Subsums of the Harmonic Series

Brian Lubeck and Vadim Ponomarenko

Abstract. We consider subsums of the harmonic series, and determine conditions for their convergence. We apply these conditions to determine convergence for a family of series that generalizes Kempner's series.

The harmonic series, of course, diverges; however, if we remove enough terms, what remains will converge. This problem has a long history with this MONTHLY. In 1914 (see [**7**]), Kempner proved that removing those terms whose denominators contain a digit 9 anywhere, the resulting series converges. In 1916 (see [**6**]), Irwin extended this result. He proved that if we choose any natural number N, and remove those terms whose denominators contain more than N 9's, the resulting series still converges. Since that time various authors (see [**3, 4, 5, 8, 9**]) have looked at variations of this idea, determining convergence of similar subsums of the harmonic series, and calculating or estimating the sums when convergent.

While Kempner allowed no 9's at all, and Irwin allowed only a miserly N 9's, we propose to generously allow arbitrarily many 9's, so long as the proportion of 9's remains below a fixed parameter λ. We will prove that the series will converge if and only if $\lambda < \frac{1}{10}$. Consequently, our result subsumes most of theirs. Our proofs use ideas from Riemann–Stieltjes integration, statistics, and combinatorics.

1. NATURAL DENSITY. We fix $A \subseteq \mathbb{N}$, a subset of the naturals. We assume A is infinite. We order this set sequentially, naming its elements as $A = \{a_k\}_{k\geq 1}$, where $a_k < a_{k+1}$ for all $k \in \mathbb{N}$. We call A *reciprocally convergent* or *r. convergent*, if $\sum_{k\geq 1} \frac{1}{a_k}$ converges. We call A *r. divergent* otherwise. For $x \in \mathbb{R}$, we define $A(x) = |\{a \in A : a \leq x\}|$, the number of elements of A less than or equal to x. A commonly used measure of the "size" of A is its *asymptotic density*, defined as

$$d(A) = \lim_{x\to\infty} \frac{A(x)}{x}.$$

For example, if A is the set of positive even (or odd) numbers, then $A(x)$ is within a fixed constant of $\frac{x}{2}$; hence $d(A) = \frac{1}{2}$. Now, $d(A)$ need not exist, since the limit will not exist for certain sets A. For example, suppose that A contains just those m-digit numbers where m is even. Now, for $k \in \mathbb{N}$, $\frac{A(10^{2k+1})}{10^{2k+1}} \geq \frac{9\cdot 10^{2k}}{10^{2k+1}} = 0.9$ but $\frac{A(10^{2k})}{10^{2k}} \leq \frac{10^{2k-1}+1}{10^{2k}}$, which is barely over 0.1. However, $d(A)$ will exist in the following important case.

Theorem 1. *If A is r. convergent, then $d(A)$ exists and equals 0.*

Proof. Let $\epsilon > 0$. Since A is r. convergent, there is some $N > 0$ such that $\sum_{a_k\geq N} \frac{1}{a_k} < \frac{\epsilon}{2}$. Set $A' = A \cap [N, \infty)$. We will now prove that $\frac{A'(x)}{x} < \frac{\epsilon}{2}$ for all $x > 0$. Suppose otherwise; then there is some $x > 0$ such that $A'(x) \geq \frac{\epsilon}{2}x$. But then $\sum_{a_k\geq N} \frac{1}{a_k} \geq \sum_{N\leq a_k\leq x} \frac{1}{a_k} \geq A'(x)\frac{1}{x} \geq \frac{\epsilon}{2}$, a contradiction. We now compute $\frac{A(x)}{x} \leq \frac{N+A'(x)}{x} < \frac{N}{x} + \frac{\epsilon}{2}$. For all $x > \frac{2N}{\epsilon}$, we have $\frac{N}{x} < \frac{\epsilon}{2}$ and so $0 \leq \frac{A(x)}{x} < \epsilon$. ∎

The converse of Theorem 1 does not hold. The most famous example is due to Euler, who proved that the set of primes $\mathbb{P}$ is r. divergent. We can see that $d(\mathbb{P}) = 0$ by the prime number theorem.

We turn now to Riemann–Stieltjes integration. This is a generalization of familiar Riemann integration that can also be used for some discrete functions. For a lovely introduction to the subject, see Chapter 7 of [**2**]; however we will only need the following specific result. Its statement uses only a standard Riemann integral. It has a short proof using Riemann–Stieltjes integration; instead we include a slightly longer, elementary, proof. Similarly, our later results on Kempner-type series were originally done with Riemann-Stieltjes integration, but have been rewritten to use alternative methods.

