# CHAPTER 1: FINITE SUMS

MIKE ZABROCKI - 2017 VERSION

## 1. Telescoping sums

Say that you want to prove an identity of the form

$$p(1) + p(2) + p(3) + \cdots + p(n) = q(n)$$

where $p(x)$ and $q(x)$ are expressions and $n$ is a non-negative integer. We must have that $q(0) = 0$ since the left hand side of the equation will be the empty sum when $n = 0$. One way to go about this is to show that

$$q(r) - q(r-1) = p(r)$$

for all $r \geq 0$, and then write down

$$\begin{aligned} q(1) - q(0) &= p(1) \\ q(2) - q(1) &= p(2) \\ q(3) - q(2) &= p(3) \\ &\vdots \\ q(n-1) - q(n-2) &= p(n-1) \\ q(n) - q(n-1) &= p(n) \end{aligned}$$

Now the sum of the expressions on the right hand side of this equation is

$$p(1) + p(2) + p(3) + \cdots + p(n-1) + p(n)$$

and the sum of the expressions on the left hand side of this equation is

$$(q(1)-q(0))+(q(2)-q(1))+(q(3)-q(2))+\cdots+(q(n-1)-q(n-2))+(q(n)-q(n-1)) = q(n)-q(0) = q(n)\ .$$

We conclude therefore that

$$p(1) + p(2) + p(3) + \cdots + p(n-1) + p(n) = q(n)\ .$$

**Example** There are lots of ways of proving the following identity.

$$1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2}\ .$$

Since $\frac{r(r+1)}{2} - \frac{(r-1)r}{2} = r$, we have

$$\frac{1\cdot 2}{2} - \frac{0\cdot 1}{2} = 1$$
$$\frac{2\cdot 3}{2} - \frac{1\cdot 2}{2} = 2$$
$$\frac{3\cdot 4}{2} - \frac{2\cdot 3}{2} = 3$$
$$\vdots$$
$$\frac{(n-1)\cdot n}{2} - \frac{(n-2)\cdot(n-1)}{2} = n-1$$
$$\frac{n\cdot(n+1)}{2} - \frac{(n-1)\cdot n}{2} = n$$

The sum of the terms on the left hand side of these equations is $\frac{n(n+1)}{2}$ and the sum of the terms on the right hand side of these equation is $1+2+3+\cdots+(n-1)+n$, therefore they are equal.

**Example** Define the Fibonacci sequence by $F_0 = 1$, $F_1 = 1$, and for $n \geq 0$, $F_{n+2} = F_{n+1} + F_n$. Say that we want to show that

$$F_0 + F_2 + F_4 + \cdots + F_{2n} = F_{2n+1},$$

or in words "The sum of the first $n$ Fibonacci numbers indexed by even $n$ is the next Fibonacci number indexed by odd $n$." So we know that for $r \geq 1$, $F_{2r+1} - F_{2r-1} = F_{2r} + F_{2r-1} - F_{2r-1} = F_{2r}$. Therefore

$$F_3 - F_1 = F_2$$
$$F_5 - F_3 = F_4$$
$$F_7 - F_5 = F_6$$
$$\vdots$$
$$F_{2n-1} - F_{2n-3} = F_{2n-2}$$
$$F_{2n+1} - F_{2n-1} = F_{2n}$$

Since the sum of the left hand side of these equations is $F_{2n+1} - F_1 = F_{2n+1} - F_0$ and the sum of the right hand side of this equation is $F_2 + F_4 + F_6 + \cdots + F_{2n}$, we conclude that

$$F_0 + F_2 + F_4 + \cdots + F_{2n} = F_{2n+1}\ .$$

## 2. An application of combinatorics to computing the sum $1^r + 2^r + \cdots + n^r$

What we would like to do is start with

$$1 + 2 + 3 + \cdots + n = n(n+1)/2$$

and then to generalize this and get to

$$1^r + 2^r + \cdots + n^r = ??? \tag{1}$$

**Example 1.** *You can easily look online and see what the formula is for the first 10 of these equations (however guessing or deriving the right hand side of these equations is not easy). Here are the first 6:*

$$1 + 2 + 3 + \cdots + n = n(n+1)/2 \tag{2}$$

$$1^2 + 2^2 + 3^2 + \cdots + n^2 = n(n+1)(2n+1)/6 \tag{3}$$

$$1^3 + 2^3 + \cdots + n^3 = n^2(n+1)^2/4 \tag{4}$$

$$1^4 + 2^4 + 3^4 + \cdots + n^4 = \frac{n(n+1)(2n+1)(3n^2+3n-1)}{30} \tag{5}$$

$$1^5 + 2^5 + 3^5 + \cdots + n^5 = \frac{[n(n+1)]^2(2n^2+2n-1)}{12} \tag{6}$$

$$1^6 + 2^6 + 3^6 + \cdots + n^6 = \frac{n(n+1)(2n+1)(3n^4+6n^3-3n+1)}{42} \tag{7}$$

As you can see it is difficult to perceive if there is a pattern between the formulae on the right hand side of equations (2)–(7).

