Certificates of Integrality for Linear Binomials

DAVID CALLAN
Department of Statistics
University of Wisconsin-Madison
1210 W. Dayton St
Madison, WI 53706-1693
callan@stat.wisc.edu

September 9, 1999

1 Introduction and Statement of Main Theorem

Everyone knows that the familiar binomial coefficients are integers. But it is not so obvious that quotients of binomial coefficients whose parameters are linear in n by factors linear in n also sometimes yield sequences of integers. For example, $\{\frac{1}{n+1}\binom{2n}{n}\}_{n\geq 0} = \{1, 1, 2, 5, 14, 42, 132, \ldots\}$ is the well known sequence of Catalan numbers. In the same vein, $\{\frac{3}{n}\binom{2n}{n-3}\}_{n\geq 3} = \{1, 6, 27, 110, 429, \ldots\}$ is sequence M4177 in Sloane and Plouffe's *Encyclopedia of Integer Sequences* [3], $\{\frac{3}{n+2}\binom{2n}{n-1}\}$ is sequence M2809, $\{\frac{4}{(3n+2)(3n+1)}\binom{3n+2}{n}\}$ is M1660, $\{\frac{5}{n+3}\binom{2n}{n-2}\}$ is M3904. There are at least another dozen such sequences listed in the *Encyclopedia*, including M1782, M2243, M2926, M2946, M2997, M3483, M3542, M3587, M4198, M4214, M4529, M4721. Incidentally, the smallest-parameter such sequence of integers *not* listed seems to be $\{\frac{1}{n}\binom{3n}{n+1}\} = \{2\binom{3n-1}{n} - \binom{3n-1}{n+1}\} = \{3, 10, 42, 198, 1001, \ldots\}$.

Why are these sequences integral while similar sequences such as $\frac{k}{n}\binom{2n}{n}$ and $\frac{k}{2n+1}\binom{2n}{n}$ are not, no matter what the integer k is? Here we attempt to shed some light on this question. Each of the above sequences is an integer multiple of a sequence of the form $\mathbf{w} = \frac{1}{P(n)}\binom{an+b}{cn+d}$ where $P(n)$ is a product of one or more factors linear in n with integral coefficients and a, b, c, d are integers with $a > c > 0$. Let us call such a sequence $\mathbf{w}$ *linear binomial.* In this paper, we establish a simple and intuitively appealing criterion for a linear binomial sequence $\mathbf{w}$ to have bounded denominators, equivalently, for the existence of an integer k such that $k\mathbf{w}$ is a sequence of integers. Furthermore, when the criterion is met, the proof consists of verification of an algorithm that produces not only a suitable multiplier k, but also a "Certificate of Integrality" for $k\mathbf{w}$ in the form of an identity expressing it as an integral linear combination of binomial coefficients. For example, the algorithm yields that the Catalan number $\frac{1}{n+1}\binom{2n}{n}$ is equal to

$\binom{2n}{n} - \binom{2n}{n-1}$. For $\frac{1}{n}\binom{2n}{n-3}$ the algorithm returns the identity $\frac{3}{n}\binom{2n}{n-3} = \binom{2n-1}{n-3} - \binom{2n-1}{n-4}$. A small Mathematica package, DecomposeBinomial, implementing this algorithm, is available from the author's home page at **http://www.stat.wisc.edu/˜callan/**.

The criterion for bounded denominators revolves around cancellation of the factors in $P(n)$ with factors in what might be called the symbolic numerator of $\binom{an+b}{cn+d}$. Here cancellation refers to proportional polynomials or, equivalently, division in the polynomial ring $\mathbb{Q}[n]$. Set $e = a - c$ and $f = b - d$. Thus for any particular n,

$$\binom{an+b}{cn+d} = \frac{(an+b)(an+b-1)\dots(en+f+1)}{(cn+d)!}. \tag{1}$$

Now define the *numerator set* N of this binomial coefficient (considered symbolically) to be $N = U \cup V$ where $U = \{an+b-i\}_{i\geq 0}$ and $V = \{en+f+j\}_{j\geq 1}$. Thus N contains both "ends" of the range of factors in the numerator in (1) but not the "middle". For example, for $\binom{6n}{2n}$, the numerator set consists of $\{6n, 6n-1, 6n-2, \dots\} \cup \{4n+1, 4n+2, \dots\}$ (but not any term of the form $5n \pm i$). Similarly, define the *denominator set* $D = \{cn+d-i\}_{i\geq 0}$. The desired criterion can now be expressed as follows. *Each linear factor in $P(n)$ must divide a factor in N and if a factor in D is proportional to one in $P(n)$, it too must divide a factor in N (always taking multiplicity into account).*

For example, $\frac{1}{2n+1}\binom{2n}{n}$ fails to meet this criterion because $2n+1$ does not divide any term in $N = \{2n, 2n-1, \dots\} \cup \{n+1, n+2, \dots\}$. And $\frac{1}{n}\binom{2n}{n}$ also fails to meet the criterion because $D = \{n, n-1, \dots\}$ includes n, giving two n's that need to divide factors in $N = \{2n, 2n-1, \dots\} \cup \{n+1, n+2, \dots\}$ but only one term in N is divisible by n. On the other hand, $\frac{1}{n}\binom{5n}{2n+1}$ does meet the criterion because, although here again D includes a factor proportional to n, namely $2n$, the numerator set $N = \{5n, 5n-1, \dots\} \cup \{3n, 3n+1, \dots\}$ contains *two* terms proportional to n, and so both offending factors can be cancelled. Clearly, no two factors in U (resp. V, resp. D) can be proportional. It follows that the criterion cannot be met if $P(n)$ has two proportional (or repeated) factors. This is because the only way N can contain two proportional factors is if one of them is in U (say in the ith position) and the other in V (say in the jth position). But then a simple calculation shows that the $(i+j)$th term in D would also be proportional to both, and "three into two won't go".

