Sums and integrals of $f(j)$ for j an integer

John Coffey, Cheshire, UK.

2020

1 Sums of j^k for j and k integers

This note is about the sum to n terms of powers of j where j is an integer: $\sum_0^n j^k$. This will be compared with the integral of x^k and the difference between the sum and integral quantified. In the course of this analysis we will come across the Bernoulli numbers and polynomials, and the Euler-Maclaurin summation formula which generalises from j^k to any smooth $f(x)$. There is nothing new here, though the account is my own, derived initially without much reference to textbooks.

At school we learn the sum of the arithmetic series ${}^1S_n = 1+2+3+4+...+(n-1)+n$. Reverse this and add corresponding terms. Each such pair adds to $n+1$ and there are n pairs, so the sum is $n(n+1)/2$. This is a polynomial in n of degree 2: $n^2/2+n/2$. I use the notation $\sum_{j=0}^n j^k = {}^kS_n$. Clearly ${}^0S_n = n$.

Now take $k=2$: ${}^2S_n = 1^2+2^2+3^2+4^2...+n^2$. The table below lists the first few values.

j	1	2	3	4	5	6	7	8
j^2	1	4	9	16	25	36	49	64
${}^2S_n$	1	5	14	30	55	91	140	204

Suspecting that this could also be a polynomial in n, and that its degree might also be one more that the exponent, we try $a_0+a_1n+a_2n^2+a_3n^3$. By inspection $a_0=0$. I can see no trick which will allow easy calculation of the coefficients, so set up three simultaneous equations and solve for the coefficients. In matrix form

$$\begin{pmatrix}1&1&1\\1&2&4\\1&3&9\end{pmatrix}\begin{pmatrix}a_1\\a_2\\a_3\end{pmatrix}=\begin{pmatrix}1\\5\\14\end{pmatrix}.$$

Inverting the matrix, ${}^2S_n = \frac{n^3}{3}+\frac{n^2}{2}+\frac{n}{6} = n(n+1)(2n+1)/6$. Continuing in a similar way, the coefficients of polynomials for k up to 11 are in the table below. This is read as in this example for $k=4$: $\sum_0^n x^4 = (6n^5+15n^4+10n^3-n)/30$.

	power denominator	n^{12}	n^{11}	n^{10}	n^9	n^8	n^7	n^6	n^5	n^4	n^3	n^2	n
$k=1$	2											1	1
$k=2$	6										2	3	1
$k=3$	4									1	2	1	
$k=4$	30								6	15	10		−1
$k=5$	12							2	6	5		−1	
$k=6$	42						6	21	21		−7		1
$k=7$	24					3	12	14		−7		2	
$k=8$	90				10	45	60		−42		20		−3
$k=9$	20			2	10	15		−14		10		−3	
$k=10$	66		6	33	55		−66		66		−33		5
$k=11$	24	2	12	22		−33		44		−33		10	

You might be struck by the seeming lack of pattern here. One feature is that all polynomials are divisible by $n(n+1)$ and the ones with odd power of k such as $\sum x^5$ are divisible by $n^2(n+1)^2$. Here they are factored in full, with the odd and even k separated. The denominators are on the right.

$k=1:$	$n(n+1)$	$/2$
$k=2:$	$n(n+1)(2n+1)$	$/6$
$k=4:$	$n(n+1)(2n+1)(3n^2+3n-1)$	$/30$
$k=6:$	$n(n+1)(2n+1)(3n^4+6n^3-3n+1)$	$/42$
$k=8:$	$n(n+1)(2n+1)(5n^6+15n^5+5n^4-15n^3-n^2+9n-3)$	$/90$
$k=10$	$n(n+1)(2n+1)(n^2+n-1)(3n^6+9n^5+2n^4-11n^3+3n^2+10n-5)$	$/66$
$=$	$n(n+1)(2n+1)(3n^8+12n^7+8n^6-18n^5-10n^4+24n^4+2n^2-15n+5)$	$/66.$
$k=3:$	$n^2(n+1)^2$	$/4$
$k=5:$	$n^2(n+1)^2(2n^2+2n-1)$	$/12$
$k=7:$	$n^2(n+1)^2(3n^4+6n^3-n^2-4n+2)$	$/24$
$k=9:$	$n^2(n+1)^2(n^2+n-1)(2n^4+4n^3-n^2-3n+3)$	$/20$
$=$	$n^2(n+1)^2(2n^6+6n^5+n^4-8n^3+n^2+6n-3)$	$/20$
$k=11:$	$n^2(n+1)^2(2n^8+8n^7+4n^6-16n^5-5n^4+26n^3-3n^2-20n+10)$	$/24$

This shows some pattern, but in each there is a polynomial factor of degree $k-2$ for k even, $k-3$ for k odd, whose coefficients seem to have little pattern. However, we can find a factor of $1/(k+1)$ as follows. For the even k sums, remove the $n(n+1)$, take the leading term of the remaining polynomial factor and divide by the integer denominator. For the odd k sums, remove the $n^2(n+1)^2$, and similarly divide the leading term by the denominator. The results are as follows:

k :	1	2	3	4	5	6	7	8	9	10	11
	1/2	2/6	1/4	6/30	2/12	6/42	3/24	10/90	2/20	6/66	2/24
	1/2	1/3	1/4	1/5	1/6	1/7	1/8	1/9	1/10	1/11	1/12

From this it is fair to propose that

$$\text{for } k \text{ even:} \quad {}^kS_n \approx n(n+1)\left[\frac{n^{k-1}}{k+1} + Cn^{k-2} + \ldots.\right], \tag{1a}$$

$$\text{for } k \text{ odd:} \quad {}^kS_n \approx n^2(n+1)^2\left[\frac{n^{k-3}}{k+1} + Cn^{k-4} + \ldots.\right] \tag{1b}$$

where C is some fraction. In both cases the leading term is $n^{k+1}/(k+1)$.

2 Integration of x^k

One of the first things we learn in school integral calculus is that

$$\int x^k\,dx \;=\; \frac{x^{k+1}}{k+1}\,, \qquad \int_0^n x^k\,dx \;=\; \frac{n^{k+1}}{k+1}\,,. \tag{2}$$

This is a 'primitive', the converse of the differentiation formula $dy/dx = k\,x^{k-1}$ if $y = x^k$. This formula is proved from first principles using the binomial theorem which is valid for all exponents k, even algebraic and transcendental numbers:

$$\frac{d(x^k)}{dx} \;=\; \lim_{h\to 0}\frac{1}{h}\left[(x+h)^k - x^k\right] \;=\; \lim_{h\to 0}\frac{1}{h}\left[x^k\left(1 + k\frac{h}{x} + \frac{k(k-1)}{2!}\left(\frac{h}{x}\right)^2 + \ldots\right) \;-\; x^k\right]$$
$$=\; \lim_{h\to 0}\;\frac{1}{h}\;x^k\;k\,\frac{h}{x} + \mathcal{O}h^2 \;=\; kx^{k-1}\,.$$

Eq 2 gives the leading term of ${}^kS_n$.

We are seeking to quantify the difference between this leading term and the complicated polynomial expressions for ${}^kS_n$ in §1. The subsequent terms in the polynomials must arise from the difference between the smooth curve of x^k and the stepped increase of j^k for integer j illustrated in Figure 1. The first level of approximation in numerical calculation of the integral of x^k would be to use the trapezium rule, and this can be used now, given the exact integral, to estimate the stepwise integer sum. The rectangles have been drawn to start at integer j and extend to $j+1$. The turquoise

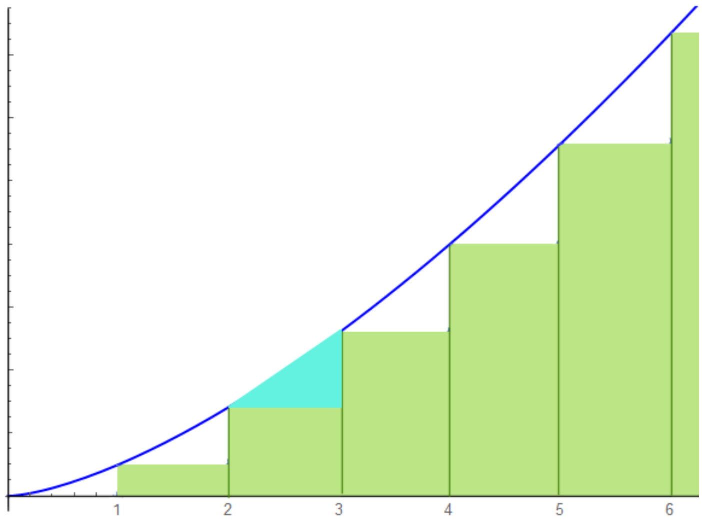

Figure 1: The smooth curve x^k (blue) and the staircase function j^k for j an integer.

triangle sitting on top of the rectangle between 2 and 3 approximately covers the area between the curve and that rectangle. Its area is $[(j+1)^k - j^k]/2$. Eq 2 gives the area under the curve from 0 to $n+1$ to be $(n+1)^{k+1}/(k+1)$, whilst the rectangles up to this position have area ${}^kS_n$. The combined area of the triangles is one half of

$$1^k \;+\; (2^k - 1^k) \;+\; (3^k - 2^k) \;+\; \ldots. \;+\; [(n+1)^k - n^k] \;=\; (n+1)^k\,. \tag{3}$$

Because the curve x^k lies inside the turquoise triangles if $k > 1$, the trapezium rule would over estimate its integral. This means that our estimate is ${}^kS_n$ will be an under-estimate. We have

$${}^kS_n \;>\; \frac{(n+1)^{k+1}}{k+1} - \frac{(n+1)^k}{2} \;=\; (n+1)^k\left[\frac{n+1}{k+1} - \frac{1}{2}\right]\,, \qquad |k| > 1\,. \tag{4}$$

As a numerical example with $k = 8$, $n = 9$, ${}^8S_9 = 6,773,133$ and Eq 4 gives 6,111,111, a 10% under-estimate. ${}^5S_{14} = 1,539,825$ while the formula gives 1,518,750, only 1·4% under the true value. Since a piecewise straight line has been fitted to the curve, we expect Eq 4 to give the exact result for $k = 1$, and indeed it does reduce to $n(n+1)/2$. There is not a strong similarity of Eq 4 to either Eq 1a or 1b.

