Discrete Calculus and Finite Sums

David Herrera 4/29/2016

Our goal is to calculate sums. Finite sums. Of course, finite sums lead to infinite sums, thought that is not what we are getting at here and the word "infinite" will likely not be used any more here.

Review and Introduction to Sums and Derivatives

So, we will begin by reviewing some notation. A function f is a rule that assigns an integer n to a real (or complex) number $f(n)$. Often f is given in terms of some formula, like $f(n) = n^2$ or $f(n) = e^n$, but other times it can be stated recursively with some initial conditions, like the fibonacci numbers $f(n+2) = f(n+1) + f(n)$, where $f(0) = 0$ and $f(1) = 1$. We will deal mainly with the first type of functions[1] and we will often consider only inputs of positive integers, though everything that we do also applies when the input for our functions are normal integers.

One of the best ways to view a function $f(n)$ is as a table. For example, for $f(n) = 2n$ and $g(n) = n^2$ the tables for inputs in the range 1 to 5 are:

$f(n) = 2n$

n	$f(n)$
1	2
2	4
3	6
4	8
5	10

Table 1

$g(n) = n^2$

n	$g(n)$
1	1
2	4
3	9
4	16
5	25

Table 2

Definition. If f is a function, we define $\sum_{k=a}^{b} f(k)$ by

$$\sum_{k=a}^{b} f(k) = f(a) + f(a+1) + \cdots + f(b-1) + f(b)$$

using what is called **Sigma notion**.

In the corner cases when, say, $a = b$ this sum is just $f(a) = f(b)$ and in the case when $a > b$ the sum is empty, which means that we are adding nothing so the answer

[1] functions that are defined recursively that have nice formulas (like linear recursions like the fibonacci sequence) can be expressed in a concise formula. Here is as link: [8] to a derivation of the linear recursion case, though it uses some basis linear algebra terminology and ideas.

is 0. The k in the notation is what we call a dummy variable; it does not actually have any significance except for telling us what we are doing with f. This means that we could have used any variable (other than a, b, or f) to indicate that we are adding the things from a to b: $\sum_{k=a}^{b} f(k) = \sum_{i=a}^{b} f(i)$. We could also have used n for the dummy variable, but we will not since often $a = 1$ and $b = n$ so we want to be consistent. Note that $\sum_{b=a}^{b} f(b)$ does not make much sense since the "$b = a$" means that we are adding $f(b)$ for every b between a and b.[2] Also, if we are going to be adding $f(k)$ for all the k's with a certain property P then we often write the sum like $\sum_{k \text{ satisfies } P} f(k)$, which means that we have the following equivalence of notations:

$$\sum_{k=a}^{b} f(k) = \sum_{a \le k \le b} f(k).$$

A very helpful way to visualize the sigma notation for a sum is, again, by using a table:

$$\begin{aligned}
\sum_{k=a}^{a-2} f(k) &= 0 \\
\sum_{k=a}^{a-1} f(k) &= 0 \\
\sum_{k=a}^{a} f(k) &= f(a) \\
\sum_{k=a}^{a+1} f(k) &= f(a) + f(a+1) \\
\sum_{k=a}^{a+2} f(k) &= f(a) + f(a+1) + f(a+2) \\
\sum_{k=a}^{a+3} f(k) &= f(a) + f(a+1) + f(a+2) + f(a+3) \\
\sum_{k=a}^{a+4} f(k) &= f(a) + f(a+1) + f(a+2) + f(a+3) + f(a+4)
\end{aligned}$$

Table 3

[2] You could probably make sense of it if you like, but it is not what we want it to mean.

Some important properties of the summation is if f and g are functions and c is a constant (it does not depend on k) then

$$\sum_{k=a}^{b}(f(k)+g(k)) = \sum_{k=a}^{b} f(k) + \sum_{k=a}^{b} g(k)$$

and

$$\sum_{k=a}^{b}(cf(k)) = c\sum_{k=a}^{b} f(k).$$

The first is true since it it does not matter the $f(k)$'s and $g(k)$'s and the LHS[3] says add the $f(k)$'s and $g(k)$'s with the same k first then add all of thses sums together and the RHS[4] says add the $f(k)$'s first then add the $g(k)$'s second and add the results together. For example,

$$\begin{aligned}\sum_{k=1}^{2}(f(k)+g(k)) &= (f(1)+g(1)) + (f(2)+g(2)) = (f(1)+f(2)) + (g(1)+g(2)) \\ &= \sum_{k=1}^{2} f(k) + \sum_{k=1}^{2} g(k).\end{aligned}$$

The second is true since this is just the distributive property. For example,

$$\sum_{k=1}^{3}(cf(k)) = cf(1)+cf(2)+cf(3) = c(f(1)+f(2)+f(3)) = c\sum_{k=1}^{3} f(k).$$

Another important, and almost dual, concept is that instead of adding everything, we could subtract things. Well, if we are going to subtract things then we need to figure out what to subtract and why we are subtracting. When we added things we did so because this was the goal of this article.[5] We could have instead have defined adding to be adding only the even numbers or the odd number or the prime numbers or... . We could try to phrase these problems using the alternate form of summation notation using a condition for the things that we are adding, but we are going to stick with the standard notation.

If we are going to define some sort of subtraction we want it to have some nice properties and the nice properties that we are after are going to be that they "undo" summation. That is, adding and subtracting are inverses, which is what we know from elementary school arithmetic: $(b+a)-a=b$.

[3]Left hand side
[4]Right hand side
[5]It's in the name!

