A Generalization of Gosper's Algorithm

Marko Petkovšek

Department of Mathematics, University of Ljubljana

Jadranska 19, 61111 Ljubljana, Slovenia

Abstract

We present a derivation of Gosper's algorithm which permits generalization to higher-order recurrences with constant least and most significant coefficients. Like Gosper's algorithm, the generalized algorithm requires only "rational" operations (such as gcd and resultant computations) but no factorization.

1 Introduction

Let F be a field of characteristic zero. We will denote by $\mathcal{F}$ the ring of all sequences over F, with addition and multiplication defined term-wise. Following Stanley [6] we identify two sequences if they agree from some point on. Formally, we are working in the quotient ring $\mathcal{F}/J$ where J is the ideal of sequences with only finitely many non-zero terms. Hence, all equalities of the form $lhs(n) = rhs(n)$ are to be interpreted as valid for all large enough n. Accordingly, a sequence is considered to be non-zero if and only if it has infinitely many non-zero terms.

Definition 1 A sequence $a \in \mathcal{F}$ will be called *rational over* F if there is a rational function r over F such that

$$a_n = r(n)\,.$$

A non-zero sequence $a \in \mathcal{F}$ will be called *hypergeometric over* F if there is a rational function r over F such that

$$a_{n+1}/a_n = r(n)\,. \tag{1}$$

The well-known Gosper's algorithm [3] (see also [4]) decides for a given hypergeometric sequence a_n whether the first-order recurrence

$$s_{n+1} - s_n = a_n \tag{2}$$

has a hypergeometric solution s_n, and if so, finds one. In this note we generalize Gosper's algorithm to non-homogeneous recurrences with rational coefficients of arbitrary order. We start in Section 2 by deriving Gosper's algorithm from uniqueness of a certain factorization of rational functions. In Section 3, we show how to generalize this derivation to non-homogeneous recurrences with rational coefficients of any order, provided their least and most significant coefficients are constant.

We assume that F is *computable*, meaning that the elements of F can be represented by finite strings over some finite alphabet, and that there exist algorithms for carrying out the field operations. We also assume that there exists an algorithm for finding integer roots of polynomials over F.

2 Gosper's algorithm revisited

If s_n is hypergeometric and satisfies (2) where a_n is a given hypergeometric sequence then the quotient $s_n/a_n = s_n/(s_{n+1} - s_n) = 1/(s_{n+1}/s_n - 1)$ is a rational function of n. So let

$$\frac{s_n}{a_n} = R(n)$$

where R is an unknown rational function over F. From (1) and (2) it follows that $R(n)$ satisfies

$$r(n)R(n+1) - R(n) = 1 \tag{3}$$

and the problem of finding *hypergeometric* solutions of (2) has been reduced to that of finding *rational* solutions of (3). Gosper's achievement was to reduce it further to the problem of finding *polynomial* solutions of yet another first-order recurrence (see (7) below).

Lemma 1 *Let F be a field of characteristic zero, and r a rational function over F. Then there exist polynomials A, B, C over F such that*

$$r(n) = \frac{A(n)}{B(n)} \frac{C(n+1)}{C(n)} \tag{4}$$

and

$$\gcd(A(n), B(n+k)) = 1$$

for all nonnegative integers k.

For a (constructive) proof, see [3]. The algorithm to compute A, B, and C given r uses gcd and resultant computations on polynomials over F.

We remark that if we additionally require that A, B, C are *monic* polynomials such that $\gcd(A(n), C(n)) = \gcd(B(n), C(n+1)) = 1$, then a factorization of the form

$$r(n) = Z \frac{A(n)}{B(n)} \frac{C(n+1)}{C(n)} \tag{5}$$

where $Z \in F$ is a constant, exists and is unique (see [5]).

Let r be related to A, B, and C as in Lemma 1. In [3], a rational solution of (3) is sought in the form

$$R(n) = \frac{B(n-1)p(n)}{C(n)} \tag{6}$$

where p is an unknown rational function. Miraculously, it turns out that p is a polynomial. Substitution of (4) and (6) into (3) shows that p satisfies the first-order recurrence

$$A(n)p(n+1) - B(n-1)p(n) = C(n)\,. \tag{7}$$

It is not difficult to determine from (7) a finite set of possible degrees which p might have, and then to find all polynomial solutions of (7). For details, see [3] or [4].

Here we want to derive Gosper's ansatz (6).

Lemma 2 *Let F be a field of characteristic zero. Let a, b, and c be polynomials over F such that* $\gcd(a(n), c(n)) = \gcd(b(n), c(n+1)) = 1$. *If*

$$r(n) = \frac{a(n)}{b(n)} \frac{c(n+1)}{c(n)}\,, \tag{8}$$

and C is related to r as in Lemma 1, then c divides C.

