Math Club Notes: 2005 October 27

1 Sum of every third binomial coefficient

Today we took care of one of our outstanding problems: expressing

$$\sum_k \binom{n}{3k}$$

in closed form as a function of n.

This problem is, incidentally, one of a family of problems. The first of this family is to compute the sum of every binomial coefficient in a row; we know that this is

$$\sum_k \binom{n}{k} = 2^n \,.$$

The second is to compute the sum of every other binomial coefficient in a row; it's not hard to spot that

$$\begin{aligned}
\sum_k \binom{n}{2k} &= \sum_k \left(\binom{n-1}{2k} + \binom{n-1}{2k-1} \right) \\
&= \sum_k \binom{n-1}{2k} + \sum_k \binom{n-1}{2k-1} \\
&= \sum_{k \text{ even}} \binom{n-1}{k} + \sum_{k \text{ odd}} \binom{n-1}{k} \\
&= \sum_k \binom{n-1}{k} \\
&= 2^{n-1}
\end{aligned}$$

(Well, this derivation doesn't work for $n = 0$, and indeed, in that row the sum isn't $2^{0-1} = \frac{1}{2}$; it's 1.)

1.1 Formulating the problem well

As we have observed before, there is no blindingly obvious pattern in the first few values of this sum:

n	0	1	2	3	4	5	6	7
$\sum_k \binom{n}{3k}$	1	1	1	2	5	11	22	43

But maybe six months from now you will have a dream in which a cactus tells you a closed form for this sequence. When you wake up, how will you prove the cactus's formula correct? Perhaps by induction on n.

So imagine you are trying to prove

$$\sum_k \binom{n}{3k} = f(n)$$

(where f is the function the cactus gave you) by induction on n. In the inductive step, you've supposed this true for n and wish to show it holds for $n+1$. This calls for some way to relate binomial coefficients from the $(n+1)$th row of Pascal's triangle to the ones from the nth row; we normally use the identity

$$\binom{n+1}{k} = \binom{n}{k} + \binom{n}{k-1}$$

for this purpose. So let's see what we could do:

$$\begin{aligned}\sum_k \binom{n+1}{3k} &= \sum_k \left(\binom{n}{3k} + \binom{n}{3k-1}\right) \\ &= \sum_k \binom{n}{3k} + \sum_k \binom{n}{3k-1} \\ &= f(n) + \sum_k \binom{n}{3k-1} \qquad \text{(inductive hypothesis)}\end{aligned}$$

A problem. Our inductive hypothesis tells us something about $\sum_k \binom{n}{3k}$, but we need to know something about $\sum_k \binom{n}{3k-1}$ as well.

Well, let's suppose the mystical dream cactus gave us a closed form for that sum too, and we are proving that closed form by induction at the same time. Then, again, we will have to prove that that closed form also holds at $n+1$; and repeating the above will bring in $\sum_k \binom{n}{3k-2}$ as well. Repeating it again introduces $\sum_k \binom{n}{3k-3}$; but since we are summing over all integers k, we can replace k with $k+1$, and that turns this sum into $\sum_k \binom{n}{3k}$, which we already had.

So it starts to look as if we want to find closed forms for three sums:

$$\sum_k \binom{n}{3k} \qquad \sum_k \binom{n}{3k-1} \qquad \sum_k \binom{n}{3k-2}$$

Negatives are a little annoying, so let's replace k with $k+1$ in the second and third sums (so that $3k-1$ becomes $3k+2$, and $3k-2$ becomes $3k+1$) and talk instead about

$$\sum_k \binom{n}{3k} \qquad \sum_k \binom{n}{3k+1} \qquad \sum_k \binom{n}{3k+2}$$

I'm getting pretty tired of writing these expressions over and over again, so let's also adopt a terser notation for them:

$$S_0(n) = \sum_k \binom{n}{3k} \qquad S_1(n) = \sum_k \binom{n}{3k+1} \qquad S_2(n) = \sum_k \binom{n}{3k+2}$$

We have seen that the value of one of these sums at $n+1$ is related to the value of two of these sums at n; specifically,

$$\begin{aligned} S_0(n+1) &= S_0(n) + S_2(n) \\ S_1(n+1) &= S_0(n) + S_1(n) \\ S_2(n+1) &= S_1(n) + S_2(n) \end{aligned}$$

With these three equations, we can easily calculate the sums for the first few values of n:

n	0	1	2	3	4	5	6	7
$S_0(n)$	1	1	1	2	5	11	22	43
$S_1(n)$	0	1	2	3	5	10	21	43
$S_2(n)$	0	0	1	3	6	11	21	42

(This is just Pascal's triangle, but wrapped around a cylinder and with overlapping entries summed.)

