Finding a Formula for the
Sum of Square Numbers

Alister McKinley

We start the sequence with $1^2 = 1$. The next term is $1^2 + 2^2 = 5$, the 3rd term is $1^2 + 2^2 + 3^2 = 14$, and so forth:

$$1, 5, 14, 30, 55, 91, 140, 204, 285, 385, \ldots$$

We can define the sequence as follows:

$$S = \{S_n\}_{n=1}^{n=\infty}$$
$$\text{where } S_n = \sum_{k=1}^{n} k^2$$

If the sequence can be represented by a polynomial, then repeated application of the difference operator on the terms will eventually reveal a constant difference. The number of applications required to arrive at the constant difference will be the degree of the polynomial.

The difference operator works by generating a new sequence, made from the difference between successive terms. For example, the difference operator applied to $\{1, 3, 5\}$ will yield $\{2, 2\}$ because the difference between 1 and 3 is 2, and the difference between 3 and 5 is 2.

Repeatedly applying the difference operator to our sequence:

$$\begin{aligned}
S &= \{1, 5, 14, 30, 55, 91, \ldots\} \\
\Delta S &= \{5 - 1, 14 - 5, 30 - 14, 55 - 30, 91 - 55, \ldots\} \\
&= \{4, 9, 16, 25, 36, \ldots\} \\
\Delta^2 S &= \{9 - 4, 16 - 9, 25 - 16, 36 - 25, 36, \ldots\} \\
&= \{5, 7, 9, 11, 13, \ldots\} \\
\Delta^3 S &= \{7 - 5, 9 - 7, 11 - 9, 13 - 11, 13, \ldots\} \\
&= \{2, 2, 2, 2, 2, \ldots\}
\end{aligned}$$

Three applications of the difference operator has created a constant difference, implying that the terms of our sequence S can be represented with a cubic polynomial.

$$\text{Assume: } S_n = an^3 + bn^2 + cn + d$$

Applying the difference operator three times to our polynomial:

$$S_n = an^3 + bn^2 + cn + d$$
$$\Delta S_n = a(3n^2 + 3n + 1) + b(2n + 1) + c(1)$$
$$= a(3n^2 + 3n + 1) + b(2n + 1) + c$$
$$\Delta^2 S_n = a(3(2n + 1) + 3(1)) + b(2(1))$$
$$= a(6n + 6) + 2b$$
$$\Delta^3 S_n = a(6(1))$$
$$= 6a$$

Evaluating the difference operator equations at the first element:

$$S_1 = a(1)^3 + b(1)^2 + c(1) + d$$
$$1 = a + b + c + d$$
$$\Delta S_1 = a(3(1)^2 + 3(1) + 1) + b(2(1) + 1) + c$$
$$4 = 7a + 3b + c$$
$$\Delta^2 S_1 = a(6(1) + 6) + 2b$$
$$5 = 12a + 2b$$
$$\Delta^3 S_1 = 6a$$
$$2 = 6a$$

This produces a system of equations that we can solve to determine our unknown coefficients:

$$\begin{cases} a + b + c + d = 1 \\ 7a + 3b + c = 4 \\ 12a + 2b = 5 \\ 6a = 2 \end{cases} \implies \begin{cases} a = \frac{1}{3} \\ b = \frac{1}{2} \\ c = \frac{1}{6} \\ d = 0 \end{cases}$$

This suggests that each term in our sequence can be represented with the following formula:

$$S_n = \sum_{k=0}^{n} k^2$$
$$= an^3 + bn^2 + cn + d$$
$$= \frac{1}{3}n^3 + \frac{1}{2}n^2 + \frac{1}{6}n$$
$$= \frac{n}{6}(2n^2 + 3n + 1)$$
$$= \frac{n}{6}(n + 1)(2n + 1)$$

Since we started by assuming our terms can be represented with a cubic polynomial, so now we need to verify that the result holds. We will use induction to prove that the determined polynomial describes each term in our sequence.

First we start with the base case $n = 1$:

$$S_1 = \frac{1}{6}(1 + 1)(2 \cdot 1 + 1)$$
$$= \frac{1}{6} \cdot 2 \cdot 3$$
$$= 1$$
$$= 1^2$$

Next we make the induction hypothesis (I. H.):

$$\text{Suppose } S_k = \sum_{i=1}^{k} i^2 = \frac{k}{6}(k + 1)(2k + 1) \quad \forall k \geq 1$$

Now we take the induction step by considering the $(k + 1)$ case:

Want to show:
$$S_{k+1} = \frac{(k + 1)}{6}((k + 1) + 1)(2(k + 1) + 1)$$

Consider:
$$S_{k+1} = \sum_{i=1}^{k+1} i^2$$
$$= \sum_{i=1}^{k} i^2 + (k + 1)^2$$
$$= \frac{k}{6}(k + 1)(2k + 1) + (k + 1)^2 \qquad \text{(by I. H.)}$$
$$= \frac{1}{6}k(k + 1)(2k + 1) + \frac{6}{6}(k + 1)^2$$
$$= \frac{k + 1}{6}(k(2k + 1) + 6(k + 1))$$
$$= \frac{k + 1}{6}(2k^2 + 7k + 6)$$
$$= \frac{k + 1}{6}(k + 2)(2k + 3)$$
$$= \frac{(k + 1)}{6}((k + 1) + 1)(2(k + 1) + 1) \quad \text{(as desired)}$$

This proves that the polynomial we found describes the terms in our sequence, so we can conclude that:

$$\sum_{k=1}^{n} k^2 = 1^2 + 2^2 + 3^2 + \cdots + n^2 = \frac{n}{6}(n + 1)(2n + 1)$$