However, it is not so clear how to define subtraction. Fortunately people [2][1] have already worked this out, so we do not have to be too original.[6]

What we need to notice is that if think about the differences between the rows (row − row before it) of Table 3 we get the following table:

$$\begin{array}{cc} & \text{NA} \\ \sum_{k=a}^{a-1} f(k) - \sum_{k=a}^{a-2} f(k) = & 0 \\ \sum_{k=a}^{a} f(k) - \sum_{k=a}^{a-1} f(k) = & f(a) \\ \sum_{k=a}^{a+1} f(k) - \sum_{k=a}^{a} f(k) = & f(a+1) \\ \sum_{k=a}^{a+2} f(k) - \sum_{k=a}^{a+1} f(k) = & f(a+2) \\ \sum_{k=a}^{a+3} f(k) - \sum_{k=a}^{a+2} f(k) = & f(a+3) \\ \sum_{k=a}^{a+4} f(k) - \sum_{k=a}^{a+3} f(k) = & f(a+4) \end{array}$$

Table 3

It does not take too much work to see that we have recovered our original sequence starting at a and it is zero before a.

We now define the notion of difference that we want:

Definition. If $f(n)$ is a sequence we define the **discrete/finite derivative** $\Delta_n f(n)$ to be $f(n+1) - f(n)$. Δ_n is called the **forward-difference operator** since it takes the difference between f at n and at the next number $n+1$.

Example. $\Delta_n(n^2) = (n+1)^2 - n^2 = n^2 + 2n + 1 - n^2 = 2n + 1$.

The discrete derivative has some of the nice properties that the summation does. If f and g are functions and c is a constant then

$$\Delta_n(f(n) + g(n)) = \Delta_n(f(n)) + \Delta_n(g(n))$$

[6]Did you think that I was citing the people who worked it out!? Of course not. I was citing people who are citing people who are citing people who either worked it out or are in the same position as the people as I am citing except one step closer to people who actually worked it out.

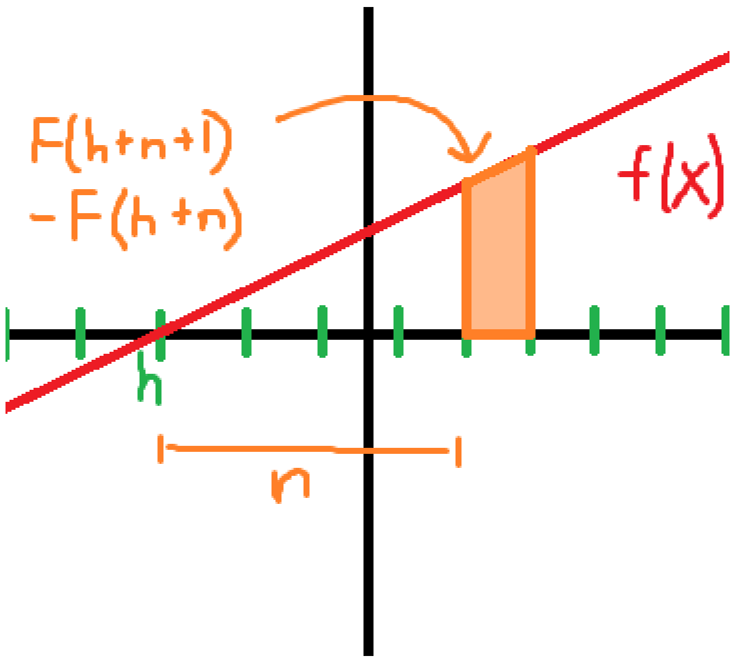

Figure 1: Picture 1

and

$$\Delta_n(cf(n)) = c\Delta_n(f(n)),$$

since

$$\begin{aligned}\Delta_n(f(n)+g(n)) &= (f(n+1)+g(n+1)) - (f(n)+g(n)) \\ &= (f(n+1)-f(n)) + (g(n+1)-g(n)) = \Delta_n(f(n)) + \Delta_n(g(n))\end{aligned}$$

and

$$\Delta_n(cf(n)) = cf(n+1) - cf(n) = c(f(n+1)-f(n)) = c\Delta_n(f(n)).$$

The discrete derivative measures how much f is changing at the nth step from $f(n)$ to $f(n+1)$.

Tangent.

In early high school / late middle school I was very interested in the following problem:

Consider a parabola: $f(x) = ax^2 + bx + c$. When plotting f starting at its vertex and going one unit out in each direction, how much does $f(x)$ change when x goes increases by one?

Put in a slightly different way, it says that if the parabola for $f(x)$ as a vertex at (h, k) then how much does f change from h to $h+1$, from $h+1$ to $h+2$, ... I had reasoned by experience that if $a = 1$ then the sequence of increases was the sequence of odd numbers. As you might imagine this might be useful knowledge to have. When graphing a parabola, you just find the vertex using the formula which will remain unnamed and unwritten then you plot it out like that. It is slightly annoying when h is not an integer, but you get the main idea.

When I learned calculus I had reasoned that this was true because of the fundamental theorem of calculus:

$$\int_a^b f(x)dx = F(b) - F(a) \text{ if } F'(x) = f(x),$$

since I could take $F(x) = ax^2 + bx + c$ then differentiate it to get $2ax + b$ and then get I could calculate the different by calculating the area, $\int_n^{n+1} f(x)dx$, which was a great result since it amounted to calculating the area of a trapezoid whose width was 1 and height given by a linear equation so I could calculate it and verify that my assertion was correct. I am not sure what I was trying to calculate, but this is how I did it. (I start at h, which is given by that formula again which is just the point h where $F'(h) = f(h) = 0$. See Picture 1 for a depiction of this.

Now, I did not realize that there is this elementary way of doing it, by just applying the discrete difference. I mean, this is of course the most obvious way of computing it in retrospect since if I want to know what $f(x+1) - f(x)$ is for some function f, then I should just calculate it, which is

$$(a(x+1)^2 + b(x+1) + c) - (ax^2 + bx + c) = a(x^2 + 2x + 1 - x^2) + b(1) = 2ax + a + b,$$

which explains why I got that the differences for $a = 1$ away from the center is the sequence of odd numbers since it is $2x+1$ (I also do not think that I knew or thought of that every parabola $ax^2 + bx + c$ is just a horizontally shifted version of $ax^2 + c$ by the vertex form of a quadratic.)