But there is a sequence of equations that continues in a way that it is easy to conjecture a formula:

$$1 + 2 + 3 + \cdots + n = \frac{(n+1)n}{2} \tag{8}$$

$$1 \cdot 2 + 2 \cdot 3 + \cdots + n(n-1) = \frac{(n+1)n(n-1)}{3} \tag{9}$$

$$1 \cdot 2 \cdot 3 + 2 \cdot 3 \cdot 4 + \cdots + n(n-1)(n-2) = \frac{(n+1)n(n-1)(n-2)}{4} \tag{10}$$

$$\vdots$$

$$1 \cdot 2 \cdots k + 2 \cdot 3 \cdots (k+1) + \cdots + n \cdot (n-1) \cdots (n-k+1) = \frac{(n+1)n(n-1)\cdots(n-k+1)}{k+1} \tag{11}$$

You should be able to prove this entire sequence of equations either by (a) induction (on $n$) or (b) telescoping sums.

By telescoping sums, you need only do the computation,

$$\frac{(n+1)n(n-1)\cdots(n-k+1)}{k+1} - \frac{n(n-1)\cdots(n-k)}{k+1} =$$
$$\frac{n(n-1)\cdots(n-k+1)}{k+1}((n+1)-(n-k)) = n(n-1)\cdots(n-k+1)$$

Therefore, by the method of telescoping sums, (11) follows and all the equations (8)–(10) are special cases of this one.

Define for $k$ and integer with $k > 0$, set:

$$(x)_k = x(x-1)(x-2)\cdots(x-k+1)$$

such that there are $k$ terms in the product.

**Example 2.** $(x)_1 = x$, $(x)_2 = x(x-1) = x^2 - x$, $(x)_3 = x(x-1)(x-2)x^3 - 3x^2 + 2x$, $(x)_4 = x(x-1)(x-2)(x-3) = x^4 - 6x^3 + 11x^2 - 6x$, ...

This is new notation that makes some of our formulas simpler. Equations (8)–(11) are now

$$(1)_1 + (2)_1 + \cdots + (n)_1 = \frac{(n+1)_2}{2} \tag{12}$$

$$(2)_2 + (3)_2 + \cdots + (n)_2 = \frac{(n+1)_3}{3} \tag{13}$$

$$(3)_3 + (4)_3 + \cdots + (n)_3 = \frac{(n+1)_4}{4} \tag{14}$$

$$\vdots$$

$$(k)_k + (k+1)_k + \cdots + (n)_k = \frac{(n+1)_{k+1}}{k+1} \tag{15}$$

But then notice that $(k-1)_k = (k-2)_k = \cdots = (1)_k = 0$ because they all have a factor of 0 in their product. Therefore we have the simple equation that

$$(1)_k + (2)_k + \cdots + (n)_k = \sum_{i=1}^{n} (i)_k = \frac{(n+1)_{k+1}}{k+1} \tag{16}$$

I norder to do find a formula for (1) we will look the change of basis between $x^r$ and $(x)_r$. We already did this in some of the examples and we see that we have an infinite system of linear equations for which we have computed the first 4 rows in Example 2:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \\ -1 & 1 & 0 & 0 & 0 & \cdots \\ 2 & -3 & 1 & 0 & 0 & \\ -6 & 11 & -6 & 1 & 0 & \\ & \vdots & & & & \ddots \end{bmatrix} \begin{bmatrix} x \\ x^2 \\ x^3 \\ x^4 \\ \vdots \end{bmatrix} = \begin{bmatrix} (x)_1 \\ (x)_2 \\ (x)_3 \\ (x)_4 \\ \vdots \end{bmatrix} \tag{17}$$

A combintorialist looks at this matrix and sees:

(1) the entries of the matrix are integers and alternating
(2) the entry in the first column of the $n^{th}$ row is $(-1)^{n-1}(n-1)!$
(3) the diagonal entries are all 1
(4) if we ignore the signs then the sums of the entries in the $n^{th}$ row is equal to $n!$
(5) if you enter the numbers into the online integer sequence database you can look them up and they are called the (signed) Stirling numbers of the first kind.