Note that nothing in this analysis places a limit on the value of k, even though the motivation has been that k be a positive integer. Since the derivation of Eq 2 relies on the binomial theorem which is valid for all exponents, Eq 4 will apply to fractional values of k. For instance, if $k = 7/2$, the sum to $n = 20$ is $177,419$ compared with $176,830$ from Eq 4. The only change to Eq 4 will be if $|k| < 1$, when the sign of the inequality will be reversed. $\sum_0^{25} \sqrt{x} = 85 \cdot 634$, whilst Eq 4 gives $85 \cdot 83$, a remarkably close over-estimate.

2.1 An improved approximation

In seeking a better approximation formula we might recall that in numerical integration the trapezium rule is surpassed by Simpson's rule, which fits a parabola through three points, and that in turn by the Newton-Cotes rule which fits a cubic through four. The main reason why Eq 4 does not itself involve a sum over powers is that in summing the areas of the triangles at Eq 3, the contributions from neighbouring points cancel. We might hope that such cancelling would occur if we look for a scheme which chains points together in a similar way. I therefore propose to fit a cubic function to x^k above each rectangle of the staircase, using the four constraints that it pass through the end points, and has the same gradients at the end points as does x^k. If the cubic is $a + bx + cx^2 + dx^3$, the matrix equation for its coefficients is

$$\begin{pmatrix} 1 & j & j^2 & j^3 \\ 0 & 1 & 2j & 3j^2 \\ 1 & j+1 & (j+1)^2 & (j+1)^3 \\ 0 & 1 & 2(j+1) & 3(j+1)^2 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} j^k \\ kj^{k-1} \\ (j+1)^k \\ k(j+1)^{k-1} \end{pmatrix}$$

The coefficients are not elegant, but each has two terms in which j and $j+1$ have opposite roles:

$$\begin{aligned} a_j &= j^2(j+1)^k(2j-k+3) \quad - \quad j^k(j+1)^2(2j+k-1) \\ b_j &= j(j+1)^{k-1}[3jk+2k-6(j+1)^2] \quad + \quad j^{k-1}(j+1)[3jk+k+6j^2] \\ c_j &= (j+1)^{k-1}[3(j+1)(2j+1)-k(3j+1)] \quad - \quad j^{k-1}[3j(2j+1)+k(3j+2)] \\ d_j &= (j+1)^{k-1}(k-2j-2) \quad + \quad j^{k-1}(k+2j)\,. \end{aligned}$$

The $_j$ subscript reminds us that the a, b, c, d differ for every step of the staircase. The area under this cubic from $x = j$ to $j+1$ is

$$a_j(j+1) + \frac{b_j}{2}(j+1)^2 + \frac{c_j}{3}(j+1)^3 + \frac{d_j}{4}(j+1)^4 - a_j j - \frac{b_j}{2}j^2 - \frac{c_j}{3}j^3 - \frac{d_j}{4}j^4$$

$$= a_j + b_j(j+\tfrac{1}{2}) + c_j(j^2+j+\tfrac{1}{3}) + d_j(j^3+\tfrac{3}{2}j^2+j+\tfrac{1}{4})\,.$$

The area A_{cv} of the curvi-triangle from j to $j+1$ is this minus j^k. This looks complicated, but much internal cancelling does occur and the area reduces to

$$\text{Area of 'triangle' on base } [j, j+1]\text{:} \quad A_{cv} = \frac{1}{12}\left[(j+1)^{k-1}(6j-k+6) - j^{k-1}(6j-k)\right]. \qquad (5)$$

Moreover, terms from adjacent triangles up the staircase do cancel as we hoped, as can be seen from the first 6 values of j and the nth term. With the denominator 12 understood,

$$
\begin{array}{ll}
j=0: & 6-k \\
j=1: & 2^{k-1}(12-k)+k-6 \\
j=2: & 3^{k-1}(18-k)+2^{k-1}(k-12) \\
j=3: & 4^{k-1}(24-k)+3^{k-1}(k-18) \\
j=4: & 5^{k-1}(30-k)+4^{k-1}(k-24) \\
 & \ldots\ldots. \\
j=n: & (1+n)^{k-1}[6(n+1)-k]+n^{k-1}(k-6n)
\end{array}
$$

The last term in any row cancels with the first term in the previous row. The sum to $j=n$ is

$$\frac{1}{12}(n+1)^{k-1}[6(n+1)-k] = \frac{(n+1)^k}{2} - \frac{k(n+1)^{k-1}}{12}. \tag{6}$$

The sum of powers ${}^kS_n$ is found by subtracting this from the integral under the curve, $(n+1)^{k+1}/(k+1)$ to be

$${}^kS_n \approx (n+1)^k\left[\frac{n+1}{k+1} - \frac{1}{2} + \frac{k}{12(n+1)}\right]. \tag{7}$$

This differs from the trapezium approximation of Eq 4 only in the last term, $k/(12(n+1)$, but gives a significantly better approximation. A few values for comparison are listed below.

k	n	exact	cubic approxn.
2.5	40	120761	120761
3	25	105625	105625
5	14	1539825	1539843
8	9	67731333	67777777
12	20	8.5534E+15	8.5549E+15
21	10	1.1192E+21	1.1773E+21

The approach has been to integrate x^k numerically, subtract j^k over each interval $[j, j+1]$, and subtract this from the exact integral given by calculus. The result, therefore, is only as good as the numerical integration scheme. Since cubic curves have been fitted, Eq 7 should give the exact result for $0<k<3$. In fact it agrees exactly only for $k=2$ and 3 as the table below shows. The exact polynomials are from §1. The formulae for $k=4$ and 5 have been expanded to show that the difference lies only in the terms of lowest degree. Since for $k>3$ x^k is more highly curved than a cubic, the formulae from Eq 7 are an over-estimate.

k	exact	Eq 7
1	$\frac{1}{2}n(n+1)$	$\frac{1}{2}n(n+1)+\frac{1}{12}$
2	$\frac{1}{6}n(n+1)(2n+1)$	$\frac{1}{6}n(n+1)(2n+1)$
3	$\frac{1}{4}n^2(1+n)^2$	$\frac{1}{4}n^2(1+n)^2$
4	$\frac{1}{30}n(n+1)(2n+1)(3n^2+3n-1)$	$\frac{1}{30}(1+n)^3(6n^2-3n+1)$
4	$=\frac{n^5}{5}+\frac{n^4}{2}+\frac{n^3}{3}-\frac{n}{30}$	$=\frac{n^5}{5}+\frac{n^4}{2}+\frac{n^3}{3}+\frac{1}{30}$
5	$\frac{n^6}{6}+\frac{n^5}{2}+\frac{5n^4}{12}-\frac{n^2}{12}$	$\frac{n^6}{6}+\frac{n^5}{2}+\frac{5n^4}{12}+\frac{n}{6}+\frac{1}{12}$

3 Further curve-fitting

The next level of approximation to ${}^kS_n$ would be to fit a 5th order polynomial to x^k over $[j, j+1]$, using as the two additional constraints the second derivatives at the end points. Then a 7th order one using the 3rd derivatives, and so on. This would add terms to Eqs 4 and 6, each bringing a further level of refinement. It may be possible to spot a pattern to these and so predict yet further terms without having to calculate them explicitly.

The method of §2 led to a relatively simple formula because the approximating function – a straight line, then a cubic – matched the staircase function j^k and first derivative at the end points of each interval. These values were chained together at touching end points of the staircase's steps, and all cancelled apart from the last. It seems natural to consider whether this concept could work with other piecewise constant functions which are indexed by an integer j – for instance $\sum^n \ln j$ or $J_j(k)$ or $J_k(j)$ where J is a Bessel function. Whether such sums would be of any practical use is, of course, another matter. Indeed, I have not examined whether Eqs 4 and 7 are valid when the power k is negative. If there are the makings of a method here for calculating a sum given its integral or *vice versa*, its use will depend on whether the chain of cancellation occurs, so that the number of quantities which have to be evaluated specifically is very small.

Since the chain seems crucial, let us examine Eq 6, which gives the sum $j = 0$ to n over the curvi-triangles. The first term on the right contains the given function x^k at $j = (n+1)$, and the second term contains its derivative kx^{k-1}. Is this a coincidence? If it is not, where do the coefficients 1/2 and 1/12 come from? Since the constraints on the fitted cubic are to match its value and its first derivative at the end points, perhaps it is not a coincidence. In this section I explore these leads.

It is obvious that the triangle-trapezium integration method used in §2 will work equally well for most smooth functions $f(x)$. Suppose $f(j) \equiv f_j$ is the value of $f(x)$ at integer j. The graph of f_j is a staircase. Draw the straight line between points (j, f_j) and $(j+1, f_{j+1})$ and calculate the area of the enclosed triangle corresponding to the turquoise one in Figure 1. Its area is $(f_{j+1} - f_j)/2$ and the sum over all such triangles from $j = 0$ to n is half of

$$(f_1 - f_0) + (f_2 - f_1) + (f_3 - f_2) + \dots \quad \dots + (f_n - f_{n-1}) + (f_{n+1} - f_n) = f_{n+1} - f_0 .$$

The generalisation of Eq 4 is

$$\sum_0^n f_j \approx \int_0^{n+1} f(x)\,dx - \frac{1}{2}(f_{n+1} - f_0) . \tag{8}$$

The denominator 2 comes from the area of the triangle being half of that of a rectangle.