This new data suggests some new conjectures; for example, it appears that the three sums are always about the same. Recalling that $S_0(n) + S_1(n) + S_2(n) = 2^n$ (since this is the sum of an entire row of Pascal's triangle), it then seems that $S_0(n) \approx 2^n/3$. We will see later that this is correct.

This table also suggests thinking of the problem as not being about the sequence of numbers

$$1, 1, 1, 2, 5, 11, 22, 43, \ldots$$

but as about the sequence of vectors

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \\ 3 \end{bmatrix}, \begin{bmatrix} 5 \\ 5 \\ 6 \end{bmatrix}, \begin{bmatrix} 11 \\ 10 \\ 11 \end{bmatrix}, \begin{bmatrix} 22 \\ 21 \\ 21 \end{bmatrix}, \begin{bmatrix} 43 \\ 43 \\ 42 \end{bmatrix}, \ldots .$$

Now that we're thinking in terms of vectors, we notice we already know that

$$\begin{bmatrix} S_0(n+1) \\ S_1(n+1) \\ S_2(n+1) \end{bmatrix} = \begin{bmatrix} S_0(n) + S_2(n) \\ S_0(n) + S_1(n) \\ S_1(n) + S_2(n) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} S_0(n) \\ S_1(n) \\ S_2(n) \end{bmatrix} .$$

This is a recurrence for these vectors, and it is trivial to solve: by applying the recurrence n times (or, more formally, by a trivial induction) we obtain

$$\begin{bmatrix} S_0(n) \\ S_1(n) \\ S_2(n) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}^n \begin{bmatrix} S_0(0) \\ S_1(0) \\ S_2(0) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}^n \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} .$$

We could, I suppose, just stop here, and give

$$\begin{bmatrix} S_0(n) \\ S_1(n) \\ S_2(n) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}^n \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

as our final result. Does it count as being in "closed form"? Depends on context, I suppose, and on taste.

In the case at hand, we can do a lot better — I prefer to think of the above equation not as a closed-form answer, but instead as merely pointing in the direction of a good formulation of the problem. We can ask it in this way: what is the effect of iterating the linear transformation associated with this 3×3 matrix?

1.2 Iterating the linear transformation

When dealing with powers of a matrix (as learned in Math 225), we often hope that the matrix is diagonalizable. For if so, we can build a basis for (in this case) $\mathbb{R}^3$ out of its eigenvectors; in the coordinate system formed by that basis, the matrix simply scales each coordinate axis by a factor of the associated eigenvalue λ. In such a case, the effect of applying the matrix n times is easy to compute: it scales each coordinate axis by λ^n (for each λ, as appropriate to the axis in question).

Even if the matrix is not diagonalizable (so its eigenvectors do not span the whole space), we might hope that at least the vector we are interested in lies in the span of the eigenvectors; for in this case we can at least compute the effect of applying the matrix n times to that specific vector.

It turns out that the matrix we are dealing with — which we'll call M:

$$M = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

— is not diagonalizable, and the vectors we are interested in — which we'll call $\vec{v}_n$:

$$\vec{v}_n = M^n \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

— do not lie in the span of its eigenvectors. We will nevertheless come to understand M's effect pretty well.

(For an example of the nicer situation where the matrix is diagonalizable, see the derivation of Binet's formula in our notes for May 22.)

We find that the characteristic polynomial of M is

$$(2-\lambda)(1-\lambda+\lambda^2) .$$

The eigenspace associated with the eigenvalue 2 is the line

$$t \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} ,$$

that is, the line $x = y = z$. The other factor in the characteristic polynomial, $1 - \lambda + \lambda^2$, has, alas, no real roots. (We could consider its complex roots; the interested reader might wish to explore this possibility.)

Still, now that we know the effect of M on a particular line, it is reasonably natural to decompose our starting vector into one portion along that line and one portion in some other direction. What other direction? Let's say, a perpendicular one.

We compute that our starting vector has the orthogonal decomposition

$$\vec{v}_0 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \frac{1}{3} \begin{bmatrix} 2 \\ -1 \\ -1 \end{bmatrix} .$$

We know the effect of M on the first vector here: we have

$$M^n \vec{v}_0 = \frac{2^n}{3} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \frac{1}{3} M^n \begin{bmatrix} 2 \\ -1 \\ -1 \end{bmatrix} . \tag{1}$$

What does M do to the second vector?

By brute force, we compute:

$$M \begin{bmatrix} 2 \\ -1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} \qquad M^2 \begin{bmatrix} 2 \\ -1 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \\ -1 \end{bmatrix}$$

$$M^3 \begin{bmatrix} 2 \\ -1 \\ -1 \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} \qquad M^4 \begin{bmatrix} 2 \\ -1 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}$$

$$M^5 \begin{bmatrix} 2 \\ -1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \qquad M^6 \begin{bmatrix} 2 \\ -1 \\ -1 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ -1 \end{bmatrix}$$

Hey — we're back where we started. So this sequence of six vectors will repeat forever; the effect of M^n on this vector depends only on $n \bmod 6$.