Motivating Examples. We will see the "telescoping sum" trick later and in fact it is the basis for all of the methods of finite calculus that we will see. The idea is simply this:
Consider the sum:

$$(2-1) + (3-2) + (4-3) + (5-4).$$

You could simply notice that each of these is 1 then add them to get the answer, but that defeats the point of the exercise. What the important observation for the telescoping sum trick is that you are adding a bunch of that cause mass-cancellation.

For instance, each number except -1 and 5 will cancel to get $5-1=4$, which is nice. This type of sum where all the middle terms cancel out is called **telescoping**.

We can see this in a more general scenario:
Consider the sum:

$$\sum_{k=1}^{n} \frac{1}{k(k+1)} = \frac{1}{2} + \frac{1}{6} + \frac{1}{12} + \frac{1}{20} + \cdots + \frac{1}{n(n+1)},$$

which is not obviously equal to anything in particular. However, if we were to calculate this for, say, $n=3$ we would find after some algebra that this is nice. The underlying reason that this sum works out so nice is that we can write $\frac{1}{k(k+1)} = \frac{1}{k} - \frac{1}{k+1}$, by using partial sums decomposition, so what we have is that this sum telescopes leaving only the first $\frac{1}{1}$ and the last $-\frac{1}{n+1}$ term. This tells us that if we let n get large then the sum approaches 1, which is not at all what we would get from the original series by just staring at it

So, now we are going to approach the general problem hoping that we can always write $f(n)$ in terms of some telescoping sum.

We now slightly change the type of adding that we will do so that the following theorem[7] looks nice.

Definition. If $f(n)$ is a sequence we define the **discrete/finite integral** from a to b, $\sum\limits_{k:a\to b} f(k)$, to be $\sum\limits_{k=a}^{b-1} f(k)$.

Clearly, the discrete integral has the above properties that we saw held for summation since the discrete integral is just a summation.

So, the example we saw earlier with taking the differences of rows of Table 3 can be generalized to the first part of the fundamental theorem of discrete/finite calculus:

Theorem (Fundamental Theorem of Discrete Calculus). If f is a sequence then if $n > a$ then

$$\Delta_n \left(\sum_{k:a\to n} f(k) \right) = f(n)$$

and

$$\sum_{k:a\to b} (\Delta_k f(k)) = f(b) - f(a)$$

The first statement says that accumulation of all the values that f has taken on from a to $n-1$ grows by $f(n)$ when you add $f(n)$ to it (of course!) and the second

[7] which is just fancy name for "important result"

says that the accumulation from a to b of the changes that f does between each step from k to $k+1$ is just the total change that f does from a to b.

These statements make sense. In fact, they are almost self-evident. The important idea behind these is that it turns the heuristical paragraph that I wrote after the theorem, which is a story about how things work, into the theorem which is a collection of mathematical tools, made precise and clear.

Proof:

$$\begin{aligned}\Delta_n\left(\sum_{k:a\to n} f(k)\right) &= \sum_{k:a\to n+1} f(k) - \sum_{k:a\to n} f(k) = \sum_{k=a}^{n} f(k) - \sum_{k=a}^{n-1} f(k)\\ &= (f(a)+f(a+1)+\cdots+f(n-2)+f(n-1)+f(n))\\ &\quad -(f(a)+f(a+1)+\cdots+f(n-2)+f(n-1)) = f(n),\end{aligned}$$

since all the other terms cancel.

$$\begin{aligned}\sum_{k:a\to b}(\Delta_k f(k)) &= \sum_{k:a\to b}(f(k+1)-f(k)) = \sum_{k=a}^{b-1}(f(k+1)-f(k))\\ &= (f(a+1)-f(a))\\ &+ (f(a+2)-f(a+1))\\ &+ (f(a+3)-f(a+2))\\ &+ \cdots\\ &+ (f(b-1)-f(b-2))\\ &+ (f((b-1)+1)-f(b-1))\\ &= f(n)-f(a),\end{aligned}$$

because all the other terms cancel. ■

The most important part of this theorem in calculating $\sum_{k=1}^{n} f(k)$ will be to find a **discrete antiderivative** F which is a function such that its discrete derivative $\Delta_k F(k) = f(k)$ so that we can just apply the second part of the FTDC to say that the sum we were inspecting is $F(n+1)-F(1)$. Note that if F is a discrete antiderivative of f then so is $F+C$ where C is any constant since

$$\Delta_n(F(n)+C) = (F(n+1)+C)-(F(n)+C) = F(n+1)-F(n) = \Delta_n F(n).$$

Calculating some sums

We will now discuss some of the "normal" ways that people calculate sums. These are usually not as helpful as the tools that we will develop using discrete calculus, though they are important ideas to understand in general and it helps give some motivation for a useful and generalizable method. We now define the principle of mathematical induction, as given in [7] with minor modifications, which is a method to prove that some given statement S is true.

Proposition (Induction). If $S(n)$ is some statement that depends on an integer n then if

1. $S(1)$ is true and
2. $S(n)$ being true implies $S(n+1)$ is true.

then $S(n)$ is true for every n.

This is just saying that if $S(1)$ is true then $S(2)$ is true then $S(3)$ is true then ... . So, $S(n)$ is true for every n.

Example. We can show that $\sum_{k=1}^{n} k = n(n+1)/2$.

This is true since it is true for $n = 1$ since $\sum_{k=1}^{1} k = 1 = (1)(2)/2$ and if we know that this formula is true for n then we will show that it is true for $n+1$:

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

So, this formula is true for every n.

