CURIOUS CONVERGENT SERIES OF INTEGERS WITH MISSING DIGITS

Melvyn B. Nathanson
Lehman College (CUNY), Bronx, New York
melvyn.nathanson@lehman.cuny.edu

Received: 9/17/20, Accepted: 1/12/21, Published: 8/23/21

Abstract

A classical theorem of Kempner states that the sum of the reciprocals of positive integers with missing decimal digits converges. This result is extended to much larger families of "missing digits" sets of positive integers with both convergent and divergent harmonic series.

– To Ron Graham

1. Kempner's Theorem

> "It is well known that the series
>
> $$\sum_{n=1}^{\infty} \frac{1}{n} = \frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \cdots$$
>
> diverges. The object of this Note is to prove that if the denominators do not include all natural numbers 1, 2, 3, . . . , but only those which do not contain any figure 9, the series converges. The method of proof holds unchanged if, instead of 9, any other figure 1, 2, . . . , 8 is excluded, but not for the figure 0."
>
> A. J. Kempner, *Amer. Math. Monthly* **21** (1914), 48–50.

A *harmonic series* is a series of the form $\sum_{a \in A} 1/a$, where A is a set of positive integers. Mathematicians have long been interested in the convergence or divergence of harmonic series. Let $c \in \{1, 2, \ldots, 9\}$, and let $A_{10}(c)$ be the set of positive integers in which the digit c does not occur in the usual decimal representation. Kempner [6] proved in 1914 that $\sum_{a \in A_{10}(c)} 1/a$ converges. He called this "a curious convergent series." More generally, for every integer $g \geq 2$, every positive integer n has a unique *g-adic representation* of the form $n = \sum_{i=0}^{k} c_i g^i$, with digits $c_i \in$

$\{0, 1, 2, \ldots, g-1\}$ for $i = 0, 1, \ldots, k$ and $c_k \neq 0$. If $A_g(c)$ is the set of integers whose g-adic representation contains no digit c, then the infinite series $\sum_{a \in A_g(c)} 1/a$ converges. This includes the case $c = 0$, which was not discussed by Kempner.

Kempner's theorem has been studied and extended by Baillie [1], Farhi [2], Gordon [3], Irwin [5], Lubeck-Ponomarenko [7], Schmelzer and Baillie [10], and Wadhwa [11, 12]. It is Theorem 144 in Hardy and Wright [4].

The g-adic representation is a special case of a more general method to represent the positive integers. A $\mathcal{G}$*-adic sequence* is a strictly increasing sequence of positive integers $\mathcal{G} = (g_i)_{i=0}^{\infty}$ such that $g_0 = 1$ and g_i divides g_{i+1} for all $i \geq 0$. The integer quotients

$$d_i = \frac{g_{i+1}}{g_i}$$

satisfy $d_i \geq 2$ and

$$g_{k+1} = g_k d_k = d_0 d_1 d_2 \cdots d_k \tag{1}$$

for all $k \geq 0$. Every positive integer n has a unique representation in the form

$$n = \sum_{i=0}^{k} c_i g_i \tag{2}$$

where $c_i \in \{0, 1, \ldots, d_i - 1\}$ for all $i \in \{0, 1, \ldots, k\}$ and $c_k \neq 0$. We call (2) the $\mathcal{G}$*-adic representation* of n. This is equivalent to de Bruijn's additive system (Nathanson [8, 9]).

Harmonic series constructed from sets of positive integers with missing $\mathcal{G}$-adic digits do not necessarily converge. In Theorem 1 we construct sets of integers with missing $\mathcal{G}$-adic digits whose harmonic series converge, and also sets of integers with missing $\mathcal{G}$-adic digits whose harmonic series diverge.

