Faulhaber's Formula and its Connection to Discrete Calculus

Abdul Musthakin

March 2025

I Prerequisites

You need very little to understand everything that will be discussed. A decent number of new concepts will be introduced, but I plan on everything being self-contained. The most important 'prerequisite' is the ability to understand new things, and to follow the simple logic that will be used to go from one equation or formula to another.

Additionally, knowledge of basic calculus (derivatives, integrals, and the relationship between them) would certainly aid in one's appreciation for the results we will derive. Whilst someone could technically understand everything without it, I do not think that there would be as much value in doing so. Connections between that which is seemingly unrelated lies at the heart of what I deem to be interesting in mathematics – which happens to be most things. There is a very particular beauty in these connections. The more you know beforehand, the more you have to link new concepts to.

Summation notation is required to be able to read, and by extension, understand, the maths presented here. You could pick it up via the simple examples given in the introduction. Regardless, prior experience is expected, and it is practically necessary to understand the way we will deal with sums in the later parts of the essay.

II Introduction

The formula for the sum of the first n positive integers is something that many people learn fairly early in school.

$$\sum_{i=1}^{n} i = 1 + 2 + 3 + \ldots + n = \frac{n(n+1)}{2}$$

The sum of the squares of the first n positive integers, as well as the sum of the cubes, have reasonably well-known formulae as well.

$$\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}$$
$$\sum_{i=1}^{n} i^3 = \frac{n^2(n+1)^2}{4}$$

These formulae are good and all, but how do we prove them? We can take an algebraic approach, a geometric one, or we may simply use induction. The latter method will always work as long as we already have an algebraic expression for the sum in terms of n. What if we do not have the end result to prove? Thus, the problem arises – deriving a formula for the sum of the k-th powers of the first n positive integers: Faulhaber's formula.

$$S_k(n) = \sum_{i=1}^{n} i^k = 1^k + 2^k + ... + n^k = \ ?$$

S_k is just another shorthand introduced; the previous three sums can be referred to as S_1, S_2, and S_3. There are several ways of deriving Faulhaber's formula, and the method that I will use is not the most common one. Whether or not it is the "best" way cannot really be answered. To me, however, it is certainly the most interesting method.

III Necessary Definitions

First, let us suppose that we have some sequence a, the n-th element of which is denoted by a_n. Now, let us define the forward difference operator, Δ, as follows.

$$\Delta_n a_n = a_{n+1} - a_n \tag{1}$$

Note that Δ applies to the sequence a, and not to n-th element of the sequence. You could write $\Delta_n a_n$ as $(\Delta a)_n$, which may technically be more correct – but it is more cumbersome to use. This is just a minor detail.

The indefinite sum, or the antidifference, of a sequence can be defined in relation to the forward difference.

$$\Delta_n \sum_n a_n = a_n \tag{2}$$

That is, if $\sum_n a_n = A_n$, then

$$A_{n+1} - A_n = a_n. \tag{3}$$

Now, you may notice that this seems familiar. Basic knowledge of calculus allows us to spot a connection. We are dealing with differences and sums, but not in the way that found in regular calculus. That is because this is not regular calculus. Instead, we have ventured into the realm of the discrete, as opposed to the continuous which is much more familiar.

The forward difference is the discrete analogue to the derivative. Whilst the latter measures the instantaneous rate of change of a function at some point,

the former measures the rate of change of a sequence between two points. Similarly, the antidifference is the discrete analogue to the antiderivative (indefinite integral).

Just as the fundamental theorem of calculus relates derivatives to integrals, we can construct a discrete version of it which relates forward differences to sums.

$$\Delta_n \sum_{i=n_0}^{n} a_i = a_{n+1} \tag{4}$$

$$\sum_{i=n_0}^{n} \Delta_i a_i = a_{n+1} - a_{n_0} \tag{5}$$

The proof of these two statements is rather simple, only requiring the use of the definitions that we have already laid out, and the properties of sums. The proof of the first statement is as follows.

$$\begin{aligned}
\Delta_n \sum_{i=n_0}^{n} a_i &= \sum_{i=n_0}^{n+1} a_i - \sum_{i=n_0}^{n} a_i \\
&= a_{n+1} + \sum_{i=n_0}^{n} a_i - \sum_{i=n_0}^{n} a_i \\
&= a_{n+1} \quad \square
\end{aligned}$$

The proof of the second statement does not require much more work than the first.