There is also the classic way of calculating the sum, which is the way to calculate a general arithmetic sum.

If you have an arithmetic sequence: $a_0, a_0+d, a_0+2d, \dots$, that is, $a_n = a_0 + nd$, then the series is

$$(a_0) + (a_0+d) + (a_0+2d) + \cdots + (a_0+(n-1)d) + (a_0+nd),$$

but after noticing that adding this series to itself, but backwards is

$$\begin{aligned}
&(a_0) + (a_0+d) + (a_0+2d) + \cdots + (a_0+(n-1)d) + (a_0+nd)\\
&+ (a_0+nd) + (a_0+(n-1)d) + \cdots + (a_0+d) + (a_0)\\
&= (2a_0+nd) + (2a_0+nd) + \cdots + (2a_0+nd) + (2a_0+nd)\\
&= 2(n+1)\left(a_0 + \frac{n}{2}d\right) = 2\left(a_0 n + a_0 + d\frac{n^2+n}{2}\right) = 2\left(\frac{d}{2}n^2 + \left(a_0 + \frac{d}{2}\right)n + a_0\right)
\end{aligned}$$

So, the total sum is $\frac{d}{2}n^2 + \left(a_0 + \frac{d}{2}\right) n + a_0$, which when we plug in $a_0 = 0$ and $d = 1$ we recover the formula we calculated above since

$$\frac{1}{2}n^2 + \frac{1}{2}n = \frac{n^2+n}{2} = \frac{n(n+1)}{2}.$$

For practice, we can calculate this sum using sigma notation:

$$\begin{aligned}\sum_{k=0}^{n}(a_0 + kd) &= \frac{1}{2}\left(\sum_{k=0}^{n}(a_0 + kd) + \sum_{k=0}^{n}(a_0 + (n-k)d)\right) \\ &= \frac{1}{2}\sum_{k=0}^{n}(2a_0 + nd) = a_0(n+1) + \frac{d}{2}n(n+1),\end{aligned}$$

which is what we got above.

Note that the sigma notation makes it much more clean and clear.

Note that we cannot (easily) do this sort of trick with a sum, like, $\sum_{k=1}^{n} k^2$. However, if we find a formula then we can prove it to be true by induction. For instance if you guess that $\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$, which is true, then you can show it by induction by showing that it holds for $n = 1$, which is does, and showing that if it holds for n then it holds for $n+1$:

$$\begin{aligned}\sum_{k=1}^{n+1} k^2 &= (n+1)^2 + \sum_{k=1}^{n} k^2 = \frac{6(n+1)(n+1)}{6} + \frac{n(n+1)(2n+1)}{6} \\ &= \frac{(n+1)(2n^2+n+6n+6)}{6} = \frac{(n+1)(2n^2+7n+6)}{6} = \frac{(n+1)(n+2)(2n+3)}{6}\end{aligned}$$

This is great, but it is unfortunate that it does not help us find a formula, just verify its validity.

So, as in "normal" calculus, we will find out some functions whose derivatives and derivatives are nice.

First, we will find the function whose derivative is itself: $\Delta_n(f(n)) = f(n)$.

So, if this is true that $f(n+1) - f(n) = f(n)$, so $f(n+1) = 2f(n)$, so of the many functions that satisfy this, the easiest is 2^n. So, $\Delta_n(2^n) = 2^n$. In general, $\Delta_n a^n = a^{n+1} - a^n = (a-1)a^n$.

We also want to find what polynomials have easy derivatives. The answer is $n^{\underline{\ell}}$, the falling product (of ℓ terms): $n^{\underline{\ell}} = n(n-1)(n-2)\cdots(n-\ell+2)(n-\ell+1)$.

We can see:

$$\begin{aligned}\Delta_n(n^{\underline{\ell}}) &= (n+1)^{\underline{\ell}} - n^{\underline{\ell}} \\ &= (n+1)(n)(n-1)\cdots(n-\ell+3)(n-\ell+2) \\ &\qquad - n(n-1)(n-2)\cdots(n-\ell+2)(n-\ell+1) \\ &= ((n+1)-(n-\ell+1))(n)(n-1)(n-2)\cdots(n-\ell+2) \\ &= \ell n(n-1)(n-2)\cdots(n-\ell+2) = \ell n^{\underline{\ell-1}}\end{aligned}$$

So, we will will now use the second part of FTDC.

Suppose that we want to compute the geometric sum (for $r \neq 1$): $\sum_{k=0}^{n} r^k = 1 + r + r^2 + \cdots + r^{n-1} + r^n$. The trick way is to note that if we multiply the whole sum by r then we only change the sum slightly to $r + r^2 + r^3 + \cdots + r^n + r^{n+1}$ and we can note that this differs from the original sum by only a few terms so

$$r\sum_{k=0}^{n} r^k - \sum_{k=0}^{n} r^k = (r+r^2+r^3+\cdots+r^n+r^{n+1}) - (1+r+r^2+\cdots+r^{n-1}+r^n) = r^{n+1}-1$$

so $\sum_{k=0}^{n} r^k = \frac{r^{n+1}-1}{r-1}$. Using only sigma notation is the not-so-simple but important if you are interested in further mathematics:

$$\begin{aligned}r\sum_{k=0}^{n} r^k - \sum_{k=0}^{n} r^k &= \sum_{k=0}^{n} r(r^k) - \sum_{k=0}^{n} r^k = \sum_{k=0}^{n} r^{k+1} - \sum_{k=0}^{n} r^k \\ &= \sum_{k=1}^{n+1} r^k - \sum_{k=0}^{n} r^k = r^{n+1} + \sum_{k=1}^{n} r^k - \sum_{k=1}^{n} r^k - 1 = r^{n+1} - 1.\end{aligned}$$

Now, we can try to solve the problem using discrete calculus. We can use the second part of the FTDC to calculate this sum. The way that we do this is to find an antiderivative F such that $\Delta_n F(n) = f(n)$ since then $\sum_{k=0}^{n-1} f(k) = F(n) - F(0)$.

