_j)}} \frac{1}{q}\right),$$

and that the O-term here is

$$\ll \sum_{1 \le i < j \le k} \frac{\log_2 x}{p_i p_j} \ll \binom{k}{2} \frac{(\log_3 x)^2}{\log_2 x} \ll \frac{(\log_3 x)^4}{\log_2 x}.$$

Moreover,

$$\sum_{i=1}^{k} \sum_{\substack{z<q\le x^{1/\log_2 x}\\ q\equiv 1 \pmod{p_i}}} \frac{1}{q} = \sum_{i=1}^{k}\left(\frac{\log_2 x}{p_i-1} + O\left(\frac{\log_3 x}{p_i}\right)\right)$$

$$= \sum_{i=1}^{k} \frac{\log_2 x}{p_i} + O\left(k\frac{(\log_3 x)^2}{\log_2 x}\right) = \sum_{i=1}^{k}\frac{\log_2 x}{p_i} + O\left(\frac{(\log_3 x)^3}{\log_2 x}\right).$$

Therefore,

$$\prod_{\substack{z<q\le x^{1/\log_2 x}\\ q\equiv 1 \pmod{p_i} \text{ for some } i}} \left(1-\frac{1}{q}\right) = \left(\prod_{i=1}^{k}\exp\left(-\frac{\log_2 x}{p_i}\right)\right)\left(1+O\left(\frac{(\log_3 x)^4}{\log_2 x}\right)\right)$$

$$= \prod_{i=1}^{k}\exp\left(-\frac{\log_2 x}{p_i}\right) + O\left(\frac{(\log_3 x)^4}{\log_2 x}\right).$$

Now collect estimates. We find that the number of $n \le x$ satisfying (a)–(c) where $p_1, \dots, p_k$ are the prime divisors of n from $(y, z]$ is

$$x\frac{e^{-\gamma}}{\log z}\left(\frac{1}{p_1\cdots p_k} - \prod_{i=1}^{k}\frac{\exp(-(\log_2 x)/p_i)}{p_i}\right) + O\left(\frac{x}{p_1\cdots p_k \exp(K_5\sqrt{\log_3 x})}\right). \tag{6}$$

Finally, we sum (6) over all sets of distinct primes $p_1, \dots, p_k \in (y, z]$. The O-terms contribute $O(x/\exp(K_5\sqrt{\log_3 x}))$. Next we look at the contribution from the $1/p_1\cdots p_k$ terms. On the one hand, the multinomial theorem immediately implies that

$$\sum_{y<p_1<p_2<\dots<p_k\le z} \frac{1}{p_1\cdots p_k} \le \frac{1}{k!}\sigma_0^k, \quad \text{where} \quad \sigma_0 := \sum_{y<p\le z}\frac{1}{p}.$$

(We have $\sigma_0 \asymp 1/\sqrt{\log_3 x}$ for large x by Mertens' theorem.) On the other hand,

$$\sum_{\substack{p_1,\dots,p_k\in(y,z]\\ \text{distinct}}} \frac{1}{p_1\cdots p_k} = \sum_{\substack{p_1,\dots,p_{k-1}\in(y,z]\\ \text{distinct}}} \frac{1}{p_1\cdots p_{k-1}} \sum_{\substack{y<p_k\le z\\ p_k\notin\{p_1,\dots,p_{k-1}\}}} \frac{1}{p_k}$$

$$\ge \left(\sigma_0 - \frac{k-1}{y}\right) \sum_{\substack{p_1,\dots,p_{k-1}\in(y,z]\\ \text{distinct}}} \frac{1}{p_1\cdots p_{k-1}}.$$

We can estimate the sum over $p_1, \dots, p_{k-1}$ in a similar way. Iterating, we find that

$$\sum_{\substack{p_1,\dots,p_k\in(y,z]\\ \text{distinct}}} \frac{1}{p_1\cdots p_k} \ge \prod_{i=0}^{k-1}\left(\sigma_0 - \frac{i}{y}\right) \ge \left(\sigma_0 - \frac{2(\log_3 x)^2}{\log_2 x}\right)^k,$$

so that

$$\sum_{y<p_1<p_2<\dots<p_k\le z} \frac{1}{p_1\cdots p_k} \ge \frac{1}{k!}\left(\sigma_0 - \frac{2(\log_3 x)^2}{\log_2 x}\right)^k.$$

