Methods of Summation and $\zeta(-1)$

Max Chien

Fall 2024

1 Introduction

These are notes written to accompany a presentation at the Albany High School Mathletes club on the "formula"

$$1 + 2 + 3 + \ldots = -\frac{1}{12} \tag{$*$}$$

A "proof" in terms of naive sums of sequences is given, followed by some discussion of the caveats of the interpretation of $(*)$. A brief discussion of the Riemann zeta function is given in relation to this formula. These notes are intended to be accessible to students with little to no formal math background, including calculus.

2 A "Proof"

About 10 years ago, the Youtube channel Numberphile released a series of videos which claimed the following formula:

$$1 + 2 + 3 + 4 + \ldots = -\frac{1}{12}$$

This generated a lot of discussion and response videos. Of course, many people objected to the unintuitive result. The video was especially controversial among mathematicians, though, who had many arguments about whether the derivation was rigorous.

Here I first present the proof from the original Numberphile video. I use the symbol $\sim$ rather than $=$ to emphasize that the below manipulations are somewhat unrigorous and should not be taken as equalities.[1]

We begin by denoting our desired sum by S:

$$S \sim 1 + 2 + 3 + \ldots \tag{$*$}$$

To calculate S, we first consider **Grandi's sum**,

$$S_1 \sim 1 - 1 + 1 - 1 + \ldots \tag{1}$$

[1] Of course, these manipulations are not completely unrigorous! There is a theoretical justification for each of these, but it is important to note that they require a very specific understanding of the equals sign, and it would be better to avoid this here. Details may be found in the appendix for the interested reader.

What value makes the most sense to assign to S_1?

There are three values that make sense here:

- Since the sum alternates between 1 and 0, we could set S_1 to be either 0 or 1.
- Since the sum oscillates, it is on average closest to $\frac{1}{2}$, so we can set $S_1 = \frac{1}{2}$.

In this case, we are going to choose to set $S_1 = \frac{1}{2}$. There are a couple reasons this choice is somewhat natural, which are discussed in the appendix. Here is one argument:

$$\begin{cases} \begin{array}{rl} S_1 \sim & 1 - 1 + 1 - 1 \dots \\ + \quad S_1 \sim & \quad\;\; 1 - 1 + 1 \dots \\ \hline 2S_1 \sim & 1 \end{array} \end{cases} \tag{2}$$

Now, consider the sum

$$S_2 \sim 1 - 2 + 3 - 4 + 5 - \dots \tag{3}$$

We can perform the following algebraic manipulation[2]:

$$\begin{cases} \begin{array}{rl} S_2 \sim & 1 - 2 + 3 - 4 + \dots \\ + \quad S_2 \sim & \quad\;\; 1 - 2 + 3 - \dots \\ \hline 2S_2 \sim & 1 - 1 + 1 - 1 + \dots \end{array} \end{cases} \tag{4}$$

But the right hand side of this is precisely Grandi's sum! Thus, combining Equations 2 and 4, we find that

$$2S_2 \sim S_1 \sim \frac{1}{2} \implies S_2 \sim \frac{1}{4} \tag{5}$$

Now, we can use S_2 to assign a value to S. Notice that if we subtract S_2 from S:

$$\begin{cases} \begin{array}{rl} S \sim & \;\;\, 1 + 2 + 3 + 4 + \dots \\ -\, S_2 \sim & -1 + 2 - 3 + 4 - \dots \\ \hline S - S_2 \sim & \;\;\, 0 + 4 + 0 + 8 + \dots \end{array} \end{cases} \tag{6}$$

we recover nothing more than 4 times S! Thus

$$S - S_2 \sim 4 + 8 + 12 + \dots \sim 4S \tag{7}$$

Thus we have arrived at our desired result:

$$3S \sim -S_2 \implies S \sim -\frac{S_2}{3} \sim -\frac{1}{12} \tag{8}$$

[2]This manipulation actually doesn't quite show what we want. What this really shows is that, for any choice of a method of summation (i.e. a way of assigning the values to these sums), a method of summation which is *linear* and *stable*, and assigns a value to S_2, will always assign the value $S_2 \sim \frac{1}{4}$. However, it could be the case that no such method exists! Luckily, the method of Abel summation does indeed work, so we are fine to make the conclusion $S_2 \sim \frac{1}{4}$. Again, see the appendix for more information.

3 Making Sense of Infinite Sums

Having now walked through the original proof, let's discuss why it was the target of so much debate. These disagreements were because mathematicians have a standardized convention for what it means for an infinite sum to "equal" a number, and this proof adopts a different idea of equality.