So, we have $f(n) = r^n$ so $\Delta_n f(n) = (r-1)f(n)$, so that means that an antiderivative of f is $\frac{1}{r-1}r^n$ so

$$\sum_{k=0}^{n-1} r^k = \frac{1}{r-1}r^n - \frac{1}{r-1}r^0 = \frac{r^n-1}{r-1},$$

but this is just the same formula we had to work for earlier... and we got it half off!

Suppose that we now want to compute $\sum_{k=1}^{n} k^2 = \sum_{k:1\to n-1} k^2$, so all we have to do is find a function F such that $\Delta_n F(n) = f(n)$. The trick is to do is algorithmically.

Since $\Delta_n(n^{\underline{\ell}}) = \ell n^{\underline{\ell-1}}$, we have that an antiderivative of $n^{\underline{\ell}}$, which is a polynomial of degree ℓ, is $\frac{1}{\ell} n^{\underline{\ell+1}}$. So, we know that if $\ell \geq 1$ then

$$\sum_{k=1}^{n-1} n^{\underline{\ell}} = \frac{n^{\underline{\ell+1}}}{\ell+1},$$

since $1^{\underline{\ell+1}} = 1 \cdots (1-\ell) = 0$ since $\ell \geq 1$.

The way that we will calculate any polynomial sum is to write it in terms of sums of these falling powers.

So, since $n^{\underline{2}} = n(n-1) = n^2 - n = n^2 - n^{\underline{1}}$, we have that $n^2 = n^{\underline{2}} + n^{\underline{1}}$ so

$$\begin{aligned}\sum_{k=1}^{n-1} k^2 &= \sum_{k=1}^{n-1} k^{\underline{2}} + \sum_{k=1}^{n-1} k^{\underline{1}} \\ &= \frac{1}{3} n^{\underline{3}} + \frac{1}{2} n^{\underline{2}} = \frac{1}{6}(2n(n-1)(n-2) + 3n(n-1)) \\ &= \frac{n(n-1)(2n-4+3)}{6} = \frac{n(n-1)(2n-1)}{6},\end{aligned}$$

so plugging in $n+1$ instead of n gives $\frac{n(n+1)(2n+2-1)}{6} = \frac{n(n+1)(2n+1)}{6}$, which is what we proved earlier.

The outcome of all this is that using induction we can prove that a formula holds... but using discrete calculus we can actually find the formula in a rigorous way and without any tricks.

Theorem. If $p(k)$ is a monic polynomial (the coefficient of the term in the polynomial of highest degree is 1) in k of degree ℓ then $\sum_{k=1}^{n} p(k)$ is a polynomial in n of degree $\ell + 1$ with leading coefficient $\frac{1}{\ell+1}$

Proof: If we show that this is true for every polynomial of the form $p(k) = n^{\ell}$ then it holds in general since if $p(k) = k^{\ell} + a_1 k^{\ell-1} + a_2 k^{\ell-2} + \cdots + a_{\ell-1} k + a_{\ell}$ then

$$\sum_{k=1}^{n} p(k) = \sum_{k=1}^{n} k^{\ell} + a_1 \sum_{k=1}^{n} k^{\ell-1} + a_2 \sum_{k=1}^{n} k^{\ell-2} + \cdots + a_{\ell-1} \sum_{k=1}^{n} k + a_{\ell} \sum_{k=1}^{n} 1$$

so the term with the highest degree will come from the $\sum_{k=1}^{n} k^{\ell}$ sum.

So, now if we want write k^{ℓ} in terms of falling polynomials, we know that $k^{\underline{\ell}} = k(k-1)\cdots(k-\ell+1)$ is an ℓ degree polynomial, so have that $k^{\ell} - k^{\underline{\ell}}$ is a polynomial of degree at most $\ell - 1$. This means that $\sum_{k=1}^{n} k^{\ell}$ and $\sum_{k=1}^{n} k^{\underline{\ell}}$ have the same highest degree ℓ term since if they did not then their difference would have a term of degree

ℓ. So, by our above argument, the highest degree term of $\sum_{k=1}^{n} k^{\ell}$ and $\sum_{k=1}^{n} k^{\underline{\ell}}$ are the same.

However, we know what the highest degree term of $\sum_{k=1}^{n} k^{\underline{\ell}}$ is.

$$\sum_{k=1}^{n} k^{\underline{\ell}} = \frac{1}{\ell+1} k^{\underline{\ell+1}} = \frac{1}{\ell+1} k(k-1)\cdots(k-\ell),$$

so we know that the highest degree term is $\frac{1}{\ell+1} k^{\ell+1}$. ■

We now proceed to some different applications, but first we need to know a little more about the discrete derivative. We actually are only going to use the first one, but it gives some intuition about the relationship between this and "normal" calculus since similar formulas hold for the "normal" derivative.