We will separate the sign from the number and call the integer in the $n^{th}$ row and $k^{th}$ column of the matrix equal to $(-1)^{n+k}s'(n,k)$ and we will refer to the $s'(n,k)$ as the 'unsigned Stirling number of the first kind.' There doesn't seem to be a complete history of these numbers readily available, but you can imagine that Stirling (who was a mathematician in the $18^{th}$ century) saw them, calculated this matrix (at least) a bit further and made the following observations:

(1) $s'(n,n) = 1$
(2) $s'(n,1) = (n-1)!$
(3) $s'(n,k) = (n-1)s'(n-1,k) + s'(n-1,k-1)$ if $2 \le k \le n-1$.

We will take this as the (recursive) definition of the integers (along with the condition that $s'(0,0) = 1$ and $s'(n,k) = 0$ if $k > n$ or if $k < 1$). That is we can create a table of rows with 1's on the diagonal, $(n-1)!$ in the first column like so:

| | $k=1$ | $k=2$ | $k=3$ | $k=4$ | $k=5$ | $k=6$ |
|---|---|---|---|---|---|---|
| $n=1$ | 1 | | | | | |
| $n=2$ | 1 | 1 | | | | |
| $n=3$ | 2 | | 1 | | | |
| $n=4$ | 6 | | | 1 | | |
| $n=5$ | 24 | | | | 1 | |
| $n=6$ | 120 | | | | | 1 |

Then using the recursion $s'(n,k) = (n-1)s'(n-1,k) + s'(n-1,k-1)$ we fill in

| | $k=1$ | $k=2$ | $k=3$ | $k=4$ | $k=5$ | $k=6$ |
|---|---|---|---|---|---|---|
| $n=1$ | 1 | | | | | |
| $n=2$ | 1 | 1 | | | | |
| $n=3$ | 2 | 3 | 1 | | | |
| $n=4$ | 6 | 11 | 6 | 1 | | |
| $n=5$ | 24 | 50 | 35 | 10 | 1 | |
| $n=6$ | 120 | 274 | 225 | 85 | 15 | 1 |

I did this on the computer and wrote precisely the conditions that we used to define them in Sage.

```
sage: def stirling_first(n,k):
....:     if n>=1 and k==1:
....:         return factorial(n-1)
....:     elif n>=1 and n==k:
....:         return 1
....:     elif n==0 and k==0:
....:         return 1
....:     elif k<=0 or n<k:
....:         return 0
....:     else:
....:         return (n-1)*stirling_first(n-1,k) + stirling_first(n-1,k-1)
sage: matrix([[stirling_first(n,k) for k in range(10)] for n in range(10)])
[     1      0      0      0      0      0      0      0      0      0]
[     0      1      0      0      0      0      0      0      0      0]
[     0      1      1      0      0      0      0      0      0      0]
[     0      2      3      1      0      0      0      0      0      0]
[     0      6     11      6      1      0      0      0      0      0]
[     0     24     50     35     10      1      0      0      0      0]
[     0    120    274    225     85     15      1      0      0      0]
[     0    720   1764   1624    735    175     21      1      0      0]
[     0   5040  13068  13132   6769   1960    322     28      1      0]
[     0  40320 109584 118124  67284  22449   4536    546     36      1]
```

Then we will prove the following:

**Proposition 3.** *For $n \geq 1$,*

$$(x)_n = \sum_{k=1}^{n} (-1)^{n+k} s'(n,k) x^k \ . \tag{18}$$

You should check that in the special cases we have calculated thus far (for $n = 1, 2, 3, 4$) that this equation describes what we have observed. However to prove it in general by induction.

*Proof.* For $n = 1$, $(x)_1 = (-1)^{1+1} s'(1,1)x = x$.

Now assume that we know $(x)_{n-1} = \sum_{k=1}^{n-1}(-1)^{n+k-1}s'(n-1,k)x^k$. Then