2. $\mathcal{G}$-adic Representations with Bounded Quotients

Define the *interval of integers*

$$[a, b] = \{n \in \mathbf{Z} : a \leq n \leq b\}.$$

Let $\mathcal{G} = (g_i)_{i=0}^{\infty}$ be a $\mathcal{G}$-adic sequence with quotients $d_i = g_{i+1}/g_i$. Let I be a set of nonnegative integers, and, for all $i \in I$, let U_i be a nonempty proper subset of $[0, d_i - 1]$. For every nonnegative integer k, let A_k be the set of integers $n \in [g_k, g_{k+1} - 1]$ whose $\mathcal{G}$-adic representation $n = \sum_{i=0}^{k} c_i g_i$ satisfies the following *missing digits condition*:

$$c_i \in [0, d_i - 1] \setminus U_i \qquad \text{for all } i \in I \cap [0, k]. \tag{3}$$

Lemma 1. *The set A_k satisfies:*

(a) *$A_k = \emptyset$ if and only if $k \in I$ and $U_k = [1, d_k - 1]$.*

(b) *If $A_k \neq \emptyset$, then*

$$|A_k| \leq \prod_{\substack{i=0 \\ i \in I}}^{k} (d_i - |U_i|) \prod_{\substack{i=0 \\ i \notin I}}^{k} d_i \leq 2|A_k|. \tag{4}$$

Proof. We use the inequality

$$x \leq 2(x-1) \qquad \text{for } x \geq 2. \tag{5}$$

If $n \in [g_k, g_{k+1} - 1]$, then n has the $\mathcal{G}$-adic representation

$$n = \sum_{i=0}^{k-1} c_i g_i + c_k g_k$$

with $c_k \neq 0$ and so $c_k \in [1, d_k - 1]$. It follows that $A_k = \emptyset$ if and only if $k \in I$ and $U_k = [1, d_k - 1]$.

For $A_k \neq \emptyset$, there are three cases.

(i) If $k \in I$ and $0 \in U_k$, then

$$|A_k| = \prod_{\substack{i=0 \\ i \in I}}^{k} (d_i - |U_i|) \prod_{\substack{i=0 \\ i \notin I}}^{k} d_i < 2|A_k|.$$

(ii) If $k \in I$ and $0 \notin U_k$, then U_k is a proper subset of $[1, d_k - 1]$ and so $|U_k| \leq d_k - 2$. Inequality (5) gives

$$d_k - |U_k| \leq 2\left(d_k - |U_k| - 1\right).$$

We obtain

$$\begin{aligned} |A_k| &= (d_k - |U_k| - 1) \prod_{\substack{i=0 \\ i \in I}}^{k-1} (d_i - |U_i|) \prod_{\substack{i=0 \\ i \notin I}}^{k} d_i \\ &< \prod_{\substack{i=0 \\ i \in I}}^{k} (d_i - |U_i|) \prod_{\substack{i=0 \\ i \notin I}}^{k} d_i \\ &\leq 2(d_k - |U_k| - 1) \prod_{\substack{i=0 \\ i \in I}}^{k-1} (d_i - |U_i|) \prod_{\substack{i=0 \\ i \notin I}}^{k} d_i \\ &= 2|A_k|. \end{aligned}$$

(iii) We have $d_k \geq 2$ and so $d_k \leq 2(d_k - 1)$ from inequality (5). If $k \notin I$, then

$$\begin{aligned}|A_k| &= (d_k - 1)\prod_{\substack{i=0\\ i\in I}}^{k}(d_i - |U_i|)\prod_{\substack{i=0\\ i\notin I}}^{k-1} d_i < \prod_{\substack{i=0\\ i\in I}}^{k}(d_i - |U_i|)\prod_{\substack{i=0\\ i\notin I}}^{k} d_i\\ &= d_k \prod_{\substack{i=0\\ i\in I}}^{k}(d_i - |U_i|)\prod_{\substack{i=0\\ i\notin I}}^{k-1} d_i \leq 2(d_k - 1)\prod_{\substack{i=0\\ i\in I}}^{k}(d_i - |U_i|)\prod_{\substack{i=0\\ i\notin I}}^{k-1} d_i\\ &= 2|A_k|.\end{aligned}$$

This completes the proof. □

Let A be a set of nonnegative integers, and let $A(n)$ be the number of elements $a \in A$ with $a \leq n$. The *upper asymptotic density* of the set A is $d_U(A) = \limsup_{n\to\infty} A(n)/n$. If $\lim_{n\to\infty} A(n)/n$ exists, then $d(A) = \lim_{n\to\infty} A(n)/n$ is called the *asymptotic density* of the set A. The set A has *asymptotic density zero* if $d(A) = d_U(A) = 0$.