Proposition. We have the following:

1. The product rule:

$$\Delta_n(f(n)g(n)) = \Delta_n(f(n))g(n+1) + f(n)\Delta_n(g(n)).$$

2. Integration by parts:

$$\sum_{n:a\to b} f(n)\Delta_n(g(n)) = (f(b)g(b) - f(a)g(a)) - \sum_{n:a\to b} \Delta_n(f(n))g(n+1).$$

Proof: In general, when showing equalities, it is useful to take the more complicated version and make it look more like the simpler. So, we will show the first equation by simplifying the right hand side of it:

$$\begin{aligned}
&\Delta_n(f(n))g(n+1) + f(n)\Delta_n(g(n)) \\
&\qquad = (f(n+1) - f(n))g(n+1) + f(n)(g(n+1) - g(n)) \\
&\qquad = f(n+1)g(n+1) - f(n)g(n+1) + f(n)g(n+1) - f(n)g(n) \\
&\qquad = \Delta_n(f(n)g(n)).
\end{aligned}$$

■

Integration by parts, just as in the normal calculus, follows immediately from the product rule:

$$\begin{aligned}\sum_{n:a\to b} f(n)\Delta_n(g(n)) &= \sum_{n:a\to b}(\Delta_n(f(n)g(n)) - \Delta_n(f(n))g(n+1))\\ &= \sum_{n:a\to b}(\Delta_n(f(n)g(n)) - \sum_{n:a\to b}\Delta_n(f(n))g(n+1))\\ &= f(b)g(b) - f(a)g(a) - \sum_{n:a\to b}\Delta_n(f(n))g(n+1))\end{aligned}$$

Integration by parts allows us to move the derivative from one function in a product to another function, but we have to add a minus sign plus there are "boundary terms" and increase the k of the function that we are moving the derivative away from.

Example. Suppose that we want to calculate $\sum_{k=1}^{n}\frac{k}{2^k}$ (we will substitute in $n-1$ so we can use the fundamental theorem). The point is that since $\frac{1}{2^k} = (1/2)^k$ it has a nice derivative and antiderivative so I can make it into a derivative: $\frac{1}{2^k} = -2\Delta_k(\frac{1}{2^k})$, so now when I put this in the sum we get that it equals $-2\sum_{k=1}^{n-1} k\Delta_k(\frac{1}{2^k})$, so we can use integration by parts to move the derivative from the exponential to the k, which gives us:

$$\begin{aligned}-2\left(\sum_{k=1}^{n-1} k\Delta_k(\frac{1}{2^k})\right) &= -2\left(\frac{k}{2^l}\bigg|_{k=1}^{n} - \sum_{k=1}^{n-1}\Delta_k(k)\frac{1}{2^{k+1}}\right)\\ &= -2\left(\frac{n}{2^n} - \frac{1}{2^1} - \frac{1}{2}\sum_{k=1}^{n-1}\Delta_k(k)\frac{1}{2^k}\right)\\ &= -\frac{2n}{2^n} + 1 + \sum_{k=1}^{n-1}\Delta_k(k)\frac{1}{2^k}\\ &= -\frac{2n}{2^n} + \sum_{k=0}^{n-1}\Delta_k(k)\frac{1}{2^k}\\ &= \frac{\left(\frac{1}{2}\right)^n - 1}{\frac{1}{2} - 1} - \frac{n}{2^{n-1}}\\ &= -2\left(\frac{1}{2}\right)^n - 1 - \frac{n}{2^{n-1}}\\ &= -\frac{1}{2^{n-1}} + 2 - \frac{n}{2^{n-1}} = 2 - \frac{n+1}{2^{n-1}}\end{aligned}$$

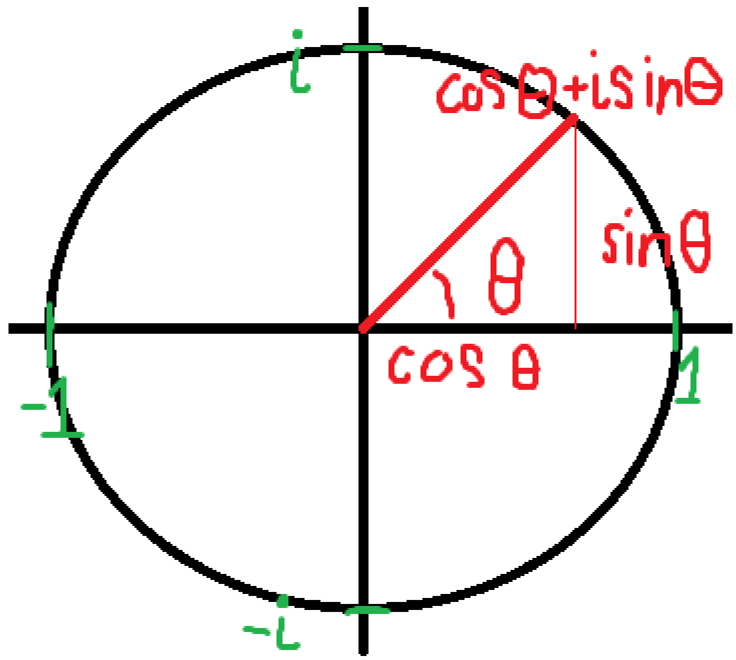

Figure 2: The Unit Circle

So, we have proved that

$$\sum_{k=1}^{n} \frac{k}{2^k} = 4 - \frac{n+2}{2^n},$$

which is nice. It gives us that if we make n larger and larger this sum approaches 2.[8]

Example.

We now tackle a harder and "real" problem in mathematics.[9]

Recall the complex number plane of points in the plane $a + ib$, where $i = \sqrt{-1}$ by definition. Recall that the length of a complex number $z = a + ib$ is $\sqrt{a^2 + b^2}$, which is the distance between points in $\mathbb{R}^2$, the plane.

Recall that the unit circle is the points that are a distance 1 from the origin. Now, using trigonometry, we can write that if the radius from a point on the unit circle has an angle θ from the positive x-axis then we can write this point as $(\cos\theta, \sin\theta)$, but

[8]When deriving this I forgot to increase k inside the function that I was moving the derivative to so when I wrote a Java program to check my work, I found out that I did it wrong (I actually did not need a program to say that I was wrong, but it would also help me check that it is right since I did not feel like calculating the sum.) So, this is a warning: Be careful.

[9]This is a joke in two ways. First, because I am calling what we have done not real mathematics. Second, we will use complex numbers.

since we are in the complex plane this point is $\cos\theta + i\sin\theta$. As a function of θ, this is a very special function, which can be abbreviated to $\text{cis}\,\theta$. (It is often written as $e^{i\theta}$ since it is an analytic extension of the function e^x to allow you to plug in complex numbers).