$$\begin{aligned}
(x)_n = (x-n+1)(x)_{n-1} &= \sum_{k=1}^{n-1}(-1)^{n+k-1}s'(n-1,k)(x-(n-1))x^k \\
&= \sum_{k=1}^{n-1}(-1)^{n+k-1}s'(n-1,k)x^{k+1} + \sum_{k=1}^{n-1}(-1)^{n+k}(n-1)s'(n-1,k)x^k \\
&= \sum_{k=2}^{n}(-1)^{n+k}s'(n-1,k-1)x^k + \sum_{k=1}^{n-1}(-1)^{n+k}(n-1)s'(n-1,k)x^k \\
&= (-1)^{n+n}s'(n-1,n-1)x^n + \sum_{k=2}^{n-1}(-1)^{n+k}(s'(n-1,k-1)+(n-1)s'(n-1,k))x^k \\
&\quad + (-1)^{n+1}(n-1)s'(n-1,1)x^1 \\
&= s'(n,n)x^n + \sum_{k=2}^{n-1}(-1)^{n+k}s'(n,k)x^k + (-1)^{n+1}s'(n,1)x^1 = \sum_{k=1}^{n}(-1)^{n+k}s'(n,k)x^k \ .
\end{aligned}$$

Hence we have by the principle of mathematical induction that equation (18) holds for all $n \geq 1$. □

We can consider only the first $n$ rows of this matrix and it is a finite calculation and it is lower triangular. There is therefore no problem in inverting it and solving for the vector $\begin{bmatrix} x \\ x^2 \\ x^3 \\ x^4 \\ \vdots \end{bmatrix}$ in terms of the vector $\begin{bmatrix} (x)_1 \\ (x)_2 \\ (x)_3 \\ (x)_4 \\ \vdots \end{bmatrix}$. In fact, I did this for the first 4 rows and we saw the following infinite matrix

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \\ 1 & 1 & 0 & 0 & 0 & \cdots \\ 1 & 3 & 1 & 0 & 0 & \\ 1 & 7 & 6 & 1 & 0 & \\ & \vdots & & & & \ddots \end{bmatrix} \begin{bmatrix} (x)_1 \\ (x)_2 \\ (x)_3 \\ (x)_4 \\ \vdots \end{bmatrix} = \begin{bmatrix} x \\ x^2 \\ x^3 \\ x^4 \\ \vdots \end{bmatrix} \tag{19}$$

where now we observe (at least to the $4^{th}$ row) all the entries are positive and the first column and the main diagonal are 1. The sums of the rows are called the Bell numbers and the entries in the matrix are called the Stirling numbers of the second kind. Some references use the notation $S(n,k)$ for the entry in the $n^{th}$ row and the $k^{th}$ column. I will use $\left\{ n \atop k \right\}$, a notation that is (at least) equally common (partly because the letter $S$ is overused in combinatorics to represent quantities and partly because the notation indicates that these coefficients are similar to binomial coefficients).

We observe/define here that $\left\{ {n \atop 1} \right\} = \left\{ {n \atop n} \right\} = 1$, $\left\{ {n \atop k} \right\} = 0$ if $k > n$ or $k < 1$ and if $2 \leq k \leq n-1$,

$$\left\{ {n \atop k} \right\} = k \left\{ {n-1 \atop k} \right\} + \left\{ {n-1 \atop k-1} \right\} . \tag{20}$$

This gives us a recursive definition of these coefficients and just as we did for the Stirling numbers of the first kind, we can compute a table of the Stirling numbers of the second kind by creating a table where the first column is 1 and the diagonal is also 1 and for the rows with $n \geq 3$ the table satisfies equation (20).

| | $k=1$ | $k=2$ | $k=3$ | $k=4$ | $k=5$ | $k=6$ |
|---|---|---|---|---|---|---|
| $n=1$ | 1 | | | | | |
| $n=2$ | 1 | 1 | | | | |
| $n=3$ | 1 | 3 | 1 | | | |
| $n=4$ | 1 | 7 | 6 | 1 | | |
| $n=5$ | 1 | 15 | 25 | 10 | 1 | |
| $n=6$ | 1 | 31 | 90 | 65 | 15 | 1 |

Similar to as we did for the Stirling numbers of the first kind, we can write a program to compute the Stirling numbers of the second kind as well by just changing a few words in our program for `stirling_first`.

```
sage: def stirling_second(n,k):
....:     if n>=1 and k==1:
....:         return 1
....:     elif n>=1 and n==k:
....:         return 1
....:     elif n==0 and k==0:
....:         return 1
....:     elif k<=0 or n<k:
....:         return 0
....:     else:
....:         return k*stirling_second(n-1,k) + stirling_second(n-1,k-1)
sage: matrix([[stirling_second(n,k) for k in range(10)] for n in range(10)])
[   1    0    0    0    0    0    0    0    0    0]
[   0    1    0    0    0    0    0    0    0    0]
[   0    1    1    0    0    0    0    0    0    0]
[   0    1    3    1    0    0    0    0    0    0]
[   0    1    7    6    1    0    0    0    0    0]
[   0    1   15   25   10    1    0    0    0    0]
[   0    1   31   90   65   15    1    0    0    0]
[   0    1   63  301  350  140   21    1    0    0]
[   0    1  127  966 1701 1050  266   28    1    0]
[   0    1  255 3025 7770 6951 2646  462   36    1]
```

