Divergent Series: why $1 + 2 + 3 + \cdots = -1/12$.

Bryden Cais

"Divergent series are the invention of the devil, and it is shameful to base on them any demonstration whatsoever."—N. H. Abel

1. Introduction

The notion of convergence of a series is a simple one: we say that the series $\sum_{n=0}^{\infty} a_n$ converges if

$$\lim_{N\to\infty} \sum_{n=0}^{N} a_n$$

exists and is finite. So for example the series

$$1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \cdots$$
$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \cdots$$

both converge (to 2 and $\log 2$, respectively). If a series $\sum a_n$ does not converge, it is said to *diverge*. Two prototypical examples of divergent series are

$$1 + 2 + 3 + 4 + 5 + \cdots$$
$$1 - 1 + 1 - 1 + 1 - \cdots,$$

where the first series diverges because the partial sums tend to $+\infty$ and the second series diverges because the partial sums s_N do not tend to any limit (even though $\lim s_{2N} = 0$ and $\lim s_{2N-1} = 1$).

One might think that not much can be said for divergent series. The goal of this these notes is to show that this is not the case, and that divergent series are in fact an interesting object of study. Historically, divergent series occur in the work of Euler, Poisson, Fourier, and Ramanjuan (among many others), and although it was not until Cauchy that the definitions of convergence were formally stated, these masters knew well enough when a series converged and when it did not. Part of the reason that divergent series were so abhorred by mathematicians after Cauchy is because no one formally *defined* what the sum of a divergent series should be. Rather, mathematicians sought for some intrinsic meaning and rapidly found themselves in rather difficult terrain. We will begin by highlighting some of the problems that arise when one tries to make sense of a divergent series without clear definitions, and this will lead naturally to two generalizations of the notion of "sum," both of which can be used to assign meaning to divergent series.

2. Problems with summing divergent series

Abel's 1828 remark that "divergent series are the invention of the devil" was not unfounded. Let us illustrate this with two examples.

First, consider the series

$$s = 1 - 1 + 1 - 1 + \cdots.$$

There are two essentially different ways in which we can make sense of this series. The first is by simple manipulations:

$$1 - s = 1 - (1 - 1 + 1 - 1 + \cdots) = 1 - 1 + 1 - 1 + \cdots = s,$$

so that $2s = 1$ or $s = 1/2$. Observe that we have used only "linear" properties, and have not rearranged any terms other than the first. We might also observe that

$$1 - 1 + 1 - 1 + \cdots = \lim_{x\to 1^-} \frac{1}{1+x} = \frac{1}{2}.$$

This is reassuring. Intuitively, the "value" of s *should* be $1/2$ as this is the average value of the partial sums, which alternate between 0 and 1.

Problems arise, however, as soon as one tries to rearrange any terms. For example, we might try to argue that

$$s = 1 - 1 + 1 - 1 + \cdots = (1-1) + (1-1) + (1-1) + \cdots = 0 + 0 + 0 + \cdots = 0,$$

but such arguments can be dismissed on the grounds that even for convergent series rearranging terms can spell disaster. Thus, for example, one can rearrange terms of the series

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \cdots$$

to obtain *any* real number as the sum! As for the second argument, one might object (as Callet did) that for $m < n$ we have

$$\frac{1 - x^m}{1 - x^n} = 1 - x^m + x^n - x^{m+n} + x^{2n} - \cdots ,$$

so that by l'hopital's rule

$$1 - 1 + 1 - 1 + \cdots = \lim_{x \to 1^-} \frac{1 - x^m}{1 - x^n} = \frac{m}{n}.$$

But here again, it is not so difficult to see what is happening: the limit as $x \to 1^-$ of $1 - x^m + x^n - x^{m+n} + \cdots$ is *not* the series

$$1 - 1 + 1 - 1 + \cdots ,$$

but rather the series

$$1 + \underbrace{0 + 0 + \cdots + 0}_{m-1 \text{ zeroes}} - 1 + \underbrace{0 + 0 + \cdots + 0}_{n-m-1 \text{ zeroes}} + 1 + \cdots ,$$

and there is no reason why these two series should have the same value.

