HARMONIC SUMS IN ARITHMETIC PROGRESSIONS

STEVE FAN

Abstract. A well-known fact in elementary number theory states that the nth harmonic number $\sum_{k=1}^{n} 1/k$ is not an integer if $n \geq 2$. More generally, one can show that the harmonic sum $\sum_{k=m}^{n} 1/k$ is not an integer if $n > m \geq 1$. In this expository note we discuss harmonic sums in arithmetic progressions of the form $\sum_{k=0}^{n} 1/(a+kd)$. More precisely, we shall follow a paper of Paul Erdős [4] to prove that if a, d, n are positive integers, then the sum $\sum_{k=0}^{n} 1/(a+kd)$ is never an integer. We shall also discuss a similar result concerning the arithmetic properties of the generalized harmonic sum $\sum_{k=m}^{n} 1/k^{\theta_k}$, where $\theta_m, \theta_{m+1} ..., \theta_n$ are positive integers. At the end of this note we point out a connection between the arithmetic properties of harmonic sums and the distribution of primes as well as potential generalizations of harmonic numbers.

1. Introduction

Throughout the paper, let $\mathbb{R}$ denote the field of real numbers, $\mathbb{R}_+$ the set of positive real numbers, $\mathbb{Q}$ the field of rational numbers, $\mathbb{Z}$ the set of integers, $\mathbb{N}_+$ the set of positive integers, $\mathbb{N} = \mathbb{N}_+ \cup \{0\}$ the set of natural numbers, and $\mathbb{P}$ the set of prime numbers. For any $x \in \mathbb{R}$, we denote by $\lfloor x \rfloor$ the integer part of x and by $\lceil x \rceil$ the least integer greater than or equal to x. For every $n \in \mathbb{N}_+$, the nth harmonic number is defined by

$$H_n := \sum_{k=1}^{n} \frac{1}{k}.$$

Here H_n can be viewed as the nth partial sum of $\zeta(1)$, where ζ is the Riemann zeta function defined by

$$\zeta(s) := \sum_{k=1}^{\infty} \frac{1}{k^s}$$

for $s \in \mathbb{C}$ with $\Re(s) > 1$. More generally, one can define the generalized harmonic number

$$H_{m,n} := \sum_{k=m}^{n} \frac{1}{k}.$$

A folklore fact in elementary number theory states that $H_n \notin \mathbb{Z}$ for all $n \geq 2$. This was proved in 1915 by Taeisinger [15, p. 1307]. The more general result that $H_{m,n} \notin \mathbb{Z}$ for all $n > m \geq 1$ was proved by Kűrschák [15, p. 1307] three years later. Kűrschák's result can be proved easily by looking at the highest power of 2 dividing the denominators $m, m+1, ..., n$. In 1932 Erdős [4], at age of 19, generalized Kűrschák's result to harmonic sums in which the denominators of the terms form an arithmetic progression. In particular, he showed that for arbitrary $a, d, n \in \mathbb{N}_+$, the sum

$$\sum_{k=0}^{n} \frac{1}{a+kd}$$

is never an integer. This is one of the earliest mathematical discoveries he made in his career.

The main purpose of this paper is to present Erdős' proof of his result in a modern and cleaner way. We follow mostly his original argument but with slight modifications and simplifications. The author hopes that an exposition of this kind will help make Erdős' result accessible to a wider audience in the math community, given that his original paper [4] was written in Hungarian. Besides Erdős' theorem, we shall discuss potential generalizations of harmonic numbers which may be interesting for further study. In Section 2 we introduce the basic lemmas needed for our proof of the main theorems. In Section 3 we prove Erdős' result in a slightly more general form. In Section 4 we prove a generalization of Kűrschák's result which is not included in Erdős' theorem. In Section 5 we discuss the connection between harmonic sums and the distribution of primes. Finally, we generalize H_n and $H_{m,n}$ to sums in algebraic number fields.

2. Elementary Lemmas

In this section, we prove some preliminary lemmas needed for the proof of Erdős' theorem. Lemmas 2.1 and 2.2 are explicitly stated and proved in [4]. Lemma 2.3 is a well-known result in elementary number theory. Lemma 2.4 and its proof are due to the author.

Lemma 2.1 (Erdős, 1932). *Let $n, \alpha \in \mathbb{N}_+$. Then the binomial coefficient*

$$\binom{2n}{n} = \frac{(2n)!}{(n!)^2}$$

is divisible by any prime p such that $\sqrt[\alpha]{n} < p \leq \sqrt[\alpha]{2n}$.

Proof. Let $p \in \mathbb{P}$ be any prime in the interval $(\sqrt[\alpha]{n}, \sqrt[\alpha]{2n}]$. By Legendre's formula, the exponent of p in $n!$ is given by

$$\sum_{k=1}^{\alpha-1} \left\lfloor \frac{n}{p^k} \right\rfloor,$$

and the exponent of p in $(2n)!$ is

$$\sum_{k=1}^{\alpha-1} \left\lfloor \frac{2n}{p^k} \right\rfloor + 1,$$

since $\lfloor 2n/p^\alpha \rfloor = 1$. Consequently, the exponent of p in $\binom{2n}{n}$ is

$$1 + \sum_{k=1}^{\alpha-1} \left(\left\lfloor \frac{2n}{p^k} \right\rfloor - 2 \left\lfloor \frac{n}{p^k} \right\rfloor \right) \geq 1,$$

since $\lfloor 2n/p^\alpha \rfloor \geq 2\lfloor n/p^\alpha \rfloor$. This shows that $\binom{2n}{n}$ is divisible by p. □

Lemma 2.2 (Erdős, 1932). *Let $n \in \mathbb{N}_+$, and let m be the largest integer for which $2^m \leq n$. For each $k = 1, ..., m$, set*

$$a_k := \left\lceil \frac{n}{2^k} \right\rceil.$$

Then

$$\prod_{k=1}^{m} \prod_{p \leq \sqrt[k]{n}} p \leq \prod_{k=1}^{m} \binom{2a_k}{a_k} < 4^n,$$

where the inner product is over all the primes not exceeding $\sqrt[k]{n}$.