$$\begin{aligned}
\sum_{i=n_0}^{n} \Delta_i a_i &= \sum_{i=n_0}^{n} (a_{i+1} - a_i) \\
&= \sum_{i=n_0}^{n} a_{i+1} - \sum_{i=n_0}^{n} a_i \\
&= \sum_{i=n_0+1}^{n+1} a_i - \sum_{i=n_0}^{n} a_i \\
&= a_{n+1} - a_{n_0} + \sum_{i=n_0}^{n} a_i - \sum_{i=n_0}^{n} a_i \\
&= a_{n+1} - a_{n_0} \quad \square
\end{aligned}$$

The relationship between the definite sum and its indefinite counterpart follows from the second statement.

$$\sum_{i=n_0}^{n} a_n = \sum_{i=n_0}^{n} \Delta_n A_n = A_{n+1} - A_{n_0} \tag{6}$$

Whilst we will not mention indefinite sums again, I thought that their inclusion here would serve to further illustrate the connection between regular old calculus, and what we have started to formulate. More connections are to be found as we continue on.

There is just one more definition that needs to be made (kind of). That is the falling factorial. There are a few different notations for it, but the notation that I will be using is $x^{\underline{k}}$. This may be read as "x to the falling k". You can see the falling factorial as just the regular factorial, but with a specified end point.

$$x^{\underline{k}} = x(x-1)(x-2)\cdots(x-k+1). \tag{7}$$

IV Formulating a Plan

What we will do next makes a lot more sense if you see everything we have done so far as creating this discrete version of calculus. The two most important concepts in calculus, the derivative and the integral, have already been translated over. Consider the fact that you can find the derivative of any polynomial very easily with the power rule. It would only make sense for there to be some discrete version of this – a similar-looking rule to evaluate the forward difference of certain expressions. Turns out there is, and it involves the falling factorial.

$$\begin{aligned}
\Delta_x x^{\underline{k}} &= (x+1)^{\underline{k}} - x^{\underline{k}} \\
&= (x+1)x(x-1)\cdots(x-k+2) - x(x-1)(x-2)\cdots(x-k+1) \\
&= x(x-1)\cdots(x-k+2)\cdot[(x+1)-(x-k+1)] \\
&= kx(x-1)\cdots(x-k+2) \\
&= kx^{\underline{k-1}}
\end{aligned} \tag{8}$$

Once again, we have an eerily familiar result. Well, we best make use of it.

It is important to always be clear on what our goal is. We have a sum to evaluate, and we have created some tools that resemble familiar concepts in calculus. If instead of a sum, we had an integral to evaluate, one way of doing so would be to express the integrand (everything inside of the integral) as the derivative of something else. Using the fact that derivatives and integrals are opposites of each other, we could then 'cancel them out' to get our answer. This is the essence of the method that we will be using.

The sum of the falling factorial can be derived with what we have. We simply have to rewrite it in terms of the forward difference of some expression, and then utilise (4).

$$\begin{aligned}
\sum_{i=1}^{n} i^{\underline{k}} &= \sum_{i=1}^{n} \Delta_i \frac{i^{\underline{k+1}}}{k+1} \\
&= \frac{(n+1)^{\underline{k+1}}}{k+1} - \frac{1^{\underline{k+1}}}{k+1} \\
&= \frac{(n+1)^{\underline{k+1}}}{k+1}
\end{aligned} \tag{9}$$

A result we used to simplify our expression above was that $x^{\underline{k}} = 0$ if $k > x$. This is because $x-k+1$ will be less than or equal to 0 (thus one of the factors in the falling factorial will always be 0). Since $k+1 > 1$ ($x^{\underline{k}}$ does not make sense otherwise), we were able to apply this. Now, if we can express any sequence in terms of falling factorials, we can evaluate the sum of that sequence.

V Simple Cases

Let us return to S_1, which we already have an expression for and can attempt to derive. Via the definition of the falling factorial, $i^{\underline{1}} = i$. Thus, formula for the sum of the first n integers can be derived as follows.

$$\begin{aligned} S_1 = \sum_{i=1}^{n} i &= \sum_{i=1}^{n} i^{\underline{1}} \\ &= \frac{(n+1)^{\underline{2}}}{2} \\ &= \frac{n(n+1)}{2} \end{aligned}$$

It works! Similarly, we can derive the formula for the sum of the squares of the first n integers, S_2. We need to find a way to express i^2 in terms of falling factorials first.

$$\begin{aligned} i^{\underline{2}} &= i(i-1) = i^2 - i \\ \rightsquigarrow i^2 &= i^{\underline{2}} + i = i^{\underline{2}} + i^{\underline{1}} \end{aligned}$$

The rest of the work simply involves using the "reverse power rule" and simplifying the expression we get.