For a different kind of "problem," consider the series

$$s = 1 + 2 + 3 + 4 + 5 + \cdots .$$

What should be the value of s? Laying aside for the moment our reservations, we have

$$\begin{aligned} -3s = (1-4)s &= (1 + 2 + 3 + 4 + 5 + \cdots) - 2(2 + 4 + 6 + \cdots) \\ &= 1 - 2 + 3 - 4 + 5 - \cdots \\ &= 1 - (2 - 3 + 4 - 5 + \cdots) \\ &= 1 - (1 - 2 + 3 - 4 + 5 - \cdots) - (1 - 1 + 1 - 1 + \cdots) \\ &= 1 + 3s - 1/2, \end{aligned}$$

we conclude that $-6s = 1/2$, that is

$$1 + 2 + 3 + 4 + 5 + \cdots = -\frac{1}{12}.$$

Such a statement obviously presents philosophical difficulties. Namely, one is forced to ask how the "sum" of a divergent series of entirely positive terms can be negative. Yet the manipulations involved in our determination of s are no more outlandish than those used in determining $1 - 1 + 1 - 1 + \cdots = 1/2$. We will see later that in a very precise sense, $-1/12$ is the *correct* value of $1 + 2 + 3 + 4 + \cdots$.

3. Definitions: Cesaro and Abel summability

Based on our manipulations of series above, two things are clear:

(1) Any useful definition of the sum of a divergent series will allow "linear" operations on the series without altering its value.
(2) One cannot hope for the rearrangement of terms in a divergent series to be an innocent operation, as this is already not the case for conditionally convergent series.

Moreover, any definition of the sum of a divergent series should be a generalization of the sum of a convergent series, so that when one tries to sum convergent series using these definitions, one obtains the same results as before.

Let us first make precise what we mean by "linear." Let (a_n), (b_n) be two sequences and suppose we have assigned the values $s = \sum a_n$ and $t = \sum b_n$. Then we want the following to hold:

(1) $\sum(a_n + b_n) = s + t$.
(2) For any real number α, we require $\sum \alpha a_n = \alpha s$.

These notions can be formulated in a precise way. Let V denote the infinite-dimensional real vector space consisting of sequences of real numbers (a_n), and let $W \subseteq V$ be the subspace of *convergent* sequences. Then we can think of

$$\sum : W \longrightarrow \mathbf{R}$$

as the linear operator defined by $\sum((a_n)) := \sum_{n=1}^{\infty} a_n$. In seeking to generalize the idea of "sum" to divergent series, we wish to extend the linear operator $\sum$ to subspaces $V' \subseteq V$ strictly containing W. We might hope to be able to extend $\sum$ to all of V, but as we shall see, this is a somewhat ambitious goal. Our generalizations are classical, and motivated by the calculations of the preceding section.

3.1. **Cesaro summability.** Cesaro summation formalizes our claim that $1 - 1 + 1 - 1 + \cdots$ *should* be $1/2$ as this is the average value of the partial sums. Concretely, let (a_n) be any sequence of real numbers and put $s_N = a_1 + a_2 + \cdots + a_N$. Let $V_{c,1}$ be the subset of V for which

$$\lim_{N\to\infty} \frac{s_1 + s_2 + \cdots + s_N}{N}$$

exists and define the map $C_1 : V_{c,1} \longrightarrow \mathbf{R}$ by

$$C_1((a_n)) = \lim_{N\to\infty} \frac{s_1 + s_2 + \cdots + s_N}{N}.$$

It is easy to see that $V_{c,1}$ is actually a subspace of V, and not just a subset. Moreover, suppose that $a_1 + a_2 + \cdots$ converges, say to L, so that for any $\epsilon > 0$ there exists $N > 0$ such that $|L - s_M| < \epsilon$ for all $M > N$. Then we have

$$\begin{aligned} C_1((a_n)) &= \lim_{k\to\infty} \frac{s_1 + s_2 + \cdots + s_N}{N+k} + \frac{s_{N+1} + \cdots + s_{N+k}}{N+k} \\ &= \lim_{k\to\infty} \frac{s_{N+1} + \cdots + s_{N+k}}{N+k}. \end{aligned}$$

But since $|s_M - L| < \epsilon$ for all $M > N$, we have the bounds

$$L - \epsilon = \lim_{k\to\infty} \frac{k(L-\epsilon)}{N+k} \le \lim_{k\to\infty} \frac{s_{N+1} + \cdots + s_{N+k}}{N+k} \le \lim_{k\to\infty} \frac{k(L+\epsilon)}{N+k} = L + \epsilon,$$

and since $\epsilon > 0$ was arbitrary, this shows that $C_1((a_n)) = \sum_{n\ge 1} a_n$ when this sum converges. Thus, $V_{c,1} \supseteq W$ and $C_1\big|_W = \sum$.