Lemma 2. *Let A be a set of positive integers. If $\sum_{a\in A} 1/a < \infty$, then $d(A) = 0$.*

Proof. We show that $d_U(A) > 0$ implies $\sum_{a\in A} 1/a = \infty$.

If $d_U(A) = \limsup_{n\to\infty} A(n)/n = \alpha > 0$, then, for every $\varepsilon > 0$, we have

$$\frac{A(n)}{n} < \alpha + \varepsilon \qquad \text{for all integers } n \geq N(\varepsilon)$$

and

$$\frac{A(n_i)}{n_i} > \alpha - \varepsilon \qquad \text{for infinitely many integers } n_i. \tag{6}$$

Let $\varepsilon < \alpha/3$. There is a sequence of positive integers $(n_i)_{i=0}^{\infty}$ satisfying inequality (6) such that $n_0 \geq N(\varepsilon)$ and $n_i > 2n_{i-1}$ for all $i \geq 1$. We have

$$\begin{aligned}A(n_i) - A(n_{i-1}) &> (\alpha - \varepsilon)n_i - (\alpha + \varepsilon)n_{i-1}\\ &> (\alpha - \varepsilon)n_i - \frac{(\alpha + \varepsilon)n_i}{2}\\ &= n_i\left(\frac{\alpha - 3\varepsilon}{2}\right)\end{aligned}$$

and so

$$\sum_{\substack{a\in A\\ n_{i-1}<a\leq n_i}} \frac{1}{a} \geq \frac{A(n_i) - A(n_{i-1})}{n_i} > \frac{\alpha - 3\varepsilon}{2} > 0.$$

It follows that

$$\sum_{\substack{a \in A \\ 1 \leq a \leq n_k}} \frac{1}{a} \geq \sum_{i=1}^{k} \sum_{\substack{a \in A \\ n_{i-1} < a \leq n_i}} \frac{1}{a} > k \left(\frac{\alpha - 3\varepsilon}{2} \right)$$

and the infinite series $\sum_{a \in A} 1/a$ diverges. Equivalently, convergence of the infinite series $\sum_{a \in A} 1/a$ implies $d(A) = 0$. This completes the proof. $\square$

The converse of Lemma 2 is false. The set of prime numbers has asymptotic density zero, but the sum of the reciprocals of the primes diverges.

Theorem 1. *Let $\mathcal{G} = (g_i)_{i=0}^{\infty}$ be a $\mathcal{G}$-adic sequence with bounded quotients, that is,*

$$d_i = \frac{g_{i+1}}{g_i} \leq d \tag{7}$$

for some integer $d \geq 2$ and all $i = 0, 1, 2, \ldots$. Let I be a set of nonnegative integers, and, for all $i \in I$, let U_i be a nonempty proper subset of $[0, d_i - 1]$.

Let $n = \sum_{i=0}^{k} c_i g_i$ be the $\mathcal{G}$-adic representation of the positive integer n. Let A be the set of positive integers n that satisfy the missing digits condition (3)*. If*

$$I(k) \geq \frac{(1+\delta) \log k}{\log(d/(d-1))} \tag{8}$$

for some $\delta > 0$ and all $k \geq k_0 = k_0(\delta)$, then the set A has asymptotic density zero and the harmonic series $\sum_{a \in A} 1/a$ converges.

If

$$I(k) \leq \frac{(1-\delta) \log k}{\log d} \tag{9}$$

for some $\delta > 0$ and all $k \geq k_1 = k_1(\delta)$, then the harmonic series $\sum_{a \in A} 1/a$ diverges.

Kempner's theorem is the special case $g_i = 10^i$, $d_i = 10$, and $U_i = \{9\}$ for all $i \in I = \mathbf{N}_0$.

Proof. For all $k \in \mathbf{N}_0$, the finite sets

$$A_k = A \cap [g_k, g_{k+1} - 1]$$

are pairwise disjoint and $A = \bigcup_{k=0}^{\infty} A_k$.