Proof. Clearly, we have $a_1 \geq a_2 \geq ... \geq a_m = 1$. Observe that

$$a_k < \frac{n}{2^k} + 1 = \frac{2n}{2^{k+1}} + 1 \leq 2a_{k+1} + 1,$$
$$a_k \geq \frac{n}{2^k} = 2\left(\frac{n}{2^{k+1}} + 1\right) - 2 > 2a_{k+1} - 2.$$

Thus $2a_{k+1} - 1 \leq a_k \leq 2a_{k+1}$. Since $2a_1 \geq n$, it follows that for any $\alpha \in \mathbb{N}_+$,

$$(1, \sqrt[\alpha]{n}] \subseteq \bigcup_{k=1}^{m} \left(\sqrt[\alpha]{a_{m+1-k}}, \sqrt[\alpha]{2a_{m+1-k}} \right].$$

Now let $q \in \mathbb{P}$ be any prime such that $\sqrt[\alpha+1]{n} < q \leq \sqrt[\alpha]{n}$, where $1 \leq \alpha \leq m$. Then the exponent of q in

$$\prod_{k=1}^{m} \prod_{p \leq \sqrt[k]{n}} p$$

equals α. Since $1 < q \leq \sqrt[k]{n}$ for any positive integer $k \leq \alpha$, there exists a positive integer $s_k \leq m$ such that $\sqrt[k]{a_{s_k}} < q \leq \sqrt[k]{2a_{s_k}}$. Note that $s_k \neq s_l$ when $1 \leq l < k \leq m$, since the inequalities $\sqrt[k]{a_s} < q \leq \sqrt[k]{2a_s}$ and $\sqrt[l]{a_s} < q \leq \sqrt[l]{2a_s}$ would imply

$$q \leq q^{k-l} < \frac{2a_s}{a_s} = 2,$$

which is absurd. From Lemma 2.1 it follows that q^α divides

$$\prod_{k=1}^{\alpha} \binom{2a_{s_k}}{a_{s_k}}.$$

Hence

$$\prod_{k=1}^{m} \prod_{p \leq \sqrt[k]{n}} p \leq \prod_{k=1}^{m} \binom{2a_k}{a_k}.$$

Now we show that

$$\prod_{k=1}^{m} \binom{2a_k}{a_k} < 4^n. \tag{2.1}$$

Simple calculation shows that (2.1) holds for all $n \leq 10$. Suppose that (2.1) holds for all integers $r < n$, where $n \geq 10$. We need to prove that it also holds for $r = n$. Take $r = 2a_2 - 1$. Since $r < 2(n/4+1) - 1 = n/2 + 1 < n$ and $2^{k-1}(a_{k+1} - 1) + 1 \leq a_2 \leq 2^{k-1}a_{k+1}$, we have

$$a_{k+1} - 1 < \frac{2^k(a_{k+1} - 1) + 1}{2^k} \leq \left\lceil \frac{r}{2^k} \right\rceil = \left\lceil \frac{2a_2 - 1}{2^k} \right\rceil < \frac{2^k a_{k+1} - 1}{2^k} + 1 < a_{k+1} + 1$$

for every $1 \leq k \leq m-1$. Hence $\lceil r/2^k \rceil = a_{k+1}$ for every $1 \leq k \leq m-1$. Our inductive hypothesis asserts that (2.1) holds with $r = 2a_2 - 1$. Thus we have

$$\prod_{k=1}^{m} \binom{2a_k}{a_k} < 4^{2a_2 - 1} \binom{2a_1}{a_1}.$$

It is not hard to show by induction that for $u \geq 5$,

$$\binom{2u}{u} = 2\left(2 - \frac{1}{u}\right)\binom{2u-2}{u-1} < 4^{u-1}.$$

Since $2a_1 \leq n+1$ and $2a_2 \leq a_1 + 1$, it follows from the above inequality that

$$\prod_{k=1}^{m} \binom{2a_k}{a_k} < 4^{2a_2 - 1 + a_1 - 1} \leq 4^{2a_1 - 1} \leq 4^n.$$

This completes the inductive proof of (2.1). □

Lemma 2.3. *Let $a, n \in \mathbb{N}_+$ with $a \leq n$, and let $x \in \mathbb{R}_+$. Denote by $N(x; n, a)$ the number of positive integers $m \leq x$ for which $m \equiv a \pmod{n}$. Then*

$$N(x; n, a) = \begin{cases} \lfloor x/n \rfloor & \text{if } a > \lfloor x \rfloor - \lfloor x/n \rfloor n, \\ \lfloor x/n \rfloor + 1 & \text{if } a \leq \lfloor x \rfloor - \lfloor x/n \rfloor n. \end{cases}$$

Proof. Given any $m \in \mathbb{Z}$, a set of n consecutive integers contains a unique element congruent to $m \pmod{n}$. Up to x there are $k := \lfloor x/n \rfloor$ complete sets of residues modulo n and possibly a partial set given by

$$E(x; n) = \big\{kn + 1, kn + 2, ..., kn + \lfloor x \rfloor - kn\big\}.$$

Hence $N(x; n, a) = k$ or $k+1$. It is clear that $N(x; n, a) = k+1$ if and only if $a \equiv m \pmod{n}$ for some $m \in E(x; n)$, or equivalently, if and only if $a \leq \lfloor x \rfloor - kn$. □