3.1 Approximation using first and higher derivatives

It is possible to fit a cubic across $[j, j+1]$ by replacing the right hand side of the matrix in §2 by

$$\begin{pmatrix} f_j \\ f'_j \\ f_{j+1} \\ f'_{j+1} \end{pmatrix}$$

where the prime denotes the first derivative. The cubic is

$$-f_j(2j - 2x - 1)(j - x + 1)^2 - (j - x)\left[f'_j(j - x + 1)^2 - (j - x)(f_{j+1}(2j - 2x + 3) + f'_{j+1}(x - j - 1))\right] .$$

This is integrated with respect to x from j to $j+1$ and the area f_j of the rectangular step subtracted. The resulting area of the curvi-triangle has the remarkably simple form

$$\frac{f_{j+1}-f_j}{2} - \frac{f'_{j+1}-f'_j}{12}\,.$$

The potential for chain cancellation as j steps from 0 to n is clear. Eq 8 becomes upgraded to

$$\sum_0^n f_j \;\approx\; \int_0^{n+1} f(x)\,dx \;-\; \frac{1}{2}(f_{n+1}-f_0) \;+\; \frac{1}{12}(f'_{n+1}-f'_0)\,. \qquad (9)$$

This is the generalisation of Eq 7. The $\frac{1}{12}$ comes from integrating a cubic: $\frac{1}{2}$ from $\int x$, $\frac{1}{3}$ from $\int x^2$ and $\frac{1}{4}$ from $\int x^3$. The last term involving the first derivatives can be understood as follows. If $f'(j+1) > f'(j)$ the curve $f(x)$ is concave as seen from above and lies inside the triangle of the trapezium rule. The correction $(f_{j+1}-f_j)/12$ is positive and must be subtracted from the area of the straight-sided triangle. The converse holds when $f'(j+1) < f'(j)$.

We press ahead to the next level of approximation by fitting a quintic curve to $f(x)$ of $[j, j+1]$. The additional constraints come from matching the second derivatives at the end points of the interval. The matrix equation is

$$\begin{pmatrix} 1 & j & j^2 & j^3 & j^4 & j^5 \\ 0 & 1 & 2j & 3j^2 & 4j^3 & 5j^4 \\ 0 & 0 & 2 & 6j & 12j^2 & 20j^3 \\ 1 & j+1 & (j+1)^2 & (j+1)^3 & (j+1)^4 & (j+1)^5 \\ 0 & 1 & 2(j+1) & 3(j+1)^2 & 4(j+1)^3 & 5(j+1)^4 \\ 0 & 0 & 2 & 6(j+1) & 12(j+1)^2 & 20(j+1)^3 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \\ g \\ h \end{pmatrix} = \begin{pmatrix} f_j \\ f'_j \\ f''_j \\ f_{j+1} \\ f'_{j+1} \\ f''_{j+1} \end{pmatrix}.$$

The result is that the area of the curvi-triangle at $[j, j+1]$ is

$$\frac{1}{120}\left[60f_{j+1} - 60f_j - 12f'_{j+1} + 12f'_j + f''_{j+1} + f''_j\right]. \qquad (10)$$

The second derivatives have the same sign and so will not cancel as j steps in integers. Moreover the coefficient of the first derivative terms is 1/10, not the 1/12 of the cubic approximation. The scheme has broken down since a chain of complete cancellation does not occur. No doubt the approximation is better than the cubic, but it cannot be presented as a simple formula because an explicit summation is required.

Out of curiosity I have pressed on to fit a 7th degree polynomial over the interval $[j, j+1]$ using the 3rd derivatives, and obtain for the area of one curvi-triangle to be

$$\frac{f_{j+1}}{2} - \frac{f_j}{2} \;-\; \frac{3f'_{j+1}}{28} + \frac{3f'_j}{28} \;+\; \frac{f''_{j+1}}{84} + \frac{f''_j}{84} \;-\; \frac{f^{(3)}_{j+1}}{1680} + \frac{f^{(3)}_j}{1680}\,. \qquad (11)$$

This formula will give the exact value when $f(x)$ is any polynomial of degree 7 or less. The surprising features are that the coefficients change again, and the terms in the third derivatives have opposite signs so they do cancel from step to step. Only the second derivatives remain uncancelled. If this were applied to ${}^kS_n$ for $(x) = x^k$, for which $f'' = k(k-1)x^{k-2}$, their sum over the staircase will involve ${}^{k-2}S_n$, the corresponding sum two degrees lower.

It is clear that we could in principle go on fitted polynomials of ever higher degree to the given curve, and we would find that these gave the exact values for all functions whose highest non-zero derivative was that of the fitted polynomial, but that errors increase steadily for functions whose higher derivatives were non-zero. Moreover, we have seen with the 5th and 7th degree fitted polynomials have some terms which do not cancel along the chain of steps, so a formula for $\sum^n f(j)$ which does not itself involve a sum of lower degree is not possible. This points to our needing to find a different approach to approximating $\sum^n f(j)$. We can, however, expect it to be a series involving successive derivatives of $f(x)$, and which starts exactly as Eq 9.

Since a Taylor series has the form required, we explore its suitability. Consider a single step of the staircase in Figure 2 and expand $f(x)$ about the corner at $x = j$. There is a forward series from j towards $j+1$ and a backward series from j towards $j-1$. They differ only in the sign of the increment h:

$$f(j+h) \;=\; f(j) \;+\; f'(j)h \;+\; \tfrac{1}{2!}f''(j)h^2 \;+\; \tfrac{1}{3!}f^{(3)}(j)h^3 \;+\; \tfrac{1}{4!}f^{(4)}(j)h^4 \;+\; .... \qquad (12)$$

There is no good reason to prefer one of these over the other in estimating the area of the curvi-triangle on $[j, j+1]$, so try using the forwards one from j to $j+\frac{1}{2}$ and the backwards one from $j+1$. Integrating Eq 12 in h from 0 to $\frac{1}{2}$ gives the area under the fitted curve and x axis to be

$$\tfrac{1}{2}f(j) \;+\; \tfrac{1}{8}f'(j) \;+\; \tfrac{1}{48}f''(j) \;+\; \tfrac{1}{384}f^{(3)}(j) \;+\; \tfrac{1}{3840}f^{(4)}(j) \;+\; ....$$

whilst the other half interval contributes

$$\tfrac{1}{2}f(j+1) \;-\; \tfrac{1}{8}f'(j+1) \;+\; \tfrac{1}{48}f''(j+1) \;-\; \tfrac{1}{384}f^{(3)}(j+1) \;+\; \tfrac{1}{3840}f^{(4)}(j) \;-\; ....$$

The estimated area of the curvi-triangle is the sum of these minus the step value, $f(j)$:

$$\tfrac{1}{2}\left[f(j+1)-f(j)\right] \;-\; \tfrac{1}{8}\left[f'(j+1)-f'(j)\right] + \tfrac{1}{48}\left[f''(j+1)+f''(j)\right]$$
$$-\; \tfrac{1}{384}\left[f^{(3)}(j+1)-f^{(3)}(j)\right] \;+\; \tfrac{1}{3840}\left[f^{(4)}(j+1)+f^{(4)}(j)\right] \;+\; .... \qquad (13)$$

As with the 5th and 7th order polynomials, the second derivatives, and now also the fourth, have the same sign so will not chain-cancel along the staircase. The coefficient of the first derivative is here $\frac{1}{8}$ whereas previously it was $\frac{1}{12}$. We might conclude that this is not the best way forwards either.

So is there another approximating series for which all derivatives will chain-cancel, not just the odd ones? Alternatively, is there an approximating series where the even order derivatives do not appear – their contribution is zero? Perhaps such a formula may be derived from a different set of ideas, such as being a series which only approximates $\int f(x)$ asymptotically, without actually converging.

In searching for clues I have looked at different functions from x^k in Eq 9. Take the case of $f(j) \equiv f_j = \ln j$. The sum being approximated by an integral is therefore $\ln 1 + \ln 2 + \ln 3 \; ... \; + \ln n = \ln(1.2.3.4.5....n) = \ln n!$. The integral is

$$\int_1^{n+1} \ln x \, dx \;=\; x\ln x - x\big|_1^{n+1} \;=\; (n+1)\ln(n+1) - n$$

$$\text{so} \qquad \ln(n!) \approx \; (n+1)\ln(n+1) - n - \tfrac{1}{2}\ln n + \tfrac{1}{12}\left(\tfrac{1}{n} - 1\right). \qquad (14)$$

This is not quite what Stirling published, but it is fairly close. $\ln 50! = 148\cdot478$ and the formula gives $148\cdot485$. $\ln 100! = 363\cdot739$ and the formula gives $363\cdot742$. $\ln 1000! = 5912\cdot1282$ and the formula gives $5912\cdot1264$. This has the qualities of an asymptotic series in that the relative error decreases as the argument n increase towards infinity. The disappointment is that it cannot be taken to the next level of precision. It is time to leave this numerical analysis approach and return to number theory.

4 From Pascal's Triangle to $\sum_j^n j^k$

We follow now in the footsteps of Blaise Pascal (1623-1662) and Jacob Bernoulli (1655-1705), two genii of the 17th century, and look at alternative methods to ${}^kS_n$ from those in §1.

4.1 ${}^kS_n$ from Pascal's Triangle using Stirling numbers

Pascal's Triangle is well known as a tabulation of the binomial coefficients, the coefficients of a and b in the expansion of $(a+b)^k$ where k is an integer. This triangle, presented below in matrix form, is alive with symmetries.