Combining the upper and lower bounds,

$$\sum_{y<p_1<p_2<\cdots<p_k\leq z}\frac{1}{p_1\cdots p_k}=\frac{1}{k!}\sigma_0^k\left(1+O\left(\frac{(\log_3 x)^3}{\log_2 x}\right)\right)^k=\frac{1}{k!}\sigma_0^k+O\left(\frac{1}{k!}\frac{(\log_3 x)^4}{\log_2 x}\right).$$

The contribution from the terms of the form $\prod_{i=1}^k \exp(-(\log_2 x)/p_i)/p_i$ can be handled similarly. Put

$$\sigma_1:=\sum_{y<p\leq z}\frac{\exp(-(\log_2 x)/p)}{p}.$$

Clearly, $\sigma_1\leq\sum_{y<p\leq z}1/p\ll 1/\sqrt{\log_3 x}$. Since $\exp(-(\log_2 x)/p)\gg 1$ when $p\geq\log_2 x$, we also have that $\sigma_1\gg\sum_{\log_2 x<p\leq z}1/p\gg 1/\sqrt{\log_3 x}$. Now a computation completely parallel to the one shown above yields

$$\sum_{y<p_1<p_2<\cdots<p_k\leq z}\prod_{i=1}^k\frac{\exp(-(\log_2 x)/p_i)}{p_i}=\frac{1}{k!}\sigma_1^k+O\left(\frac{1}{k!}\frac{(\log_3 x)^4}{\log_2 x}\right).$$

Piecing together all of our estimates, we conclude that

$$\#\mathcal{S}_k=e^{-\gamma}\frac{x}{\log z}\left(\frac{\sigma_0^k}{k!}-\frac{\sigma_1^k}{k!}\right)+O\left(\frac{x}{\exp(K_5\sqrt{\log_3 x})}+\frac{x}{k!}\frac{(\log_3 x)^4}{\log_2 x}\right). \tag{7}$$

3.3. **Denouement.** Summing (7) over positive integers $k\leq\log_3 x$, keeping in mind that $\sigma_0,\sigma_1\ll 1/\sqrt{\log_3 x}$, we find that

$$\sum_{1\leq k\leq\log_3 x}\#\mathcal{S}_k=e^{-\gamma}\frac{x}{\log z}(\exp(\sigma_0)-\exp(\sigma_1))+O\left(\frac{x}{\exp(K_6\sqrt{\log_3 x})}\right).$$

By Mertens' theorem, $\exp(\sigma_0)=\frac{\log z}{\log y}\left(1+O(1/\exp(K_7\sqrt{\log_3 x}))\right)$. So recalling (4),

$$\#\mathcal{S}_0-\sum_{1\leq k\leq\log_3 x}\#\mathcal{S}_k=e^{-\gamma}\frac{x}{\log z}\exp(\sigma_1)+O(x/\exp(K_8\sqrt{\log_3 x})).$$

By another application of the prime number theorem with the de la Vallée Poussin error term,

$$\sigma_1=\int_y^z\frac{\exp(-(\log_2 x)/t)}{t\log t}\,d\theta(t)=\int_y^z\frac{\exp(-(\log_2 x)/t)}{t\log t}\,dt+O(1/\exp(K_9\sqrt{\log_3 x})),$$

and thus

$$\#\mathcal{S}_0-\sum_{1\leq k\leq\log_3 x}\#\mathcal{S}_k=e^{-\gamma}\frac{x}{\log z}\exp\left(\int_y^z\frac{\exp(-(\log_2 x)/t)}{t\log t}\,dt\right)+O(x/\exp(K_{10}\sqrt{\log_3 x})). \tag{8}$$

We proceed to analyze the integral appearing in this last estimate. Making the change of variables $u=(\log_2 x)/t$,

$$\int_y^z\frac{\exp(-(\log_2 x)/t)}{t\log t}\,dt=\frac{1}{\log_3 x}\int_{(\log_2 x)/z}^{2\log_3 x}\frac{\exp(-u)}{u}\left(1-\frac{\log u}{\log_3 x}\right)^{-1}du.$$

Here $(\log_2 x)/z = \exp(-\sqrt{\log_3 x})$. Inside the domain of integration, $\log u \ll \sqrt{\log_3 x}$, and so for each fixed positive integer M,

$$\left(1 - \frac{\log u}{\log_3 x}\right)^{-1} = 1 + \left(\frac{\log u}{\log_3 x}\right) + \left(\frac{\log u}{\log_3 x}\right)^2 + \cdots + \left(\frac{\log u}{\log_3 x}\right)^M + O_M((\log_3 x)^{-(M+1)/2}).$$