Before we discuss this, we should first understand what is usually meant when we say that an infinite sum equals a number. Consider the following sum:

$$\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \ldots \tag{9}$$

Let's look at what happens as we add up these terms. We have:

$$\frac{1}{2} = \frac{1}{2} \tag{10}$$

$$\frac{1}{2} + \frac{1}{4} = \frac{3}{4} \tag{11}$$

$$\frac{1}{2} + \frac{1}{4} + \frac{1}{8} = \frac{7}{8} \tag{12}$$

$$\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} = \frac{15}{16} \tag{13}$$

As we add each number, we get closer and closer to the value 1. In fact, we can get as close as we want to 1, simply by adding up enough of the terms. Because the "partial sums" get closer and closer to 1, we say that the entire sum *equals* (or **converges to**) 1:

$$\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \ldots = 1 \tag{14}$$

However, a sum might not approach any number at all. For instance, the sum that we are considering just gets larger and larger and approaches infinity. We say that this series **diverges**, which we sometimes write as

$$1 + 2 + 3 + 4 + \ldots = \infty \tag{15}$$

In any case, it is not proper for us to say that this sum *equals* any finite number, in the normal sense of equality.

Sums can also fail to converge in more subtle ways. For instance, we previously discussed Grandi's sum:

$$1 - 1 + 1 - 1 + \ldots \tag{16}$$

The partial sums alternate between 0 and 1, so they don't approach a single value. In this case, we also say that the sum diverges.

Thus, we see that using the typical understanding of equality for infinite sums, none of the three sums we considered in section 2 can be said to equal any number. In order to do so, we need to use some other rule for assigning values to infinite sums. A rule like this is called a **method of summation**. Methods of summation (other than the normal one) are not commonly taught in math, but they are actually occasionally used in physics. For instance, the equation $1 + 2 + 3 + 4 + \ldots = -\frac{1}{12}$ is used in string theory, and equation $1 + 1 + 1 + \ldots = -\frac{1}{2}$ is also used elsewhere. A richer discussion of methods of summation is left to the appendix for those who are interested.

4 The Riemann Zeta Function

As another method of assigning a value to $1+2+3+4+\ldots$, we employ a common strategy in math, where we first generalize the problem in order to spot patterns. Consider the following sums:

$$\begin{gathered}
\vdots \\
1^3 + 2^3 + 3^3 + \ldots \\
1^2 + 2^2 + 3^2 + \ldots \\
1 + 2 + 3 + \ldots \\
1 + 1 + 1 + \ldots \\
\frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \ldots \\
\frac{1}{1^2} + \frac{1}{2^2} + \frac{1}{3^2} + \ldots \\
\frac{1}{1^3} + \frac{1}{2^3} + \frac{1}{3^3} + \ldots \\
\vdots
\end{gathered}$$

Generally, we can write any one of these sums as $1^n + 2^n + 3^n + \ldots$ for an integer n. When $n \geq 0$, then sum still tends to infinity, so we don't really gain anything by working with these sums over our original sum. When $n = -1$ the sum is still infinite, although this is less obvious.

However, when $n \leq -2$, the sum actually converges to a very well defined value, in the typical sense that we discussed above. For instance, if $n = -2$, then

$$\frac{1}{1^2} + \frac{1}{2^2} + \frac{1}{3^2} + \ldots = \frac{\pi^2}{6} \tag{17}$$

The above equality is known the **Basel problem**. While I will not derive this result, the point is that for $n \leq -2$, the sum is well defined and no issues will occur. We can be more general and define the following function, called the **Riemann zeta function**:

$$\zeta(t) = \frac{1}{1^t} + \frac{1}{2^t} + \frac{1}{3^t} + \ldots \tag{18}$$

for any number $t > 1$.

We can go one step further and allow t to be a complex number as well. I won't go into what it means to raise a number to a complex power here. The reason that considering complex powers is helpful is that due to a (rather magical) property of complex functions known as **analytic continuation**, we can "extend" ζ to define it even when the sum itself doesn't properly converge. Moreover, this definition is unique, meaning that there is no *other* value that we could define for ζ. When we do this, the unique extended value for -1 is

$$1 + 2 + 3 + 4 + \ldots \sim \zeta(-1) = -\frac{1}{12} \tag{19}$$

Thus, if we consider $1+2+3+4+\ldots$ as part of a family of sums of the form $1^t+2^t+3^t+\ldots$, the only value that makes sense to assign to the sum is $-\frac{1}{12}$.[3]

While this method does provide another source of justification for the statement $1 + 2 + 3 + 4 + \ldots = -\frac{1}{12}$, the Riemann zeta function on its own is one of (arguably the) most important functions in math. For instance, it is deeply connected with the distribution of prime numbers, which are central to modern cryptography.[4]

One of the most important properties of the Riemann zeta function is the values for which it is equal to zero. It is not too difficult to show that $\zeta(s)$ is zero whenever s is a negative even integer. The famous **Riemann hypothesis** asserts that the only other zeroes of the Riemann zeta function are complex numbers of the form $\frac{1}{2} + it$. This problem has been unsolved for over 150 years, and it is one of the Millenium problems, which comes with a $1 million reward for a proof (or disproof).