Theorem 2. *[2] Let $\{b_k\}$ be a sequence of real numbers, and $B(x) = \sum_{k\le x} b_k$. Let $t \ge 1$, and let $f(x)$ have a continuous derivative in the interval $[1, t]$. Then*

$$\sum_{k\le t} b_k f(k) = -\int_1^t B(x)f'(x)dx + B(t)f(t).$$

Proof. For any natural number k, and every $x \in [k, k+1)$, note that $B(x) = B(k)$. Now, fix $t \ge 1$ and set n to be the integer satisfying $n \le t < n+1$. We break up the integral and apply the the fundamental theorem of calculus, finding that

$$\begin{aligned}
\int_1^t B(x)f'(x)dx &= \int_n^t B(x)f'(x)dx + \sum_{k=1}^{n-1}\int_k^{k+1} B(x)f'(x)dx \\
&= \int_n^t B(n)f'(x)dx + \sum_{k=1}^{n-1}\int_k^{k+1} B(k)f'(x)dt \\
&= B(n)(f(t) - f(n)) + \sum_{k=1}^{n-1} B(k)(f(k+1) - f(k)) \\
&= B(n)f(t) - \sum_{k=1}^{n} B(k)f(k) + \sum_{k=2}^{n} B(k-1)f(k) \\
&= B(n)f(t) - B(1)f(1) - \sum_{k=2}^{n}(B(k) - B(k-1))f(k) \\
&= B(n)f(t) - \sum_{k=1}^{n} b_k f(k) \\
&= B(t)f(t) - \sum_{k\le t} b_k f(k).
\end{aligned}$$

∎

Theorem 2 allows us to convert sums into integrals, which we do in the following.

Theorem 3. *With notation as above, $\sum_{a_k\le t}\frac{1}{a_k} = \int_1^t \frac{A(x)}{x^2}dx + \frac{A(t)}{t}$.*

Proof. Given our sequence $A = \{a_k\}_{k\ge1}$, we compute the closely related zero-one sequence $B = \{b_k\}_{k\ge1}$, defined as $b_k = \begin{cases}1 & k \in A\\ 0 & k \notin A\end{cases}$. For any real $x \ge 0$, we define $B(x) = \sum_{k \le x} b_k$, and note that $B(x) = A(x)$. We set $f(x) = \frac{1}{x}$, observe that $\sum_{k \le t} b_k f(k) = \sum_{a_k \le t} \frac{1}{a_k}$, and apply Theorem 2. ■

We can now present our convergence characterization for harmonic subsums. Theorems 1 and 4 determine r. convergence for all A.

Theorem 4. *Suppose that* $d(A) = 0$. *Then* $\sum_{k \ge 1} \frac{1}{a_k} = \int_1^\infty \frac{A(x)}{x^2} dx$. *In particular,* A *is r. convergent if and only if* $\int_1^\infty \frac{A(x)}{x^2} dx$ *converges.*

Proof. Take $t \to \infty$ in Theorem 3. ■

This Riemann–Stieltjes method is not new, but deserves to be better known. For example, in 1915 Brun considered the set of twin primes $\mathbb{P}_2$, and proved that $\mathbb{P}_2(x) = O(\frac{x(\log \log x)^2}{(\log x)^2})$; then by ideas similar to those in Theorem 4 he concluded that $\mathbb{P}_2$ was r. convergent.

Although Theorem 4 gives an exact answer, typical sets A give a step function $A(x)$ whose steps are dispersed in some complicated way that makes it too difficult to integrate exactly. However, we can still get good bounds on the sum with the following, by computing a partial sum of the first s terms, and estimating the error by bounding the remaining integral.

Theorem 5. *Suppose that* A *is r. convergent. Let* $s \in \mathbb{N}$, *and set* $t = a_s$. *Then*

$$\sum_{k \ge 1} \frac{1}{a_k} = \sum_{a_k \le t} \frac{1}{a_k} + \int_t^\infty \frac{A(x)}{x^2} dx - \frac{s}{t}.$$

Proof. Let $T > t$. Apply Theorem 3 with t and also T to get $\sum_{a_k \le t} \frac{1}{a_k} = \int_1^t \frac{A(x)}{x^2} dx + \frac{A(t)}{t}$ and $\sum_{a_k \le T} \frac{1}{a_k} = \int_1^T \frac{A(x)}{x^2} dx + \frac{A(T)}{T}$. Subtract, and let $T \to \infty$. ■