Thus,

$$\begin{aligned}\frac{1}{\log_3 x}\int_{(\log_2 x)/z}^{2\log_3 x} \frac{\exp(-u)}{u}\left(1 - \frac{\log u}{\log_3 x}\right)^{-1} du \\ = \sum_{k=0}^{M} \frac{1}{(\log_3 x)^{k+1}} \int_{(\log_2 x)/z}^{2\log_3 x} \frac{\exp(-u)}{u} \log^k u\, du \\ + O\left(\frac{1}{(\log_3 x)^{(M+3)/2}} \int_{(\log_2 x)/z}^{2\log_3 x} \frac{\exp(-u)}{u}\, du\right).\end{aligned}$$

The O-term here is $\ll (\log_3 x)^{-\frac{1}{2}(M+3)} \int_{(\log_2 x)/z}^{2\log_3 x} du/u \ll (\log_3 x)^{-1-\frac{1}{2}M}$. To handle the main term, we integrate by parts to find that

$$\begin{aligned}\int_{(\log_2 x)/z}^{2\log_3 x} \frac{\exp(-u)}{u} \log^k u\, du = \exp(-u)\frac{\log^{k+1} u}{k+1}\Bigg|_{u=(\log_2 x)/z}^{u=2\log_3 x} \\ + \frac{1}{k+1}\int_{(\log_2 x)/z}^{2\log_3 x} \exp(-u) \log^{k+1} u\, du.\end{aligned}$$

For each $0 \le k \le M$, and all large x,

$$\exp(-u)\frac{\log^{k+1} u}{k+1}\Bigg|_{u=(\log_2 x)/z}^{u=2\log_3 x} = \frac{-1}{k+1}\left(\log\left(\frac{\log_2 x}{z}\right)\right)^{k+1} + O_M(1/\exp(K_{11}\sqrt{\log_3 x})),$$

while

$$\begin{aligned}\frac{1}{k+1}\int_{(\log_2 x)/z}^{2\log_3 x} \exp(-u)\log^{k+1} u\, du \\ = \frac{1}{k+1}\int_0^\infty \exp(-u)\log^{k+1} u\, du + O_M(1/\exp(K_{12}\sqrt{\log_3 x})) \\ = \frac{1}{k+1}\Gamma^{(k+1)}(1) + O_M(1/\exp(K_{12}\sqrt{\log_3 x})) \\ = k! C_{k+1} + O_M(1/\exp(K_{12}\sqrt{\log_3 x})).\end{aligned}$$

Assembling our results,

$$\begin{aligned}
&\int_y^z \frac{\exp(-(\log_2 x)/t)}{t\log t}\,dt \\
&\qquad = -\sum_{k=0}^{M} \frac{1}{k+1}\left(\frac{\log((\log_2 x)/z)}{\log_3 x}\right)^{k+1} + \sum_{k=0}^{M} \frac{k!C_{k+1}}{(\log_3 x)^{k+1}} + O_M((\log_3 x)^{-1-\frac{1}{2}M}) \\
&\qquad = \log\left(1 - \frac{\log((\log_2 x)/z)}{\log_3 x}\right) + \sum_{k=0}^{M} \frac{k!C_{k+1}}{(\log_3 x)^{k+1}} + O_M((\log_3 x)^{-1-\frac{1}{2}M}) \\
&\qquad = \log\frac{\log z}{\log_3 x} + \sum_{k=0}^{M} \frac{k!C_{k+1}}{(\log_3 x)^{k+1}} + O_M((\log_3 x)^{-1-\frac{1}{2}M}).
\end{aligned}$$

We now choose $M = 2N$, where N is as in Theorem 1.1. In the last displayed sum on k, the terms of the sum with $k \geq N$ may be absorbed into the error. Doing so and exponentiating,

$$\begin{aligned}
&\exp\left(\int_y^z \frac{\exp(-(\log_2 x)/t)}{t\log t}\,dt\right) \\
&\qquad\qquad = \frac{\log z}{\log_3 x}\exp\left(\sum_{1\leq k\leq N}\frac{(k-1)!C_k}{(\log_3 x)^k}\right)\left(1 + O_N((\log_3 x)^{-1-N})\right),
\end{aligned}$$

so that

$$\begin{aligned}
&e^{-\gamma}\frac{x}{\log z}\exp\left(\int_y^z \frac{\exp(-(\log_2 x)/t)}{t\log t}\,dt\right) \\
&\qquad\qquad = e^{-\gamma}\frac{x}{\log_3 x}\exp\left(\sum_{1\leq k\leq N}\frac{(k-1)!C_k}{(\log_3 x)^k}\right)\left(1 + O_N((\log_3 x)^{-1-N})\right) \\
&\qquad\qquad = e^{-\gamma}\frac{x}{\log_3 x}\exp\left(\sum_{1\leq k\leq N}\frac{(k-1)!C_k}{(\log_3 x)^k}\right) + O_N(x(\log_3 x)^{-2-N}).
\end{aligned}$$

