# THE $p$-ADIC GROWTH OF HARMONIC SUMS

KEITH CONRAD

The numbers

$$H_n = 1 + \frac{1}{2} + \cdots + \frac{1}{n} = \sum_{k=1}^{n} \frac{1}{k}$$

are called *harmonic sums*. Integrals can approximate these sums as real numbers, *e.g.*, $H_n$ and $\int_1^n \mathrm{d}t/t = \log n$ differ by a bounded amount ($H_n - \int_1^n \mathrm{d}t/t$ is positive and montonically decreasing), so the $H_n$'s slowly diverge in $\mathbf{R}$. We are interested here in the arithmetic behavior of harmonic sums, *i.e.*, their $p$-adic growth. This is definitely not an important topic, but it is interesting to compare what happens with this rational sequence across the real and different $p$-adic landscapes.

Here are some initial values:

| $n$ | $H_n$ |
|---|---|
| 1 | 1 |
| 2 | 3/2 |
| 3 | 11/6 |
| 4 | 25/12 |
| 5 | 137/60 |
| 6 | 49/20 |
| 7 | 363/140 |
| 8 | 761/280 |
| 9 | 7129/2520 |
| 10 | 7381/2520 |
| 11 | 83711/27720 |

TABLE 1. Harmonic Sums

For $n > 1$, it appears that $H_n \notin \mathbf{Z}$. More precisely, the numerator of $H_n$ is odd and the denominator of $H_n$ is even. This suggests a strategy for proving $H_n$ is not an integer.

**Theorem 1** (Taeisinger, 1915). *For any $n > 1$, $H_n \notin \mathbf{Z}$.*

*Proof.* Let $L$ be the least common multiple of $1, 2, \ldots, n$, so $H_n$ can be written as a fraction with denominator $L$. For $1 \le k \le n$, write $L = ka_k$ with $a_k \in \mathbf{Z}^+$, so $1/k = a_k/L$. Then

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

Since $n \ge 2$, $L$ is even. We will show $\sum_{k=1}^{n} a_k$ is odd, so the ratio is not integral.

Set $2^r$ to be the largest power of 2 up to $n$: $2^r \le n < 2^{r+1}$. The only integer up to $n$ divisible by $2^r$ is $2^r$ itself, since $2 \cdot 2^r > n$. Therefore $L = 2^r b$ where $b$ is odd, so $2^r b = ka_k$ for $1 \le k \le n$. When $k = 2^r$ we see that $a_k = b$ is odd. When $k \ne 2^r$, $k$ is not divisible by $2^r$, so $a_k$ must be even. Therefore in the numerator $\sum_{k=1}^{n} a_k$, one term (at $k = 2^r$) is odd and the rest are even, so the total sum is odd. $\square$

1

Theorem 1 is actually a 2-adic result: $H_n \notin \mathbf{Z}$ because $H_n \notin \mathbf{Z}_2$. Here is a second proof of Theorem 1 brings out the 2-adic features more clearly.

*Proof.* Let $2^r \leq n < 2^{r+1}$, so $r \geq 1$ and the highest power of 2 that appears in any reciprocal in the sum defining $H_n$ is $2^r$. The only reciprocal in $H_n$ with denominator divisible by $2^r$ is $1/2^r$. Indeed, any other such reciprocal would be $1/(2^r c)$ for odd $c > 1$, but if that term is in the sum then so is $1/(2^r \cdot 2) = 1/2^{r+1}$, which is false since $2^{r+1} > n$. Therefore $1/2^r$ has a more highly negative 2-adic valuation than any other term in $H_n$, so it is not cancelled out in the sum. This means $H_n \notin \mathbf{Z}_2$, so $H_n \notin \mathbf{Z}$. $\qquad\square$

This proof gives a formula for the 2-adic valuation of the harmonic sums: $\mathrm{ord}_2(H_n) = -r$ where $2^r \leq n < 2^{r+1}$. To see this formula in action, we rewrite the initial harmonic sums in Table 2 with the power of 2 in the denominators made explicit. The exponent jumps when $n$ is a power of 2.

