

Summing the First n Natural Numbers

Duncan Samson

St Andrew's College & The Diocesan School for Girls, Makhanda (Grahamstown)

d.samson@sacschool.com

INTRODUCTION

In the mid-1780s, Carl Friedrich Gauss was a young pupil at a provincial German public primary school. The schoolmaster, endeavouring to keep the class occupied for half an hour or so, gave the somewhat tedious exercise of adding up the first 100 positive integers, i.e. $1 + 2 + 3 + \dots + 98 + 99 + 100$. No sooner had the assignment been set, so the legend goes, young Gauss quickly produced the correct answer of 5050. Rather than systematically adding each whole number, a rather laborious and time-consuming process, he noticed that if he paired the first number with the last number ($1 + 100$), the second number with the second-last number ($2 + 99$), and so on, then each pair had the same sum, namely 101. And since there were 50 such pairs then the total sum was simply $50 \times 101 = 5050$.

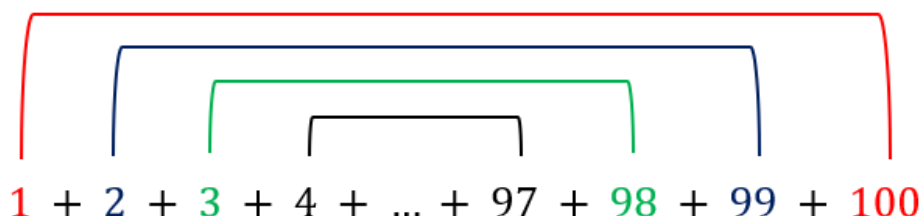


FIGURE 1: Gauss's elegant solution for summing the series.

I am sure that most of us are familiar with this tale, and no doubt many of us use it as a way of introducing arithmetic series. For a general arithmetic sequence with first term a and common difference d , the formula for the sum of the first n terms, i.e. $S_n = \frac{n}{2}[2a + (n - 1)d]$, can be derived directly from Gauss's approach, as illustrated in Figure 2.

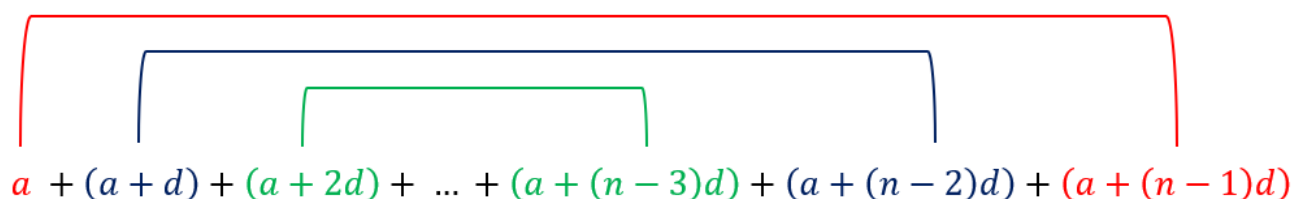


FIGURE 2: Applying Gauss's method to a general arithmetic series.

Pairing the terms as before, first with last, second with second-last, and so on, each pair has the sum $2a + (n - 1)d$. And since there are n terms in total, there are $\frac{n}{2}$ pairs. Multiplying the number of pairs by the sum of each pair gives us the standard formula for an arithmetic series, $S_n = \frac{n}{2}[2a + (n - 1)d]$. In the case of there being an odd number of terms, where the pairing technique would leave a single unpaired term in the middle of the sequence, then this approach still works if one simply pairs the middle term with itself and treats it as 'half a pair'. For example, if we wanted to apply this method to the sum of the first 9 natural numbers, $1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 9$, each pair (including the middle term paired with itself) has a sum of 10, and there are a total of $4\frac{1}{2}$ pairs.