This expression describes $\#(\mathcal{S}_0 \setminus \bigcup_{1\leq k\leq \log_3 x}\mathcal{S}_k)$, by (8), and so also describes $C(x)$, by the discussion in §3.1. Theorem 1.1 follows, along with the description of the constants c_k appearing in the introduction.

Acknowledgements

The author thanks Matthew Just and the referee for comments and corrections.

References

[Beg01] A. V. Begunts, *On an asymptotic formula for $F_1(x)$*, Vestnik Moskov. Univ. Ser. I Mat. Mekh. (2001), no. 5, 57–59, 71–72 (Russian).

[dB81] N. G. de Bruijn, *Asymptotic methods in analysis*, third ed., Dover Publications, Inc., New York, 1981.

[Dic05] L. E. Dickson, *Definitions of a group and a field by independent postulates*, Trans. Amer. Math. Soc. **6** (1905), 198–204.

[EM88] P. Erdős and M. E. Mays, *On nilpotent but not abelian groups and abelian but not cyclic groups*, J. Number Theory **28** (1988), 363–368.

[EMM87] P. Erdős, M. R. Murty, and V. K. Murty, *On the enumeration of finite groups*, J. Number Theory **25** (1987), 360–378.

[Erd48] P. Erdős, *Some asymptotic formulas in number theory*, J. Indian Math. Soc. (N.S.) **12** (1948), 75–78.

[HR74] H. Halberstam and H.-E. Richert, *Sieve methods*, London Mathematical Society Monographs, vol. 4, Academic Press, London-New York, 1974.

[May79] M. E. Mays, *Counting abelian, nilpotent, solvable, and supersolvable group orders*, Arch. Math. (Basel) **31** (1978/79), 536–538.

[Mil99] G. A. Miller, *Report on recent progress in the theory of the groups of a finite order*, Bull. Amer. Math. Soc. **5** (1899), 227–249.

[MM84] M. R. Murty and V. K. Murty, *On the number of groups of a given order*, J. Number Theory **18** (1984), 178–191.

[MV07] H. L. Montgomery and R. C. Vaughan, *Multiplicative number theory. I. Classical theory*, Cambridge Studies in Advanced Mathematics, vol. 97, Cambridge University Press, Cambridge, 2007.

[Nor76] K. K. Norton, *On the number of restricted prime factors of an integer. I*, Illinois J. Math. **20** (1976), 681–705.

[NP18] R. Nedela and C. Pomerance, *Density of singular pairs of integers*, Integers **18** (2018), paper A82, 7 pages.

[NS88] M. J. Narlikar and S. Srinivasan, *On orders solely of abelian groups. II*, Bull. London Math. Soc. **20** (1988), 211–216.

[OLBC10] F. W. J. Olver, D. W. Lozier, R. F. Boisvert, and C. W. Clark (eds.), *NIST handbook of mathematical functions*, U.S. Department of Commerce, National Institute of Standards and Technology, Washington, DC; Cambridge University Press, Cambridge, 2010.

[Pol] P. Pollack, *The number of non-cyclic Sylow subgroups of the multiplicative group modulo n*, Canad. Math. Bull., to appear.

[Pom77] C. Pomerance, *On the distribution of amicable numbers*, J. Reine Angew. Math. **293(294)** (1977), 217–222.

[PP] C. Pomerance and P. Pollack, *Phi, primorials, and Poisson*, Illinois J. Math., to appear.

[Sri87] S. Srinivasan, *On orders solely of abelian groups*, Glasgow Math. J. **29** (1987), 105–108.

[Sri91] ———, *On orders solely of abelian groups. III*, J. Number Theory **39** (1991), 175–180, corrigendum in **65** (1997), 180–181.

[Sze47] T. Szele, *Über die endlichen Ordnungszahlen, zu denen nur eine Gruppe gehört*, Comm. Math. Helv. **20** (1947), 265–267.

[War85] R. Warlimont, *On the set of natural numbers which only yield orders of abelian groups*, J. Number Theory **20** (1985), 354–362.

Department of Mathematics, University of Georgia, Athens, GA 30602

Email address: pollack@uga.edu