It is clear that C_1 is a linear map. We can say more, though. Indeed, suppose that $C_1((a_1, a_2, a_3, \ldots)) = s$. Then we claim that for any *finite* $j > 0$ we have

$$C_1((a_j, a_{j+1}, \ldots)) = s - \sum_{n=1}^{j-1} a_n. \tag{1}$$

Indeed, let s_N denote the N th partial sum of the sequence $(a_1, a_2, \ldots)$ and t_N the N th partial sum of $(a_j, a_{j+1}, \ldots)$. Then $t_N = s_{j+N}$ and

$$\begin{aligned} \lim_{N\to\infty} \frac{t_1 + t_2 + \cdots + t_N}{N} &= \lim_{N\to\infty} \frac{s_{j+1} + s_{j+2} + \cdots + s_{j+N}}{N} \\ &= \lim_{N\to\infty} \left(\frac{N+j}{N} \cdot \frac{s_1 + s_2 + \cdots + s_{j+N}}{N+j} - \frac{s_1 + s_2 + \ldots s_j}{N} \right) \\ &= \lim_{N\to\infty} \frac{s_1 + s_2 + \cdots + s_N}{N}, \end{aligned}$$

which proves our claim.

We see almost immediately that the containment $V_{c,1} \supseteq W$ is *strict*, for the sequence $(1, -1, 1, -1, \ldots)$ has partial sums

$$s_N = \begin{cases} 1 & \text{if } N \text{ is odd} \\ 0 & \text{if } N \text{ is even} \end{cases}$$

so that $s_1 + s_2 + \cdots + s_N = \lfloor (N+1)/2 \rfloor$, from which it follows that

$$C_1(1, -1, 1, -1, \ldots) = \frac{1}{2}.$$

Of course, we already *rigorously* determined this, as our manipulations above used only those properties of the operator C_1 that we have proved (namely linearity and the property (1)).

Observe, however, that the linear map C_1 does *not* extend $\sum$ to all of V. Indeed, the sequence $(1, 2, 3, 4, \ldots)$ has partial sums $s_N = N(N+1)/2$ so that $C_1((1, 2, 3, 4, \ldots)) = \lim_{N\to\infty}(N+1)(N+2)/2$, which is not finite. However, if we are to extend C_1 any further, there is only *one* value we can assign to the sum $1 + 2 + 3 + 4 + \cdots$. This follows by our argument above and linearity.

We would like to generalize the linear map C_1 to maps C_k for all $k \geq 1$. We proceed as follows: let (a_n) be any sequence and let s_N be the N th partial sum. Define

$$H_N^k = \begin{cases} \frac{1}{N}\sum_{i=1}^N s_i & \text{if } k = 1 \\ \frac{1}{N}\sum_{i=1}^N H_i^{k-1} & \text{otherwise} \end{cases}.$$

Thus we have $C_1((a_1, a_2, \ldots)) = \lim_{N\to\infty} H_N^1$ when this limit exists. We define the subset $V_{c,k} \subseteq V$ to be the set of all sequences for which the limit $\lim_{N\to\infty} H_N^k$ exists and for any $(a_1, a_2, \ldots) \in V_{c,k}$ define

$$C_k((a_1, a_2 \ldots)) := \lim_{N\to\infty} H_N^k.$$

Then it is easy to show that

(1) The subset $V_{c,l} \subseteq V$ is a subspace.
(2) $C_k : V_{c,k} \longrightarrow \mathbf{R}$ is a linear map.
(3) Fix an integer l. Then for all $k > l$ we have *strict* containment $V_{c,l} \subseteq V_{c,k}$.
(4) Let $(a_n) \in V_{c,l}$ and $k > l$. Then $C_k((a_n)) = C_l((a_n))$.

It is not difficult to see that $(1, -2, 3, -4, \ldots) \in V_2$. Indeed, the partial sums are $1, -1, 2, -2, 3, -3, \ldots$ from which it follows that

$$H_N^1 = \begin{cases} \frac{N+1}{2N} & \text{if } N \text{ is odd} \\ 0 & \text{otherwise} \end{cases},$$

and hence that

$$H_N^2 = \frac{1}{N}\sum_{i=1}^{\lfloor (N+1)/2 \rfloor}\left(\frac{1}{2} + \frac{1}{2i}\right),$$

from which we see that $\lim_{N\to\infty} H_N^2 = 1/4$, in agreement with our earlier calculation.

It can be shown by induction that if a sequence (a_n) is an element of $V_{c,k}$ then $\lim_{n\to\infty} a_n/n^k = 0$. Thus, there exist sequences that are not in $V_{c,k}$ for *any* k (one might take, for example, $1! - 2! + 3! - 4! + \cdots$).