	0	1	2	3	4	5	6	7	8	9	10
0	1										
1	1	1									
2	1	2	1								
3	1	3	3	1							
4	1	4	6	4	1						
5	1	5	10	10	5	1					
6	1	6	15	20	15	6	1				
7	1	7	21	35	35	21	7	1			
8	1	8	28	56	**70**	56	28	8	1		
9	1	9	36	84	126	126	84	36	9	1	
10	1	10	45	120	210	252	210	120	45	10	1

For our purposes note that the sum of any partial column is given by the next number down and diagonally to the right. Thus the sum of the blue column is the red bold number 70. Writing ${}^nC_r$ for the binomial coefficient, also written $\binom{n}{r}$, this is

$$\sum_{r=0}^{n} {}^nC_r = {}^{n+1}C_{r+1} = \frac{(n+1)n(n-1)....(n+1-r)}{(r+1)!} .$$

Moreover it follows that

$${}^{n+1}C_r = \frac{n+1}{n+1-r}\,{}^nC_r , \qquad {}^nC_{r+1} = \frac{n-r}{r+1}\,{}^nC_r .$$

These relations mean that is it systematic and fairly easy to evaluate sums of binomial coefficients. What is important here is that for general n these coefficients are polynomials in n, formed by multiplying out the truncated factorials. Here are the first few

$$\begin{aligned} &{}^nC_0 = 1, \qquad {}^nC_1 = n, \qquad {}^nC_2 = \frac{n^2}{2} - \frac{n}{2}, \qquad {}^nC_3 = \frac{n^3}{6} - \frac{n^2}{2} + \frac{n}{3}, \\ &{}^nC_4 = \frac{n^4}{24} - \frac{n^3}{4} + \frac{11n^2}{24} - \frac{n}{4}, \qquad {}^nC_5 = \frac{n^5}{120} - \frac{n^4}{12} + \frac{7n^3}{24} - \frac{5n^2}{12} + \frac{n}{5} . \end{aligned} \tag{15}$$

Bernoulli must have been exploring how to add these polynomials to form sums of powers, $\sum j^k$, which otherwise are much more difficult to calculate. We return, therefore, to the theme of §1 and look to express ${}^kS_n$ in terms of sums of binomial coefficients.

The first couple can be written down by inspection :

$$j = {}^jC_1 \text{ so } {}^1S_n = {}^{n+1}C_2 = \frac{1}{2}n(n+1) .$$

$$j^2 = 2\,.^jC_2 + {}^jC_1 \quad \text{so} \quad {}^2S_n = 2\,.^{n+1}C_3 + {}^{n+1}C_2 = \tfrac{1}{6}n(n+1)(2n+1)\,.$$

To find n^3 take a weighted sum of the ${}^jC_1$, ${}^jC_2$, ${}^jC_3$, solve for the weighting multiplying n^3 and then back substitute to find the other weightings:

$$a\,.^jC_1 + b\,.^jC_2 + c\,.^jC_3 = \frac{cj^3}{6} - \frac{cj^2}{2} + \frac{bj^2}{2} + \frac{cj}{3} - \frac{bj}{2} + aj$$

$$c = 6, \qquad \frac{b}{2} - 3 = 0 \;\text{ so }\; b = 6: \;\text{ then }\; aj - j = 0 \;\text{ making }\; a = 1\,.$$

$$\text{Thus} \quad \sum^n j^3 = {}^{n+1}C_2 + 6\,.^{n+1}C_3 + 6\,.^{n+1}C_4 = \frac{1}{4}n^2(n+1)^2$$

in agreement with the tables in §1. Continuing in this way to find higher j^k:

$$j^4 = {}^jC_1 + 14\,.^jC_2 + 36\,.^jC_3 + 24\,.^jC_4, \qquad j^5 = {}^jC_1 + 30\,.^jC_2 + 150\,.^jC_3 + 240\,.^jC_4 + 120\,.^jC_5\,,$$

$$j^6 = {}^jC_1 + 62\,.^jC_2 + 540\,.^jC_3 + 1560\,.^jC_4 + 1800\,.^jC_5 + 720\,.^jC_6\,.$$

and their sums ${}^kS_n$ are readily calculated. This was a clever piece of lateral thinking by Bernoulli.

The process replaces the polynomials ${}^kS_n$ of §1 with an equivalent set of polynomials in the ${}^jC_r$. Their coefficients are listed in the *On-line encyclopedia of integer sequences* at *www.oeis.org* to be rows in the triangle of $T(n,r)$ numbers defined by $r!\,\mathcal{S}(n,r)$ where $\mathcal{S}(n,r)$ are the Stirling numbers of the second kind: for $n \geq r$ they count the number of ways n distinct (labelled) objects can be partitioned into r non-empty, unlabelled sets. These $T(n,r)$ numbers increase rapidly and therefore rather detract from the ease of computation. The triangle begins:

$n\backslash\ r$	1	2	3	4	5	6	7	8	9
1 :	1								
2 :	1	2							
3 :	1	6	6						
4 :	1	14	36	24					
5 :	1	30	150	240	120				
6 :	1	62	540	1,560	1,800	720			
7 :	1	126	1,806	8,400	16,800	15,120	5,040		
8 :	1	254	5,796	40,824	126,000	191,520	141,120	40,320	
9 :	1	510	18,150	186,480	834,120	1,905,120	2,328,480	1,451,520	362,880

An explicit formula for these entries is

$$T(n,k) = \sum_{j=0}^{k} (-1)^j\ .^kC_j\,.(k-j)^n.$$

Fortunately the Stirling numbers themselves satisfy the recursion relation

$$\mathcal{S}(n+1,r) = r\,\mathcal{S}(n,r) + \mathcal{S}(n,r-1)$$

which translates into the recursion relation for the $T(n,r)$ as

$$T(n+1,r) = r[T(n,r) + T(n,r-1)]\,. \tag{16}$$

For instance $3(540+62) = 1,806$ and $6(720+1,800) = 15,120$. It is therefore straightforward to extend the table downwards. The explicit formula above is needed only for $T(n,n)$, and then only for the smallest n on the right of the table, to start a new column. The calculations are probably significantly simpler than setting up then inverting an $n \times n$ matrix, as was done in §1, so Bernoulli's approach is computationally helpful as well as giving insight. How the partitioning of n individual items into r boxes comes in, I will not pursue.

4.2 ${}^kS_n$ by summing binomial coefficients

An alternative route to calculating the sums ${}^kS_n = \sum^n j^k$ is to express the j^k as polynomials in the ${}^kS_n$ instead of in the binomial coefficients. The trick is to use chain-cancellation which we first saw in Eq 3:

$$(1^k - 0^k) + (2^k - 1^k) + (3^k - 2^k) + ...[n^k - (n-1)^k] = \sum_{j=0}^{n} [j^k - (j-1)^k] = n^k .$$

We know the $[j^k - (j-1)^k]$ values; they are the binomial coefficients ${}^kC_r$ with the end one omitted and alternate signs changed to − since we are expanding $(j-1)^k$. It is therefore straightforward to write down a matrix equation for the unknown ${}^kS_n$. For entries up to $k = 6$:

$$\begin{pmatrix} 1 & & & & & \\ -1 & 2 & & & & \\ 1 & -3 & 3 & & & \\ -1 & 4 & -6 & 4 & & \\ 1 & -5 & 10 & -10 & 5 & \\ -1 & 6 & -15 & 20 & -15 & 6 \end{pmatrix} \begin{pmatrix} {}^0S_n \\ {}^1S_n \\ {}^2S_n \\ {}^3S_n \\ {}^4S_n \\ {}^5S_n \end{pmatrix} = \begin{pmatrix} n \\ n^2 \\ n^3 \\ n^4 \\ n^5 \\ n^6 \end{pmatrix}$$

This gives simultaneously all sums ${}^kS_n$ up to the selected highest k. Inverting a large matrix except with algebraic software is computationally challenging. However, once we have a few for low values of k, higher ones can be assembled. Significantly, the matrix reveals the recursion relation between ${}^kS_n$ and the lower sums. Observe, for example, that $-1\,.{}^0S_n + 4\,.{}^1S_n - 6\,.{}^2S_n + 4\,.{}^3S_n = n^4$. Once the lower sums are known, this can be solved for ${}^3S_n$ and simplified to the correct value of $\frac{n^2}{4}(n+1)^2$. We therefore have

$${}^kS_n = \frac{1}{{}^{k+1}C_k} \left\{ n^{k+1} + (-1)^k \left[{}^{k+1}C_{k-1}.{}^{k-1}S_n - {}^{k+1}C_{k-2}.{}^{k-2}S_n - ...... - \pm{}^0S_n \right] \right\} . \qquad (17)$$

5 Bernoulli numbers and polynomials

5.1 Bernoulli numbers

I happened to spot in a book the series of fractions $\frac{1}{2}$, $\frac{1}{6}$, $\frac{1}{30}$, $\frac{1}{42}$, $\frac{1}{30}$, $\frac{5}{66}$. These seemed familiar from the denominators in the exact polynomial series of ${}^kSn$ in §1. Specifically they are the coefficients of the lowest order terms for $\sum j^k$ when k is even, listed in the table on page 2. Only the even k polynomials have a term in n^1; for $k > 1$ the odd k terms have n^2 as lowest degree. I have therefore looked up the origin and properties of these numbers, and also sought any special mention of the corresponding series for odd k polynomials. From $k = 3$ these run $\frac{1}{4}$, $\frac{1}{12}$, $\frac{1}{12}$, $\frac{3}{20}$, $\frac{5}{12}$, etc. Since Bernoulli numbers occupy a whole chapter in the *Handbook of Mathematical Functions* by Abramowitz and Stegun, they merit their own section here.

The Bernoulli numbers are the coefficients of n in ${}^kS_n$. They can be picked out from the right-most column of the first table in §1 and are :

$$B_0 = 1, \quad B_1 = \pm\frac{1}{2}, \quad B_2 = \frac{1}{6}, \quad B_3 = 0, \quad B_4 = \frac{-1}{30}, \quad B_5 = 0, \quad B_6 = \frac{1}{42},$$

$$B_7 = 0, \quad B_8 = \frac{-1}{30}, \quad B_9 = 0, \quad B_{10} = \frac{5}{66}, \quad B_{11} = 0, \quad B_{12} = -\frac{691}{2730}, \quad etc.$$

The ambiguous ± sign of B_1 will be explained shortly. These have been extensively studied over many decades and appear in many branches of mathematics. Amongst the material on the internet, the article by Nathaniel Larson, 2019, is comprehensive and readable[1]. Since so much has been written about them, I will merely pick out a few aspects. The numbers seem to have almost no pattern other than the odd ones being zero and the even ones alternating in sign. Perhaps they would have remained unnoticed and unspecial along with dozens of other sequences which occur when people play with strings of numbers. I wonder how long Bernoulli stared at the table of ${}^kS_b$ in §1 before he spotted that these coefficients of n are deeply embedded throughout this table.