What we are seeing here? M scales the line $x = y = z$ by a factor of 2, and rotates the plane perpendicular to that line by 60°. (This might seem less mysterious if you look into the aforementioned complex eigenvalues.) Our starting vector $\vec{v}_0$, being slightly off the line and so having components both in the eigenline and in the perp, stretches in the direction of the eigenline and pivots around it, always remaining the same distance from the line. The sequence of vectors forms a kind of helix that keeps getting more and more stretched out.

Extracting the first component of the vectors in (1), we obtain

$$\sum_k \binom{n}{3k} = \tfrac{1}{3}(2^n + b_n)\text{ , where } b = [\overline{2, 1, -1, -2, -1, 1}]. \tag{2}$$

This is a perfectly good closed form.

(The notation $[\overline{2, 1, -1, -2, -1, 1, 2}]$ denotes the repeating sequence

$$2, 1, -1, -2, -1, 1, 2, 1, -1, -2, -1, 1, 2, 1, -1, -2, -1, 1, \ldots .$$

These are the first components of the repeating sequence of six vectors obtained above. The notation

$$b_n = \begin{cases} 2 & \text{if } n \equiv 0 \pmod 6 \\ 1 & \text{if } n \equiv 1 \pmod 6 \\ -1 & \text{if } n \equiv 2 \pmod 6 \\ -2 & \text{if } n \equiv 3 \pmod 6 \\ -1 & \text{if } n \equiv 4 \pmod 6 \\ 1 & \text{if } n \equiv 5 \pmod 6 \end{cases}$$

is more traditional, but also more cumbersome.)

Earlier we conjectured that this sum was approximately $2^n/3$; now we see that this is correct; indeed, since $|b_n| < 3$, we can see from (2) that

$$\sum_k \binom{n}{3k} = \begin{cases} \lceil 2^n/3 \rceil & \text{if } n \equiv 0, 1, 5 \pmod 6 \\ \lfloor 2^n/3 \rfloor & \text{if } n \equiv 2, 3, 4 \pmod 6 \end{cases}$$

(Note that $2^n/3$ is never an integer, but our sum is.)

Incidentally, this result suggests a revision of our solution for the sum of every other binomial coefficient; earlier we said it was 2^{n-1}, except when $n = 0$, when the sum is not $\frac{1}{2}$, but 1. We might well write that $\sum_k \binom{n}{2k} = \lceil 2^n/2 \rceil$. And, of course, the sum of every binomial coefficient is $\sum_k \binom{n}{k} = \lceil 2^n/1 \rceil$.

1.3 A nifty variation

Back when we first looked at this problem, we noticed that each value of the sum is approximately double the one before:

$$\begin{aligned} &1 \\ 1 &= 2 \cdot 1 - 1 \\ 1 &= 2 \cdot 1 - 1 \\ 2 &= 2 \cdot 1 \\ 5 &= 2 \cdot 2 + 1 \\ 11 &= 2 \cdot 5 + 1 \\ 22 &= 2 \cdot 11 \\ 43 &= 2 \cdot 22 - 1 \end{aligned}$$

This makes sense given our formula $\frac{1}{3}(2^n + b_n)$; the sum is approximately $2^n/3$, which does indeed double when n increases by 1. And we now have plenty of reason to think that the small corrections $-1, -1, 0, +1, +1, 0, -1, \dots$ form a repeating sequence of period 6.

Let's have a look at those corrections; call them c_n.

$$\begin{aligned} c_n &= S_0(n+1) - 2S_0(n) \\ &= \tfrac{1}{3}(2^{n+1} + b_{n+1}) - \tfrac{2}{3}(2^n + b_n) \\ &= \tfrac{1}{3}(b_{n+1} - 2b_n) \end{aligned}$$

By direct calculation from the sequence b, we find that, as expected,

$$c = [\overline{-1, -1, 0, 1, 1, 0}] \ .$$

So we now have

$$S_0(n+1) = 2S_0(n) + c_n \ ,$$

which is a recurrence for $S_0(n)$. There is a common trick for changing a recurrence of this type into a sum: divide by 2^{n+1} to obtain

$$\frac{S_0(n+1)}{2^{n+1}} = \frac{S_0(n)}{2^n} + \frac{c_n}{2^{n+1}} \ .$$

Now let $T_n = S_0(n)/2^n$, so this equation is

$$T_{n+1} = T_n + \frac{c_n}{2^{n+1}} \ ,$$

a recurrence for T_n which is easily transformed into a sum:

$$T_n = T_0 + \sum_{k=0}^{n-1} \frac{c_k}{2^{k+1}} \ .$$

Replacing T_n with $S_0(n)/2^n$ and rearranging a bit, we obtain

$$S_0(n) = 2^n + 2^n \sum_{k=0}^{n-1} \frac{c_k}{2^{k+1}} \ .$$

This new sum looks a little like the binary representation of some number; it's

$$\frac{c_0}{2} + \frac{c_1}{4} + \frac{c_2}{8} + \cdots + \frac{c_{n-1}}{2^n} \ .$$