3.2. **Abel summability.** Now we touch briefly on another important method of defining the sum of a divergent series. Let $V_a \subseteq V$ be the subset of all sequences (a_n) such that the corresponding power series $\sum_{n\geq 1} a_n x^n$ has radius of convergence 1, and represents a function $f(x)$ such that $\lim_{x\to 1^-} f(x)$ exists. We observe right away that we have the containment $W \subseteq V_a$. We define a map $A : V_a \longrightarrow \mathbf{R}$ by $A((a_n)) = \lim_{x\to 1^-} f(x)$. It is clear that this is a linear map. Moreover, suppose that $\sum a_n$ is convergent. Then $f(x) := \sum a_n x^n$ has radius of convergence 1 and by a Theorem of Abel, we have $\lim_{x\to 1^-} f(x) = \sum a_n$, so that the linear operator A, when restricted to $W \subseteq V_a$ agrees with $\sum$.

It is clear from our earlier examples that the containment $W \subseteq V_a$ is strict. For we saw that

$$A((1, -1, 1, -1, \ldots)) = \lim_{x\to 1^-}\frac{1}{1+x} = \frac{1}{2}.$$

Observe that our definition makes it impossible to claim that the sum of $1 - 1 + 1 - 1 + \cdots$ could be m/n for any $m < n$ on grounds of the identity

$$\frac{1 - x^m}{1 - x^n} = 1 - x^m + x^n - x^{m+n} + x^{2n} - \cdots,$$

because as we have pointed out, this identity does not correspond to the sequence $(1,-1,1,-1,\ldots)$, but to the sequence $(1,\underbrace{0,0,\ldots,0}_{m-1 \text{ zeroes}},-1,\underbrace{0,0,\ldots,0}_{n-m-1 \text{ zeroes}},1,\ldots)$

In the next section we will reconstruct Euler's derivation of the functional equation for the Riemann zeta function using the linear operator A.

4. Euler and the functional equation of $\zeta(s)$

The Riemann zeta function is a function of the complex variable s defined by

$$\zeta(s) = \frac{1}{1^s} + \frac{1}{2^s} + \frac{1}{3^s} + \frac{1}{4^s} + \cdots.$$

The series expansion we have written down is convergent (even absolutely) for any s with $\Re s > 1$. However, like the function $1/(1-s) = 1 + s + s^2 + s^3 + \cdots$, whose power series expansion converges only for $|s| < 1$, but which nonetheless defines an analytic function on $\mathbf{C} - \{1\}$, the Riemann zeta function can be analytically continued to a function of s for all $s \in \mathbf{C} - \{1\}$ (ζ has a simple pole at $s = 1$).

Although named for Riemann, the zeta function was studied extensively by Euler, who lived about 100 years before Riemann. In particular, Euler proved that $\zeta(s)$ can be written as an infinite product

$$\zeta(s) = \prod_p \left(1 - \frac{1}{p^s}\right)^{-1},$$

with the product being over all primes, and used this to give an analytic proof of the infinitude of primes. This viewpoint ultimately enabled Dirichlet to prove the infinitude of primes in arithmetic progressions.

One of the most important properties of the zeta function is that it satisfies a "functional equation" that relates $\zeta(s)$ to $\zeta(1-s)$:

$$\zeta(1-s) = 2(2\pi)^{-s} \cos\left(\frac{s\pi}{2}\right) \Gamma(s)\zeta(s), \tag{2}$$

where $\Gamma(s)$ is a function of the complex variable s satisfying $\Gamma(s) = (s-1)!$ for integers $s \geq 1$. This functional equation—which exhibits the symmetry of the ζ-function about the line $s = 1/2$—played a large role in Riemann's formulation of his infamous "hypothesis." What is fascinating is that Euler conjectured (2) based on calculations with divergent series over 100 years before Riemann wrote his influential paper on the zeta function. We now reconstruct Euler's arguments.

Euler works exclusively with the linear operator A. He begins with the power series

$$e^{-y} - e^{-2y} + e^{-3y} - e^{-4y} + \cdots = \frac{1}{e^y + 1},$$

which converges for all $y > 0$. By differentiating n times, he obtains

$$1^n e^{-y} - 2^n e^{-2y} + 3^n e^{-3y} - 4^n e^{-4y} + \cdots = (-1)^n \frac{d^n}{dy^n}\left(\frac{1}{e^y + 1}\right),$$

which again converges for any $y > 0$.