Lemma 2.4. *For any positive integer $n \geq 2$,*

$$\prod_{k=1}^{n} \left(\frac{n+1}{k} + 2\right) > 4^n. \tag{2.2}$$

Proof. Note that

$$\left[\prod_{k=1}^{n} \left(\frac{n+1}{k} + 2\right)\right]^2 = \prod_{k=1}^{n} \left(\frac{n+1}{k} + 2\right) \prod_{k=1}^{n} \left(\frac{n+1}{n+1-k} + 2\right) = \prod_{k=1}^{n} \left(4 + \frac{3(n+1)^2}{k(n+1-k)}\right).$$

Since $n \geq 2$ and $k(n+1-k) \leq (n+1)^2/4$ with equality only when $k = (n+1)/2$, we have

$$\left[\prod_{k=1}^{n} \left(\frac{n+1}{k} + 2\right)\right]^2 > \prod_{k=1}^{n} \left(4 + \frac{3(n+1)^2}{(n+1)^2/4}\right) = 4^{2n}.$$

We obtain (2.2) by taking positive square root of both sides of the above inequality. □

3. Harmonic Sums in Arithmetic Progressions

Now we are ready to prove Erdős' theorem on harmonic sums in arithmetic progressions. We will see shortly that it follows easily from the following theorem formulated based on [4].

Theorem 3.1 (Erdős, 1932)**.** *Let $a, d, n \in \mathbb{N}_+$.*

(1) If d is odd, then there exists $\alpha \in \mathbb{N}_+$ such that 2^α divides exactly one of $a, a+d, ..., a+nd$.

(2) If $n \geq a$, then there exists $\alpha \in \mathbb{N}_+$ such that 3^α divides exactly one of $a, a+2, ..., a+2n$.

(3) *If $d = 2$ and $a \geq n+1$, or if $d \geq 4$, then there exist $\alpha \in \mathbb{N}_+$ and $p \in \mathbb{P}$ such that $p^\alpha > n$ divides exactly one of $a, a+d, ..., a+nd$.*

Proof. Without loss of generality, we may assume $\gcd(a, d) = 1$. Let us first prove (1). Let α be the largest integer for which 2^α divides at least one of $a, a+d, ..., a+nd$. Since d is odd, either a or $a+d$ is even. Thus $\alpha \geq 1$. Let $k \leq n$ be the smallest natural number such that $a + kd = 2^\alpha x$, where x is odd. If $k' \leq n$ were the second smallest natural number such that $a + k'd = 2^\alpha y$ with y odd, then $2^{\alpha+1} \mid (k' - k)$. This implies $k' - k \geq 2^{\alpha+1}$. But then 2^α would divide $a + (k+2^\alpha)d$ with $k + 2^\alpha < k'$, which contradicts the definition of k'. Therefore, no such k' can exist. In other words, 2^α divides precisely one of $a, a+d, ..., a+nd$.

The proof of (2) goes along the same lines with slight complications. Since the case $n = a = 1$ is trivial, we may suppose that $n \geq 2$. Observe that a must be odd. As above, let α be the largest integer for which 3^α divides at least one of $a, a+2, ..., a+2n$. Since $n \geq 2$, exactly one of a, $a+2$, and $a+4$ is divisible by 3, which implies $\alpha \geq 1$. Let $k \leq n$ be the smallest natural number for which $3^\alpha \mid (a+2k)$. If $k' \leq n$ were the second smallest natural number for which $3^\alpha \mid (a+2k')$, then $3^\alpha \mid (k'-k)$. According to the definition of k', we would have $k' = k + 3^\alpha$. We may write

$$\begin{aligned} a + 2k \ &= 3^\alpha x, \\ a + 2k' &= 3^\alpha (x+2), \end{aligned}$$

where $x \in \mathbb{N}_+$. The assumption that $\gcd(a, 2) = 1$ implies that x is odd. Moreover, since $3^{\alpha+1} \nmid (a+2k)$ and $3^{\alpha+1} \nmid (a+2k')$, we must have $x \equiv 2 \pmod 3$ and $x \geq 5$. If $k > 3^\alpha$, then $a + 2(k - 3^\alpha) = 3^\alpha(x-2)$ would be divisible by $3^{\alpha+1}$; if $k \leq 3^\alpha$, then $n \geq a \geq 3^\alpha \cdot 5 - 2k = k + 3^\alpha \cdot 5 - 3k \geq k + 3^\alpha \cdot 2$, and thus $a + 2(k + 3^\alpha \cdot 2) = 3^\alpha(x+4)$ would be divisible by $3^{\alpha+1}$. In either case, we would derive a contradiction. Therefore, 3^α divides exactly one of $a, a+2, ..., a+2n$.

It remains to prove (3). If $d = 2$, then a and $a+2$ are both odd and relatively prime to each other, since $\gcd(a, 2) = 1$. Thus every prime factor $p > 2$ of $a+2$ does not divide a. This verifies the case $(d, n) = (2, 1)$. Suppose now that $(d, n) \neq (2, 1)$. Let m be the largest integer for which $2^m \leq n$. If there were no prime power $p^\alpha > n$ dividing at least one of $a+d, a+2d, ..., a+nd$, then for any $p \in \mathbb{P}$ such that $\sqrt[k+1]{n} < p \leq \sqrt[k]{n}$ with $1 \leq k \leq m$, one would have $p^{k+1} \nmid (a + ld)$ for all $1 \leq l \leq n$. By Lemma 2.3, the congruence $a + xd \equiv 0 \pmod{p^s}$ has at most $\lfloor n/p^s \rfloor + 1$ solutions for $1 \leq x \leq n$, where $s \geq 1$. Since $p^{k+1} > n$, the exponent v_p of p in $(a+d)(a+2d)\cdots(a+nd)$ satisfies

$$v_p \leq \sum_{s=1}^{k} \left(\left\lfloor \frac{n}{p^s} \right\rfloor + 1 \right).$$

