Summing The Curious Series of Kempner and Irwin

Robert Baillie

1. INTRODUCTION.

In 1914, Kempner proved [3] that if we delete from the harmonic series all terms whose denominators have the digit 9 (that is, 1/9, 1/19, 1/29, ...), then the remaining series converges. This is surprising because it *appears* that this process removes only every 10^{th} term from the harmonic series, so the sum of what remains would seem to diverge.

Nonetheless, Kempner's series converges to about 22.92067661926415034816 [1]. The convergence is so slow that the sum of all terms with denominators < 10^{30} is still less than 22. The following generalization also holds: if we delete from the harmonic series all terms whose denominators have any set of numbers with one or more digits, we also get convergent series. Schmelzer and Baillie [5] show how to compute sums of these slowly-converging series.

In 1916, Irwin [2] generalized Kempner's result in a different way. He showed that we also get convergent series if we sum over those *n* that have at most a fixed number of occurrences of one or more digits. It seems that, until now, no one has computed sums of these series. This note shows how to do so.

For example, we can calculate that the sum of *1/n* where *n* has exactly one 9 is about 23.04428708074784831968. It is surprising that this is larger than the sum with no 9, given above, because the series with one 9 begins 1/9 + 1/19 + ..., while the series with no 9 begins 1/1 + 1/2 + ... + 1/8 + 1/10 + ... The partial sums of the series with one 9 remain less than the full sum with no 9 until we reach terms whose denominators have 70 digits. There are $N_i = 9^{i-1} + 8(i-1)9^{i-2}$ integers with *i* digits that have exactly one 9, so the 'one 9' series needs more than $N_1 + ... + N_{69} \approx 5 \cdot 10^{66}$ terms to exceed the sum of the 'no 9' series.

A more complicated example: the sum of *1/n* where *n* has *at most* one 1, two 2's, three 3's, four 4's and five 5's (with no conditions placed on the other digits) is about 27.56008294889636705754. The sum where *n* has *exactly* one 1, two 2's, three 3's, four 4's, and five 5's is about 0.00465390225405638155564. (All sums are rounded to the number of places shown).

2. THE ALGORITHM.

It's easy to add terms having denominators up to, say, 7 digits, but we must use extrapolation to get much beyond that point. It turns out that we can use sums over denominators with *i* digits to compute the needed sums over denominators with *i* + 1 digits. Then we repeat the process. The basic idea goes back to Kempner, and it is the key to accurately computing these sums. We will illustrate the method when there are two conditions for *1/n* to be included in the series: we show how to compute the sum of 1/n where *n* has (exactly) n_1 occurrences of the digit d_1 and n_2 occurrences of d_2. The idea extends in the obvious way to conditions on three or more digits, and easily generalizes to bases other than 10.

Define $S(i,k_1,k_2)$ to be the set of positive integers with i digits that have k_1 occurrences of d_1 and k_2 occurrences of d_2. We can generate the set $S(i+1,k_1,k_2)$ from $S(i,k_1-1,k_2)$, $S(i,k_1,k_2-1)$, and $S(i,k_1,k_2)$, as follows:

(a) For each x in $S(i,k_1-1,k_2)$ multiply by 10, then add d_1.
(b) For each x in $S(i,k_1,k_2-1)$, multiply by 10, then add d_2.
(c) For each x in $S(i,k_1,k_2)$, multiply by 10, then add d = 0, 1, 2, ..., 9, except for d_1 and d_2.

Step (a) starts with an i-digit number having k_1 - 1 occurrences of d_1 and appends d_1 as the final digit, forming an (i+1) digit number having k_1 occurrences of d_1. Step (b) does the same for k_2 and d_2. In step (c), the i-digit numbers already have the desired number of d_1 and d_2, so in this step, we create (i+1)-digit numbers by appending all other digits except d_1 and d_2. Together, these steps generate $S(i+1,k_1,k_2)$. If k_1 is 0, we omit step (a). If k_2 is 0, we omit step (b).