For all $i \in I$, we have

$$1 \leq |U_i| \leq d_i - 1$$

and

$$\frac{1}{d} \leq \frac{1}{d_i} \leq \frac{d_i - |U_i|}{d_i} = 1 - \frac{|U_i|}{d_i} \leq 1 - \frac{1}{d} < 1.$$

Let $I(k)$ satisfy inequality (8). We obtain

$$\left(1-\frac{1}{d}\right)^{I(k)} \leq \left(\frac{d-1}{d}\right)^{\frac{(1+\delta)\log k}{\log(d/(d-1))}} = \frac{1}{k^{1+\delta}}. \tag{10}$$

If $a \in A_k$, then $a \geq g_k = d_0 d_1 \cdots d_{k-1}$. By Lemma 1,

$$\begin{aligned}
\sum_{\substack{a\in A\\ a\geq g_{k_0}}} \frac{1}{a} = \sum_{k=k_0}^{\infty} \sum_{a\in A_k} \frac{1}{a} \leq \sum_{k=k_0}^{\infty} \frac{|A_k|}{g_k} &\leq \sum_{k=k_0}^{\infty} \frac{d_k}{\prod_{i=0}^{k} d_i} \prod_{\substack{i=0\\ i\in I}}^{k} (d_i - |U_i|) \prod_{\substack{i=0\\ i\notin I}}^{k} d_i \\
&\leq d \sum_{k=k_0}^{\infty} \prod_{\substack{i=0\\ i\in I}}^{k} \frac{d_i - |U_i|}{d_i} \\
&\leq d \sum_{k=k_0}^{\infty} \left(1-\frac{1}{d}\right)^{I(k)} \\
&\leq d \sum_{k=k_0}^{\infty} \frac{1}{k^{1+\delta}} < \infty.
\end{aligned}$$

Thus, the harmonic series converges. By Lemma 2, the set A has asymptotic density zero.

Let $I(k)$ satisfy inequality (9). We obtain

$$\left(\frac{1}{d}\right)^{I(k)} \geq \left(\frac{1}{d}\right)^{\frac{(1-\delta)\log k}{\log d}} = \frac{1}{k^{1-\delta}}. \tag{11}$$

If $a \in A_k$, then $a < g_{k+1} = d_0 d_1 \cdots d_{k-1} d_k$. By Lemma 1,

$$\begin{aligned}
\sum_{\substack{a\in A\\ a\geq g_{k_1}}} \frac{1}{a} = \sum_{k=k_1}^{\infty} \sum_{a\in A_k} \frac{1}{a} &\geq \sum_{k=k_1}^{\infty} \frac{|A_k|}{g_{k+1}} \\
&\geq \frac{1}{2} \sum_{k=k_1}^{\infty} \frac{1}{\prod_{i=0}^{k} d_i} \prod_{\substack{i=0\\ i\in I}}^{k} (d_i - |U_i|) \prod_{\substack{i=0\\ i\notin I}}^{k} d_i \\
&= \frac{1}{2} \sum_{k=k_1}^{\infty} \prod_{\substack{i=0\\ i\in I}}^{k} \frac{d_i - |U_i|}{d_i} \geq \frac{1}{2} \sum_{k=k_1}^{\infty} \prod_{\substack{i=0\\ i\in I}}^{k} \frac{1}{d_i} \\
&\geq \frac{1}{2} \sum_{k=k_1}^{\infty} \left(\frac{1}{d}\right)^{I(k)} \\
&\geq \frac{1}{2} \sum_{k=k_1}^{\infty} \frac{1}{k^{1-\delta}}
\end{aligned}$$

and the harmonic series diverges. This completes the proof. □

Corollary 1. *Let I be a set of nonnegative integers, and let $(v_i)_{i\in I}$ be a sequence of 0s and 1s. Let A be the set of integers n such that, if $n \in \left[2^k, 2^{k+1}-1\right]$ has the 2-adic representation $n = \sum_{i=0}^{k} c_i 2^i$, then $c_i = v_i$ for all $i \in I \cap [0,k]$. If*

$$I(k) \geq (1+\delta)\log_2 k$$

for some $\delta > 0$ and all $k \geq k_0(\delta)$, then the harmonic series $\sum_{a\in A} 1/a$ converges. If

$$I(k) \leq (1-\delta)\log_2 k$$

for some $\delta > 0$ and all $k \geq k_1(\delta)$, then the harmonic series $\sum_{a\in A} 1/a$ diverges.