We leave as an exercise the following result:

**Proposition 4.** *For $n \geq 1$,*

$$x^n = \sum_{k=1}^{n} \left\{ {n \atop k} \right\} (x)_k \ . \tag{21}$$

Now that we have the coefficients which change between $\{(x)_n\}_{n\geq 1}$ and $\{x^n\}_{n\geq 1}$, we have a way of answering the question we posed at the beginning of this section. Because for a fixed $r \geq 1$, by equation (21) and then by equation (16),

$$1^r + 2^r + \cdots + n^r = \sum_{i=1}^{n} i^r = \sum_{i=1}^{n}\sum_{k=1}^{r} \left\{ {r \atop k} \right\} (i)_k = \sum_{k=1}^{r} \left\{ {r \atop k} \right\} \sum_{i=1}^{n} (i)_k = \sum_{k=1}^{r} \left\{ {r \atop k} \right\} \frac{(n+1)_{k+1}}{k+1} \ .$$

What we've done is changed a sum over $n$ terms on the left (involving a power of $r$) into a sum over $r$ terms on the right. So we see for example if $r = 1$, then

$$\begin{aligned}
1 + 2 + \cdots + n &= \frac{(n+1)n}{2} \\
1^2 + 2^2 + \cdots + n^2 &= \left\{ {2 \atop 1} \right\} \frac{(n+1)_2}{2} + \left\{ {2 \atop 2} \right\} \frac{(n+1)_3}{3} \\
&= (n+1)n \left( \frac{1}{2} + \frac{(n-1)}{3} \right) \\
1^3 + 2^3 + \cdots + n^3 &= \left\{ {3 \atop 1} \right\} \frac{(n+1)_2}{2} + \left\{ {3 \atop 2} \right\} \frac{(n+1)_3}{3} + \left\{ {3 \atop 3} \right\} \frac{(n+1)_4}{4} \\
&= (n+1)n \left( \frac{1}{2} + (n-1) + \frac{(n-1)(n-2)}{4} \right)
\end{aligned}$$

## 3. A step towards advanced proofs: $A = B$

I am going to give you an advanced proof technique for proving certain types of formulas. Just an FYI, this is one technique that I don't ever use with my first year students because it goes against almost everything I want them to first learn about proofs. It uses the following lemma that you might run across in an algebra class that introduces polynomials:

**Lemma 5.** *Let $p(x) = a_r x^r + a_{r-1} x^{r-1} + \cdots + a_1 x + a_0$ with $a_i$ are complex numbers, then there are at most $r$ values $x_1, x_2, \ldots, x_r$ (we call roots of the polynomial) such that $p(x_i) = 0$.*

*Proof.* Since we are not going into details about polynomials, I will only sketch the proof here (you can easily look it up in any number of references). The proof goes something like if $p(x_i) = 0$, then $(x - x_i)$ divides $p(x)$ so $p(x) = (x - x_i)q(x)$ for some polynomial $q(x)$ which is of degree at most $r - 1$ and by induction $q(x)$ has at most $r - 1$ roots and hence $p(x)$ has at most $r$ roots. □

As a consequence if $p(x)$ is a polynomial of degree at most $r$ and $p(x_i) = 0$ for $r+1$ (or more) values $x_i$, then $p(x)$ must be the 0 polynomial.

Lets look at this a different way. If $p(x)$ and $q(x)$ are polynomials of degree at most $r$ then their difference $p(x) - q(x)$ is a polynomial of degree at most $r$. Therefore if $p(x_i) = q(x_i)$ for $r+1$ values, then $p(x) - q(x)$ is the zero polynomial and it implies that $p(x) = q(x)$.

The reason this is interesting is that sometimes you have an identity and you know that it will be of the form

$$\text{some expression involving a parameter } n = p(n)$$

where $p(n)$ is some polynomial. Lets say that this polynomial is of of degree at most $r$ (for instance we were able to calculate enough to know that $1^{r-1} + 2^{r-1} + \cdots + n^{r-1}$ will be a polynomial of degree at most $r$). If this is the case, then *just by checking that the identity holds on $r$ different values* we can conclude that the identity must be true. The reason is that there is only one polynomial of degree $r$ that will vanish on the $r$ different points.

ToDo `explain this better, give an example`