Now the function $1/(e^y + 1)$ can be expanded as a Taylor series about $y = 0$. Specifically, we have

$$\frac{1}{e^y + 1} = \sum_{k=0}^{\infty} a_k y^k \tag{3}$$

for some real numbers a_k. Using this information, we find that

$$\begin{aligned} A((1^n, -2^n, 3^n, -4^n, \ldots)) &= \lim_{y \to 0+} 1^n e^{-y} - 2^n e^{-2y} + 3^n e^{-3y} - 4^n e^{-4y} + \cdots \\ &= \lim_{y \to 0+} (-1)^n \frac{d^n}{dy^n}\left(\frac{1}{e^y + 1}\right) \\ &= (-1)^n n! a_n. \end{aligned}$$

On the other hand, consider the function $\pi \tan\left(\frac{\pi z}{2}\right)$. This function blows up at each odd integer $z = 2m - 1$ (because the denominator has a zeros there). However, from elementary calculus we know that

$$\lim_{z \to (2m-1)} (z - (2m-1))\pi \tan\left(\frac{\pi z}{2}\right) = -2.$$

This makes the series expansion

$$\pi \tan\left(\frac{\pi z}{2}\right) = -2 \sum_{m=0}^{\infty} \left(\frac{1}{z - (2m-1)} + \frac{1}{z + (2m-1)} \right) \tag{4}$$

plausible, and in fact this is correct (as can be shown using techniques from complex analysis). Differentiating term-wise as before, we find that

$$\frac{d^n}{dz^n}\left(\pi \tan\left(\frac{\pi z}{2}\right)\right) = 2(-1)^{n+1} n! \sum_{m=0}^{\infty} \left(\frac{1}{(z - (2m-1))^{n+1}} + \frac{1}{(z + (2m-1))^{n+1}} \right) \tag{5}$$

Since

$$\pi \tan\left(\frac{\pi z}{2}\right) = \frac{\pi}{i}\left(1 - \frac{2}{e^{\pi i z} + 1}\right),$$

we can use (3) to obtain

$$\pi \tan\left(\frac{\pi z}{2}\right) = \frac{\pi}{i}\left(1 - 2\sum_{k=0}^{\infty} a_k (\pi i)^k z^k\right). \tag{6}$$

We observe that since $\tan z$ is an *odd* function of z we necessarily have $a_0 = 1/2$ and $a_{2k} = 0$ for all $k > 0$.

Combining (5) and (6) and setting $z = 0$, we get, for *odd* values of n

$$2\sum_{m=0}^{\infty} \frac{1}{(2m-1)^{n+1}} = a_n(-\pi i)^{n+1}. \tag{7}$$

But

$$2(1 - 2^{-n-1})\left(1 + \frac{1}{2^{n+1}} + \frac{1}{3^{n+1}} + \cdots\right) = 2\sum_{m=0}^{\infty} \frac{1}{(2m-1)^{n+1}} = a_n(-\pi i)^{n+1},$$

so that

$$\zeta(1+n) = \frac{a_n(-\pi i)^{n+1}}{2(1 - 2^{-n-1})},$$

which since n is odd we may write as

$$\zeta(1+n) = \frac{a_n(-\pi)^{n+1}}{2(1 - 2^{-n-1})} \cos\left(\pi \frac{n+1}{2}\right). \tag{8}$$

On the other hand, we computed

$$\begin{aligned}(1 - 2 \cdot 2^n)\zeta(-n) &= A((1^n, -2^n, 3^n, -4^n, \ldots)) \\ &= (-1)^n n! a_n,\end{aligned}$$

whence

$$\zeta(-n) = \frac{(-1)^n n! a_n}{1 - 2^{n+1}}, \tag{9}$$

from which we can immediatly conclude that $\zeta(-2k) = 0$ for $k > 0$ since we saw above that $a_{2k} = 0$ for $k > 0$. Thus, combining (8) and (9) we get

$$2n!\zeta(n+1) = (2\pi)^{n+1} \cos\left(\pi \frac{n+1}{2}\right) \zeta(-n),$$

or by replacing n with $n - 1$ and rearranging,

$$\zeta(1-n) = 2(2\pi)^{-n} \cos\left(\frac{n\pi}{2}\right)(n-1)!\zeta(n). \tag{10}$$

Observe that even though this *a priori* holds only for *even* n, we can see that it is also true for odd $n \geq 3$ as both sides are zero in these cases. Now Euler reasons as follows: if we replace $(n-1)!$ with $\Gamma(n)$ then both sides of (10) make sense for non-integral values of n. He thus conjectures that for all complex s for which both sides make sense, one has

$$\zeta(1-s) = 2(2\pi)^{-s}\cos\left(\frac{s\pi}{2}\right)\Gamma(s)\zeta(s). \tag{11}$$

He then proceeds to verify this equation for $s = 1/2$ by using the fact that $\Gamma(1/2) = \sqrt{\pi}$.

References

[1] Hardy, G.H. *Divergent Series.* Clarendon Press, Oxford. 1949.