Proof. For all $i \in I$, let $u_i = 1 - v_i$ and $U_i = \{u_i\}$. Apply Theorem 1. □

It is an open problem to determine the convergence or divergence of $\sum_{a\in A} 1/a$ if $I(k) \sim \log_2 k$.

3. $\mathcal{G}$-adic Representations with Unbounded Quotients

Let $\mathcal{G} = (g_i)_{i=0}^{\infty}$ be a $\mathcal{G}$-adic sequence with quotients $d_i = g_{i+1}/g_i$. Let I be an infinite set of nonnegative integers, and, for all $i \in I$, let U_i be a nonempty proper subset of $[0, d_i - 1]$. If the sequence $\mathcal{G} = (g_i)_{i=0}^{\infty}$ has bounded quotients $d_i \leq d$, then

$$\frac{|U_i|}{d_i} \geq \frac{1}{d}$$

for all $i \in I$ and the infinite series $\sum_{i\in I} \frac{|U_i|}{d_i}$ diverges. Equivalently, the convergence of this series implies that $\mathcal{G}$ has unbounded quotients.

Let $n = \sum_{i=0}^{k} c_i g_i$ be the $\mathcal{G}$-adic representation of the positive integer n. Let A be the set of positive integers whose $\mathcal{G}$-adic representations satisfy the missing digits condition (3). The missing digits set A is finite if and only if I is a cofinite set of nonnegative integers and $U_i = [1, d_i - 1]$ for all sufficiently large i. The harmonic series of a finite set of positive integers converges.

Theorem 1 shows that harmonic series constructed from infinite sets of integers with missing $\mathcal{G}$-adic digits do not always converge. It follows from Theorem 1 that if

$$I(k) \geq (\log k)^{1+\delta}$$

for some $\delta > 0$ and all $k \geq k_0(\delta)$, and if $\sum_{a\in A} 1/a$ diverges, then the sequence $\mathcal{G}$ must have *unbounded quotients*, that is,

$$\limsup d_i = \infty.$$

Theorem 2 gives a sufficient condition for the divergence of harmonic series of sets of positive integers constructed from $\mathcal{G}$-adic sequences with unbounded quotients.

We use the following inequality, which is easily proved by induction: If $0 \leq x_i < 1$ for $i = 1, \ldots, n$, then

$$\prod_{i=1}^{n}(1 - x_i) \geq 1 - \sum_{i=1}^{n} x_i. \tag{12}$$

Theorem 2. *Let $\mathcal{G} = (g_i)_{i=0}^{\infty}$ be a $\mathcal{G}$-adic sequence, and let $n = \sum_{i=0}^{k} c_i g_i$ be the $\mathcal{G}$-adic representation of the positive integer n. Let I be an infinite set of nonnegative integers, and, for all $i \in I$, let U_i be a nonempty proper subset of $[0, d_i - 1]$. Let A be the set of positive integers whose $\mathcal{G}$-adic representations satisfy the missing digits condition* (3)*. If the set A is infinite and if*

$$\sum_{i \in I} \frac{|U_i|}{d_i} < \infty \tag{13}$$

then the sequence $\mathcal{G} = (g_i)_{i=0}^{\infty}$ has unbounded quotients and the harmonic series $\sum_{a \in A} 1/a$ diverges.

For example, the "missing digits" set constructed from $\mathcal{G} = (g_i)_{i=0}^{\infty}$ with $g_i = 2^{i(i+1)/2}$ and $d_i = 2^{i+1}$ and with $I = \mathbf{N}_0$ and $U_i = \{0\}$ for all $i \in I$ has a divergent harmonic series.

