Factorial Polynomials in Computer Algebra Problems Related to Symbolic Summation

Eugene Zima
Wilfrid Laurier University
Waterloo, Canada
ezima@wlu.ca

1. Introduction and definitions
2. Factorial Polynomials
3. Simple application (Polynomial Normal Forms)
4. Two Gosper's methods for indefinite hypergeometric summation
5. Factorial Polynomials in Gosper's algorithm

1. Intro

$\mathbb{K}$ – field of characteristic zero,

x – an independent variable,

E – shift operator with respect to x, i.e.

$$Ef(x) = f(x+1)$$

for an arbitrary expression $f(x)$,

all $f/g \in \mathbb{K}(x)$, with $f, g \in \mathbb{K}[x] \setminus \{0\}$ coprime and $\deg f < \deg g$ (i.e., *proper*),

ρ – the *dispersion* of f/g (Abramov 1971), the maximal integer distance between roots of the denominator g, and

$\sum$ and $\sum_x$ are the same everywhere

The problem of indefinite summation in general is:
given a closed form expression $F(x)$ to find a closed form expression $G(x)$, which satisfies the first order linear difference equation

$$(E-1)G(x) = F(x). \tag{1}$$

If found, write $\sum F(x) = G(x) + c$, where c is an arbitrary constant or any periodic function of x with period 1.

$$F(x) = (E-1)[G(x) + c],$$

$$\sum F(x) = \sum (E-1)[G(x) + c],$$

and since $\sum_x (E-1) = (E-1)\sum_x = 1$

$$\sum F(x) = G(x) + c.$$

If not found one can try to solve the *additive decomposition problem*:
construct two closed form expressions $R(x)$ and $H(x)$, such that

$$F(x) = (E - 1)R(x) + H(x), \tag{2}$$

where $H(x)$ is simpler than $F(x)$ is some sense (or as variation, as simple as possible), and write result of summation as

$$\sum_x F(x) = R(x) + \sum_x H(x)$$

The measure of simplicity can be different for different classes of functions. For example if $F(x)$ is a rational function over $\mathbb{K}$ one requires both $R(x)$ and $H(x)$ to be rational with $H(x)$ having the denominator of the lowest possible degree ($\rho = 0$).

2. Factorial Polynomials

Following Moenck (1977) define the *factorial polynomial* (a generalization of the falling factorial) for $p(x) \in \mathbb{K}[x]$ as

$$[p(x)]_k = p(x) \cdot p(x-1) \cdot \ldots \cdot p(x-k+1) \tag{3}$$

for $k > 0$ and $[p(x)]_0 = 1$.

Note, that the left hand side of (3) offers succinct (most compact) representation of the product in the right hand side for large values of k, as it requires $\Theta(\log k)$ bits to represent the polynomial $p(x) \cdot p(x-1) \cdot \ldots \cdot p(x-k+1)$ assuming the degree of $p(x)$ is fixed.

Factorial polynomials satisfy many obvious identities, which capture their multiplicative nature and allow manipulate them without expanding. We list only few of them for illustration purposes:

$$[p(x)]_k = [p(x)]_{k-1} \cdot p(x-k+1),$$

$$[p(x+1)]_k = p(x+1) \cdot [p(x)]_{k-1}\,, \quad \text{for } k > 0,$$

$$[p_1(x) \cdot p_2(x)]_k = [p_1(x)]_k\, [p_2(x)]_k$$

Based on these it is easy to implement lazy evaluation rules, such as for example,

$$A \cdot [p(x)]_k \pm B \cdot [p(x+1)]_k = [A \cdot p(x-k+1) \pm B \cdot p(x+1)] \cdot [p(x)]_{k-1}\,, \tag{4}$$

which holds for arbitrary expressions A and B.

Another set of rules involves computation of gcd and cancelations. For example, given natural h, k, and l:

$$\gcd([p(x)]_k\,,[p(x+h)]_l) = \begin{cases} 1 & \text{if } l-h \le 0, \\ [p(x)]_{\min(k,l-h)} & \text{otherwise.} \end{cases}$$

...

The ultimate goal of lazy manipulation rules is to avoid complete expansion of the involved factorial polynomials as much as possible.

3. Simple application (Polynomial Normal Forms)

Consider $R \in \mathbb{K}(x)$. If $z \in \mathbb{K}$ and monic polynomials $A, B, C \in \mathbb{K}[x]$ satisfy

(i) $R = z \cdot \frac{A}{B} \cdot \frac{EC}{C}$,

(ii) $A \perp E^k B$ for all $k \in \mathbb{N}$,

then (z, A, B, C) is a *polynomial normal form* (PNF) of R. If in addition,

(iii) $A \perp C$ and $B \perp EC$,

then (z, A, B, C) is a *strict* (PNF) of R (see S. A. Abramov and M. Petkovšek, 2002 for details).