Interestingly, most formal proofs of this formula found in school textbooks take the approach of writing out the terms of the sequence (expressed in terms of a , d and n), then reversing the sequence and writing it out again below the first sequence. Adding in columns shows that each pair of terms sums to $2a + (n - 1)d$, thus $2S_n = n[2a + (n - 1)d]$ from which the standard formula immediately follows by dividing through by 2. I suspect this latter approach is used to avoid the seemingly problematic issue of dealing with series containing an odd number of terms. As previously discussed, one can work around this problem by pairing the middle term with itself and treating it as half a pair, but one can see why this latter approach may be considered a more ‘convincing’ derivation of the formula. Nonetheless, Gauss’s pairing method gives a very simple way of remembering both forms of the S_n formula for an arithmetic series, $S_n = \frac{n}{2}[2a + (n - 1)d]$ and $S_n = \frac{n}{2}[a + l]$, the second of which is no longer given on the formula sheet provided in the matric examination.

SUMMING THE FIRST n NATURAL NUMBERS

Let us now move to the main focus of this article, namely calculating the sum of the first n natural numbers. Exploring different ways of doing this, using a specific number of terms as a generic example (chosen for convenience), is a valuable classroom exercise, not only because it allows for varied and differentiated approaches, but because it provides a useful platform for developing a sense of generality by moving from the specific to the general. As a result, it also allows pupils to arrive at different yet algebraically equivalent expressions of generality. Proving the equivalence of such expressions is also a useful way to practise algebraic manipulation skills in a meaningful context. In general, the sum of the first n natural numbers is given by the formula $S_n = \frac{n(n+1)}{2}$.

APPROACH 1: PAIRING TERMS

The first approach is the Gaussian pairing method previously described. Using this approach to determine the sum of the first 10 natural numbers (for example) is illustrated in Figure 3.

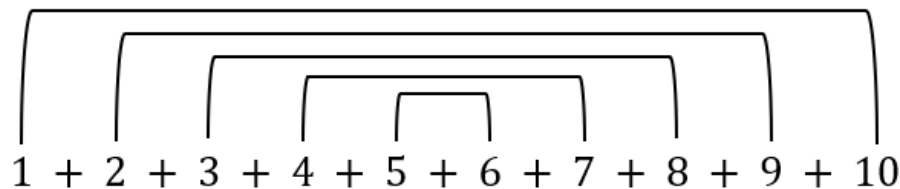


FIGURE 3: Using a pairing technique to sum the first 10 natural numbers.

Since there are 5 pairs, and each pair adds to 11, the sum of the first 10 natural numbers is $5 \times 11 = 55$. More generally, if we were adding the first n natural numbers there would be $\frac{n}{2}$ pairs each with a sum of $(n + 1)$. The sum of the first n natural numbers is thus $S_n = \frac{n}{2}(n + 1)$.

What if the series contained an odd number of terms? How might we determine the sum of the first 9 natural numbers for example? One approach might be to work with the first 8 numbers, using the pairing technique as before, and then adding on the final term at the end. This approach is illustrated in Figure 4. In general, for the sum of the first n terms (where n is odd), there are $\frac{n-1}{2}$ pairs, each with a sum of n . This gives a total of $\frac{n-1}{2} \times n$ for the pairs. Adding the final term to this gives $S_n = \frac{n-1}{2} \times n + n$ which simplifies to the standard formula $S_n = \frac{n(n+1)}{2}$.

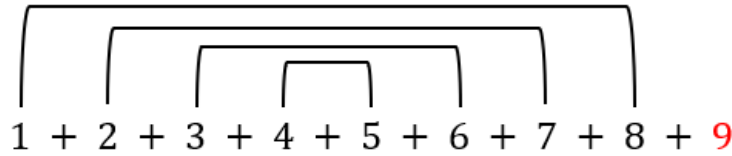


FIGURE 4: Separating the final term and pairing the first 8 terms.

An alternative approach might be to pair the first term with the last, the second term with the second-last, and so on as before, but to leave the middle term unpaired. This modified approach for the sum of the first 9 natural numbers is illustrated in Figure 5.

