Symbolic Summation

Victor Adamchik
Carnegie Mellon University

Wilf-Zeilberger's algorithm

The art of doing mathematics consists in finding that special case which contains all the germs of generality -

David Hilbert

Forethoughts

$F(n, k)$ is Gosper-summable if there is a rational function $G$ such that

$$F(n, k) = G(n, k + 1) - G(n, k)$$

Moreover, $G$ is a rational multiple of $F$:

$$G(n, k) = R(n, k) F(n, k)$$

Indefinite summation:

$$\sum_{k} F(n, k) = \sum_{k} [G(n, k + 1) - G(n, k)] = G(n, k)$$

Definite summation:

$$\sum_{k=1}^{n} F(n, k) = \sum_{k=1}^{n} [G(n, k + 1) - G(n, k)] = G(n, n + 1) - G(n, 1)$$

What do we do if the summand $F(n, k)$ is not Gosper-summable? Doron Zeilberger observed that Gosper's algorithm of indefinite summation could be used in a non-obvious and nontrivial way, namely for PROVING combinatorial identities.
**Definition.** $F(n, k)$ has a finite support if $F(n, k) \neq 0$ only for finitely many $k \in \mathbb{Z}$ and fixed $n \in \mathbb{N}_0$.

In other words, all such series

$$\sum_{k=-\infty}^{\infty} F(n, k)$$

are actually finite sums. For example,

$$\sum_{k=-\infty}^{\infty} \binom{n}{k} = \sum_{k=0}^{\infty} \binom{n}{k}$$

Now let look at definite series (1) where $F(n, k)$ is Gosper-summable. Since it's telescoping, we have

$$\sum_{k=-\infty}^{\infty} F(n, k) = \sum_{k=-\infty}^{\infty} \left[ G(n, k + 1) - G(n, k) \right] = 0$$

assuming that $G(n, k)$ has no singularities. Thus, we deduced that

$$F(n, k) \text{ is Gosper-summable} \implies \sum_{k=-\infty}^{\infty} F(n, k) = 0$$

Conversely, if $F(n, k)$ has a finite support and it is a hypergeometric term, then if

$$\sum_{k=-\infty}^{\infty} F(n, k) \neq 0 \implies F(n, k) \text{ is not Gosper-summable}$$

**Wilf-Zeilberger's algorithm**

Wilf-Zeilberger (or in short WZ) method is an application of Gosper's algorithm to definite summation, namely proving identities of the form

$$S_n = \sum_{k=-\infty}^{\infty} F(n, k) = 1$$

where $F$ is a hypergeometric term with a finite support. As we saw in the section above, the summand $F$ is not Gosper-summable. But what can we say about its difference wrt to $n$?

$$F(n + 1, k) - F(n, k)$$

Suppose it's Gosper-summable, therefore there exists such $G(n, k)$ that

$$F(n + 1, k) - F(n, k) = G(n, k + 1) - G(n, k)$$

(3)
where $G$ is a rational multiple of $F$. Summing (3) over all $k$, yields

$$
\sum_{k=-\infty}^{\infty} [F(n + 1, k) - F(n, k)] = \sum_{k=-\infty}^{\infty} [G(n, k + 1) - G(n, k)]
$$

The right-hand side is telescoping to zero, while the left hand side becomes $S_{n+1} - S_n$. This gives a linear recurrence equation of the first order

$$S_{n+1} - S_n = 0$$

Hence, $S_n$ is a constant. Lastly, we need to make sure that $S_0 = 1$.

The trend in mathematics these days is started to go from computer-assisted conjectures to computer-generated conjectures and then proofs.

**computer-assisted conjectures.** Pythagoras, Archimedes, Euler, Gauss, Riemann and all the other giants, who did extensive experimentation to find conjectures...

**computer-generated conjectures.** There exist powerful software packages that automatically finds conjectures, but without proving them. We will consider them later in the course.

**computer-assisted proofs.** Many proofs nowadays are computer-assisted, but in most of them computers are not mentioned. For example, Coq system.

**computer-generated proofs.** The first full-fledged computer-generated proofs started with WZ theory. In 1931, Kurt Godel proved that every consistent system of axioms is necessarily incomplete.