Next, define $t(i,j,k1,k2) = \sum_{x \in S(i,k_1,k_2)} \frac{1}{x^j}$. We will show how to compute $t(i+1,j,k_1,k_2)$ by using the values of $t(i,j,k_1-1,k_2)$, $t(i,j,k_1,k_2-1)$, and $t(i,j,k_1,k_2)$ for j = 1, 2, 3, ...

If x is an i-digit number and d is any digit, then the reciprocal of the corresponding (i+1)-digit number $10x+d$ can be expanded in powers of $1/x$:

$$\frac{1}{10x+d} = \frac{1}{10x(1+\frac{d}{10x})} = \frac{1}{10x}\sum_{n=0}^{\infty}(-1)^n\left(\frac{d}{10x}\right)^n = \sum_{n=0}^{\infty}(-1)^n\frac{d^n}{(10x)^{n+1}}$$

This expansion is valid because $\left|\frac{d}{10x}\right| < 1$. A similar expansion holds for higher powers:

$$\frac{1}{(10x+d)^j} = \frac{1}{(10x)^j(1+\frac{d}{10x})^j} = \frac{1}{(10x)^j}\sum_{n=0}^{\infty}(-1)^n\frac{(j+n-1)!}{n!(j-1)!}\left(\frac{d}{10x}\right)^n$$

$$= \sum_{n=0}^{\infty}(-1)^n\frac{(j+n-1)!}{n!(j-1)!}\frac{d^n}{(10x)^{j+n}}$$

Now, recalling step (a), we sum these expansions for all x in $S(i,k_1-1,k_2)$. Call this sum A.

$$A = \sum_{x \in S(i,k_1-1,k_2)} \frac{1}{(10x+d_1)^j} = \sum_{x \in S(i,k_1-1,k_2)} \sum_{n=0}^{\infty}(-1)^{n+1}\frac{(j+n-1)!}{n!(j-1)!}\frac{d_1^n}{(10x)^{j+n}} \tag{1}$$

$$= \sum_{n=0}^{\infty} (-1)^{n+1} \frac{(j+1-n)!}{n!(j-1)!} \frac{d_1^{\ n}}{10^{j+n}} \sum_{x \in S(i,k_1-1,k_2)} \frac{1}{x^{j+n}}$$

Likewise, summing over the sets described in steps (b) and (c), we get

$$B = \sum_{x \in S(i,k_1,k_2-1)} \frac{1}{(10x+d_2)^j} = \sum_{n=0}^{\infty} (-1)^{n+1} \frac{(j+1-n)!}{n!(j-1)!} \frac{d_2^{\ n}}{10^{j+n}} \sum_{x \in S(i,k_1,k_2-1)} \frac{1}{x^{j+n}} \qquad (2)$$

$$C = \sum_{d} \sum_{x \in S(i,k_1,k_2)} \frac{1}{(10x+d)^j} = \sum_{n=0}^{\infty} (-1)^{n+1} \frac{(j+1-n)!}{n!(j-1)!} \frac{0^n + 1^n + \ldots + 9^n - d_1^{\ n} - d_2^{\ n}}{10^{j+n}} \sum_{x \in S(i,k_1,k_2)} \frac{1}{x^{j+n}} \qquad (3)$$

The sum in (3) is over digits $d \neq d_1, d_2$. Together, the sets generated by steps (a) - (c) comprise $S(i+1,k_1,k_2)$. Also, $A+B+C = t(i+1,j,k_1,k_2)$. We have just computed a needed sum over (i+1) digit numbers by using sums over i-digit numbers. Programming detail: we take $0^0 = 1$.

In order to compute $t(i+1,j,k_1,k_2)$, we used the values of $t(i,j,k_1-1,k_2)$ and $t(i,j,k_1,k_2-1)$. But $t(i,j,k_1-1,k_2)$, in turn, was computed using $t(i-1,j,k_1-2,k_2-1)$ and $t(i-1,j,k_1-1,k_2-2)$. This means that, for each i and j, in order to compute $t(i+1,j,n_1,n_2)$, we must compute all $(n_1+1)(n_2+1)$ values of $t(i,j,k_1,k_2)$ for $0 \le k_1 \le n_1$ and $0 \le k_2 \le n_2$.