Proof. Because the infinite series (13) converges, there is an integer $i_0 \in I$ such that

$$\sum_{\substack{i=i_0 \\ i \in I}}^{\infty} \frac{|U_i|}{d_i} < \frac{1}{2}.$$

Inequality (12) implies that, for all $k \in \mathbf{N}_0$,

$$\prod_{\substack{i=i_0 \\ i \in I}}^{k} \left(1 - \frac{|U_i|}{d_i}\right) \geq 1 - \sum_{\substack{i=i_0 \\ i \in I}}^{k} \frac{|U_i|}{d_i} > \frac{1}{2}$$

and so

$$\begin{aligned} \prod_{\substack{i=0 \\ i \in I}}^{k} \left(1 - \frac{|U_i|}{d_i}\right) &= \prod_{\substack{i=0 \\ i \in I}}^{i_0-1} \left(1 - \frac{|U_i|}{d_i}\right) \prod_{\substack{i=i_0 \\ i \in I}}^{k} \left(1 - \frac{|U_i|}{d_i}\right) \\ &> \frac{1}{2} \prod_{\substack{i=0 \\ i \in I}}^{i_0-1} \left(1 - \frac{|U_i|}{d_i}\right) = \delta > 0. \end{aligned}$$

Let $A_k = A \cap [g_k, g_{k+1} - 1]$. The set A is infinite if and only if $A_k \neq \emptyset$ for infinitely many k. Applying inequality (4) of Lemma 1, we obtain

$$\begin{aligned}\sum_{a\in A}\frac{1}{a} &= \sum_{\substack{k=0\\ A_k\neq\emptyset}}^{\infty}\sum_{a\in A_k}\frac{1}{a} \geq \sum_{\substack{k=0\\ A_k\neq\emptyset}}^{\infty}\frac{|A_k|}{g_{k+1}}\\ &\geq \frac{1}{2}\sum_{\substack{k=0\\ A_k\neq\emptyset}}^{\infty}\frac{1}{\prod_{i=0}^{k} d_i}\prod_{\substack{i=0\\ i\in I}}^{k}(d_i-|U_i|)\prod_{\substack{i=0\\ i\notin I}}^{k} d_i\\ &= \frac{1}{2}\sum_{\substack{k=0\\ A_k\neq\emptyset}}^{\infty}\prod_{\substack{i=0\\ i\in I}}^{k}\left(1-\frac{|U_i|}{d_i}\right)\end{aligned}$$

and so the harmonic series $\sum_{a\in A}\frac{1}{a}$ diverges. This completes the proof. □

References

[1] R. Baillie, Sums of reciprocals of integers missing a given digit, *Amer. Math. Monthly* **86** (1979), no. 5, 372–374.

[2] B. Farhi, A curious result related to Kempner's series, *Amer. Math. Monthly* **115** (2008), no. 10, 933–938.

[3] R. A. Gordon, Comments on "Subsums of the harmonic series", *Amer. Math. Monthly* **126** (2019), no. 3, 275–279.

[4] G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, 6th ed., Oxford University Press, Oxford, 2008.

[5] F. Irwin, A curious convergent series, *Amer. Math. Monthly* **23** (1916), no. 5, 149–152.

[6] A. J. Kempner, A curious convergent series, *Amer. Math. Monthly* **21** (1914), no. 2, 48–50.

[7] B. Lubeck and V. Ponomarenko, Subsums of the harmonic series, *Amer. Math. Monthly* **125** (2018), no. 4, 351–355.

[8] M. B. Nathanson, Additive systems and a theorem of de Bruijn, *Amer. Math. Monthly* **121** (2014), no. 1, 5–17.

[9] M. B. Nathanson, Limits and decomposition of de Bruijn's additive systems, *Combinatorial and Additive Number Theory II* (New York), Springer Proc. Math. Stat., vol. 220, Springer, 2017, pp. 255–267.

[10] T. Schmelzer and R. Baillie, Summing a curious, slowly convergent series, *Amer. Math. Monthly* **115** (2008), no. 6, 525–540.

[11] A. D. Wadhwa, An interesting subseries of the harmonic series, *Amer. Math. Monthly* **82** (1975), no. 9, 931–933.

[12] A. D. Wadhwa, Some convergent subseries of the harmonic series, *Amer. Math. Monthly* **85** (1978), no. 8, 661–663.