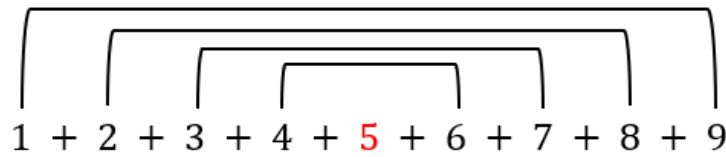


FIGURE 5: Isolating the middle term.

In general, for the sum of the first n terms (where n is odd), there are $\frac{n-1}{2}$ pairs, each with a sum of $n+1$. This gives a total of $\frac{n-1}{2} \times (n+1)$ for the pairs. Because of the symmetry of the series, the middle term will be the arithmetic mean of the first and last terms, i.e. the middle term can be expressed as $\frac{n+1}{2}$. Adding this middle term to the sum of the pairs gives $S_n = \frac{n-1}{2} \times (n+1) + \frac{n+1}{2}$ which simplifies to the standard formula $S_n = \frac{n(n+1)}{2}$.

A third alternative might be to pair the middle term with itself, and then to subtract the 'extra' middle term at the end to correct for the overcount. This approach is illustrated in Figure 6.

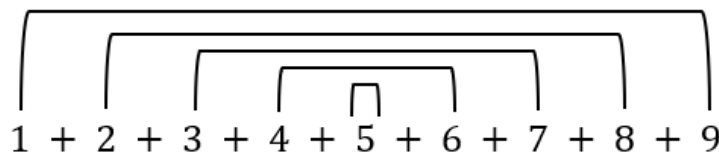


FIGURE 6: Pairing the middle term with itself.

Using this approach, then in general there would be $\frac{n+1}{2}$ pairs, each with a sum of $n+1$, giving a total of $\frac{n+1}{2} \times (n+1)$ for the pairs. In order to correct for the overcount we need to subtract the 'extra' middle term. As before, using the symmetry of the series, the middle term can be expressed as $\frac{n+1}{2}$. Subtracting this from the sum of the pairs gives $S_n = \frac{n+1}{2} \times (n+1) - \frac{n+1}{2}$ which, as before, simplifies to the standard formula $S_n = \frac{n(n+1)}{2}$.

What I particularly like about these last two approaches is that it really directs one to engage with the generality of the context. Showing the algebraic equivalence of the various expressions, while simple enough, also allows for multiple approaches, for example multiplying out and factorising, or more directly (and more elegantly) taking out a binomial bracket as a common factor.

APPROACH 2: USING THE AVERAGE OF THE SERIES

Since an arithmetic series is by definition a sequence of equally spaced terms, the symmetry of the series means that the arithmetic mean will be the same value as the median. For an odd number of terms, the median value is simply the middle term. For an even number of terms, the median is the value halfway between the middle two terms, i.e. the arithmetic mean of the middle two terms. Note that since the terms of an arithmetic series are equally spaced, the median value, and consequently the mean value, will also be the arithmetic mean of the first and last terms (or indeed any two terms equidistant from the ‘middle’). We now have two very simple ways of determining the mean of an arithmetic series – we can either determine the median value, or simply determine the average of the first and last terms. Knowing the mean value of a series is very powerful if one also knows how many terms there are in the series, since the sum of the series is simply the product of the mean and the number of terms.

Let us now consider this approach for the sum of the first 10 natural numbers as illustrated in Figure 7.

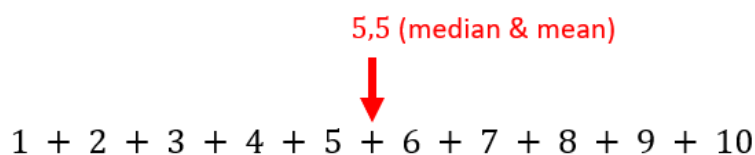


FIGURE 7: Summing using the arithmetic mean.