Trying to reconstruct his journey of discovery, we first multiply each row by $k+1$, the denominator in Eq 1a, b and Eq 2. This normalises the leading diagonal's values to 1.

	n^{13}	n^{12}	n^{11}	n^{10}	n^9	n^8	n^7	n^6	n^5	n^4	n^3	n^2	n
$k=1:$												1	1
$k=2:$											1	$\frac{3}{2}$	$\frac{1}{2}$
$k=3:$										1	2	1	
$k=4:$									1	$\frac{5}{2}$	$\frac{5}{3}$		$-\frac{1}{6}$
$k=5:$								1	3	$\frac{5}{2}$		$-\frac{1}{2}$	
$k=6:$							1	$\frac{7}{2}$	$\frac{7}{2}$		$-\frac{7}{6}$		$\frac{1}{6}$
$k=7:$						1	4	$\frac{14}{3}$		$-\frac{7}{3}$		$\frac{2}{3}$	
$k=8:$					1	$\frac{9}{2}$	6		$-\frac{21}{5}$		2		$-\frac{3}{10}$
$k=9:$				1	5	$\frac{15}{2}$		-7		5		$-\frac{3}{2}$	
$k=10:$			1	$\frac{11}{2}$	$\frac{55}{6}$		-11		11		$-\frac{11}{2}$		$\frac{5}{6}$
$k=11:$		1	6	11		$-\frac{33}{2}$		22		$-\frac{33}{2}$		5	
$k=12:$	1	$\frac{13}{2}$	13		$-\frac{143}{6}$		$\frac{286}{7}$		$\frac{429}{10}$		$\frac{65}{3}$		$-\frac{691}{210}$

Now look at the next diagonal to the right, reading 1, $\frac{3}{2}$, 2, $\frac{5}{2}$, 3, $\frac{7}{2}$, ... These are all $(k+1)/2$. What about the next diagonal? The denominators are 1, 2, 3 or 6. The table below translates a few of these in multiples of 1/6.

$$\begin{array}{rcccccccc} & \frac{1}{2} & 1 & \frac{5}{3} & \frac{5}{2} & \frac{7}{2} & \frac{14}{3} & 6 & \frac{15}{2} \\ =\frac{1}{6}\times & 3 & 6 & 10 & 15 & 21 & 28 & 36 & 45 \end{array}$$

They advance in steps which increase by 1, and form the column labelled '2' in Pascal's Triangle, §4.1. They are binomial coefficients. What about the fourth diagonal? The lowest common denominator is 30. Expressing the entries as thirtieths:

$$\begin{array}{rcccccccc} & -\frac{1}{6} & -\frac{1}{2} & -\frac{7}{6} & -\frac{7}{3} & -\frac{21}{5} & -7 & -11 & -\frac{33}{2} \\ =-\frac{1}{30}\times & 5 & 15 & 35 & 70 & 126 & 210 & 330 & 495 \end{array}$$

This is the next but one column in Pascal's Triangle. Amazing! So what do be multiply the fifth diagonal by? Could it be 42?

$$\begin{array}{rccccccc} & \frac{1}{6} & \frac{2}{3} & 2 & 5 & 11 & 22 & \frac{286}{7} \\ =\frac{1}{42}\times & 7 & 28 & 84 & 210 & 462 & 924 & 1716 \end{array}$$

Again these are the next but one column (or diagonal) in Pascal's Triangle. The fact that the pattern jumps columns and uses only those in §4.1 labelled with even index corresponds to alternate Bernoulli

[1] https://www.whitman.edu/documents/Academics/Mathematics/2019/Larson-Balof.pdf

numbers being zero. So now the numbers which were first seen as the denominators of n in ${}^kS_n$ are seen as multiplying all the diagonals in turn. Bernoulli and later Euler had found deep patterns in the sums of powers of integers. The coefficients of n^r is the table of ${}^kS_n$ are binomial coefficients modulated by the Bernoulli numbers, which appear as the coefficients of n^1. The coefficient of $n^r,\ 1 \le r \le k+1$ is

$$\frac{1}{k+1} B_{k+1-r} \,.\, {}^{k+1}C_r \tag{18}$$

with the proviso that B_1 is positive. This represents a major step forwards in understanding the sum of powers. Yet the Bernoulli numbers themselves seem so haphazard, even mysterious. Can we find them in some other area of mathematics and thereby see them more clearly? If they were integers, I might look at the partial quotients of continued fractions, which can be very irregular. But they are fractions.

5.2 Power series of some trigonometric functions

When I was at school learning about power series representations of elementary functions, I was struck by the highly regular patterns in expanding e^x, $\sin x$ and $\cos x$ as Taylor-Maclaurin series, but equally struck by the knotty, ungainly series[2] for $\tan x$ and $\cot x$:

$$\tan x = x + \frac{x^3}{3} + \frac{2x^5}{15} + \frac{17x^7}{315} + \frac{62x^9}{2835} + \frac{1382x^{11}}{155925} + \frac{21844x^{13}}{6081075} + O\left(x^{15}\right).$$

$$\cot x = \frac{1}{x} - \frac{x}{3} - \frac{x^3}{45} - \frac{2x^5}{945} - \frac{x^7}{4725} - \frac{2x^9}{93555} - \frac{1382x^{11}}{638512875} - \frac{4x^{13}}{18243225} + O\left(x^{15}\right).$$

The hyperbolic functions $\tanh x$ and $\coth x$ differ only in that alternate signs are changed. Like the Bernoulli numbers, every other one is zero. If we are to find a pattern in these two series, it will be useful to multiply out the factorials from each denominator. We obtain just the successive derivatives of $\tan x$ and $x\cot x$ evaluated at $x = 0$. Ignoring the signs and the alternate values which are zero, this table lists the derivatives of $x\cot x$ in the limit $x \to 0$. In the second row are the reciprocals of the Bernoulli numbers, and the bottom row lists their product. These are all powers of 2.

$\frac{2}{3}$	$\frac{8}{15}$	$\frac{32}{21}$	$\frac{128}{15}$	$\frac{2560}{33}$	$\frac{1415168}{1365}$	$\frac{57344}{3}$
$\times 6$	$\times 30$	$\times 42$	$\times 30$	$\times \frac{66}{5}$	$\times \frac{2730}{691}$	
$= 4$	16	64	256	1024	4096	

This allows us to write

$$x\cot x = 1 - \frac{(2x)^2}{2!}B_2 + \frac{(2x)^4}{4!}B_4 - \frac{(2x)^6}{6!}B_6 + \frac{(2x)^8}{8!}B_8 - \ldots\ldots \tag{19}$$

So in principle higher Bernoulli numbers could be calculated from the derivatives of $x\cot x$.

Performing the same type of detective work on $\tan x$, the table below lists the successive derivatives at $x = 0$:

index, r :	1	2	3	4	5	6	7	8	9	10	11	12	13
derivative :	1	0	2	0	16	0	272	0	7936	0	353792	0	22368256
factors :	1		2^1		2^4		$2^4.17$		$2^8.31$		$2^9.691$		$2^{12}.43.127$

[2] Because $\cot 0 \to \infty$, the $\cot x$ series is obtained from $\tan x$ by dividing the $\tan x$ Taylor series by x, expanding its reciprocal by the binomial theorem, then restoring the factor x. Alternatively use the Taylor series for $x\cot x$.

This does not look promising, though there are some faint clues: $17 = 2^4+1$, $31 = 2^5-1$, $127 = 2^7-1$, and 691, a prime, is the numerator of B_{12}. Just on the off chance that there may be something here, divide the derivatives by B_{r+1}.

$index, r:$	3	5	7	9	11	13
	$2^2.3.5$	$2^5.3.7$	$2^5.15.17$	$2^9.31.33$	$2^{10}.15.7.13$	$2^{13}.127.129.\frac{1}{7}$

There are elements of a pattern here, with factors of 2^a, 2^b-1 and 2^c+1. Also, since $(a+1)(a-1) = a^2-1$, $3\times5 = 2^4-1$, $15\times17 = 2^8-1$, $31\times33 = 2^{10}-1$ and $127\times129 = 2^{14}-1$. These are all of the form $2^{r+1}-1$. The other factor seems to be 2^r. The table below divides the rth derivatives of $\tan x$, $x\to0$ by $2^r(2^{r+1}-1)$. In the bottom row are the Bernoulli numbers B_{r+1}.

$index, r:$	1	3	5	7	9	11	13	15
$derivative,\ \delta^{(r)}:$	1	2	16	272	7936	353792	22368256	1903757312
$\frac{\delta^{(r)}}{2^r(2^{r+1}-1)}$	$\frac{1}{6}$	$\frac{1}{60}$	$\frac{1}{126}$	$\frac{1}{120}$	$\frac{1}{66}$	$\frac{691}{16380}$	$\frac{1}{6}$	$\frac{3617}{4080}$
B_{r+1}	$\frac{1}{6}$	$\frac{1}{30}$	$\frac{1}{42}$	$\frac{1}{30}$	$\frac{5}{66}$	$\frac{691}{2730}$	$\frac{7}{6}$	$\frac{3617}{510}$

Ah! Yes. We spot that $126 = 3\times42$, $120 = 4\times30$, $16380 = 6\times2730$ and $4080 = 8\times510$, and arrive at the formula

$$\tan x = \sum_{r=0}^{\infty} (-1)^\beta\, 2^{r+1}(2^{r+1}-1)\, B_{r+1}\frac{x^r}{(r+1)!} \qquad (20)$$

where $\beta = 0$ unless $(r+1) \mod 4 = 0$ in which case $\beta = 1$. The expansion of $\tanh x$ is the same with $\beta = 0$ throughout.