| $n$ | $H_n$ |
|---|---|
| 1 | 1 |
| 2 | $3/2$ |
| 3 | $11/2 \cdot 3$ |
| 4 | $25/2^2 \cdot 3$ |
| 5 | $137/2^2 \cdot 15$ |
| 6 | $49/2^2 \cdot 5$ |
| 7 | $363/2^2 \cdot 35$ |
| 8 | $761/2^3 \cdot 35$ |
| 9 | $7129/2^3 \cdot 315$ |
| 10 | $7381/2^3 \cdot 2520$ |
| 11 | $83711/2^3 \cdot 3465$ |

TABLE 2. Harmonic Sums, 2-adically

Since $|H_n|_2 = 2^r$ where $2^r \leq n < 2^{r+1}$, we see that $|H_n|_2 \to \infty$ as $n \to \infty$ (explicitly, $n/2 < |H_n|_2 \leq n$). Does $|H_n|_p \to \infty$ for odd primes $p$? We will address this question later.

**Theorem 2** (Kürschák, 1918). *For $m < n$, $H_n - H_m \notin \mathbf{Z}_2$. In particular, $H_n - H_m \notin \mathbf{Z}$.*

Taking $m = 1$ recovers Theorem 1, since $H_1 \in \mathbf{Z}$.

*Proof.* Writing
$$H_n - H_m = \sum_{k=m}^{n-1} \frac{1}{k},$$
we will show there is a unique term in the sum with the most negative 2-adic valuation. Let $r = \max_{m \leq k < n} \mathrm{ord}_2(k)$. Since $n > m$, the sum $H_n - H_m$ has at least two terms in it, so one of the $k$'s is even and therefore $r \geq 1$.

We will show there is only one integer from $m$ to $n-1$ with 2-adic valuation $r$. Suppose there are two such numbers. Write them as $2^r c$ and $2^r d$ with odd $c < d$. Then $c + 1$ is even and $2^r c < 2^r (c+1) < 2^r d$, so $1/2^r(c+1)$ appears in $H_n - H_m$. But $\mathrm{ord}_2(2^r(c+1)) \geq r+1$ since $c$ is odd. This contradicts the definition of $r$. Therefore there is only one term in $H_n - H_m$ with 2-adic valuation $-r$, so $\mathrm{ord}_2(H_n - H_m) = -r$. $\qquad\square$

We now turn to the $p$-adic behavior of $H_n$ for $p \neq 2$. The material below is based on [3].

The harmonic sums $H_1, H_2, \ldots, H_{p-1}$ are all $p$-adically integral, since the denominators appearing in them are prime to $p$ and thus are $p$-adic integers (even $p$-adic units). Since $1/p$

is outside the ring $\mathbf{Z}_p$, $H_p = H_{p-1} + 1/p$ is outside $\mathbf{Z}_p$: $\mathrm{ord}_p(H_p) = -1$. It need not be true (as it is for $p = 2$) that $H_n \notin \mathbf{Z}_p$ for all $n \geq p$. For example, $H_3, H_4$ and $H_5$ are outside $\mathbf{Z}_3$ (all with 3-adic valuation $-1$), but $H_6, H_7$, and $H_8$ are in $\mathbf{Z}_3$. Then $H_9, H_{10}, \ldots, H_{20}$ are outside of $\mathbf{Z}_3$. The harmonic sums return to $\mathbf{Z}_3$ for $H_{21}, H_{22}$, and $H_{23}$, then leave $\mathbf{Z}_3$ and come back in again for $H_{66}, H_{67}$, and $H_{68}$. Then the harmonic sums leave $\mathbf{Z}_3$ and never return. In fact, $|H_n|_3 \to \infty$ as $n \to \infty$. We will not discuss the proof of this, but the next theorem gives a simple connection between $p$-adic integrality and $p$-adic divergence of the harmonic sums.