**Example 1** (Vandermonde's identity)

Prove

$$
\sum_{k=0}^{n} \binom{n}{k} \binom{n}{k} = \binom{n+a}{a}
$$
**Proof.**

We rewrite this identity in the form

\[
\sum_{k=-\infty}^{\infty} F(n, k) = 1
\]

where

\[
F(n, k) = \binom{a}{k} \binom{n}{k} \binom{n+a}{a}
\]

Introduce

\[
s_k = F(n+1, k) - F(n, k)
\]

and apply Gosper’s algorithm to \(s_k\). If it’s Gosper-summable, we find such \(G(n, k)\) that

\[
F(n+1, k) - F(n, k) = G(n, k+1) - G(n, k)
\]

First we show that \(s_k\) is a hypergeometric term

\[
\frac{s_{k+1}}{s_k} = \frac{(k-a)(k-n-1)(1+k+a+n-a+k+n)}{(k+1)^2(-a+k+a-k-an+k+n)}
\]

We choose

\[
p_k = -a+k+a+n+k
\]
\[
q_{k+1} = (k-a)(k-n-1)
\]
\[
r_{k+1} = (k+1)^2
\]

A difference equation

\[
q_{k+1} f_k - r_k f_{k-1} = p_k
\]

\[
(k-a)(k-n-1) f_k - k^2 f_{k-1} = -a+k+a+n+k+n
\]

Its polynomial solution is a constant \(f_k = -1\). Therefore, our function \(G(n, k)\) is

\[
G(n, k) = z_k = \frac{r_k}{p_k} f_{k-1} s_k = -(a!^2 n!^2) / ((k-1)!(a-k)! (a+n+1)! (n-k+1)!
\]

or

\[
G(n, k) = \frac{k-a-1}{(a+n+1)} F(n, k-1)
\]
Finally, we need to prove the initial case, namely that the sum

$$\sum_{k=0}^{n} \binom{a}{k} \binom{n}{k} = 1$$

is 1 for $n = 0$, which is indeed so:

$$\sum_{k=0}^{n} \binom{n}{k} = \binom{n}{0} = 1$$

- **Certificate**

This pair of function $(G, F)$ is a WZ-pair

$$F(n + 1, k) - F(n, k) = G(n, k + 1) - G(n, k) \quad (4)$$

The rational function $R(n, k)$

$$R(n, k) = \frac{G(n, k)}{F(n, k)}$$

is called a certificate. Knowing $R(n, k)$ we can restore $G(n, k)$ and then verify identity (5). The latter is done by dividing (4) by $F(n, k)$

$$\frac{F(n + 1, k)}{F(n, k)} - 1 = \frac{G(n, k + 1)}{F(n, k)} - \frac{G(n, k)}{F(n, k)}$$

and making use

$$G(n, k) = R(n, k) F(n, k)$$

we obtain

$$\frac{F(n + 1, k)}{F(n, k)} - 1 = \frac{F(n, k + 1)}{F(n, k)} R(n, k + 1) - R(n, k)$$

This defines a meaning of the certificate - we need to verify the above identity.
Example

Prove

\[
\sum_{k=1}^{n} k \binom{n}{k} \binom{m}{k} = \frac{mn}{m+n} \binom{m+n}{m}
\]

First we define the summand as a Mathematica function

\[
F[n_, k_] := (k (m+n) \text{Binomial}[n, k] \text{Binomial}[m, k]) /
(m n \text{Binomial}[n+m, m])
\]

The problem is reduced to proving

\[
\sum_{k=1}^{n} \frac{k (m+n) \binom{n}{k} \binom{m}{k}}{mn \binom{m+n}{m}} = 1
\]

or

\[
S(n) \equiv \sum_{k=1}^{n} F(n, k) = 1
\]

This can be also written as

\[
S(n) \equiv \sum_{k=\infty}^{k=-\infty} F(n, k) = 1
\]

Check a single value \(S(1)\)

\[
F[1, 1]
\]

\[1\]

Therefore, if we can show that

\[
S(n+1) - S(n) = 0 \tag{5}
\]

then by induction we prove that \(S(n) = 1\) for all \(n\). Thus, we prove the original identity. By definition of \(S(n)\), we have

\[
S(n+1) - S(n) = \sum_{k=1}^{n+1} F(n+1, k) - \sum_{k=1}^{n} F(n, k) =
\]
\[ \sum_{k=-\infty}^{\infty} F(n + 1, k) - \sum_{k=-\infty}^{\infty} F(n, k) = \sum_{k=-\infty}^{\infty} [F(n + 1, k) - F(n, k)] \]

Note, \( F(n + 1, n + 1) \) is not zero

```math
<table>
<thead>
<tr>
<th>F[n + 1, n + 1]</th>
</tr>
</thead>
<tbody>
<tr>
<td>(((1 + m + n) \text{Binomial}[m, 1 + n]) / (m \text{Binomial}[1 + m + n, m]))</td>
</tr>
</tbody>
</table>
```

and therefore, we will have to make an additional step. We will take care of this at the end. Let us assume that \( F(n + 1, k) - F(n, k) \) is Gosper-summable, i.e. exist function \( G \) such that

\[ F(n + 1, k) - F(n, k) = G(n, k + 1) - G(n, k) \]

and

\[ \sum_{k=-\infty}^{\infty} [F(n + 1, k) - F(n, k)] = \sum_{k=-\infty}^{\infty} [G(n, k + 1) - G(n, k)] = 0 \]  \hspace{1cm} (6)

since the right hand side is telescoping under the additional assumptions

\[ \lim_{x \to \infty} G(n, x) = 0 \]

\[ \lim_{x \to -\infty} G(n, x) = 0 \]