We can visually see that the mean value is 5,5. We could also calculate it by taking the average of the two middle terms, $\frac{5+6}{2} = 5,5$, or by taking the average of the first and last terms, $\frac{1+10}{2} = 5,5$. The sum of the series can then be calculated by multiplying this mean with the number of terms. The sum of the first 10 natural numbers is thus $5,5 \times 10 = 55$. Expressing this in general for the sum of the first n natural numbers, the mean value is simply $\frac{n+1}{2}$ (the average of the first and last terms), and multiplying this by the number of terms (n) gives the standard formula directly: $S_n = \frac{(n+1)}{2} \times n$.

This approach also allows us to re-interpret the S_n formula of a general arithmetic series. For a general arithmetic series with first term a and common difference d , the n th term is $T_n = a + (n - 1)d$. The mean value of the first n terms is thus $\frac{a + [a+(n-1)d]}{2}$, the average of the first and last terms. The sum of the first n terms of the series is then simply this mean value multiplied by the number of terms, n . This gives us the expression $S_n = \frac{a + [a+(n-1)d]}{2} \times n$, which is of course equivalent to $S_n = \frac{n}{2}[2a + (n - 1)d]$. The alternative S_n formula of a general arithmetic series, $S_n = \frac{n}{2}[a + l]$, can also be interpreted using this approach, where $\frac{[a + l]}{2}$ is simply the average of the first and last terms.

APPROACH 3: USING PARTIAL SUMS

An alternative approach is to consider the partial sums of the series. We again consider the sum of the first n natural numbers, but this time we look at the pattern formed by the partial sums S_1 , S_2 , S_3 and so on.

$$S_1 = 1$$

$$S_2 = 1 + 2 = 3$$

$$S_3 = 1 + 2 + 3 = 6$$

$$S_4 = 1 + 2 + 3 + 4 = 10$$

$$S_5 = 1 + 2 + 3 + 4 + 5 = 15$$

The series of partial sums forms a quadratic sequence, i.e. a sequence which has a constant second difference. This is illustrated in Figure 8.

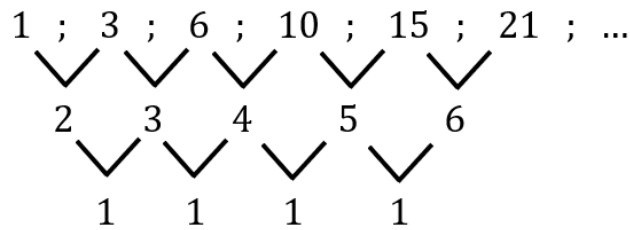


FIGURE 8: Looking for a pattern in the partial sums.

We can now determine the formula for the quadratic sequence in the form $T_n = an^2 + bn + c$ using any of the standard approaches. The one that pupils are perhaps most familiar with is setting the second difference equal to $2a$, the first term of the row of first differences equal to $3a + b$, and the first term of the original sequence equal to $a + b + c$.

$$2a = 1 \rightarrow a = \frac{1}{2}$$

$$3a + b = 2 \rightarrow 3\left(\frac{1}{2}\right) + b = 2 \rightarrow b = \frac{1}{2}$$

$$a + b + c = 1 \rightarrow \frac{1}{2} + \frac{1}{2} + c = 1 \rightarrow c = 0$$

The quadratic formula for the sequence of partial sums, and hence for the sum of the first n natural numbers, is thus $T_n = \frac{1}{2}n^2 + \frac{1}{2}n$, which is of course equivalent to $S_n = \frac{n(n+1)}{2}$ as before.

APPROACH 4: USING A VISUAL ANALOGUE

Another way of determining the sum of the first n natural numbers is by using the technique commonly used for deriving the formula for the sum of a general arithmetic sequence. As previously alluded to, this entails writing out the original series, reversing the series and writing this below the original series, and then adding the columns (i.e. the pairs of terms) to get twice the sum. Dividing by 2 then gives the final tally. This process is illustrated in Figure 9 for the sum of the first 7 natural numbers.

$$\begin{array}{r} S = 1 + 2 + 3 + 4 + 5 + 6 + 7 \\ S = 7 + 6 + 5 + 4 + 3 + 2 + 1 \\ \hline 2S = 8 + 8 + 8 + 8 + 8 + 8 + 8 \end{array}$$

FIGURE 9: Reversing the original series to create columns (pairs) with equal sums.