$$\begin{aligned} S_2 &= \sum_{i=1}^{n} i^2 \\ &= \sum_{i=1}^{n} \left(i^{\underline{2}} + i^{\underline{1}}\right) \\ &= \frac{(n+1)^{\underline{3}}}{3} + \frac{(n+1)^{\underline{2}}}{2} \\ &= \frac{n(n+1)(n-1)}{3} + \frac{n(n+1)}{2} \\ &= \frac{2n(n+1)(n-1) + 3n(n+1)}{6} \\ &= \frac{n(n+1)[2(n-1)+3]}{6} \\ &= \frac{n(n+1)(2n+1)}{6} \end{aligned}$$

The exact same process works for S_3.

$$\begin{aligned} i^{\underline{3}} &= i(i-1)(i-2) = i^3 - 3i^2 + 2i \\ \rightsquigarrow i^3 &= i^{\underline{3}} + 3i^2 - 2i = i^{\underline{3}} + 3i^{\underline{2}} + i^{\underline{1}} \end{aligned}$$

$$\begin{aligned} S_3 &= \sum_{i=1}^{n} i^3 \\ &= (i^{\underline{3}} + 3i^{\underline{2}} + i^{\underline{1}}) \\ &= \frac{(n+1)^{\underline{4}}}{4} + 3\frac{(n+1)^{\underline{3}}}{3} + \frac{(n+1)^{\underline{2}}}{2} \end{aligned}$$

$$
\begin{aligned}
&= \frac{n(n+1)(n-1)(n-2)}{4} + n(n+1)(n-1) + \frac{n(n+1)}{2} \\
&= \frac{n(n+1)(n-1)(n-2)}{4} + \frac{4n(n+1)(n-1)}{4} + \frac{2n(n+1)}{4} \\
&= \frac{n(n+1)(n-1)(n-2) + 4n(n+1)(n-1) + 2n(n+1)}{4} \\
&= \frac{n(n+1)[(n-1)(n-2) + 4(n-1) + 2]}{4} \\
&= \frac{n(n+1)(n^2+n)}{4} \\
&= \frac{n^2(n+1)^2}{4}
\end{aligned}
$$

VI Generalisation

We now have a method to find S_n for any n. The amount of work required would increase with time, but the method itself does not change. To generalise what we have done, first notice that expanding the falling factorial gives us a polynomial with certain coefficients.

$$
\begin{aligned}
x^{\underline{1}} &= x \\
x^{\underline{2}} &= x^2 - x \\
x^{\underline{3}} &= x^3 - 3x^2 + 2x \\
x^{\underline{4}} &= x^4 - 6x^3 + 11x^2 - 6x \\
x^{\underline{5}} &= x^5 - 10x^4 + 35x^3 - 50x^2 + 24x \\
&\vdots
\end{aligned}
$$

When we rearrange these polynomials in order to get x^k in terms of falling factorials, we get equations that resemble polynomials.

$$
\begin{aligned}
x &= x^{\underline{1}} \\
x^2 &= x^{\underline{2}} + x^{\underline{1}} \\
x^3 &= x^{\underline{3}} + 3x^{\underline{2}} + x^{\underline{1}} \\
x^4 &= x^{\underline{4}} + 6x^{\underline{3}} + 7x^{\underline{2}} + x^{\underline{1}} \\
x^5 &= x^{\underline{5}} + 10x^{\underline{4}} + 25x^{\underline{3}} + 15x^{\underline{2}} + x^{\underline{1}} \\
&\vdots
\end{aligned}
$$

This is the final thing we need. We can evaluate S_n simply by expanding and evaluating the sum of falling factorials. There is just one problem: those coefficients. They are called the Stirling numbers of the second kind, denoted as $S(k,i)$ or $\left\{ {k \atop i} \right\}$, and they can be defined via the above expansion.

$$
x^k = \sum_{i=1}^{k} \left\{ {k \atop i} \right\} x^{\underline{i}} \tag{10}
$$

For example,

$$x^5 = \left\{ {5 \atop 5} \right\} x^{\underline{5}} + \left\{ {5 \atop 4} \right\} x^{\underline{4}} + \left\{ {5 \atop 4} \right\} x^{\underline{3}} + \left\{ {5 \atop 3} \right\} x^{\underline{2}} + \left\{ {5 \atop 1} \right\} x^{\underline{1}}.$$

So far, that might feel a bit unsatisfying – understandably so. We had some coefficients we wanted to know more about (specifically how to calculate them in general), and we just gave them a name. How does one actually go about calculating $\left\{ {k \atop i} \right\}$?