An alternative look at this is to invert the formula for $\cot x$. Take the reciprocal of $\cot x$ from Eq 19 and expand it as a Taylor series in x. This gives a correct but inelegant series for $\tan x$ in terms of the Bernoulli numbers:

$$\frac{1}{x}\tan x = 1 + \frac{B_2(2x)^2}{2!} + \left(\left(\frac{B_2}{2!}\right)^2 - \frac{B_4}{4!}\right)(2x)^4 + (2x)^6\left(\left(\frac{B_2}{2!}\right)^3 - \frac{2B_2B_4}{2!.4!} + \frac{B_6}{6!}\right) + \ldots$$

the terms involving products of $B_r/r!$, r = 2, 4, 6, 8, etc. Let $Q_r = B_r/r!$.

$$\frac{1}{x}\tan x = 1 + Q_2(2x)^2 + \left(Q_2^2 - Q_4\right)(2x)^4 + \left(Q_2^3 - 2Q_2Q_4 + Q_6\right)(2x)^6$$
$$+ \left(Q_2^4 + Q_4^2 - 3Q_2^2Q_4 + 2Q_2Q_6 - Q_8\right)(2x)^8 + \ldots.$$

Comparison with Eq 20 points to some relations amongst the Bernoulli numbers, though they do not seem simple. I will not explore this further.

5.3 Euler's generating function

It is remarkable and surprising that sums of powers of integers should be related in some way to the derivatives of $\cot x$. Our explorations so far make us ask even more earnestly 'What are these Bernoulli numbers really?'. The great Leonard Euler must have been asking himself the same question at a time when he was exploring series, such as for the gamma function, a generalisation of $n!$ to non-integers. A generating function $G(x)$ is for a sequence S is a function whose power series expansion in x has the sequence S as its coefficients. The analogy is with a machine which, when we turn the handle, will crack out the next term in the required sequence. For instance, the so-called ‘cyclotomic polynomial’ is generated by the binomial expansion of $(1-x)^{-1}$. The sequence is 1, 1, 1, 1, 1, ...

$$\frac{1}{1-x} \;=\; 1+x+x^2+x^3+x^4+x^5+x^6+x^7+x^8+O\left(x^9\right), \quad |x|<1\,.$$

A sequence with gaps can be created by replacing x by a power of x, as in

$$\frac{1}{1-ax^2} \;=\; 1+ax^2+a^2x^4+a^3x^6+a^4x^8+a^5x^{10}+O\left(x^{11}\right)\,.$$

It is sometimes possible to deduce a generating function from a series with a regular pattern. For instance, if the one above is multiplied by $-ax^2$ and added to itself, the sum is 1: $\sigma - ax^2\sigma = 1$, so the sum, σ, is $1/(1-ax^2)$.

We almost have in Eq 19 a generating function for the Bernoulli numbers. As noted above, the coefficients in expanding $x\coth x$ are the same as for $x\cot x$ except for some changes of sign. $\coth x$ is expressed in terms of the exponential function by

$$\coth x \;=\; \frac{e^x+e^{-x}}{e^x-e^{-x}} \;=\; \frac{e^{2x}+1}{e^{2x}-1}\,.$$

This is an even function of x, symmetric about the vertical axis. Euler may have been looking for a function similar to this which would generate the Bernoulli numbers in sequence without the factorials and powers of 2 in Eq 19. The powers of 2 can be removed by replacing x by $x/2$. In §5.2 I introduced $Q_r = B_r/r!$. Then Eq 19 is

$$\frac{x}{2}\cot\frac{x}{2} \;=\; 1-Q_2x^2+Q_4x^4-Q_6x^6+Q_8x^8-\;\ldots\ldots\ldots$$

which we might feel is good enough. Euler perhaps suspected that the Bernoulli numbers also hide in a slightly simpler function corresponding essentially to one half of the symmetric $\coth x$. He discovered that

$$\frac{x}{e^x-1} \;=\; 1-\frac{x}{2}+\frac{x^2}{12}-\frac{x^4}{720}+\frac{x^6}{30240}-\frac{x^8}{1209600}+\frac{x^{10}}{47900160}-\;\ldots. = \sum_{r=0}^{\infty}\frac{B_r\,x^r}{r!}\,. \tag{21}$$

The ambiguity over the ± sign of B_1 comes from Euler adopting Eq 21 as the definition of the Bernoulli numbers. It makes B_1 negative. They all agree with the coefficient of n in ${}^kS_n$ except B_1 which is positive in the table in §1. See how close this series is to

$$\frac{x}{2}\cot\frac{x}{2} \;=\; 1-\frac{x^2}{12}-\frac{x^4}{720}-\frac{x^6}{30240}-\frac{x^8}{1209600}-\frac{x^{10}}{47900160}+\;\ldots..$$

$$\text{and}\;\; \frac{x}{2}\coth\frac{x}{2} \;=\; 1+\frac{x^2}{12}-\frac{x^4}{720}+\frac{x^6}{30240}-\frac{x^8}{1209600}+\frac{x^{10}}{47900160}+\;\ldots...$$

There is a hint that the Bernoulli numbers are linked to the singularity at $x=0$, arising when a singular function in the exponential-trigonometric family is expanded as a Taylor series. All the trigonometric functions are cousins of the exponential, so we may suspect that the B_n feature in expansions of other trig functions than $\cot x$ and $\coth x$, which also have a singularity at $x=0$. These could include $1/\sin x = 1/(x-\frac{x^3}{3!}+\frac{x^5}{5!}-\ldots)$ and $1/(1-\cos x)$. In fact Abramowitz and Stegun give at §4.3.68 in their *Handbook* a series for the singular $1/\sin x$ in terms of Bernoulli numbers. I conjecture that the Bernoulli numbers may play a role in the Taylor expansion of some singular functions similar to the role of the factorial in expanding e^x, $\sin x$ and other non-singular functions. To test this I examined the singular functions $1/(1-\cos x)$ and $1/(1-\cosh x)$ and found

$$\frac{x^2}{2(1-\cos x)} \quad \text{or} \quad \frac{x^2}{2(1-\cosh x)} = \sum_{r=0}^{\infty}(-1)^{\beta}\,\frac{(2r-1)}{(2r)!}\,B_{2r}\,x^{2r}\,. \tag{22}$$

$\beta = 0$ for the cosh function and $r+1$ for cos. I have not seen these in the published literature. On the other hand, Abramowitz and Stegun list series for $\tan x$, $\ln(\sin x/x)$, $\ln\cos x$ and $\ln(\tan x/x)$ which are not singular at $x = 0$ and yet contain Bernoulli numbers. The conjectured link, therefore, is tenuous. One further thought: in Taylor's theorem the factorial arises from repeated differentiation of x^n: thus nx^{n-1}, $n(n-1)x^{n-2}$, $n(n-1)(n-2)x^{n-3}$. We may wonder whether the B_n have some similar relationship. Their role seems to be confined to the exponential family of functions; I have found no evidence of them in series for Bessel functions, for example.

5.4 Euler and the Bernoulli polynomials

Euler was on a roll. He turned his generating function for fractions into one for polynomials by incorporating a further factor of e^{ux} into Eq 21:

$$\frac{x\,e^{ux}}{e^x-1} = (1+ux+\frac{1}{2!}u^2x^2+\frac{1}{3!}u^3x^3+\;....)(1-\frac{x}{2}+B_2\frac{x^2}{2!}+B_4\frac{x^4}{4!}+B_6\frac{x^6}{6!}+\;....)$$

$$= 1 + \left(u-\tfrac{1}{2}\right)x + \left(\frac{u^2}{2!}-\frac{u}{2}+\frac{B_2}{2!}\right)x^2 + \left(\frac{u^3}{3!}-\frac{u^2}{2.2!}+\frac{B_2\,u}{2!}\right)x^3 + \left(\frac{u^4}{4!}-\frac{u^3}{2.3!}+\frac{B_2u^2}{(2!)^2}+\frac{B_4}{4!}\right)x^4\,....$$

$$= 1 + \left(u-\tfrac{1}{2}\right)x + \tfrac{1}{12}\left(6u^2-6u+1\right)x^2 + \left(\tfrac{u^3}{6}-\tfrac{u^2}{4}+\tfrac{u}{12}\right)x^3$$

$$+ \left(\frac{u^4}{24}-\frac{u^3}{12}+\frac{u^2}{24}-\frac{1}{720}\right)x^4 + \left(\frac{u^5}{120}-\frac{u^4}{48}+\frac{u^3}{72}-\frac{u}{720}\right)x^5$$

$$+ \left(\frac{u^6}{720}-\frac{u^5}{240}+\frac{u^4}{288}-\frac{u^2}{1440}+\frac{1}{30240}\right)x^6 + \left(\frac{u^7}{5040}-\frac{u^6}{1440}+\frac{u^5}{1440}-\frac{u^3}{4320}+\frac{u}{30240}\right)x^7 + ...$$

$$= \sum_{r=0}^{\infty}Q_r(u)\,x^r = \sum_{r=0}^{\infty}B_r(u)\,\frac{x^r}{r!} \tag{23}$$

where I have generalised the Q quantity, introduced in §5.2, so that $Q_r(u)$ is the coefficient of x^r. The Bernoulli polynomials $B_r(u)$ are the coefficients of $x^r/r!$ in this expansion and have the Bernoulli numbers at their lowest powers of u. Explicitly

$$B_0(u) = 1, \qquad B_1(u) = u-\tfrac{1}{2}, \qquad B_2(u) = u^2-u+\tfrac{1}{6}, \qquad B_3(u) = u^3-\tfrac{3}{2}u^2-\tfrac{1}{2}u,$$

$$B_4(u) = u^4-2u^3+u^2-\tfrac{1}{30}, \qquad B_5(u) = u^5-\tfrac{5}{2}u^4+\tfrac{5}{3}u^3-\tfrac{1}{6}u\,.$$

$$B_6(u) = u^6-3u^4+\tfrac{5}{2}u^3-\tfrac{1}{2}u+\tfrac{1}{42}, \qquad B_7(u) = u^7-\tfrac{7}{2}u^6+\tfrac{7}{2}u^5-\tfrac{7}{6}u^3+\tfrac{1}{6}u\,,\ \textit{etc.} \tag{24}$$

He could have chosen to multiply the generating function of Eq 21 with many other functions and so spawned dozen of sets of polynomials, so why e^{ux}? The series above does contain the Bernoulli numbers as the lowest coefficients, but this is hardly surprising. The polynomials have the property that $\frac{d}{du}B_r(u) = rB_{r-1}(u)$. In fact the $Q(u)$ polynomials in the expansion at Eq 23 are the derivatives of each other, without the multiplying r, as a couple of examples show:

$$Q_1(u) = u-\tfrac{1}{2}. \qquad Q_2(u) = \tfrac{1}{12}\left(6u^2-6u+1\right). \quad \tfrac{d}{du}Q_2(u) = Q_1(u)\,.$$

$$Q_7(u) = \frac{u^7}{5040} - \frac{u^6}{1440} + \frac{u^5}{1440} - \frac{u^3}{4320} + \frac{u}{30240};$$

$$\frac{d}{du}Q_7(u) = \frac{7u^6}{5040} - \frac{6u^5}{1440} + \frac{5u^4}{1440} - \frac{3u^2}{4320} + \frac{1}{30240} = \frac{u^6}{720} - \frac{u^5}{240} + \frac{u^4}{288} - \frac{u^2}{1440} + \frac{1}{30240} = Q_6(u)\,.$$

Interesting. But it gets better. A glance back at the table in §5.1 and you will see that the B_{k-1} are very similar to the normalised polynomials for ${}^kS_n = \sum^n j^k$. There are two differences:

1. sign of the second highest term, in n^k, should be +, not −. This comes from Euler's definition of the Bernoulli numbers.
2. the ${}^kS_n$ for odd k have an extra constant term which should be deleted. The ${}^kS_n$ terminate in either n or n^2.