If we add the $t(i,1,n_1,n_2)$ values over all i, then we get the sum of $1/n$ where n has *exactly* n_1 occurrences of d_1 and *exactly* n_2 occurrences of d_2. If we add the $t(i,1,k_1,k_2)$ values over all i, and over all k_1 and k_2 with $0 \le k_1 \le n_1$ and $0 \le k_2 \le n_2$, then we get the sum of $1/n$ where n has *at most* n_1 occurrences of d_1 and *at most* n_2 occurrences of d_2.

If there are three conditions (n_1, d_1), (n_2, d_2), (n_3, d_3), the procedure is similar, except that we have three steps in place of (a) and (b) above, and that for each i and j, we must compute $(n_1+1)(n_2+1)(n_3+1)$ values for the t array. The time and memory requirements increase accordingly. Still, on a personal computer, we can specify a condition for each of the ten digits, provided the product $(n_1+1)(n_2+1)\cdot \ldots \cdot (n_{10}+1)$ is not too large.

To summarize, our calculation goes as follows. We start by calculating $t(i,j,k_1,k_2)$ for i = 1, 2, and 3, by explicitly adding the terms whose denominators are in the sets $S(i,k_1,k_2)$, for all $0 \le k_1 \le n_1$ and $0 \le k_2 \le n_2$. For i = 1 and 2, we need only j = 1. For i = 3, we compute all sums $t(i,j,k_1,k_2)$ for $j \le J$, where J depends on the number of decimal places desired. Then we use equations (1) - (3), along with the $t(3,j,k_1,k_2)$ values to successively compute the needed $t(i+1,j,k_1,k_2)$ for i = 3, 4, 5, ... We continue until the $t(i,j,k_1,k_2)$ values become small enough to be neglected. Of course, the more decimal places we want, the larger the range of i and j values we will need.

3. CONFIRMING THE CALCULATIONS.

Generally, the series considered here converge very slowly. However, there are a few special cases where they converge rapidly. These special cases can serve as a check on the algorithm. For example, the sum of *1/n* where *n* has no 0 in base 2 is

$$\frac{1}{1}+\frac{1}{3}+\frac{1}{7}+\frac{1}{15}+\ldots=\sum_{n=1}^{\infty} 1/(2^n-1).$$

This series converges rapidly to 1.60669 51524 ... . Likewise, the sum of *1/n* where *n* has a single 1 in base 2, is equal to 2. One can also construct special cases in other bases.

We can also test the extrapolation procedure as follows. First, for a given set of conditions (n_1, d_1), (n_2, d_2), ..., we explicitly compute the sums over denominators of, say, 1, 2, ..., 7 digits. Then, we explicitly compute the sums through, say, 4 digits, then use equations (1) - (3) to estimate the sums for 5, 6, and 7 digits. In all cases, we get good agreement. This algorithm produces good agreement with Weisstein's result in Example 1 (b) below. Finally, when the conditions specify only the *absence* of one or more digits (that is, $n_1 = \ldots = n_m = 0$), the results match those obtained by the algorithms in [1] and [5].

4. MORE EXAMPLES.

Example 1 (a). Set $n_1 = n_2 = \ldots = n_{10} = 1$. Because we are limiting the number of occurrences of every digit, this series has only a finite number of terms. In every denominator of this series, each digit occurs exactly once. This means the denominators have exactly 10 digits, all distinct. There are $10!\cdot 9/10 = 3265920$ such numbers. The sum of this finite series is about 0.00082589034791925293861.

Example 1 (b). In order obtain the sum in 1 (a), we had to compute the 2^{10} sums over those *n* which have either 0 or 1 occurrence of each of the ten digits. Together, these *n* comprise the positive integers that have distinct digits. There are 8877690 of them between 1 and 9876543210. The sum of their reciprocals is about 8.92994817475544342417. This is a more precise answer to Monthly Problem E2533 than the bounds given in [4]. Interestingly, Weisstein [7] used *Mathematica* to compute the *exact* value of this finite sum, a fraction whose numerator and denominator have 14816583 and 14816582 digits.