We have the following property:

Proposition (DeMoivre's Theorem). For any natural number $(\text{cis}\,\theta)^n = \cos(n\theta)$.

You can prove this by induction using the addition properties of sine and cosine. We will not do this (consider it an exercise).

This is a manifestation of the way that complex numbers multiply in general. When you apply the rules of multiplying complex numbers you are actually making a new complex number whose length is the product of the lengths of the origin complex numbers ($|zw| = |z||w|$) and you are adding the angles. This is explained in that $i(-i) = 1$ since if you rotate 90^o then rotate another 270^o you get 260^o, which is not rotating at all, so $i(-i) = 1$.

To not lose sight, what we are going to prove is this:

Proposition. For any θ where $\cos(\theta) \neq 1$, $\sum_{k=1}^{n} \cos(k\theta)$ is bounded by some number that depends on θ but not on n. That is, there is a positive number M_θ that depends on θ but not on n such that $|\sum_{k=1}^{n} \cos(k\theta)| \leq M_\theta$ for every n.

Before I prove this, I want to hand-wave why this feels right. First off, if we just add any function then we might get odd and increasing behavior. Cosine is not like that. It is periodic and uniform so that it feels like the positives values in this sum should be around as many and as much as the negatives so they basically cancel each other out. A better justification of this intuition is based on two scenarios:

If θ is a multiple of 2π then cos

If θ is a rational multiple of 2π then we have that this sequence will be completely periodic and the sum will be periodic so of course it is bounded in a way that does not depends on n. It depends on θ in that the period could be long or short and if it is long then it might spend a lot of time where $\cos(k\theta)$ is positive so the sum could get very large.

If θ is not a rational multiple of 2π then there is an important result in ergotic theory that tells us what happens here. It says that if we look at the angles $k\theta$ that are possible and think about what points they describe on the unit circle then if we keep track of all the points that we meet by rotating by θ then rotating again by θ, then ... , we get a set of points that basically fills out the entire circle, but not quite since it does not hit many points (like the ones that are rational multiples of 2π) but it gets very close to every point. Not only that, but it spreads out around the unit circle uniformly.

So, what we get as a result of that is that the cosine function, which just takes an angle and spits out the x-coordinates should have the property that there are about the same number of positives and negatives and they are all around the same size when added together, so this should be bounded regardless of how large n is, but how bounded it is, just as before, certainly depends on θ for the same reason as before.

We now prove this.

Proof: The idea is to write $\cos(k\theta)$ in terms of the cis function which we know behaves like an exponential.

So, we have that $\text{cis}(k\theta) = \cos(k\theta) + i\sin(k\theta)$, so how do we "extract" the cosine part of it. The trick is to look at the complex conjugate: $\cos(k\theta) - i\sin(k\theta)$. First, we notice that since cosine is even and sine is odd that this equals $\cos(-k\theta) + i\sin(-k\theta) = \text{cis}(-k\theta)$.

So, now, using some elementary algebra we get:

$$\cos(k\theta) = \frac{\text{cis}(k\theta) + \text{cis}(-k\theta)}{2} = \frac{\text{cis}(\theta)^k + \text{cis}(-\theta)^k}{2}$$

Now, we have something that we could plug into our geometric sum!

Since $\cos(0\theta) = \cos(0) = 1$, we can extend the sum that we are looking at to equal $\sum_{k=0}^{n} \sin(k\theta)$, by only increasing its value by 1. So, evaluating this using the decomposition above we get

$$\begin{aligned} &\frac{1}{2}\left(\sum_{k=0}^{n} \text{cis}(\theta)^k + \sum_{k=0}^{n} \text{cis}(-\theta)^k\right) \\ &\quad = \frac{1}{2}\left(\frac{\text{cis}(\theta)^{n+1} - 1}{\text{cis}(\theta) - 1} + \frac{\text{cis}(-\theta)^{n+1} - 1}{\text{cis}(-\theta) - 1}\right) \end{aligned}$$

Now, we just need to say that this is bounded. Now, we will use the triangle inequality: if z and w are complex numbers then $|z+w| \leq |z| + |w|$. So, we have that the absolute value of our sum is at most

$$
\begin{aligned}
&\frac{1}{2}\left|\frac{\operatorname{cis}(\theta)^{n+1}-1}{\operatorname{cis}(\theta)-1}+\frac{\operatorname{cis}(-\theta)^{n+1}-1}{\operatorname{cis}(-\theta)-1}\right| \\
\leq\, &\frac{1}{2}\left|\frac{\operatorname{cis}(\theta)^{n+1}-1}{\operatorname{cis}(\theta)-1}\right|+\left|\frac{\operatorname{cis}(-\theta)^{n+1}-1}{\operatorname{cis}(-\theta)-1}\right| \\
\leq\, &\frac{|\operatorname{cis}(\theta)^{n+1}|+1}{2|\operatorname{cis}(\theta)-1|}+\frac{|\operatorname{cis}(-\theta)^{n+1}|+1}{2|\operatorname{cis}(-\theta)-1|}
\end{aligned}
$$

So, noting that for any ψ, $|\operatorname{cis}(\phi)| = |\cos(\phi)+i\sin(\phi)| = \sqrt{\cos^2(\phi)+\sin^2(\phi)} = 1$, we have that the entire thing that we got does not depend on n, so we are done. ■

In fact, we have the following result which comes right from the proof above.

Proposition. For any θ with $\cos\theta \neq 1$, we have that

$$
\left|\sum_{k=1}^{n}\cos(k\theta)\right| \leq \frac{2}{\sqrt{2(1-\cos\theta)}}+1
$$

Proof: We just need to continue messing with the ugly thing we got at the end of the last proof.