Proof: Use (4) to rewrite (8) as

$$A(n)b(n)c(n)C(n+1) = a(n)B(n)C(n)c(n+1)\,. \tag{9}$$

This equation holds for all large enough integers n, but since the characteristic is zero it follows that it also holds as an equality between polynomials. Let

$$\begin{aligned} g(n) &:= \gcd(c(n), C(n))\,, && (10)\\ d(n) &:= c(n)/g(n)\,, && (11)\\ D(n) &:= C(n)/g(n)\,. && (12) \end{aligned}$$

Then $\gcd(d(n), D(n)) = \gcd(a(n), d(n)) = \gcd(b(n), d(n+1)) = 1$, and canceling $g(n)g(n+1)$ from (9) gives

$$A(n)b(n)d(n)D(n+1) = a(n)B(n)D(n)d(n+1)\,.$$

It follows that

$$\begin{aligned} d(n) &\mid B(n)d(n+1)\,,\\ d(n+1) &\mid A(n)d(n)\,. \end{aligned}$$

Using this repeatedly we obtain

$$\begin{aligned} d(n) &\mid B(n)B(n+1)\cdots B(n+k-1)d(n+k)\,,\\ d(n) &\mid A(n-1)A(n-2)\cdots A(n-k)d(n-k)\,, \end{aligned}$$

for all $k \in \mathbb{N}$. Since F is of characteristic zero, $\gcd(d(n), d(n+k)) = \gcd(d(n), d(n-k)) = 1$, for all large enough k. It follows that $d(n)$ divides both $B(n)B(n+1)\cdots B(n+k-1)$ and $A(n-1)A(n-2)\cdots A(n-k)$ for all large enough k. But these two polynomials are relatively prime by the assumptions of the Lemma, so $d(n)$ is constant. Hence $c(n)$ divides $C(n)$, by (11) and (10). $\square$

We remark that uniqueness of factorization (5) is an immediate consequence of Lemma 2.

Write

$$R(n) = \frac{f(n)}{g(n)}$$

where $f(n)$, $g(n)$ are two unknown, relatively prime polynomials. Then from (3) we obtain

$$r(n) = \frac{f(n) + g(n)}{f(n+1)} \frac{g(n+1)}{g(n)} \tag{13}$$

where $\gcd(g(n+1), f(n+1)) = \gcd(f(n) + g(n), g(n)) = 1$. By Lemma 2 it follows that g divides C where C is related to r as in Lemma 1. Therefore C can be used as the denominator of the unknown rational function R. So write

$$R(n) = \frac{q(n)}{C(n)} \tag{14}$$

where $q(n)$ is an unknown polynomial. Inserting (4) and (14) into (3) gives

$$A(n)q(n+1) = (q(n) + C(n))B(n)\,,$$

which shows that $q(n)$ is divisible by $B(n-1)$. The final form in which we seek $R(n)$ is thus (6), with $p(n)$ an unknown polynomial satisfying (7).

3 Gosper's algorithm for recurrences of arbitrary order

In this section we consider the recurrence

$$\sum_{k=0}^{d} p_k(n) s_{n+k} = a_n \tag{15}$$

where a_n is a given non-zero sequence, and $p_0(n)$, $p_1(n)$, ..., $p_d(n)$ are given polynomials, with the additional restriction that p_0 and p_d are constant. Does (15) have a hypergeometric solution? If this is the case then the left-hand side of (15) is a rational multiple of s_n, so a_n is hypergeometric, and

$$\frac{s_n}{a_n} = R(n)$$

is a rational function of n. This reduces the search for hypergeometric solutions s_n of (15) to the search for rational solutions $R(n)$ of

$$\sum_{k=0}^{d} p_k(n) R(n+k) \prod_{j=0}^{k-1} r(n+j) = 1 \tag{16}$$

where $r(n) = a_{n+1}/a_n$ is a rational function of n.

Let $r(n)$ be related to $A(n)$, $B(n)$, and $C(n)$ as in Lemma 1. Write

$$R(n) = \frac{f(n)}{g(n)C(n)} \tag{17}$$

where $f(n)$ and $g(n)$ are unknown, relatively prime polynomials. Using (4) and (17) in (16) gives

$$\begin{aligned} &\sum_{k=0}^{d} p_k(n) f(n+k) \prod_{\substack{j=0 \\ j\neq k}}^{d} g(n+j) \prod_{j=0}^{k-1} A(n+j) \prod_{j=k}^{d-1} B(n+j) \\ &= C(n) \prod_{j=0}^{d-1} B(n+j) \prod_{j=0}^{d} g(n+j) \end{aligned} \tag{18}$$

after some manipulation. All terms in (18) except the one with $k = 0$ are divisible by $g(n)$, so

$$g(n) \ \mid\ p_0(n) \prod_{j=1}^{d} g(n+j) \prod_{j=0}^{d-1} B(n+j)\,. \tag{19}$$