Of course, it's not really a binary representation, since c_k takes the values 0, 1, and -1, instead of just the values 0 and 1, as in a true binary representation.

So let's separate out the positive and the negative parts of c; let

$$\begin{aligned} d &= [\overline{0, 0, 0, 1, 1, 0}] \\ \text{and } e &= [\overline{1, 1, 0, 0, 0, 0}] \ , \end{aligned}$$

so that $c_n = d_n - e_n$. Then we have

$$S_0(n) = 2^n + 2^n \sum_{k=0}^{n-1} \frac{d_k}{2^{k+1}} - 2^n \sum_{k=0}^{n-1} \frac{e_k}{2^{k+1}} \; .$$

These new sums really are binary representations.

Let's look at the first one, the sum involving d_k:

$$2^n \sum_{k=0}^{n-1} \frac{d_k}{2^{k+1}}$$

What computation does this expression represent? In binary, the sum would simply look like

$$(0.d_0 d_1 d_2 \ldots d_{n-1})_2 \; .$$

(In this notation, the d_k represent the bits of the number.) Then we multiply this value by 2^n; this moves the binary point n places to the right, so we obtain the integer

$$(d_0 d_1 d_2 \ldots d_{n-1})_2 \; .$$

Here's another way to get the same effect: take *all* the bits from the sequence d, put them after the binary point, shift the binary point n places to the right, and then throw away the bits after the binary point.

How do we throw away bits after the binary point? Or, indeed, after the decimal point, etc.? That operation is just rounding down to the nearest integer — which is also known as taking the floor.

What this (admittedly somewhat informal) reasoning establishes is that

$$2^n \sum_{k=0}^{n-1} \frac{d_k}{2^{k+1}} = \left\lfloor 2^n \sum_{k=0}^{\infty} \frac{d_k}{2^{k+1}} \right\rfloor \; .$$

The left-hand side is the expression we started with, where we just take the first n bits of d; the right-hand side expresses the new computation, where we take all the bits of d and throw the excess away at the end. (You might want to take a stab at proving this with less handwaving.)

Making a similar manipulation of the sum with e_k as well, we obtain

$$S_0(n) = 2^n + \left\lfloor 2^n \sum_{k=0}^{\infty} \frac{d_k}{2^{k+1}} \right\rfloor - \left\lfloor 2^n \sum_{k=0}^{\infty} \frac{e_k}{2^{k+1}} \right\rfloor \; .$$

The new sums can be easily evaluated; they are infinite geometric sums. Let

$$x = \sum_{k=0}^{\infty} \frac{d_k}{2^{k+1}} = (0.d_0 d_1 d_2 \ldots)_2 \; .$$

Then

$$\begin{aligned}
2^6 x &= (d_0 d_1 d_2 d_3 d_4 d_5 . d_6 d_7 d_8 \dots)_2 \\
&= (d_0 d_1 d_2 d_3 d_4 d_5)_2 + (0.d_6 d_7 d_8 \dots)_2 \\
&= (000110)_2 + (0.d_6 d_7 d_8 \dots)_2 \\
&= 6 + (0.d_6 d_7 d_8 \dots)_2 \\
&= 6 + x
\end{aligned}$$

since d is periodic with period 6. Solving for x, we obtain

$$\sum_{k=0}^{\infty} \frac{d_k}{2^{k+1}} = \frac{2}{21} .$$

Similarly,

$$\sum_{k=0}^{\infty} \frac{e_k}{2^{k+1}} = \frac{16}{21} .$$

Thus

$$S_0(n) = 2^n + \left\lfloor 2^n \cdot \frac{2}{21} \right\rfloor - \left\lfloor 2^n \cdot \frac{16}{21} \right\rfloor .$$

Simplifying a bit (and returning to the original notation),

$$\sum_k \binom{n}{3k} = \left\lfloor 2^n \cdot \frac{23}{21} \right\rfloor - \left\lfloor 2^n \cdot \frac{16}{21} \right\rfloor ,$$

which is pretty nifty, if you ask me.

(If the cactus had given you this closed form, would you have believed it?)

2 PostScript

See our notes for 2006 January 12 for an alternative method for this problem.