Since the exponent of p in $n!$ equals $\sum_{s=1}^{k} \lfloor n/p^s \rfloor$, it follows that p would divide the reduced form of

$$\frac{(a+d)(a+2d)\cdots(a+nd)}{n!}$$

to a power no greater than

$$v_p - \sum_{s=1}^{k} \left\lfloor \frac{n}{p^s} \right\rfloor \leq k.$$

Hence we have

$$\frac{(a+d)(a+2d)\cdots(a+nd)}{n!} \leq \prod_{k=1}^{m} \prod_{\sqrt[k+1]{n}<p\leq \sqrt[k]{n}} p^k = \prod_{k=1}^{m} \prod_{p\leq \sqrt[k]{n}} p.$$

If $d=2$, $n\geq 2$, and $a\geq n+1$, then it follows from Lemma 2.4 that

$$\frac{(a+2)(a+4)\cdots(a+2n)}{n!} = \prod_{k=1}^{n}\left(\frac{a}{k}+2\right) \geq \prod_{k=1}^{n}\left(\frac{n+1}{k}+2\right) > 4^n;$$

if $d\geq 4$, then

$$\frac{(a+d)(a+2d)\cdots(a+nd)}{n!} = \prod_{k=1}^{n}\left(\frac{a}{k}+d\right) > d^n \geq 4^n.$$

In both cases, we would have

$$\prod_{k=1}^{m} \prod_{p\leq \sqrt[k]{n}} p > 4^n,$$

which contradicts Lemma 2.2. We have proved that there must exist a prime power $p^\alpha > n$ dividing at least one of $a+d, a+2d, ..., a+nd$. To finish the proof of (3), we need to show that such a prime power necessarily divides only one of $a, a+d, ..., a+nd$. To this end, suppose that $p^\alpha > n$ divides at least one of $a+d, a+2d, ..., a+nd$. Since $\gcd(a,d)=1$, we have $\gcd(d,p)=1$. If $0\leq k,k'\leq n$ were two distinct integers such that $a+kd$ and $a+k'd$ are both divisible by p^α, then $p^\alpha \mid (k-k')$ and thus $|k-k'|\geq p^\alpha$. But this is impossible, since we have $|k-k'|\leq n<p^\alpha$. Therefore, we conclude that p^α divides precisely one of $a, a+d, ..., a+nd$. □

Now it is an easy matter to show that for arbitrary $a,d,n\in\mathbb{N}_+$, the sum $\sum_{k=0}^{n} 1/(a+kd)$ is never an integer. In fact, we have the following more general theorem [4].

Theorem 3.2 (Erdős, 1932)**.** *Let $a,d,\theta\in\mathbb{N}_+$ and $m,n\in\mathbb{N}$ with $n>m$. If $a_m, a_{m+1}, ..., a_n\in\mathbb{Z}$ are integers such that $\gcd(a_k, a+kd)=1$ for all $m\leq k\leq n$, then the sum*

$$\sum_{k=m}^{n}\frac{a_k}{(a+kd)^\theta}$$

is never an integer.

Proof. Let $b:=a+md$. In view of Theorem 3.1, there exist $\alpha\in\mathbb{N}_+$ and $p\in\mathbb{P}$ such that p^α divides exactly one of $b, b+d, ..., b+(n-m)d$. Since $\gcd(a_k, a+kd)=1$ for all $m\leq k\leq n$, we may write

$$\sum_{k=m}^{n}\frac{a_k}{(a+kd)^\theta} = \frac{px+y}{p^{\theta\alpha}z},$$

where $z\in\mathbb{N}_+$ and $x,y\in\mathbb{Z}$ with $\gcd(y,p)=1$. Clearly, the number on the right-hand side cannot be an integer. □

4. The Generalized Harmonic Sum

In this short section, we discuss a simple result of the same nature concerning the generalized harmonic sum $\sum_{k=m}^{n} 1/k^{\theta_k}$, where $\theta_m, \theta_{m+1}, ..., \theta_n \in \mathbb{N}_+$. We shall give a proof of this result similar to that of Theorem 3.2. In particular, we need a substitute for Theorem 3.1. In the proof, we shall use Bertrand's postulate which asserts that for any $n \in \mathbb{N}_+$ there exists a prime $n < p \le 2n$. This result is sometimes referred to as Chebyshev's theorem for the reason that Chebyshev [3] provided the first complete proof of it. For a modern proof of Bertrand's postulate, see [7, Theorem 418]. We shall also need a generalization of Bertrand's postulate due to Sylvester [14], which states that for any $n, k \in \mathbb{N}_+$ with $n \ge 2k$, there exists a prime $p > k$ dividing precisely one of the numbers $n-k+1, n-k+2, ..., n$. Bertrand's postulate is clearly a special case of Sylvester's theorem with $n = 2k$. The interested reader is referred to [5] for an elementary proof of Sylvester's theorem due to Erdős.

Theorem 4.1. *Let $n > m \ge 1$ be positive integers. If $\vartheta_m, \vartheta_{m+1}, ..., \vartheta_n \in \mathbb{N}_+$, and if $a_m, a_{m+1}, ..., a_n \in \mathbb{Z}$ are integers such that $\gcd(a_k, k) = 1$ for all $m \le k \le n$, then*

$$\sum_{k=m}^{n} \frac{a_k}{k^{\vartheta_k}}$$

is never an integer.