I will add an asterisk to the $B_n(u)$ notation to indicate that these two changes as to be imposed. With these

$$^kS_n = \frac{B^*_{k+1}(n)}{k+1}\,. \tag{25}$$

Comparing Eq 25 with Eq 18 and using the Euler definition of B_1 the polynomials are

$$B_k(u) = \sum_{r=0}^{k} B_{k-r}\,.{}^kC_r\,.u^r \quad \text{and} \quad {}^kS_n = \frac{1}{k+1}\sum_{r=1}^{k+1} B^*_{k+1-r}\,.{}^{k+1}C_r\,.n^r\,. \tag{26}$$

Note the different range of r. Moreover from Eq 21

$$B_k = \lim_{x\to 0}\frac{d^k}{dx^k}\frac{x}{e^x-1}\,. \tag{27}$$

Surely this is totally remarkable – awesome!, as some now say. This is perhaps the high point of this article. It means that exact formulae for the sums of integer powers can be written down in terms of the derivatives of several functions (mostly singular) in the exponential-trigonometric family.

A few related points can now be made.

- Almost the same polynomials are generated by $e^{ux}/(1-e^{-x})$. They differ only in some terms having the opposite signs.
- $B_n(1) - B_n = 0$. That is, the coefficients of $n^j,\ j > 0$ (with B_1 negative) sum to zero.
- In his classic book on series by Bromwich[3] that author uses a slightly different generating function for the Bernoulli polynomials. He uses

  $$x\,\frac{e^{ux}-1}{e^x-1}\,.$$

  This subtracts the Bernoulli numbers from the polynomials. This definition seems to have been superseded by the one above.
- Euler proposed another set of polynomials, named after himself, using a similar generating function:

  $$\frac{2\,e^{ux}}{e^x+1} = \sum_{r=0}^{\infty} E_r(u)\,\frac{t^r}{r!}\,.$$

  However, they seem to have fewer applications than the Bernoulli ones.

[3] T. J. l'A. Bromwich: 'Introduction to the Theory of Infinite Series' 1908, 1925

6 The Euler-Maclaurin sum formula

We now return to the issue in §2 on estimating the difference between the staircase sum of f_j and the integral of $f(x)$ where f is a smooth but otherwise fairly general function. You will recall that I tried using numerical integration to quantify the area of each curvi-triangle on the step over the interval $[j, j+1]$, with limited success. The method could not be extended because chain-cancellation fails for the second, fourth and higher even derivatives of $f(x)$.

We start by supposing the function we are concerned with the monomial function $M(x) = x^k$. We wish to deal with this function only, and with no other. Write out ${}^kS_n$ in full from Eq 26:

$$\sum_{j=0}^{n} j^k \equiv {}^kS_n = \frac{n^{k+1}}{k+1} + \frac{n^k}{2} + B_2\frac{k}{2!}n^{k-1} + B_3\frac{k(k-1)}{3!}n^{k-2} + B_4\frac{k(k-1)(k-2)}{4!}n^{k-3} + ...$$

This can be cast in terms of $M(n)$ only by expressing every term as a function of x^n, for which we use the formula for the derivatives of x^k. Remembering that $B_3 = B_5 = B_7 = ... = 0$,

$$ {}^kS_n = \int_0^n x^k\,dx + \frac{n^k}{2} + B_2\frac{1}{2!}\frac{d}{dx}(n^k) + B_4\frac{1}{4!}\frac{d^3}{dx^3}(n^k) + B_6\frac{1}{6!}\frac{d^5}{dx^5}(n^k) + ... \qquad (28)$$

The number of terms is $3 + (k-2)DIV2$. This is the most elementary form of the Euler-Maclaurin summation formula, discovered independently by these two mathematicians of the early 18th century. Euler (Swiss) and Maclaurin (Scot) were contemporaries of Benjamin Franklin, the composer Thomas Arne, and Thomas Simpson, inventor of Simpson's rule which I mentioned in §2.

The next step is to take the function of interest to be a polynomial, $P(x)$. Since this is a linear combination of monomials, the sum of $P(j)$ over integers j must be the sum of expressions like Eq 28, weighted appropriately. Suppose the polynomial is $P(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + ... + a_kx^k$. The formula is intended to evaluate

$$P(0) + P(1) + P(2) + ... + P(n) = (n+1)a_0 + {}^1S_na_1 + {}^2S_na_2 + {}^3S_na_3 + ... + {}^kS_na_k$$
$$= a_0 + \left(a_0 + \frac{a_1}{2} + \frac{a_2}{6} - \frac{a_3}{30}\right)n + \left(\frac{a_1 + a_2}{2} + \frac{a_3}{4} + ..\right)n^2 + ...$$

Taking Eq 28 at face value, what does each term contribute? The integral gives $a_0n+\frac{1}{2}a_1n^2+\frac{1}{3}a_2n^3+....$ The value of the integral at its lower limit is zero. The second term would contribute $\frac{1}{2}P(n)$ which has a constant term of $a_0/2$. No higher derivative contributes anything in a_0 so the factor weighting a_0 from the first two terms would be only $n + \frac{1}{2}$, whereas $n + 1$ is the correct value. The correction is to add the value $P(0)/2 = a_0/2$.

A converse situation occurs with the higher terms of Eq 28 involving the derivatives of $P(x)$. Since each derivative creates a constant term which is in general non-zero, this must be subtracted in order to give the correct weighting to a_1, a_2, etc. in turn. The constant term arising from $\frac{d^r}{dx^r}P(x)$ is $\frac{d^r}{dx^r}P(0)$, the substitution of $x = 0$ being performed after the differentiation. By these arguments we arrive at the following formula which will correctly sum a polynomial over integral values of its argument. The result is exact.

$$\sum_{j=0}^{n} P(j) = \int_0^n P(x)\,dx + \tfrac{1}{2}[P(n) + P(0)] + B_2\tfrac{1}{2!}\left[\tfrac{d}{dx}P(n) - \tfrac{d}{dx}P(0)\right]$$
$$+ B_4\frac{1}{4!}\left[\frac{d^3}{dx^3}P(n) - \frac{d^3}{dx^3}P(0)\right] + B_6\frac{1}{6!}\left[\frac{d^5}{dx^5}P(n) - \frac{d^5}{dx^5}P(0)\right] + ... \qquad (29)$$

Through expanding functions as series we are used to seeing smooth functions, possessing many high derivatives, as capable of being approximated by truncated power series, which are just polynomials. The summation formula for a general smooth function $f(x)$ is therefore conjectured to be

$$\sum_{j=0}^{n} f(j) = \int_0^n f(x)\,dx + \tfrac{1}{2}[f(n)+f(0)] + B_2\,\tfrac{1}{2!}[f'(n)-f'(0)]$$
$$+ B_4\,\frac{1}{4!}[f^{(3)}(n)-f^{(3)}(0)] + B_6\,\frac{1}{6!}[f^{(5)}(n)-f^{(3)}(0)] + \;.... \quad (30)$$

This is the Euler-Maclaurin summation formula. Though it seems plausible, it is still a significant step to write this series since, without some investigation, we cannot even be sure that it converges. I will not give a formal proof of Eq 30 since it is given in standard textbooks. It is probably sufficient to observe that it is based on the Taylor-Maclaurin expansion of $f(x)$ about 0 and we have specified that it must have a sufficient number of continuous derivatives. We could be sure that the series would converge if the Bernoulli numbers which multiply each of the later terms decreased in size. Unfortunately the contrary is the case – they increase rapidly: $B_{10} = 0\cdot076$, $B_{14} = 1\cdot17$, $B_{18} = 55\cdot0$, $B_{22} = 6192$, $B_{26} = 1,425,517$, etc., and may swamp the factorials on the denominators. To use the formula, you have to stop adding terms before they start increasing. For a general function f the summation formula is an valuable approximation to be used with care.

7 Comparison with numerical integration

It is fair to ask how the Euler-Maclaurin summation formula (EMSF) compares with the formulae developed in §2 which are essentially numerical integration schemes based on fitted polynomials of increasing degree. For a polynomial of degree k EMSF uses a polynomial of degree $k+1$ and so fits the given function exactly. The numerical integration schemes, in contrast, have fixed degree and cannot exactly sum polynomials of higher degree. However where the degree does not out-strip the fitted curve, the EMSF and numerical integration should agree exactly. I showed at Eq 7 that the cubic approximation gives the exact value for ${}^kS_n$ when $k \le 3$.