The first thing that we can say is that since $\operatorname{cis}(\theta)$ and $\operatorname{cis}(-\theta)$ have the same real part and the imaginary parts are just negatives of each other, $|\operatorname{cis}(\theta)+1| = |\operatorname{cis}(-\theta)+1|$, even though they are not necessarily equal. So, with adding in the concluding remark of the previous proof, we can simply the bound we found to get:

$$
\frac{2}{2|\operatorname{cis}(\theta)-1|}+\frac{2}{2|\operatorname{cis}(\theta)-1|}=\frac{2}{|\operatorname{cis}(\theta)-1|},
$$

now we just want a way to deal with the denominator.

So,

$$
\begin{aligned}
|\operatorname{cis}(\theta)-1| = |-1+\cos(\theta)+i\sin\theta| &= \sqrt{(1-\cos(\theta))^2+\sin^2(\theta)} \\
&= \sqrt{1-2\cos(\theta)+\cos^2(\theta)+\sin^2(\theta)} = \sqrt{2(1-\cos(\theta))},
\end{aligned}
$$

so we are done once we recall that the $+1$ comes from the fact that we added in the $\cos(0\theta) = 1$ in the previous theorem, so we include it in the bound. ■

Note that the bound which depends only on θ increases as $\cos(\theta)$ gets closer to 1, which is exactly, what we discussed since if θ is very small, that is, $\cos(\theta)$ is very closed to 1 then the sum can be very big.

Here is a table for some of the values of the sum and our upper bound

Exercises.

1. Verify our calculation of $\sum_{k=1}^{n} \frac{1}{k(k+1)}$ by induction.
2. Calculate the series $\sum_{k=1}^{n} n^3$.
3. Prove that your formula holds by induction.
4. Verify the triangle inequality. (Hint: Show that the inequality holds when you have squared both sides.)
5. Prove De Moivre's Theorem.
6. (Challenge) Derive a similar inequality for an upper bound of the sum $\sum_{k=1}^{n} \sin k\theta$.

Conclusion

This actually ends our discussion on the topic. I hope that you enjoyed reading and hopefully took me serious even despite the third-grade-looking depicts on path in paint. We introduced the notation and ideas to address finite sums in general and then we introduced the ideas behind and proofs for the fundamental theorem of finite calculus, which helped us apply the telescoping sum trick to solve many different problems by finding a nice collection of functions that behave well with the discrete derivative and integral.

We introduced the method of induction and the problem of calculating a series of polynomials. We proved that the result of a series formed by adding the values of a polynomial is a polynomial of one degree higher (unless it is the zero polynomial). Once you know that the answer is a polynomial you can either use finite calculus to solve the problem by solving the sums of all smaller degree then putting them together to get the whole sum. At the end of the day, it is not best to do this. It is best to note that we have proved that it is a polynomial, so we just have to find the values of its coefficients. We can do this by writing $\sum_{\ell=1}^{n} p(\ell) = P(n) = \frac{1}{k+1}n^{k+1} + a_k n^k + a_{k-1}n^{k-1} + \cdots + a_1 n + a_0$.

Then by evaluating the sum $P(n)$ on the first (or just some) $k+1$ different values we can $k+1$ equations which we can solve using the method of elimination to solve this system of $k+1$ linear equations in $k+1$ variables that you may have learned about in high school algebra I. Linear Algebra, a great subject btw, has good methods of attacking these sort of problems and they also imply that you can always solve this system of equation when you are looking for the polynomial $P(n)$. (

For more about the theory of discrete calculus, please see [1] and [2] and if you are into the second part of the article please look into mathematical analysis. It is a great subject.

I just wanted to say that it was nice making these notes to accompany the lecture, so I hope that you enjoyed[10] this as much as I did. Unfortunately, I believe that I will not get to some of the later stuff, so this is just for future reference for you, I suppose. :)

God bless and I hope this was helpful/useful.

[10] I guess "will enjoy" is better grammar, but in future-spect this should be seen in retrospect.

References

[1] Gleich, David. *Finite Calculus: A Tutorial for Solving Nasty Sums*. 2005. https://www.cs.purdue.edu/homes/dgleich/publications/Gleich%202005%20-%20finite%20calculus.pdf

[2] Hamrick, Brian. *Discrete Calculus*. http://homepages.math.uic.edu/~kauffman/DCalc.pdf

[3] https://en.wikipedia.org/wiki/Finite_difference

[4] http://www.artofproblemsolving.com/wiki/index.php?title=LaTeX:Layout#Tables for code.

[5] http://tex.stackexchange.com/questions/2832/how-can-i-have-two-tables-side-by-side

[6] http://tex.stackexchange.com/questions/23650/when-should-we-use-begincenter-instead-of-centering

[7] Lehman, Eric et. al. *Mathematics for Computer Science*. http://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-042j-mathematics-for-computer-science-fall-2010/readings/

[8] Conrad, Keith. "Solving linear recursions over all fields". http://www.math.uconn.edu/~kconrad/blurbs/linmultialg/linearrecursion.pdf.

[9] Conway, John. "Functions of One Complex Variable". 1978. Springer Science+Business Media, Inc. 2nd Edition.

[10] Stillwell, "Elements of Number Theory". 2003. Springer Science+Business Media New York.

[11] Tao, Terence. "Terence Tao: Nilsequences and the Primes: UCLA". YouTube.com. January 30, 2009. May 30, 2016. https://www.youtube.com/watch?v=XpocOKj0lxs .

[12] Wikipedia contributors. "Vandermonde matrix." Wikipedia, The Free Encyclopedia. Wikipedia, The Free Encyclopedia, 11 Feb. 2016. Web. 30 Apr. 2016. https://en.wikipedia.org/wiki/Vandermonde_matrix