Proof. If $m \le \lceil n/2 \rceil$, it follows from Bertrand's postulate that there exists $q \in \mathbb{P}$ with $n/2 < q \le n$. Since $n < 2q$ and $m \le q$, we have $\gcd(k, q) = 1$ for all $m \le k \le n$ except for $k = q$. If $m \ge \lceil n/2 \rceil + 1$, then $n \ge 2(n-m+1)$. By Sylvester's theorem, there exists a prime $q' > n-m+1$ such that q' divides exactly one of $m, m+1, ..., n$. In either case, there exists $p \in \mathbb{P}$ dividing exactly one of $m, m+1, ..., n$, denoted by l. As in the proof of Theorem 3.2, we may write

$$\sum_{k=m}^{n} \frac{a_k}{k^{\vartheta_k}} = \frac{px + y}{p^{\vartheta_l} z},$$

where $z \in \mathbb{N}_+$ and $x, y \in \mathbb{Z}$ with $\gcd(y, p) = 1$. Hence, this sum is never an integer. □

5. Harmonic Sums and the Distribution of Primes

Let $f \in \mathbb{Z}[x]$ be a polynomial of degree at least 1 with integer coefficients such that $0 \notin f(\mathbb{N}_+)$. One may look for sufficient conditions under which it is true that $\sum_{k=1}^{n} a_k/(f(k))^{\theta_k} \notin \mathbb{Z}$, where $a_1, ..., a_n \in \mathbb{Z}$ are integers such that $\gcd(a_k, f(k)) = 1$ for all $1 \le k \le n$ and $\theta_1, ..., \theta_n \in \mathbb{N}_+$ are arbitrary positive integers. Theorems 3.2 and 4.1 assert that $\sum_{k=1}^{n} a_k/f(k) \notin \mathbb{Z}$ if f is the power of a linear polynomial with non-negative integer coefficients. If f is a polynomial of degree at least 2 with non-negative integer coefficients, then it is easy to see that $\sum_{k=1}^{n} 1/(f(k))^{\theta_k} \notin \mathbb{Z}$ for all $n > 1$. Indeed, we have $f(k) \ge k^2$ for all $k \ge 1$, which implies

$$1 < \sum_{k=1}^{n} \frac{1}{(f(k))^{\theta_k}} \le \sum_{k=1}^{n} \frac{1}{k^2} < 1 + \sum_{k=2}^{n} \frac{1}{k(k-1)} = 2 - \frac{1}{n} < 2.$$

In this case we do not need to know anything about the arithmetic properties of the sequence $\{f(k)\}_{k \ge 1}$. The general case is tricky. Note that if $f(x)$ can take negative values, then it is not necessarily true in general that $\sum_{k=1}^{n} 1/f(k) \notin \mathbb{Z}$. For instance, consider the case where $n = 2$ and $f(x) = 2x - 3$. As we saw above, the distribution of primes and prime powers in arithmetic progressions plays an essential role in the proofs of these two theorems. The idea is that if one can show that there exists a prime power p^α dividing precisely one of $f(1), ..., f(n)$, then $\sum_{k=1}^{n} a_k/(f(k))^\theta \notin \mathbb{Z}$, where $\theta \in \mathbb{N}_+$. Similarly, if one can prove that there exists a unique prime $q \in \mathbb{P}$ dividing precisely one of $f(1), ..., f(n)$, then $\sum_{k=1}^{n} a_k/(f(k))^{\theta_k} \notin \mathbb{Z}$.

Formally, let $P(k)$ denote the largest prime factor of k for every $k \in \mathbb{Z} \setminus \{0, \pm 1\}$ and set $P(\pm 1) := 1$. Define

$$P_f(n) := P\left(\prod_{k=1}^{n} f(k)\right).$$

If $f(x) = a + d(x-1)$ is a linear polynomial with $d \neq 0$ and $\gcd(a,d) = 1$, and if $P_f(n) > \max(n-1, |d|)$, then $P_f(n)$ divides precisely one of $f(1), ..., f(n)$, and hence $\sum_{k=1}^{n} a_k/(f(k))^{\theta_k} \notin \mathbb{Z}$. Sylvester [14] proved that if $a, d, n \geq 1$, then $P_f(n) > n$ whenever $a \geq d+n$. Langevin [10] replaced the assumption $a \geq d+n$ by $a > n$. Later Shorey and Tijdeman [13] showed that $P_f(n) > n$ for all $a \geq 1$, $d \geq 2$, and $n \geq 3$ but $(a,d,n) = (2,7,3)$. From this result it follows immediately that $\sum_{k=1}^{n} a_k/(f(k))^{\theta_k} \notin \mathbb{Z}$ whenever $a \geq 1$ and $n > d \geq 2$.

For the general case, Nagell [12] showed that if $f \in \mathbb{Z}[x]$ is not a product of linear factors in $\mathbb{Z}[x]$, then $P_f(n) > c_1 n \log n$ for some constant $c_1 = c_1(f) > 0$. Erdős [6] improved this result by showing that under the same assumption, the inequality

$$P_f(n) > n(\log n)^{c_2 \log\log\log n}$$

holds for some $c_2 = c_2(f) > 0$. As an application, let $f(x) = ax^2 + bx + c \in \mathbb{Z}[x]$ be an irreducible quadratic polynomial with $ab \geq 0$. By Nagell's theorem, there exists $N = N(f) \in \mathbb{N}_+$ such that $P_f(n) > (2n-1)|a| + |b|$ for all $n \geq N$. Note that if $f(k) \equiv f(l) \pmod{p}$ for some $1 \leq k < l \leq n$, where $p \in \mathbb{P}$ is any prime, then $0 < |a(k+l)+b| \leq (2n-1)|a| + |b|$ and $p \mid [a(k+l)+b]$. It follows that $P_f(n)$ divides precisely one of $f(1), ..., f(n)$ whenever $n \geq N$. This implies that $\sum_{k=1}^{n} a_k/(f(k))^{\theta_k} \notin \mathbb{Z}$ for all sufficiently large n. It is clear that lower bounds of higher orders for $P_f(n)$ will allow us to handle the sum $\sum_{k=1}^{n} a_k/(f(k))^{\theta_k}$ when $\deg f > 2$.

There is no doubt that we can say more about general harmonic sums if we are able to gain a better understanding of the distribution of primes in polynomial sequences. Given an irreducible polynomial $f \in \mathbb{Z}[x]$ of degree at least 1 with a positive leading coefficient, one may wonder how frequently f hits primes as n varies. One of the most famous conjectures concerning the the distribution of primes in polynomial sequences is the following conjecture proposed by Paul T. Bateman and Roger A. Horn [2].