To illustrate, consider the series $\sum_{k \ge 1} \frac{1}{k^2}$, which corresponds to $a_k = k^2$. This is known as the Basel problem. Euler determined its sum to be $\frac{\pi^2}{6} \approx 1.6449$, but we will estimate it using Theorem 5. We compute the first ten terms, i.e., $s = 10$ with $t = 100$, and get $\sum_{a_k \le 100} \frac{1}{a_k} = \frac{1968329}{1270080} \approx 1.5498$. Hence the full series has sum $\frac{1968329}{1270080} - \frac{10}{100} + \int_{100}^\infty \frac{A(x)}{x^2} dx$. The function $A(x)$ is a step function (whose steps have height 1) that satisfies $\sqrt{x} - 1 \le A(x) \le \sqrt{x}$. Hence

$$\int_{100}^\infty \frac{\sqrt{x} - 1}{x^2} dx \le \int_{100}^\infty \frac{A(x)}{x^2} dx \le \int_{100}^\infty \frac{\sqrt{x}}{x^2} dx.$$

The bounding integrals are easy to compute, so we find $0.19 \le \int_{100}^\infty \frac{A(x)}{x^2} dx \le 0.2$. We conclude that $\sum_{k \ge 1} \frac{1}{k^2}$ is in the range $[1.6398, 1.6498]$.

2. KEMPNER-TYPE SERIES. We are now ready to generalize Kempner's series. Fix $\lambda \in [0, 1]$, and define

$$A^\lambda = \{n \in \mathbb{N} : (\#9'\text{s in } n) \le \lambda(\#\text{ digits in } n)\}.$$

Special cases include A^0, Kempner's original series, and A^1, the harmonic series.

Fix $m \in \mathbb{N}$, and consider numbers with m digits. Intuitively, we select one at random. The quantity of 9's we get is modeled by a binomial distribution, with m experiments and probability of success (i.e., getting a 9), of $p = \frac{1}{10}$. Chernoff's bounds tell us that the number of successes will usually cluster near the expected number of successes, which is $pm = \frac{m}{10}$. That is, we expect that a "typical" m-digit number will have $\frac{m}{10}$ 9's. Thus, if $\lambda \geq \frac{1}{10}$, then A^λ will include "typical" numbers, and we might expect A^λ to be not too different from A^1, and thus r. divergent. On the other hand, if $\lambda < \frac{1}{10}$, then we might expect the opposite. This intuition will turn out to be correct.

We now recall a version of Chernoff's bound (see, e.g., [**1**]) on the lower tail of a binomial (or Poisson) distribution. A similar bound holds for the upper tail.

Theorem 6 (Chernoff). *Let X denote the sum of n independent 0-1 random variables, and let μ denote the expected value of X. Let δ be fixed in the interval $(0, 1)$. Then*

$$Pr(X \leq (1 - \delta)\mu) \leq e^{-\frac{\delta^2 \mu}{2}}.$$

Theorem 7. *With notation as above, if $\lambda < \frac{1}{10}$, then A^λ is r. convergent.*

Proof. We first estimate $|A^\lambda \cap [10^{m-1}, 10^m)|$, the number of elements of A with exactly m digits. Consider choosing each of the m digits uniformly at random, and letting X denote the total number of 9's among the m digits chosen. If $X > \lambda m$, then that yields a number outside of A^λ, as it has too many 9's. If $X \leq \lambda m$, then that will yield a number with an allowable quantity of 9's, but it might still be outside of A^λ if the first digit selected happens to be 0. Hence $|A^\lambda \cap [10^{m-1}, 10^m)| \leq Pr(X \leq \lambda m)(10^m - 10^{m-1})$. Now, set $\delta = 1 - 10\lambda$. Note that $\lambda m = (1 - \delta)\frac{m}{10} = (1 - \delta)\mu$. By Chernoff's bound, we have

$$|A^\lambda \cap [10^{m-1}, 10^m)| \leq s^{-m}(10^m - 10^{m-1}),$$

for $s = e^{\frac{(1-10\lambda)^2}{20}}$. Since $\lambda < \frac{1}{10}$, we must have $s > 1$. We now calculate

$$\begin{aligned}\sum_{a \in A^\lambda} \frac{1}{a} &= \sum_{m \geq 1} \sum_{a \in A^\lambda \cap [10^{m-1}, 10^m)} \frac{1}{a} \leq \sum_{m \geq 1} |A^\lambda \cap [10^{m-1}, 10^m)| \frac{1}{10^{m-1}} \\ &\leq \sum_{m \geq 1} s^{-m}(10^m - 10^{m-1})10^{1-m} = 9 \sum_{m \geq 1} s^{-m} = \frac{9}{s - 1} < \infty.\end{aligned}$$

∎

For $\lambda < \frac{1}{10}$, Theorem 7 gives an upper bound of $\frac{9}{s-1}$ for the series sum; however this isn't particularly tight. With some care, and computation of partial sums, this could be much improved. We leave the search for such tighter bounds to others, and turn to the case of $\lambda \geq \frac{1}{10}$. We will use different tools to prove r. divergence.