We thus have $2S = 7 \times 8$ from which $S = \frac{7 \times 8}{2}$. What is pleasing about this particular approach is that it has a very compelling visual analogue, as illustrated in Figure 10. Representing each term of the original series by a corresponding number of blocks (or similar), and then doing the same for the ‘reversed’ series, allows the two series visually to combine to form a 7 by 8 rectangular array. The original series, which is half the rectangular array, thus has a sum of $\frac{7 \times 8}{2}$. Note that this approach can also be thought of as a visual extension of Approach 3, since each triangular structure represents a partial sum.

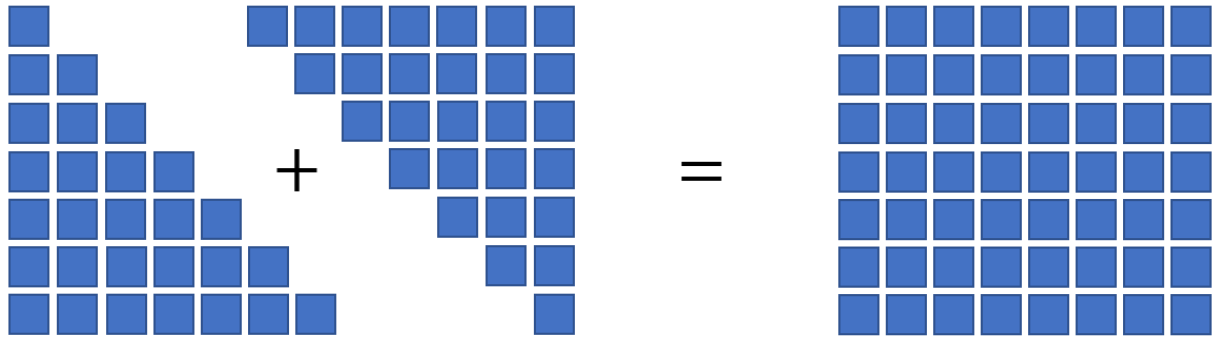


FIGURE 10: Combining two series to form a rectangular array.

If we use this visual analogue more generally for the sum of the first n natural numbers, then the addition of the original series with the ‘reversed’ series results in an n by $n + 1$ rectangular array (Figure 11).

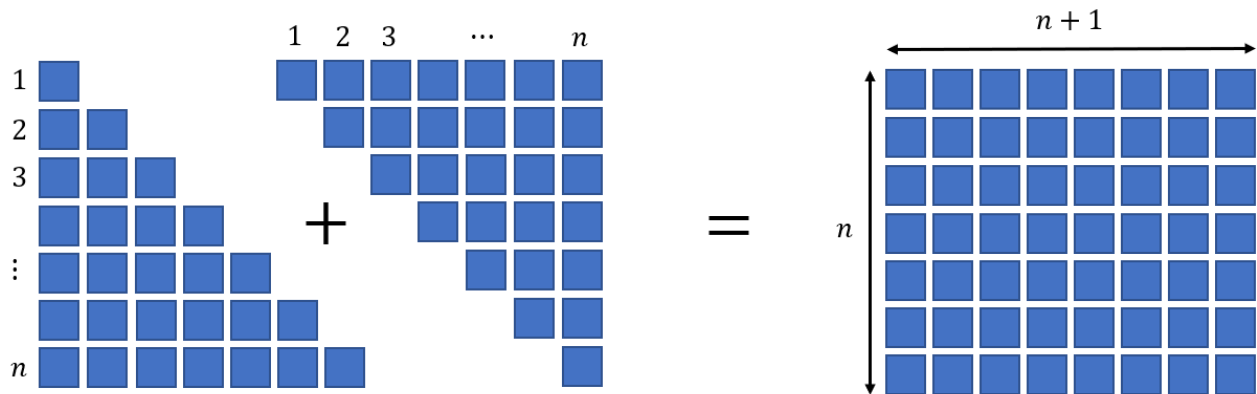


FIGURE 11: Generalising the visual analogue.