Conjecture 5.1 (Bateman-Horn, 1962). *Let $f_1, ..., f_k \in \mathbb{Z}[t]$ be distinct non-constant irreducible polynomials with positive leading coefficients, and let $f := \prod_{i=1}^{k} f_i$. Define*

$$\pi(x; f_1, ..., f_k) := \#\{n \leq x \colon f_i(n) \text{ is a prime for every } 1 \leq i \leq k\}.$$

Suppose that f does not reduce to 0 modulo p for any prime p. Denote by $\omega_p(f)$ the number of roots of f in $\mathbb{Z}/p\mathbb{Z}$. Then

$$\pi(x; f_1, ..., f_k) \sim \frac{C(f)}{\prod_{i=1}^{k} \deg f_i} \int_2^x \frac{dt}{(\log t)^k},$$

where

$$C(f) = \prod_p \left(1 - \frac{1}{p}\right)^{-k} \left(1 - \frac{\omega_p(f)}{p}\right).$$

As a simple example, consider the linear polynomail $f(t) = a + qt$, where $a \in \mathbb{Z}$ and $q \in \mathbb{N}_+$ with $\gcd(a, q) = 1$. Then f is irreducible over $\mathbb{Z}$. Moreover, we have $\omega_p(f) = 1$ if $p \nmid q$ and $\omega_p(f) = 0$ if $p \mid q$. Thus we have

$$C(f) = \prod_{p \mid q} \left(1 - \frac{1}{p}\right)^{-1} = \frac{q}{\varphi(q)},$$

where φ is Euler's totient function. Since

$$\int_2^{(x-a)/q} \frac{dt}{\log t} \sim \frac{(x-a)/q}{\log((x-a)/q)} \sim \frac{1}{q} \cdot \frac{x}{\log x} \sim \frac{1}{q} \int_2^x \frac{dt}{\log t},$$

as $x \to \infty$, the Bateman-Horn conjecture, if true, would imply

$$\pi(x; q, a) := \#\{p \le x \colon p \equiv a \pmod q\} = \pi((x-a)/q; f) \sim \frac{1}{\varphi(q)} \int_2^x \frac{dt}{\log t}.$$

This is the prime number theorem for arithmetic progressions. Similarly, one can show, by taking $f_1(t) = t$ and $f_2(t) = t + 2$, that the Bateman-Horn conjecture, if true, would imply

$$\pi_2(x) := \#\{p \le x \colon p \text{ and } p + 2 \text{ are both primes}\} \sim 2C_2 \int_2^x \frac{dt}{(\log t)^2},$$

where

$$C_2 = \prod_{p > 2} \left(1 - \frac{1}{(p-1)^2}\right).$$

This is the twin prime conjecture of Hardy and Littlewood. It is worth noting that the series which defines the Bateman-Horn constant $C(f)$ always converges and hence $C(f) > 0$. For a proof of this (highly nontrivial) fact and more on the Bateman-Horn conjecture as well as its far-reaching implications in number theory, see [1]. The only known case of the conjecture is the case where $k = 1$ and $\deg f_1 = 1$, which is equivalent to the prime number theorem for arithmetic progressions, as we saw above. We don't even know if the polynomial $t^2 + 1$ takes prime values infinitely often, though Iwaniec [8] showed that there are infinitely many $n \in \mathbb{N}_+$ for which $n^2 + 1$ has at most two prime factors.

Now we discuss the general harmonic sum $\sum_{k=1}^n a_k/(f(k))^{\theta_k} \notin \mathbb{Z}$ assuming the Bateman-Horn conjecture. Let $f(x) = \sum_{r=0}^m a_r x^r \in \mathbb{Z}[x]$ be an irreducible polynomial of degree $m \ge 1$ with $a_m > 0$ and $a_r \ge 0$ for all $1 \le r < m$. The Bateman-Horn conjecture implies that there exists a constant $c_3 = c_3(f) > 0$ such that

$$\pi(n; f) - \pi(n/2; f) > c_3 \frac{n}{\log n}$$

holds for all sufficiently large n. Since $f(x)$ is strictly increasing when n is sufficiently large, it follows that $P_f(n) > f(n/2)$ for all sufficiently large n. Note that if $f(k) \equiv f(l) \pmod p$

for some $1 \leq k < l \leq n$, where $p \in \mathbb{P}$ is any prime, then p divides

$$\sum_{r=1}^{m} a_r \sum_{s=0}^{r-1} k^s l^{r-1-s},$$

which is positive and bounded above by

$$\sum_{r=1}^{m} a_r r n^{r-1}.$$

This implies that

$$p \leq \sum_{r=1}^{m} a_r r n^{r-1}.$$

Of course for sufficiently large n, we have

$$P_f(n) > f(n/2) > \sum_{r=1}^{m} a_r r n^{r-1}.$$

Thus if n is sufficiently large, then $P_f(n)$ divides precisely one of $f(1), ..., f(n)$. We can conclude that if the Bateman-Horn conjecture is true, then $\sum_{k=1}^{n} a_k/(f(k))^{\theta_k} \notin \mathbb{Z}$ for all sufficiently large n.