Fortunately, we have recurrence formulae to work with. Before we start deriving it, note that all x^k can be expressed as the following.

$$x^k = \cdots + \left\{ {k \atop k+2} \right\} x^{\underline{k+2}} + \left\{ {k \atop k+1} \right\} x^{\underline{k+1}} + \left\{ {k \atop k} \right\} x^{\underline{k}} + \cdots + \left\{ {k \atop 1} \right\} x^{\underline{1}} + \left\{ {k \atop 0} \right\} x^{\underline{0}}$$

I have just expanded out the definition above, whilst introducing a few extra terms. This is just to show that this is valid to do, but any term with $\left\{ {k \atop i} \right\}$, where $i > k$, equals zero. Additionally, $\left\{ {k \atop 0} \right\} = 0$ for all k, as can be seen by the absence of a $x^{\underline{0}}$ term in any of our above expansions. $\left\{ {k \atop k} \right\} = 1$ for all k, and that is also fairly obvious from our above expansions.

With that out of the way, note the following.

$$x \cdot x^{\underline{k}} = x^{\underline{k}}(x-k) + kx^{\underline{k}} = x^{\underline{k+1}} + kx^{\underline{k}}$$

Using the definition of the Stirling numbers of the second kind, we have:

$$\begin{aligned}
x^{k+1} &= x \cdot x^k \\
&= x \sum_{i=1}^{k} \left\{ {k \atop i} \right\} x^{\underline{i}} \\
&= \sum_{i=1}^{k} \left\{ {k \atop i} \right\} x^{\underline{i+1}} + \sum_{i=1}^{k} i \left\{ {k \atop i} \right\} x^{\underline{i}} \\
&= \sum_{i=2}^{k+1} \left\{ {k \atop i-1} \right\} x^{\underline{i}} + \sum_{i=1}^{k} i \left\{ {k \atop i} \right\} x^{\underline{i}} \\
&= \sum_{i=1}^{k+1} \left\{ {k \atop i-1} \right\} x^{\underline{i}} + \sum_{i=1}^{k+1} i \left\{ {k \atop i} \right\} x^{\underline{i}} \\
&= \sum_{i=1}^{k+1} \left(\left\{ {k \atop i-1} \right\} + i \left\{ {k \atop i} \right\} \right) x^{\underline{i}}.
\end{aligned}$$

We can express x^{k+1} using a more direct approach.

$$x^{k+1} = \sum_{i=1}^{k+1} \left\{ {k+1 \atop i} \right\} x^{\underline{i}}$$

These two results can be equated to arrive at our recurrence formula.

$$\sum_{i=1}^{k+1} \left\{ {k+1 \atop i} \right\} x^{\underline{i}} = \sum_{i=1}^{k+1} \left(\left\{ {k \atop i-1} \right\} + i \left\{ {k \atop i} \right\} \right) x^{\underline{i}}$$

$$\left\{ {k+1 \atop i} \right\} x^{\underline{i}} = \left\{ {k \atop i-1} \right\} + i \left\{ {k \atop i} \right\} \tag{11}$$

With this formula, and some initial values – which we have – we can find any Stirling number! A possibly arduous process, sure, but certainly not more than working with increasingly long polynomials.

Now, we finally have everything we need. Faulhaber's formula is ready to be presented in all its glory. The sum of the k-th powers of the first n positive integers can be written in terms of the Stirling numbers of the second kind and falling factorials.

$$\begin{aligned} S_k &= \sum_{i=1}^{n} i^k \\ &= \sum_{i=1}^{n} \sum_{j=1}^{k} \left\{ {k \atop j} \right\} i^{\underline{j}} \\ &= \sum_{j=1}^{k} \left\{ {k \atop j} \right\} \sum_{i=1}^{n} i^{\underline{j}} && \text{(Swapped order of sums)} \\ &= \sum_{j=1}^{k} \left\{ {k \atop j} \right\} \frac{(n+1)^{\underline{j+1}}}{j+1} \end{aligned}$$

With a swap of index variables, we get Faulhaber's formula.

$$S_k = \sum_{i=1}^{n} i^k = \sum_{i=1}^{k} \left\{ {k \atop i} \right\} \frac{(n+1)^{\underline{i+1}}}{i+1} \tag{12}$$

VII Conclusion and Additional Remarks

Well, that is it. We are done. In the end, it only took a few lines to derive the formula which was our main goal the entire time. The set-up was by far the lengthier portion, so perhaps that was time wasted. Obviously, as I spent the time and effort required to write this, I do not believe that to be the case at all.