Expressed symbolically:

$$\sum_{i=1}^n i + \sum_{i=1}^n i = n(n+1)$$

The sum of the original series is thus simply half this answer, i.e. $S_n = \frac{n(n+1)}{2}$ as before.

APPROACH 5: USING COMBINATORICS

One of the great joys of mathematics is finding connections between different areas of mathematics – particularly when they are unexpected. As a case in point, the formula for the sum of the first n natural numbers is equivalent to the number of ways of choosing two distinct items from a set of $n + 1$ different items. In other words, $S_n = {}^{n+1}C_2$. The standard formula for choosing r items from n different items is the combination ${}^nC_r = \frac{n!}{r!(n-r)!}$. Using this formula for the case of choosing two distinct items from a set of $n + 1$ different items gives:

$${}^{n+1}C_2 = \frac{(n+1)!}{2!(n-1)!} = \frac{(n+1) \times n \times (n-1)!}{2! (n-1)!} = \frac{n(n+1)}{2}$$

Let us try to make a bit more sense of why this combinatorial approach gives the correct formula. Imagine that we need to choose two distinct natural numbers from the first $n + 1$ natural numbers. If one of our chosen numbers is 1, then there are n possible pairs: $(1, 2); (1, 3); (1, 4); \dots; (1, n + 1)$. These are all the possible pairings that include 1. If one of our chosen numbers is 2, then there are only $n - 1$ possible pairs, since we cannot choose 1 again: $(2, 3); (2, 4); (2, 5); \dots; (2, n + 1)$. Continuing this process we see that there are $n - 2$ possible pairings where one of the numbers is 3, $n - 3$ possible pairings where one of the numbers is 4, and so on. We can summarise this as follows:

$(1, 2); (1, 3); (1, 4); (1, 5); \dots; (1, n + 1)$	n pairs
$(2, 3); (2, 4); (2, 5); \dots; (2, n + 1)$	$n - 1$ pairs
$(3, 4); (3, 5); \dots; (3, n + 1)$	$n - 2$ pairs
$\vdots$	$\vdots$
$(n - 1, n); (n - 1, n + 1)$	2 pairs
$(n, n + 1)$	1 pair

The total number of distinct pairs is thus $1 + 2 + 3 + \dots + (n - 2) + (n - 1) + n$, i.e. the sum of the first n natural numbers. Also notice that each collection of pairs can be seen as a row/column of the visual analogue illustrated in Figure 11, which is rather pleasing.

CONCLUDING COMMENTS

The purpose of this article was to explore different ways of calculating the sum of the first n natural numbers, using numeric examples as a route to explaining the general expression for the sum. The first approach involved grouping numbers in pairs, where each pair has the same sum. This proved an interesting approach as it allows for subtle variations depending on whether the series contains an odd or even number of terms. The second approach involved using the mean value of the series and has the potential to open up interesting avenues of discussion based on the inherent symmetry of general arithmetic series. The third approach made use of partial sums, and links directly to the standard technique of determining the general term of a quadratic sequence. The fourth approach resonates with the standard derivation of the S_n formula for a general arithmetic series, and is powerfully supported by a compelling visual analogue. The fifth approach, intended more as an extension, highlights the interconnectedness of mathematical ideas and topics.

What I hope each of these approaches highlights is the development of a sense of generality as one moves from a specific instantiation to the general case. It is this progression from the specific to the general that lies at the heart of mathematics, and as a mathematical process of logical reasoning it is something that we should nurture at every possible opportunity. Finally, although each of the five methods discussed is different in its own way, there are nonetheless strong connections between them. It is this aspect of mathematical interconnectedness that I hope comes through in these different approaches.