6. Harmonic Sums in Algebraic Number Fields

Let K be an algebraic number field and denote by $\mathcal{O}_K$ the ring of integers of K. One can define the nth harmonic number $H_n(K)$ of K by

$$H_n(K) := \sum_{\substack{0 \neq I \subseteq \mathcal{O}_K \\ \|I\| \leq n}} \frac{1}{\|I\|},$$

where I ranges through all the non-zero ideals of $\mathcal{O}_K$ with absolute norm $\|I\| := [\mathcal{O}_K : I]$ not exceeding n. Here $H_n(K)$ can be viewed as the nth partial sum of $\zeta_K(1)$, where ζ_K is the Dedekind zeta function of K defined by

$$\zeta_K(s) := \sum_{0 \neq I \subseteq \mathcal{O}_K} \frac{1}{\|I\|^s}$$

for $s \in \mathbb{C}$ with $\Re(s) > 1$. In the case $K = \mathbb{Q}$, every non-zero ideal $I \subseteq \mathcal{O}_{\mathbb{Q}} = \mathbb{Z}$ is a principal ideal of the form (k) for some $k \in \mathbb{N}_+$. Thus we have

$$H_n(\mathbb{Q}) = \sum_{\substack{k \in \mathbb{N}_+ \\ \|(k)\| \leq n}} \frac{1}{\|(k)\|} = \sum_{k=1}^{n} \frac{1}{k} = H_n.$$

This provides a way of generalizing H_n to sums in general number fields. More generally, we may define the generalized harmonic number $H_n(K)$ of K by

$$H_{m,n}(K; \boldsymbol{a}, \boldsymbol{\theta}) := \sum_{\substack{0 \neq I \subseteq \mathcal{O}_K \\ m \leq \|I\| \leq n}} \frac{a_I}{\|I\|^{\theta_I}},$$

where $\boldsymbol{\theta} = \{\theta_I\}_{m\le\|I\|\le n}$ with each $\theta_I \in \mathbb{N}_+$ and $\boldsymbol{a} = \{a_I\}_{m\le\|I\|\le n}$ with each $a_I \in \mathbb{Z}$ satisfying $\gcd(a_I, \|I\|) = 1$. The following example illustrates $H_n(K)$ for $K = \mathbb{Q}(i)$, where $i = \sqrt{-1}$ is the imaginary unit.

Example 6.1. Let $K = \mathbb{Q}(i)$. Then $\mathcal{O}_K = \mathbb{Z}[i]$ is a PID. Note that $\|(\alpha)\| = |N_{K/\mathbb{Q}}(\alpha)|$ for all $\alpha \in K \setminus \{0\}$ and that $(\alpha) = (\beta)$ implies $\alpha = u\beta$ for some unit $u \in \{\pm 1, \pm i\}$. So we have

$$H_n(K) = \frac{1}{4} \sum_{\substack{a,b\in\mathbb{Z}\\ 1\le a^2+b^2\le n}} \frac{1}{a^2+b^2} = \frac{1}{4}\sum_{k=1}^{n} \frac{r_2(k)}{k},$$

where $r_2(k)$ denotes the number of ways of writing k as a sum of two squares of integers, i.e.,

$$r_2(k) := \#\{(a,b) \in \mathbb{Z}^2 \colon a^2 + b^2 = k\}.$$

To see how $H_n(K)$ grows as n increases, we define

$$r(x) := \sum_{k\le x} r_2(k)$$

for $x \ge 1$. A classical result in number theory states that $r(x) = \pi x + O(\sqrt{x})$ [7, Theorem 339]. By partial summation we obtain

$$H_n(K) = \frac{r(n)}{4n} + \frac{1}{4}\int_1^n \frac{r(t)}{t^2}\,dt = \frac{\pi}{4}\log n + \frac{c}{4} + O(n^{-1/2}),$$

where

$$c = \pi + \int_1^\infty \frac{r(t) - \pi t}{t^2}\,dt.$$

It is easy to show that $H_n(K) \notin \mathbb{Z}$ for all $n \ge 2$. To see this, let $m \in \mathbb{N}_+$ be the largest integer for which $2^m \le n$. By [7, Theorem 278], we have $r_2(k) = 4\delta(k)$ with $\delta(k) \in \mathbb{N}$ and $\delta(2^m) = 1$. Writing

$$H_n(K) = \sum_{\substack{k=1\\ \delta(k)\ne 0}}^{n} \frac{\delta(k)}{k},$$

we see that 2^m divides precisely one of the denominators of the terms in the sum on the right side. This implies that $H_n(K) \notin \mathbb{Z}$.

It is well known that $\mathcal{O}_K$ is a Dedekind domain so that every non-zero ideal $I \subseteq \mathcal{O}_K$ has a unique representation as a product of prime ideals of $\mathcal{O}_K$ [11, Theorem 16]. In this setting prime numbers are replaced by prime ideals and the prime number theorem is replaced by the following more general theorem due to Landau [9].

Theorem 6.1 (Prime Ideal Theorem)**.** *Let K be an algebraic number field and denote by $\mathcal{O}_K$ the ring of integers of K. Define*

$$\pi_K(x) := \#\{I \subseteq \mathcal{O}_K \text{ is prime} \colon \|I\| \le x\}.$$

Then

$$\pi_K(x) \sim \mathrm{Li}(x) := \int_2^x \frac{dt}{\log t}.$$

As an application of Theorem 6.1, we show that if $K=\mathbb{Q}(i)$ and $1\le a_I<\|I\|^{\theta_I}/2$ for all prime ideals $I\subseteq\mathcal{O}_K$ with $2<\|I\|\le n$, then $H_{1,n}(K;\boldsymbol{a},\boldsymbol{\theta})\notin\mathbb{Z}$ for all sufficiently large n. As in the preceding section, we see that Theorem 6.1 implies that for all sufficiently large n, there exists a prime ideal $P\subseteq\mathcal{O}_K$ such that $2\le n/2<\|P\|\le n$. By [7, Theorem 252], we have either $P=(\alpha)$ for some $\alpha\in\mathcal{O}_K$ such that $N_{K/\mathbb{Q}}(\alpha)=p\in\mathbb{P}$ is congruent to 1 modulo 4, or $P=(q)$ for some $q\in\mathbb{P}$ congruent to 3 modulo 4. In the former case, we have $n/2<\|P\|=p\le n$. If $p\mid\|I\|$ for some non-zero ideal $I\subseteq\mathcal{O}_K$ with $\|I\|\le n$, then $\|I\|=p$. This implies that I is a prime ideal with $\|I\|=p$. By [7, Theorem 252], this is possible if and only if $I=P$ or $I=\bar{P}$, where $\bar{P}:=(\bar{\alpha})$ with $\bar{\alpha}$ being the complex conjugate of α. So we may write