Appendix: Methods of Summation

In this section, I elaborate a bit more for those interested on methods of summation, which were briefly mentioned in section 3. As defined previously, a method of summation is some way of assigning values to infinite sums. Note that a method of summation is not required to assign a value to every infinite sum.

In our derivations of the value of S_1, S_2, S, we made use of term-by-term addition and shifting of sequences. In other words, we implicitly assumed that our method of summation respects these two operations. Methods that do are called linear and stable, respectively.

A technical point is also that linearity and stability only require that the properties hold when a method of summation actually assigns a value to all the sequences involved. Therefore, when we earlier "proved" that $1-1+1-1+\ldots = \frac{1}{2}$, or that $1+2+3+4+\ldots = -\frac{1}{12}$, we were actually just showing that any method of summation which is linear, stable,[5] *and* assigns a value to those sums, must assign precisely the values we derived. However, it might be the case that no such method exists!

For S_1 and S_2, it turns out that this is not a problem, since Cesaro summation works for S_1 and Abel summation works for both (see sections A.1 and A.2, respectively).

However, it can be shown that there is no linear and stable method of summation which

[3]This is called zeta function regularization.

[4]For those curious about other applications of modern math to cryptography, it might be of interest to read about elliptic curve cryptography.

[5]Technically we also need the summation to be *regular*; this just means it assigns the correct values for sums converging in the normal sense.

assigns a value to $1+2+3+4+\ldots$! This is actually a major problem for the proof that we presented above, and it means that the last step is a total failure. Instead, we must use a method like zeta function regularization (which is nonlinear) or Ramanujan summation (which is not stable).

A.1 Cesaro Sums

When we considered Grandi's sum,

$$1-1+1-1+\ldots$$

We observed that the sum is, "on average," about $\frac{1}{2}$. Cesaro summation is a method of summation which formalizes this idea.

For a sequence $a_1, a_2, a_3, \ldots$, the **Cesaro mean** of the sequence is defined to be the limit of the running averages:

$$C = \lim_{n\to\infty} \frac{a_1 + a_2 + \ldots + a_n}{n}$$

It can be shown that if the sequence converges, then the Cesaro mean is the same as the limit of the sequence. So if we define the **Cesaro sum** of a sum $a_1 + a_2 + a_3 + \ldots$ to be the Cesaro mean of the partial sums:

$$C^{\Sigma} = \lim_{n\to\infty} \frac{s_1 + s_2 + \ldots + s_n}{n} = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \sum_{i=1}^{k} a_i$$

then C^{Σ} is indeed regular. It can also be shown that C^{Σ} is linear and stable.

Using Cesaro summation, it is easy to see that $S_1 \sim \frac{1}{2}$. The partial sums alternate between 0 and 1, so the Cesaro sum is

$$S_1 = \lim_{n\to\infty} \frac{1+0+1+0+\ldots+0}{n} = \frac{1}{n}\left\lceil \frac{n}{2} \right\rceil = \frac{1}{2}$$

A.2 Abel Sums

While Cesaro summation does help us assign a value to Grandi's series, it fails when we consider the series

$$1-2+3-4+\ldots$$

Abel summation is a method of summation which uses power series from calculus. If you have learned about power series, consider the MacLaurin expansion of $\frac{1}{(1+x)^2}$. This is given by

$$\frac{1}{(x+1)^2} = \sum_{n=0}^{\infty} (n+1)(-1)^n x^n = 1 - 2x + 3x^2 - 4x^3 + \ldots$$

Notice that if we naively plug in $x = 1$, we get

$$\frac{1}{4} \sim 1-2+3-4+\ldots \sim S_2$$

However, we cannot simply plug in $x = 1$. The radius of convergence of this power series is 1, but it is divergent at 1. Instead, we can let x get closer and closer to 1, and see what value it approaches.

Given a series $a_0 + a_1 + a_2 + \ldots$, the Abel sum A^Σ is defined to be this limiting value:

$$A^\Sigma = \lim_{x \to 1^-} \sum_{n=0}^{\infty} a_n x^n$$

whenever the limit exists. As with Cesaro summation, it can be shown that if the series converges in the usual sense, then the Abel sum is equal to the normal sum. Moreover, any Cesaro summable series is also Abel summable, and it has the same sum. By the properties of limits and infinite sums from calculus, A^Σ is linear and stable.

Using Abel summation, we can avoid the issues with just plugging in $x = 1$ by taking the limit:

$$S_2 = \lim_{x \to 1^-} \sum_{n=0}^{\infty} (n+1)(-1)^n x^n = \lim_{x \to 1^-} \frac{1}{(x+1)^2} = \frac{1}{4}$$