Theorem 8. *With notation as above, if $\lambda \geq \frac{1}{10}$, then A^λ is r. divergent.*

Proof. Let $m > 1$, and set T_m to be the set of all m-digit numbers whose leading digit is not 9. For each $x \in T_m$, we apply the cyclic digit permutation $0 \to 1 \to \cdots \to 8 \to 9 \to 0$, simultaneously to every digit but the leading one. For example, $30772 \to 31883 \to 32994 \to 33005 \to \cdots \to 39661 \to 30772$. There are ten numbers in this cycle, all distinct. Call these ten elements of T_m equivalent. Fix

$x \in T_m$, and count how many times each of $0, 1, \dots, 9$ appears in x, ignoring the leading digit. For example, 31883 contains one 1, two 8's, and one 3. By a version of the pigeonhole principle, if we choose $m-1$ times from ten options, then some option must have been selected at most $\lfloor \frac{m-1}{10} \rfloor$ times. Hence, among the options $0, 1, \dots, 9$, there must be some option that appears in x at most $\lfloor \frac{m-1}{10} \rfloor \leq \frac{m}{10} \leq \lambda m$ times. So there is some digit in x, not necessarily 9, appearing at most λm times. Hence there is some $y \in T_m$, equivalent to x, where the digit 9 appears at most λm times. Therefore, this $y \in A^\lambda$. Each equivalence class in T_m must therefore contain at least one member of A^λ. Note that the leading $m-1$ digits (all but the last) of any $x \in T_m$ determine its equivalence class. Therefore, T_m contains $8 \cdot 10^{m-2}$ equivalence classes, and so at least $8 \cdot 10^{m-2}$ elements of A^λ. Hence

$$A^\lambda(10^m - 1) \geq 8 \cdot 10^{m-2} + 8 \cdot 10^{m-3} + \cdots + 8 \cdot 10^0 = \frac{8}{9}(10^{m-1} - 1)$$

and $\frac{A^\lambda(10^m-1)}{10^m-1} \geq \frac{8}{9}\frac{10^{m-1}-1}{10^m-1}$. This lower bound for $\frac{A^\lambda(x)}{x}$, for an infinite family of values of x, is very close to $\frac{8}{90}$. Hence either $d(A^\lambda) \geq \frac{8}{90}$ or $d(A^\lambda)$ doesn't exist. In either case, $d(A^\lambda) \neq 0$, so by Theorem 1, A is r. divergent. ∎

Similar methods to those of Theorems 7 and 8 will work for other digits, and in bases other than base 10. We invite the reader to generalize to longer substrings, like 89, and determine which λ will lead to convergent subsums. We also invite the reader to apply these methods to other subsums of the harmonic, for example keeping no more than 9.99^m of the integers with m digits.

ACKNOWLEDGMENT. The authors would like to thank the anonymous referees for many helpful suggestions, and for the elementary proof of Theorem 2.

REFERENCES

1. D. Angluin, L. G. Valiant, Fast probabilistic algorithms for Hamiltonian circuits and matchings, *J. Comput. System Sci.* **18** no. 2 (1979) 155–193.
2. T. M. Apostol, *Mathematical Analysis*. Addison-Wesley Publishing Co., Reading, Mass.–London–Don Mills, Ont., second edition, 1974.
3. R. Baillie, Sums of reciprocals of integers missing a given digit, *Amer. Math. Monthly* **86** (1979) 372–374.
4. G. H. Behforooz, Thinning out the harmonic series, *Math. Mag.* **68** no. 4 (1995) 289–293.
5. B. Farhi, A curious result related to Kempner's series, *Amer. Math. Monthly* **115** (2008) 933–938.
6. F. Irwin, A curious convergent series, *Amer. Math. Monthly* **23** (1916) 149–152.
7. A. J. Kempner, A curious convergent series, *Amer. Math. Monthly* **21** (1914) 48–50.
8. T. Schmelzer, R. Baillie, Summing a curious, slowly convergent series, *Amer. Math. Monthly* **115** (2008) 525–540.
9. A. D. Wadhwa, Some convergent subseries of the harmonic series, *Amer. Math. Monthly* **85** (1978) 661–663.

BRIAN LUBECK (blubeck7@gmail.com)

VADIM PONOMARENKO (vponomarenko@mail.sdsu.edu)
San Diego State University, 5500 Campanile Dr., San Diego CA 92182-7720