Example 2 (a). Set $n_1 = n_2 = \ldots = n_{10} = 2$. Here, each denominator has *exactly* two of every digit, so all denominators in this finite series have 20 digits. This sum is about 0.00005440621942909909 1465.

Example 2 (b). As part of the calculation in 2 (a), we also computed the 3^{10} sums of *1/n* where *n* has exactly 0, 1, or 2 occurrences of each of the ten digits. Together, these sums comprise a finite sum that terminates after 20-digit denominators. Adding these sums together gives the sum of *1/n* where *n* has *at most* two of every digit. This sum is about 20.58988677491808564961.

Example 3. Wadhwa [6] considers s_k = sum of $1/n$ where n has exactly k 0's. He shows that s_k is strictly decreasing and that $s_k > 19.28$. We calculate that $s_0 \approx 23.10344790942054161603$, $s_1 \approx 23.02673534156912696109$, $s_2 \approx 23.02586068273551997642$, $s_3 \approx 23.02585103714853825460$, ..., $s_{10} \approx 23.0258509299404568401819892$, ... Notice also that $s_{10} - 10\ln(10) \approx 2 \times 10^{-21}$. Query: why is this difference so small?

Example 4. Example 3 suggests how to construct nontrivial, convergent subseries of the harmonic series that have arbitrarily large, but computable, sums. For example, the sum of $1/n$ where n has at most 43 0's is about 1013.21593 21696 83236 58704. Replacing "43" in the last sentence with any larger number yields a convergent series with an even bigger sum.

Example 5. Here we display, for each digit d, the sum of $1/n$ where n has zero, one, or two d.

d	Sum for Zero Occurrences	Sum for One Occurrence	Sum for Two Occurrences
0	23.10344 79094 20541 61603	23.02673 53415 69126 96109	23.02586 06827 35519 97642
1	16.17696 95281 23444 26658	23.16401 85942 72832 04085	23.02727 62863 56005 71224
2	19.25735 65328 08072 22453	23.08826 06627 56342 39334	23.02648 59737 68470 65598
3	20.56987 79509 61230 37108	23.06741 08819 30230 10242	23.02627 31906 67935 05960
4	21.32746 57995 90036 68664	23.05799 24133 81824 39576	23.02617 78853 92600 17317
5	21.83460 08122 96918 16341	23.05272 88945 30117 49904	23.02612 48753 15647 60861
6	22.20559 81595 56091 88417	23.04940 99732 95500 55704	23.02609 15498 64887 12587
7	22.49347 53117 05945 39818	23.04714 61901 98641 85083	23.02606 88649 14415 07436
8	22.72636 54026 79370 60283	23.04551 39079 82155 53342	23.02605 25308 45693 67648
9	22.92067 66192 64150 34816	23.04428 70807 47848 31968	23.02604 02659 61243 78845

This table also shows, for example, that the sum of $1/n$ where n has at most two 9's is about 22.92067 ... + 23.04428 ... + 23.02604 ... ≈ 68.99100 ... .

A *Mathematica* package implementing this algorithm can be downloaded from [8].

References.

[1] R. Baillie, Sums of Reciprocals of Integers Missing a Given Digit, this MONTHLY **86** (1979) 372-374.

[2] F. Irwin, A Curious Convergent Series, this MONTHLY **23** (1916) 149-152.

[3] A. J. Kempner, A Curious Convergent Series, this MONTHLY **21** (1914) 48-50.

[4] E. S. Pondiczery, Elementary Problem E2533, this MONTHLY **82** (1975) 401. Solution: **83** (1976) 570.

[5] T. Schmelzer and R. Baillie, Summing a Curious, Slowly Convergent Series, this MONTHLY **115** (June/July 2008).

[6] A. D. Wadhwa, Some Convergent Subseries of the Harmonic Series, this MONTHLY **85** (1978) 661-663.

[7] Weisstein, Eric W. "Digit." From MathWorld--A Wolfram Web Resource. http://mathworld.wolfram.com/Digit.html.

[8] The code will be submitted to http://library.wolfram.com/infocenter/MathSource. Meanwhile, the code is available from the author upon request.

Remcom, Inc., 315 South Allen Street, State College, PA 16801; email: rjbaillie@frii.com