For example, for the rational function $\frac{x+1000}{x+1}$ PNF is

$$1, 1, 1, (x + 999) \cdot (x + 998) \cdot (x + 997) \cdot \cdots \cdot (x + 2) \cdot (x + 1),$$

with polynomial C of degree 999.

The notion of Rational Normal Form (RNF) was introduced in the context of hypergeometric summation (see S. A. Abramov and M. Petkovšek, 2002 for details). The main steps in computation of RNF for a given rational function $R \in \mathbb{K}(x)$ involve construction of two PNFs, computation of gcd, and divisions:

```
(z,a,b,c) := PolynomialNormalForm(R,n);
(z1,a1,b1,c1) := PolynomialNormalForm(b/a,n);
g := gcd(c,c1);
return (z,b1,a1,c/g,c1/g)
```

The gain from using succinct representation is transparent from the following example: for $R = \frac{x(x+1000)}{(x+3)(x+1003)}$ the first call to PolynomialNormalForm produces $(1, x, x+1003, [x+999]_{997})$ with polynomial c of degree 997, the second call produces $(1, 1, 1, [x+1002]_{1003})$ with polynomial $c1$ of degree 1003, and $\gcd(c, c1) = [x+999]_{997}$. The RNF for R is

$$(1, 1, 1, 1, (x+1001)\,(x+1002)\,(x+1)\,(x+2)\,x\,(x+1000)),$$

with most of the terms from PNFs cancelled. Our prototype produces this result in 0.012 seconds, while standard Maple implementation requires 3.5 second on the same computer.

4. Two Gosper's methods for indefinite hypergeometric summation

Recall that a nonzero expression $F(x)$ is called a hypergeometric term over $\mathbb{K}$ if there exists a rational function $r(x) \in \mathbb{K}(x)$ such that $F(x+1)/F(x) = r(x)$. Usually $r(x)$ is called the rational *certificate* of $F(x)$.

The problem of indefinite hypergeometric summation (anti-differencing) is: given a hypergeometric term $F(x)$ to find a hypergeometric term $G(x)$, which satisfies the first order linear difference equation

$$(E - 1)G(x) = F(x). \tag{5}$$

Gosper's approach is based on simple observation that if a given hypergeometric term $F(x)$ is summable, then the terms $G(x)$ and $F(x)$ are *similar*: i.e., there exists $Y(x) \in \mathbb{K}(x)$ such that $G(x) = Y(x)F(x)$.

This reduces the original summation problem to the problem of finding a rational solution of

$$Y(x+1)r(x) - Y(x) = 1, \tag{6}$$

where $r(x)$ is the rational certificate of the summand.

5. Two Gosper's methods...

1. R. W. Gosper, Jr. Indefinite hypergeometric sums in MACSYMA. In *Proceedings of the 1977 MACSYMA Users' Conference*, pages 237–251, 1977.

2. R. W. Gosper, Jr. Decision procedure for indefinite hypergeometric summation. *Proc. Nat. Acad. Sci. U.S.A.*, 75(1):40–42, 1978.

Comminicated by Donald E. Knuth, September 26, 1977

In order to solve (6), the rational certificate is transformed to so-called Gosper form [a] , i.e. one finds polynomials P, Q, R such that

$$r(x) = \frac{P(x+1)Q(x)}{P(x)R(x)},$$

where $Q(x)$ and $R(x+h)$ are co-prime for all non-negative integers h. This reduces the search for a rational solution of (6) to the search of a polynomial $y(x)$ solving the *key equation*:

$$Q(x)y(x+1) - R(x-1)y(x) = P(x). \qquad (7)$$

[a] Sometimes also called Gosper-Petkovšek form, if an extra conditions on P,Q,R is used in this step. Modern term for this representation – *Polynomial Normal Form* – can be found for example in [1].

If $y(x)$ is found, then the rational multiple of the summand is $Y(x) = \frac{R(x-1)y(x)}{P(x)}$ and

$$G(x) = F(x)\frac{R(x-1)y(x)}{P(x)}. \tag{8}$$

5. Factorial Polynomials in Gosper's algorithm

Example. Consider the application of Gosper's algorithm to the following hypergeometric summand

$$F(x) = \frac{\left(27\,x^3 + 819\,x^2 + 246\,x - 194\right)(2\,x)!}{(3\,x+91)\,(3\,x+1)\,(x+1)\,(3\,x+94)\,(3\,x+4)\,(x!)^2}. \qquad (9)$$

The rational certificate of $F(x)$ is

$$r(x) = 2\,\frac{(3\,x+1)\,(3\,x+91)\,(2\,x+1)\left(27\,x^3 + 900\,x^2 + 1965\,x + 898\right)}{(27\,x^3 + 819\,x^2 + 246\,x - 194)\,(3\,x+7)\,(3\,x+97)\,(x+2)},$$

which has dispersion of the numerator and the denominator equal to 32.