**Theorem 3.** *For a prime $p$, the following conditions are equivalent:*

(1) $\{n \geq 1 : H_n \in \mathbf{Z}_p\}$ *is finite,*
(2) $|H_n|_p \to \infty$ *as $n \to \infty$.*

*Proof.* Trivially the second condition implies the first. To prove the (more interesting) reverse implication, pick $n \geq 1$ and write $n = pq + r$ with $0 \leq r \leq p - 1$. Since $1/k \in \mathbf{Z}_p$ when $(p, k) = 1$, in the quotient group $\mathbf{Q}_p/\mathbf{Z}_p$ a harmonic sum is equal to the sum of its terms whose denominators are divisible by $p$, so

$$H_n \equiv \frac{1}{p} + \frac{1}{2p} + \cdots + \frac{1}{pq} = \frac{1}{p}H_q,$$

so $H_n - (1/p)H_q \in \mathbf{Z}_p$.

We are assuming that $|H_n|_p \leq 1$ only finitely many times, so there is some $N_0 \geq 1$ such that $|H_n|_p > 1$ for $n \geq N_0$. We will show this $N_0$ controls the $p$-adic divergence of the harmonic sums: for any $k \geq 0$,

$$n \geq p^k N_0 \implies |H_n|_p > p^k.$$

For $k = 0$ this is true by the definition of $N_0$. Assuming it is true for $k$, suppose $n \geq p^{k+1}N_0$. Writing $n = pq + r$ with $0 \leq r \leq p - 1$, from $pq + r \geq p^{k+1}N_0$ we must have $q \geq p^k N_0$, so $|H_q|_p > p^k$. Therefore $|(1/p)H_q|_p > p^{k+1}$. Since $H_n - (1/p)H_q \in \mathbf{Z}_p$,

$$\left| H_n - \frac{1}{p}H_q \right|_p \leq 1 < p^{k+1} = \left| \frac{1}{p}H_q \right|_p,$$

so by the nonarchimedean property $|H_n|_p = |(1/p)H_q|_p > p^{k+1}$. $\qquad\square$

The same reasoning shows $|H_n|_p \to \infty$ as $n \to \infty$ if and only if the set $\{n \geq 1 : H_n \in p\mathbf{Z}_p\}$ is finite. This set, introduced in [2], is denoted

$$J(p) = \{n \geq 1 : H_n \in p\mathbf{Z}_p\}.$$

To say $n \notin J(p)$ means $H_n$ has no $p$ in its numerator. For instance, Theorem 1 says $J(2) = \emptyset$. In [2] it is proved that $J(3) = \{2, 7, 22\}$, $J(5) = \{4, 20, 24\}$, and

$$J(7) = \{6, 42, 48, 295, 299, 337, 341, 2096, 2390, 14675, 16731, 16735, 102728\}.$$

The set $J(11)$ has 638 elements [1]. It is conjectured that $J(p)$ is finite for all $p$, which by Theorem 3 is equivalent to saying $|H_n|_p \to \infty$ as $n \to \infty$ for all $p$.

The rest of our discussion will focus on the particular sum $H_{p-1}$.

**Theorem 4** (Babbage, 1819). *For any odd prime $p$, $H_{p-1} \in p\mathbf{Z}_p$. Therefore $J(p) \neq \emptyset$ for $p > 2$.*

*Proof.* In the sum

$$H_{p-1} = 1 + \frac{1}{2} + \cdots + \frac{1}{p-1},$$

each term is a unit in $\mathbf{Z}_p$, so we can reduce modulo $p$ and check that $H_p \equiv 0 \bmod p\mathbf{Z}_p$. The integers $\{1, 2, \ldots, p-1\}$ represent the units modulo $p$, so their inverses do as well. Thus, after replacing each $1/k$ with the integer from 1 to $p-1$ which is equal to it in $\mathbf{Z}_p/(p)$,

$$
\begin{aligned}
H_{p-1} &\equiv 1 + 2 + \cdots + p - 1 \bmod p \\
&\equiv \frac{p(p-1)}{2} \bmod p \\
&\equiv 0 \bmod p,
\end{aligned}
$$

where in the last step we needed $p \neq 2$.                                                  $\square$