We proceed with Gosper’s algorithm

```math
<table>
<thead>
<tr>
<th>s[k_] := F[n + 1, k] - F[n, k]</th>
</tr>
</thead>
<tbody>
<tr>
<td>s[k + 1] / FunctionExpand / Factor</td>
</tr>
<tr>
<td>s[k]</td>
</tr>
<tr>
<td>(((k - m) (-1 + k - n) (k m + n + k n - m n)) / (k (1 + k) (-m + k m + k n - m n)))</td>
</tr>
</tbody>
</table>
```

Since it’s a hypergeometric term, we find a triple \((p_k, q_k, r_k)\) such that

\[ \frac{s_{k+1}}{s_k} = \frac{p_{k+1}}{p_k} \frac{q_{k+1}}{r_{k+1}} \]

```math
| p[k_] := -m + k m + k n - m n |
| q[k_] := (k - m - 1) (k - n - 2) |
| r[k_] := k (k - 1) |
```
Here is a difference equation for $f_k$

$$p_k = q_{k+1} f_k - r_k f_{k-1}$$

$$p[k] = q[k+1] f[k] - r[k] f[k-1]$$

$$-m - m n + k (m + n) = -(-1 + k) k f[-1 + k] + (k - m) (-1 + k - n) f[k]$$

The solution is obvious $f_k = -1$

$$\text{Collect}[p[k] - q[k+1] f[k] + r[k] f[k-1] /. f[k_] \to -1, k, \text{Factor}]$$

$$0$$

Thus,

$$G(n, k) = \frac{r_k}{p_k} f_{k-1} s_k$$

$$G[n_, k_] := \frac{r[k]}{p[k]} (-1) s[k]$$

We need to show that $G(n, \pm \infty) = 0$.

$$G[n, x] /\text{FunctionExpand} /\text{FullSimplify}$$

$$-\left(\frac{m n (-1 + x) \text{Gamma}[m]^2 \text{Gamma}[n]^2}{(\text{Gamma}[1 + m + n] \text{Gamma}[1 + m - x] \text{Gamma}[2 + n - x] \text{Gamma}[x]^2)}\right)$$

$$\text{Series}[%22, \{x, \infty, 0\}]$$

$$\left(\frac{1}{x}\right)^{m+n} \left(\frac{m n \text{Gamma}[m]^2 \text{Gamma}[n]^2}{\pi^2 \text{Gamma}[1 + m + n]} + O\left(\frac{1}{x}\right)^1\right) \text{Sin}[m \pi - \pi x] \text{Sin}[n \pi - \pi x]$$

To find a certificate we do the following
\[ \frac{G[n, k]}{F[n, k]} \quad \text{// FunctionExpand // Simplify} \]

\[ \frac{(-1 + k) \, k}{(-1 + k - n) \, (m + n)} \]

\[ R[n\_, k\_] := \frac{(-1 + k) \, k}{(-1 + k - n) \, (m + n)} \]

Verification of the proof by using the certificate

\[ \frac{F[n + 1, k]}{F[n, k]} - 1 - \left( \frac{F[n, k + 1] \, R[n, k + 1]}{F[n, k]} - R[n, k] \right) \quad \text{// FunctionExpand //} \]

\text{FullSimplify}

\[ 0 \]

We noted at the beginning of this example that \( F(n + 1, n + 1) \neq 0 \), so formally

\[ S(n + 1) - S(n) = F(n + 1, n + 1) + \sum_{k=1}^{n} (F(n + 1, k) - F(n, k)) \]

\[ = F(n + 1, n + 1) - G(n, 1) + G(n, n + 1) \]

\[ -G[n, 1] + F[n + 1, n + 1] \text{ + FunctionExpand}[G[n, k]] \text{ /. } k \rightarrow n + 1 \]

\( ((1 + m + n) \text{ Binomial}[m, 1 + n]) / (m \text{ Binomial}[1 + m + n, m]) - \)
\[ (m \, n \, (1 + n) \, \text{Gamma}[m]^2) / \]
\[ ((-m - m \, n + m \, (1 + n) + n \, (1 + n)) \, \text{Gamma}[m - n] \, \text{Gamma}[1 + m + n]) \]

\text{FullSimplify[\%]}

\[ 0 \]
Do-it-yourself

1) Prove

\[
\sum_{k=0}^{n} \frac{n+2}{2k+1} \binom{n}{2k} \binom{2k+1}{k} 2^{n-2k-1} = \binom{2n+1}{n}
\]

2) Show that the WZ algorithm fails on this identity:

\[
\sum_{k=0}^{n} (-1)^k \binom{3k}{n} \binom{n}{k} = (-3)^n
\]

References