$$H_{1,n}(K;\boldsymbol{a},\boldsymbol{\theta})=\frac{a_Pp^{\theta_{\bar{P}}}+a_{\bar{P}}p^{\theta_P}}{p^{\theta_P+\theta_{\bar{P}}}}+\sum_{\substack{P,\bar{P}\neq I\subseteq\mathcal{O}_K\\1\le\|I\|\le n}}\frac{a_I}{\|I\|^{\theta_I}}$$

and note that $p\nmid\|I\|$ for all non-zero $I\neq P,\bar{P}$ with $\|I\|\le n$. If $\theta_P\neq\theta_{\bar{P}}$, then the first term on the right side of the above equality is in its reduced form with denominator divisible by p. If $\theta_P=\theta_{\bar{P}}=\theta$, then $2\le a_P+a_{\bar{P}}<p^\theta$ and so $p^\theta\nmid(a_P+a_{\bar{P}})$. It follows that p divides the denominator of the reduced form of $(a_P+a_{\bar{P}})/p^\theta$. Hence $H_{1,n}(K;\boldsymbol{a},\boldsymbol{\theta})\notin\mathbb{Z}$ for all sufficiently large n. Suppose now that $P=(q)$. Then $n/2<\|P\|=q^2\le n$. If $q\mid\|I\|$ for some non-zero ideal $I\subseteq\mathcal{O}_K$ with $\|I\|\le n$, then I has a prime factor Q for which $q\mid\|Q\|$. Since $q\equiv 3\pmod 4$, it follows that $Q=P$ and so $\|P\|=q^2\mid\|I\|$. This implies that $I=P$, since $2q^2>n\ge\|I\|$. Hence q divides precisely one of the members of $\{\|I\|\}_{1\le\|I\|\le n}$. As before, we conclude that $H_{1,n}(K;\boldsymbol{a},\boldsymbol{\theta})\notin\mathbb{Z}$ for all sufficiently large n.

Note that in the above argument, one may resort to the prime number theorem for arithmetic progressions instead of Theorem 6.1. Besides, there is no difficulty adapting the argument to handle the general quadratic field $K=\mathbb{Q}(\sqrt{d})$, where $d\in\mathbb{Z}$ is square-free. Indeed, for sufficiently large n there exists a prime ideal $P\subseteq\mathcal{O}_K$ such that $2\le n/2<\|P\|\le n$. By a well known result [11, Theorem 25] on splitting of primes in quadratic number fields, we know that either $\|P\|=p$ for some odd $p\in\mathbb{P}$ such that d is a square modulo p, or $\|P\|=q^2$ for some odd $q\in\mathbb{P}$ such that d is not a square modulo q. The rest of the argument remains the same. However, new methods may be needed for achieving full generality.

Acknowledgment. The author is grateful to Prof. Carl Pomerance for providing insightful comments and pointing out a typo in the proof of Theorem 3.2.

References

[1] S. L. Aletheia-Zomlefer, L. Fukshansky, and S. R. Garcia, *The Bateman-Horn conjecture: Heuristics, history, and applications*, Expo. Math. **38** (2020), 430–479.

[2] P. T. Bateman and R. A. Horn, *A heuristic asymptotic formula concerning the distribution of prime numbers*, Math. Comp. **16** (1962), 363–367.

[3] P. L. Chebyshev, *Mémoire sur les nombres premiers*, J. Math. Pures. Appl. **17** (1852), 366–390.

[4] P. Erdős, *Egy Kürschák-féle elemi számelméleti tétel általánosítása*, Mat. Fiz. Lapok **39** (1932), 17–24.

[5] P. Erdős, *A theorem of Sylvester and Schur*, J. Lond. Math. Soc. **9** (1934), 282–288.

[6] P. Erdős, *On the greatest prime factor of* $\prod_{k=1}^{x} f(k)$, J. Lond. Math. Soc. **27** (1952), 379–384.

[7] G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, 6th. ed., Oxford Univ. Press, Oxford, 2008. Revised by D. R. Heath-Brown and J. H. Silverman; With a forward by A. J. Wiles.
[8] H. Iwaniec, *Almost-primes represented by quadratic polynomials*, Invent. Math. **47** (1978), 171–188.
[9] E. Landau, *Neuer Beweis des Primzahlsatzes und Beweis des Primidealsatzes*, Math. Annalen **56** (1903), 645–670.
[10] M. Langevin, *Plus grand facteur premier d'entiers en progression arithmétique*, Sém. Delange-Pisot-Poitou, 18e année, 1976/77, Exp. No. 3, 6 pp., Paris, 1977.
[11] D. A. Marcus, *Number Fields*, 2nd. ed., Universitext, Springer Nature, New York, 2018.
[12] T. Nagell, *Généralisation d'un théorème de Tchebycheff*, J. Math. Pures Appl. **4** (8) (1921), 343–356.
[13] T. N. Shorey and R. Tijdeman, *On the greatest prime factor of an arithmetical progression*, A Tribute to Paul Erdős, ed. A. Baker, B. Bollobas and A. Hajnal, Cambridge Univ. Press (1990), 385–389.
[14] J. J. Sylvester, *On arithmetical series*, Messenger Math. **21** (1892), 1–19, 87–120.
[15] E. W. Weisstein, *Concise Encyclopedia of Mathematics*, 2nd. ed., Chapman & Hall/CRC, 2003.

Department of Mathematics, Dartmouth College, Hanover, NH 03755, USA
Email address: steve.fan.gr@dartmouth.edu