After computing the Gosper-Petkovšek form the key equation becomes

$$4\,(x+1/2)\,(x+1/3)\,y\,(x+1)-(x+94/3)\,(x+1)\,y\,(x)=$$
$$=\left(x^3+\frac{91}{3}\,x^2+\frac{82}{9}\,x-\frac{194}{27}\right)\left[x+\frac{88}{3}\right]_{28}$$

with the right-hand side of degree 31.

This equation has a polynomial solution $y(x)$ of degree 29:

$$\frac{1}{3}\,\left[x+\frac{88}{3}\right]_{29},$$

which after substitution in (8) forces the denominator $P(x)$ to cancel completely, and final result of summation is

$$\sum_x \frac{\left(27\,x^3+819\,x^2+246\,x-194\right)(2\,x)!}{(3\,x+91)\,(3\,x+1)\,(x+1)\,(3\,x+94)\,(3\,x+4)\,(x!)^2}=$$

$$=\frac{(2\,x)!}{(3\,x+91)\,(3\,x+1)\,(x!)^2}.$$

In what follows let $[p(x)]_k$ be one of the factors of $P(x)$ in (7). Our approach is based on a succinct representation of the factorial polynomials appearing in the Gosper-Petkovšek form, lazy evaluation of consecutive values of $y(x)$ in (7) and very simple properties of the components of the equation (7):

1. Polynomial Normal Form (PNF) has a "local" property [4]: PNF of the product of polynomials from different shift equivalence classes is the product of PNFs of those polynomials.

2. Factorial polynomials appear only in the right-hand side of the key equation (7) and they are the only candidates for cancelation. Moreover, the number of these factorial polynomials is bounded by the degrees of the numerator and the denominator of the rational certificate $r(x)$ and does not depend on the value of the dispersion. On the other hand, each term of the form $[p(x)]_k$ contributes the value of k towards the upper bound of the degree of the solution $y(x)$, which can be as

large as the value of the dispersion.

3. The term $[p(x)]_k$ vanishes at any root α of $p(x)$ and also at $\alpha+1,\ldots,\alpha+k-1$.

4. If a solution $y(x)$ of (7) is equal to zero at any of $\alpha,\alpha+1,\ldots,\alpha+k$ (where α is a root of $p(x)$), then $y(x)$ is equal to zero at all these points. This means that $[p(x)]_{k+1}$ is a factor of $y(x)$ and the factorial polynomial term $[p(x)]_k$ in $P(x)$ cancels after substituting $y(x)$ into (8).

5. Any shift equivalent to $p(x)$ factor of $Q(x)$ or $R(x-1)$ from (7) provides initial value for a solution of $y(x)$ at a root β of this factor. If neither $Q(x)$ nor $R(x-1)$ contains a factor shift equivalent to $p(x)$, then the term $[p(x)]_k$ is present in the summand $F(x)$.

6. The evaluations required to detect equality or non-equality of $y(x)$ to zero at the consecutive points starting at β can be done

lazily using (4). The expanded form of $[p(x)]_k$ is not required for this test. Moreover, every nonzero value of $y(x)$, computed at the consecutive points $\beta, \beta+1, ...$ (or $\beta, \beta-1, ...$) during the test, is represented by a nontrivial factor in the rational certificate $r(x)$.

These properties allow us to incorporate simple and efficient changes into Gosper's decision procedure, which do not worsen the total asymptotic complexity of the procedure, but can lead to tremendous savings in the running time for summable terms with large dispersion of the rational certificate. Returning to the example above, two evaluation points ($x = -91/3$ and $x = -88/3$) are sufficient to find out that the term $\left[x + \frac{88}{3}\right]_{28}$ will cancel, and the substitution of $y(x) = \left[x + \frac{88}{3}\right]_{29} y(x)$ into the key equation gives reduced key equation:

$$4\,(x + 1/2)\,(x + 1/3)\,(x + 91/3)\,y\,(x + 1) - (x + 94/3)\,(x + 4/3)\,(x + 1)\,y\,(x) =$$

$$= \left(x^3 + \frac{91}{3}\,x^2 + \frac{82}{9}\,x - \frac{194}{27}\right)$$

with the degree of the solution < 3. The solution $y(x) = 1/3$ of the last equation produces the desired result in reduced form.

References

[1] S. A. Abramov and M. Petkovšek. Rational normal forms and minimal decompositions of hypergeometric terms. *Journal of Symbolic Computation*, 33(5):521–543, 2002. Computer algebra (London, ON, 2001).

[2] G. Boole. *A Treatise on the Calculus of Finite Differences*. Cambridge Library Collection. Cambridge University Press, Cambridge, 2009. Reprint of the 1860 original.

[3] A. Bostan, F. Chyzak, B. Salvy, and T. Cluzeau. Low complexity algorithms for linear recurrences. In *Proceedings of the 2006 International Symposium on Symbolic and Algebraic Computation*, ISSAC '06, pages 31–38, New York, NY, USA, 2006. ACM.

[4] R. Pirastu and V. Strehl. Rational summation and Gosper-Petkovsšek representation. *J. Symb. Comput.*, 20(5-6):617–635, Nov. 1995.