Faulhaber's formula is relatively simple. It is not the most useful formula in the world, nor is it the prettiest. However, putting everything aside, one can appreciate it for its simplicity. Just a few concepts had to be introduced, and it ties them all in beautifully. If the only benefit in doing this was to see that end result, then I would still consider it worth it.

The main value that was attained from this, I would argue, would be the connections that were made. From what I have learnt about calculus in the past, I never would have expected it to link to something that is so 'not continuous'. Yet, discrete calculus is in a world of its own, and there is so much to the subject that cannot possibly be discussed in a brief essay such as this.

The Stirling numbers of the second kind have so much more to them as well. Firstly they have a clear combinatorial interpretation. $\left\{ {k \atop i} \right\}$ is the number of

ways to partition a set of k elements into i nonempty subsets. The following diagram illustrates this well for the case where $k = 4$.

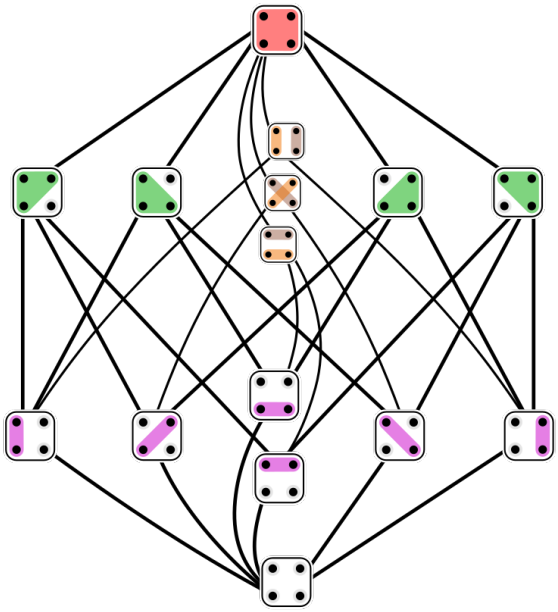


Figure 1: Combinatorial Interpretation of S(4, i) Shown in a Hasse Diagram
(Credits: Tilman Piesk, CC BY 3.0, via Wikimedia Commons)

Each colour is its own subset; uncoloured dots are within their own individual subsets. $\left\{ {4 \atop 1} \right\}$ is on the top of the diagram and $\left\{ {4 \atop 4} \right\}$ is on the bottom. Combinatorics can be used to show every result that we have shown via falling factorials [1].

The previously cited article gives us a one-sum formula for the Stirling numbers of the second kind using the inclusion-exclusion principle.

$$\left\{ {n \atop k} \right\} = \frac{1}{k!} \sum_{r=0}^{k} (-1)^r \binom{k}{r} (k-r)^n$$

Using the substitution $j = k - r$, we get:

$$\left\{ {n \atop k} \right\} = \frac{1}{k!} \sum_{j=0}^{k} (-1)^{k-j} \binom{k}{j} j^n$$

$$\left\{ {k \atop i} \right\} = \frac{1}{i!} \sum_{j=0}^{i} (-1)^{i-j} \binom{i}{j} j^k \qquad \text{(Swapping variable names)}$$

$$= \frac{1}{i!}\sum_{j=1}^{i}(-1)^{i-j}\binom{i}{j}j^k$$
$$= \sum_{j=1}^{i}\frac{(-1)^{i-j}j^k}{(i-j)!j!}$$

Faulhaber's formula may then be expressed explicitly as the following (just for fun).

$$\sum_{i=1}^{n} i^k = \sum_{i=1}^{k}\sum_{j=1}^{i}\frac{(-1)^{i-j}j^k(n+1)^{\underline{i+1}}}{(i+1)(i-j)!j!}$$

Moreover, Stirling numbers are related to probability and the Poisson distribution [2]. These connections are everywhere, and most of them you simply could not predict. I find immense satisfaction in searching for these connections, and I hope that anyone reading this essay will share some of that satisfaction. Deriving Faulhaber's formula was nice and all, but everything brought up along the way is where the true beauty lies.

References

[1] Cynthia Tian and Chahat Kalia. "Stirling Numbers of the Second Kind". In: *MIT PRIMES Circle* Spring (2023). URL: https://math.mit.edu/research/highschool/primes/circle/documents/2023/Cynthia_and_Chahat.pdf.

[2] José A. Adell and Alberto Lekuona. "A probabilistic generalization of the Stirling numbers of the second kind". In: *Journal of Number Theory* 194 (2019), pp. 335–355. ISSN: 0022-314X. DOI: https://doi.org/10.1016/j.jnt.2018.07.003. URL: https://www.sciencedirect.com/science/article/pii/S0022314X18302221.