In Table 3, we factor the initial harmonic sums into primes, and can see in this data (for $p = 5, 7$, and 11) that not only is $H_{p-1}$ divisible by $p$, but by $p^2$.

| $n$ | $H_n$ |
|-----|-------|
| 1 | 1 |
| 2 | $3/2$ |
| 3 | $11/2 \cdot 3$ |
| 4 | $5^2/2^2 \cdot 3$ |
| 5 | $137/2^2 \cdot 3 \cdot 5$ |
| 6 | $7^2/2^2 \cdot 5$ |
| 7 | $3 \cdot 11^2/2^2 \cdot 5 \cdot 7$ |
| 8 | $761/2^3 \cdot 5 \cdot 7$ |
| 9 | $7129/2^3 \cdot 3^2 \cdot 5 \cdot 7$ |
| 10 | $11^2 \cdot 61/2^3 \cdot 3^2 \cdot 5 \cdot 7$ |
| 11 | $97 \cdot 863/2^3 \cdot 3^2 \cdot 5 \cdot 7 \cdot 11$ |

TABLE 3. Harmonic Sums factored

**Theorem 5** (Wolstenholme, 1862). *For any prime $p \geq 5$, $H_{p-1} \in p^2\mathbf{Z}_p$.*

*Proof.* We collect terms in $H_{p-1}$ which are equidistant from the middle of the sum:

$$
\begin{aligned}
H_{p-1} &= 1 + \frac{1}{2} + \cdots + \frac{1}{p-1} \\
&= \left(1 + \frac{1}{p-1}\right) + \left(\frac{1}{2} + \frac{1}{p-2}\right) + \cdots + \left(\frac{1}{(p-1)/2} + \frac{1}{(p+1)/2}\right) \\
&= \sum_{k=1}^{(p-1)/2} \left(\frac{1}{k} + \frac{1}{p-k}\right) \\
&= \sum_{k=1}^{(p-1)/2} \frac{p}{k(p-k)} \\
&= p \sum_{k=1}^{(p-1)/2} \frac{1}{k(p-k)}.
\end{aligned}
$$

Since a $p$ has been pulled out, we want to show this last sum is in $p\mathbf{Z}_p$. The terms in the sum are $p$-adic units, and reducing the terms modulo $p$ yields

$$
\sum_{k=1}^{(p-1)/2} \frac{1}{k(p-k)} \equiv - \sum_{k=1}^{(p-1)/2} \frac{1}{k^2} \bmod p\mathbf{Z}_p.
$$

The numbers $1^2, \ldots, ((p-1)/2)^2$ represent all the nonzero squares modulo $p$, so their reciprocals also represent the nonzero squares modulo $p$. Therefore

$$\sum_{k=1}^{(p-1)/2} \frac{1}{k^2} \equiv \sum_{k=1}^{(p-1)/2} k^2 \bmod p\mathbf{Z}_p.$$

Using the formula $\sum_{k=1}^n k^2 = n(n+1)(2n+1)/6$ with $n = (p-1)/2$,

$$\sum_{k=1}^{(p-1)/2} k^2 = \frac{p(p^2-1)}{24}.$$

Since $p > 3$, $(p, 24) = 1$, so this sum is in $p\mathbf{Z}_p$ and we're done. $\qquad \square$

Theorem 5 says $\operatorname{ord}_p(H_{p-1}) \geq 2$. If one calculates $H_{p-1}$ for primes $3 < p < 10000$, always $\operatorname{ord}_p(H_{p-1}) = 2$. But at $p = 16843$ and at $p = 2124679$, $\operatorname{ord}_p(H_{p-1}) = 3$. No further examples where $\operatorname{ord}_p(H_{p-1}) > 2$ are known.

## REFERENCES

[1] D. W. Boyd, *A p-adic study of the partial sums of the harmonic series*, Experiment. Math **3** (1994), 287–302.
[2] A. Eswarathasan and E. Levine, *p-integral harmonic sums*, Discrete Math. **91** (1991), 249–257.
[3] N. Rogers, "Harmonics and Primes," `http://math.arizona.edu/~nfrogers/talks/harmonic.pdf`.