The approach in numerical integration, §2, has been to estimate the area of the curvi-triangle over $[n, n+1]$ then add then over integer n. This area is

$$\int_n^{n+1} f(x)\,dx - f(n)$$

$$= \tfrac{1}{2}[f(n+1)-f(n)] - \tfrac{B_2}{2!}\,[f'(n+1)-f'(n)] - \tfrac{B_4}{4!}\left[f^{(3)}(n+1)-f^{(3)}(n)\right] - ...$$
$$= \tfrac{1}{2}[f_{n+1}-f_n] - \tfrac{1}{12}\,[f_{n+1}-f'_n] + \tfrac{1}{720}\left[f^{(3)}_{n+1}-f^{(3)}_n\right] - \tfrac{1}{30240}\left[f^{(5)}_{n+1}-f^{(5)}_n\right] + ...$$

The corresponding formulae from §2 are

Linear: The trapezium of Eq 8 gives Area $= \frac{1}{2}[f_{n+1}-f_n]$.

Cubic : Eq 9 gives Area $= \frac{1}{2}[f_{n+1}-f_n] - \frac{1}{12}\,[f_{n+1}-f'_n]$.

5th degree : Eq 10 gives Area $= \frac{1}{2}[f_{n+1}-f_n] - \frac{1}{10}\,[f'_{n+1}-f'_n] + \frac{1}{120}\,[f''_{n+1}+f''_n]$.

7th degree : Eq 11 gives Area $= \frac{1}{2}[f_{n+1}-f_n] - \frac{3}{28}\,[f'_{n+1}-f'_n] + \frac{1}{84}\,[f''_{n+1}+f''_n] - \frac{1}{1680}\left[f^{(3)}_{n+1}-f^{(3)}_n\right]$.

There is only a superficial similarity amongst these 5 formulae. In detail most of the coefficients differ in magnitude, and some terms also change their signs. Note again how for the 5th and 7th degree fitted curves, the terms involving the second and fourth derivatives have a + sign within the bracket, which is why they will not chain-cancel. In contrast, all the terms of the EMSF have a − sign within the bracket [..]. One striking difference is that the EMSF does not involve the second derivative at all, or indeed any even derivatives. A further difference is the upper limits of integration of $f(x)$; n for the EMSF and $n+1$ for the numerical integrations.

It is fair to ask how good these are at summing non-polynomial series. As one test, we can use all five approximations to evaluate the sine integral at $x = 10$ given values and derivatives at $x = 0, 1, 2, \ldots 9, 10$. This is an oscillatory function defined by

$$\mathrm{Si}(x) \;=\; \int_0^x \frac{\sin t}{t}\, dt\,.$$

So $f(t) = (\sin t)/t$. The correct numerical value is Si(10)= $1\cdot658347594\ldots$. The sum $\sum_{j=0}^{9} \sin t/t = 2\cdot17892092447$ so clearly a large correction from the curvi-triangles is expected.

The derivatives are

$$f'(x) \;=\; \frac{\cos t}{t} - \frac{\sin t}{t^2}, \qquad f''(x) \;=\; \frac{2\sin t}{t^3} - \frac{2\cos t}{t^2} - \frac{\sin t}{t},$$

$$f^{(3)}(x) \;=\; -\frac{6\sin t}{t^4} + \frac{6\cos t}{t^3} + \frac{3\sin t}{t^2} - \frac{\cos t}{t}\,.$$

The powers of t in the denominators might suggest that $f(t)$ and these derivatives increase to infinity as $x \to 0$. However, by expanding as Taylor-Maclaurin series about $x = 0$, it is straightforward to show that $f(0) = 1$, $f'(0) = 0$, $f''(0) = -\frac{1}{3}$, $f^{(3)}(x) = 0$ and $f^{(4)}(x) = \frac{1}{5}$. These conditions make the function suitable to apply all five approximation formulae. To aid this account the table below lists values at integer increments from $t = 0$ to 10, and their values are plotted in Figure 2.

t	$\frac{\sin t}{t}$	$f'(t)$	$f''(t)$	$f^{(3)}(t)$
0	1	0	−0·33333	0
1	0·84147	−0·30117	−0·23913	0·17710
2	0·45465	−0·43540	−0·01925	0·23695
3	0·04704	−0·34568	0·18341	0·14659
4	−0·18920	−0·11611	0·24726	−0·02203
5	−0·19178	0·09509	0·15375	−0·14898
6	−0·04657	0·16779	−0·00936	−0·15535
7	0·09386	0·09429	−0·12080	−0·05593
8	0·12367	−0·03365	−0·11526	0·06141
9	0·04579	−0·10632	−0·02216	0·10862
10	−0·05440	−0·07847	0·07010	0·06288

Taking $n = 9$ in Eq 8 the trapezium rule gives Si(10) $\approx 1\cdot6517$, an $0\cdot4\%$ under-estimate. Considering how little effort goes into this calculation, the result is remarkably good. It may have been helped by the oscillations in the function, which would produce some positive errors, some negative, over the unit intervals.

The cubic approximation from Eq 9 takes us a large step forwards in accuracy, giving 1·65826, an under-estimate of only 9×10^{-5}. Calculation with the fifth order polynomial using Eq 10 involves

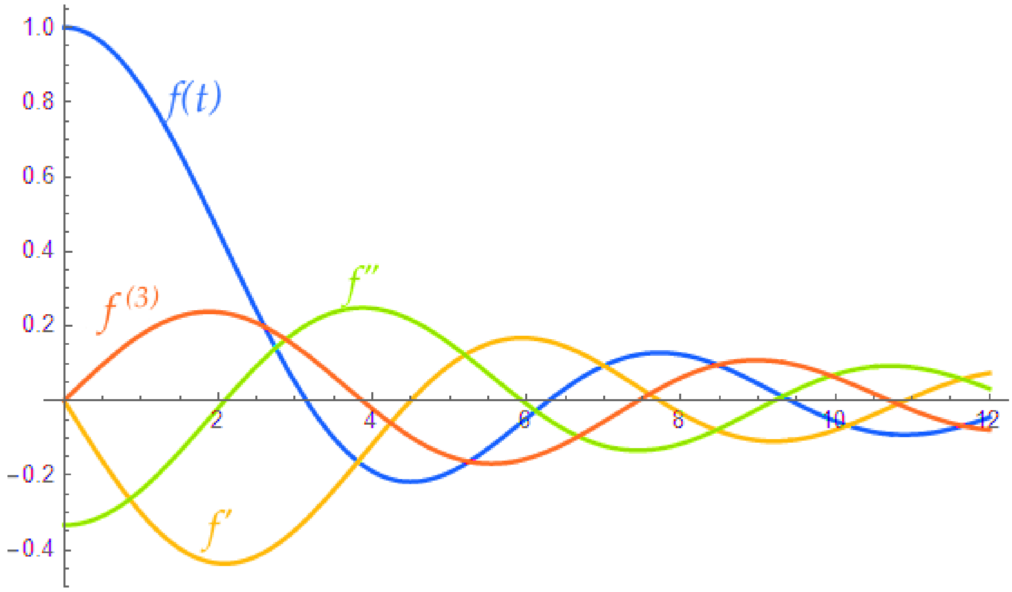

Figure 2: $\frac{\sin t}{t}$ and its derivatives.

summing over the second derivatives, so has become rather unwieldy. It gives a result accurate to $4 \cdot 5 \times 10^{-7}$: $1 \cdot 65834715$ compared with the true value $1 \cdot 65834759$. This error is only 1/200 of that using the cubic. Having accepted that a sum of second derivatives is necessary, little extra effort is needed to use the seventh order polynomial at Eq 11 since the only further information needed is the value of the 3rd derivative at $t = 10$. The result of $1 \cdot 658347593$ is accurate to 1×10^{-9}, an improvement of almost 400 in precision.

Turning to the EMSF, Eq 30, the value using derivatives up to the 3rd is $1 \cdot 65834757$, correct to $-2 \cdot 3 \times 10^{-8}$. In this one example it is has an error 20 times that of the 7th order fitted polynomial, but the saving in effort has been that the sum over 2nd derivatives has not been needed. If the next term in the EMSF is added (the one with $B_8/8!$), the error shrinks to -3×10^{-10}. It clearly is a powerful invention.

8 Postscript

This article opened with the nominally simple question, asked by many mathematicians over the centuries, 'What is the sum of the first n integers when each is raised to the same power k?'. The sum is a polynomial ${}^kS_n$ of degree $k + 1$ in n. We have traced two types of approach to finding its value, one using approximate numerical integration, and the other using insights into the numbers themselves, which involve binomial coefficients and the mysterious Bernoulli numbers. These Bernoulli numbers are generated by the Taylor-Maclaurin series expansion of $1/(e^x - 1)$, and occur in expansions of several other function in the exponential-trigonometric family, several of which have a singularity at the origin. The generation of Bernoulli numbers was extended by Euler into the generation of Bernoulli polynomials, B_n, which have the Bernoulli numbers as their constant terms. These Bernoulli polynomials are very close to the polynomials ${}^kS_n$. Eq 26,

$$ {}^kS_n = \frac{1}{k+1} \sum_{r=1}^{k+1} B^*_{k+1-r} \,.\, {}^{k+1}C_r \,.\, n^r \,, \tag{26} $$

is perhaps the high point of this study. The * means that B_1 has a positive sign and all polynomials terminate without a constant term.

§4.1 showed a further link with the Stirling numbers of the second kind, which count the number of partitions of n labelled objects into r unlabelled boxes. This suggests that the Bernoulli numbers, the exponential family of functions and the sum of powers are in some remote way linked to an allocation of certain quantities into categories, though I have not explored what this link might be. Towards the end of the article I have derived the Euler-Maclaurin summation formula and demonstrated its accuracy with one example function. Its main advantage over the numerical integration schemes is that it does not require summations of any derivatives – only the values of derivatives at the end points are needed. The EMSF can be used either to evaluate a sum when the integral is fairly easy to evaluate, or *vice versa* to find an integral given an easily calculated summation. Euler used it one way round, and Maclaurin the other. One application, Eq 14 of §3.1, is to approximate $\ln(n!)$. The result is similar to Stirling's classic asymptotic formula, and using the EMSF more terms could be appended.

John Coffey, September 2020