Similarly, looking at the term with $k = d$ and substituting $n - d$ for n, we find that

$$g(n) \ \mid\ p_d(n-d) \prod_{j=1}^{d} g(n-j) \prod_{j=1}^{d} A(n-j)\,. \tag{20}$$

Using (19) and (20), and the fact that $p_0(n)$ and $p_d(n)$ are constant, one can show by induction that for every $m \in \mathbb{N}$, $g(n)$ divides a product of factors of the form $g(n+j)$ and $B(n+i)$ where $j \geq m$ and $i \geq 0$, as well as a product of factors of the form $g(n-j)$ and $A(n-i)$ where $j \geq m$ and $i \geq 1$. As the characteristic is zero there is an m such that $g(n)$ is relatively prime with $g(n+j)$ for $j \geq m$ and for $j \leq -m$. From the properties of $A(n)$ and $B(n)$ it follows that $g(n)$ is a constant. Therefore we may write $R(n) = q(n)/C(n)$ where $q(n)$ is a polynomial satisfying

$$\sum_{k=0}^{d} p_k(n) q(n+k) \prod_{j=0}^{k-1} A(n+j) \prod_{j=k}^{d-1} B(n+j) = C(n) \prod_{j=0}^{d-1} B(n+j)\,. \tag{21}$$

Equation (21) implies that $q(n)$ is divisible by $B(n-1)$, so we seek $R(n)$ in the form $R(n) = B(n-1)p(n)/C(n)$ where $p(n)$ is a polynomial satisfying

$$\sum_{k=0}^{d} p_k(n)p(n+k) \prod_{j=0}^{k-1} A(n+j) \prod_{j=k-1}^{d-2} B(n+j) = C(n) \prod_{j=0}^{d-2} B(n+j) \,. \tag{22}$$

Algorithms for finding polynomial solutions of recurrences with rational coefficients are given in [1] and [5]. These algorithms require finding integer roots of polynomials over F.

Example 1 Let $a_n = 4\binom{2n}{n+2} - 5\binom{2n}{n}$. Then $r(n) = \frac{a_{n+1}}{a_n} = \frac{A(n)C(n+1)}{B(n)C(n)}$ where

$$\begin{aligned} A(n) &= 4n+2\,, \\ B(n) &= n+3\,, \\ C(n) &= n^2+19n+10\,. \end{aligned}$$

To find all hypergeometric solutions of

$$2s_{n+2} - 8s_{n+1} - s_n = a_n\,, \tag{23}$$

let $s_n = B(n-1)p(n)a_n/C(n)$. Then, by (22), $p(n)$ is a polynomial which satisfies

$$\begin{aligned} 8(2n+1)(2n+3)p(n+2) - 16(2n+1)(n+4)p(n+1) - (n+3)(n+4)p(n) \\ = (n^2+19n+10)(n+3)(n+4)\,. \end{aligned}$$

Using the algorithms of [1] or [5] it can be shown that the only polynomial solution of this equation is $p(n) = -(n+1)(n+2)$. Thus

$$s_n = \frac{p(n)B(n-1)}{C(n)}\,a_n = \binom{2n}{n}$$

is the only hypergeometric solution of (23). □

Alternatively, the algorithm of [2] can be used on (16) to find all its rational solutions. This algorithm poses no additional restrictions on the coefficients of the recurrence.

An algorithm which will find all solutions of (15) (without restrictions on the coefficients) that are *linear combinations* of hypergeometric terms is given in [5]. Unlike the algorithm presented here, it requires factorization of polynomials (namely, of the least and most significant coefficients of the recurrence).

References

[1] S.A. Abramov, Problems in computer algebra that are connected with a search for polynomial solutions of linear differential and difference equations, *Moscow Univ. Comput. Math. Cybernet.* (1989) no. 3, 63 – 68. Transl. from *Vestn. Moskov. univ. Ser. XV. Vychisl. mat. kibernet.* (1989) no. 3, 56 – 60.

[2] S.A. Abramov, Rational solutions of linear differential and difference equations with polynomial coefficients, *U.S.S.R. Comput. Maths. Math. Phys.* **29** (1989) 7 – 12. Transl. from *Zh. vychisl. mat. mat. fiz.* **29** 1611 – 1620.

[3] R.W. Gosper, Jr., Decision procedure for indefinite hypergeometric summation, *Proc. Natl. Acad. Sci. USA* **75** (1975) 40 – 42.

[4] R.L. Graham, D.E. Knuth, and O. Patashnik, *Concrete Mathematics*, Addison-Wesley, Reading, Massachusetts, 1989.

[5] M. Petkovšek, Hypergeometric solutions of linear difference equations with polynomial coefficients, *J. Symb. Comput.*, to appear.

[6] R.P. Stanley, Differentiably finite power series, *European J. Combin.* **1** (1980) 175 – 188.