To state the criterion (and our main result) succinctly, we make two definitions. Say a linear factor *appears* in a set if it is proportional to a term in the set. Thus $2n+1$ appears in the numerator set of $\binom{4n+3}{n}$. Also, say a linear binomial sequence $\frac{1}{P(n)}\binom{an+b}{cn+d}$ is *normalized* if each linear factor $gn+h$ in $P(n)$ has relatively prime coefficients g, h.

Using this terminology, our main result can be formulated as follows.

Theorem 1 *Suppose* $\mathbf{w} = \frac{1}{P(n)}\binom{an+b}{cn+d}$ *is a normalized linear binomial sequence. Then* $\mathbf{w}$ *has bounded denominators if and only if $P(n)$'s linear factors are distinct and each such factor* *appears in the numerator set N of the binomial coefficient (as defined above), and appears there twice if it also appears in the denominator set D.*

Furthermore, if a linear binomial sequence $\mathbf{w}$ *has bounded denominators, then there is a positive integer k such that $k\mathbf{w}$ is an integral linear combination of a fixed number (independent of n) of binomial coefficients with parameters linear in n.*

Remark Bearing in mind that a factor can appear at most twice in N, an equivalent but more pithy formulation of the criterion for bounded denominators is: if and only if $P(n)$'s linear factors are distinct, and each appears more often in N than in D.

The "only if" part is proved in §2. It relies on Dirichlet's classic theorem on primes in arithmetic progressions [1, Chap. 7], and Kummer's pretty rule for finding the exact power of a prime p that divides a binomial coefficient: the number of carries when its parameters are subtracted in base p. See [2, Ex. 5.36, p. 245] for a proof of Kummer's rule (in an equivalent formulation in terms of addition in base p). The "if" part is proved in §4. It relies on a neat determinant expansion, of interest in its own right, that is presented in §3. Finally, §5 contains a mild extension of the main theorem, some further remarks, and a conjecture.

2 Main Theorem: Proof of "Only If"

We will show that infinitely many primes occur among the denominators in $\frac{1}{P(n)}\binom{an+b}{cn+d}$ when the criterion of Theorem 1 is not met. Let $gn+h$ be a factor in $P(n)$. Suppose $p = gn+h$ is prime (as it will be for infinitely many n by Dirichlet's theorem, since g and h are relatively prime). Write $a = q_1 g + r_1$ with $0 \le r_1 \le g$ and $c = q_2 g + r_2$ with $0 \le r_2 \le g$ (division algorithm). Expressed in base p, the two parameters of the binomial coefficient are then (for sufficiently large n)

$$an+b = \begin{array}{|c|c|}\hline p & 1 \\ \hline q_1 & r_1 n + b - q_1 h \\ q_1 & b - q_1 h \\ q_1 - 1 & p - (q_1 h - b) \\ \hline \end{array} \quad \begin{array}{l} \\ \text{if } r_1 \neq 0, \\ \text{if } r_1 = 0 \text{ and } b \ge q_1 h, \\ \text{if } r_1 = 0 \text{ and } b < q_1 h. \end{array}$$

and similarly,

$$cn+d = \begin{array}{|c|c|}\hline p & 1 \\ \hline q_2 & r_2 n + d - q_2 h \\ q_2 & d - q_2 h \\ q_2 - 1 & p - (q_2 h - d) \\ \hline \end{array} \quad \begin{array}{l} \\ \text{if } r_2 \neq 0, \\ \text{if } r_2 = 0 \text{ and } d \ge q_2 h, \\ \text{if } r_2 = 0 \text{ and } d < q_2 h. \end{array}$$

In particular, since $an+b$ has only two digits in base p, at most one carry can occur in subtracting $cn+d$ from $an+b$ in base p. Thus $p^2 \nmid \binom{an+b}{cn+d}$ and if $gn+h$ is a repeated factor in $P(n)$, then p will occur among the denominators in $\mathbf{w}$ (for infinitely many primes p) and $\mathbf{w}$ will have unbounded denominators. Also, no carries occur in the subtraction—and hence $p \nmid \binom{an+b}{cn+d}$—if and only if $(an+b) \mod p \geq (cn+d) \mod p$. It is straightforward to verify that $gn+h$ appears (i) in U iff $r_1 = 0$ and $b \geq q_1 h$, (ii) in V iff $r_1 = r_2$ and $(q_1 - q_2)h > b - d$, (iii) in D iff $r_2 = 0$ and $d \geq q_2 h$. Except for one wrinkle, it is now simply a matter of checking cases to verify $p \nmid \binom{an+b}{cn+d}$ unless $gn+h = p$ appears in the numerator set N at least once, and twice if it appears in the denominator set D. This will show that infinitely many primes occur among the denominators in $\mathbf{w}$, as desired. The one wrinkle is that when $0 < r_1 < r_2$ (a subcase where $gn+h$ does not appear in N at all), p *does* divide $\binom{an+b}{cn+d}$ and we proceed as follows. Set $n = (g-1)m - h$ with m variable; thus $\frac{1}{gn+h}\binom{an+b}{cn+d} = \frac{1}{g-1}\frac{1}{gm-h}\binom{a(g-1)m-ah+b}{c(g-1)m-ch+d}$. Here $r_1' := (a(g-1)) \mod g = g - r_1$ and $r_2' := (c(g-1)) \mod g = g - r_2$. Since $r_1' > r_2'$, the case $r_1 > r_2$ applies with m in place of n, $a(g-1)$ in place of a, and the rôle of p played by $gm - h$. This completes the proof of the "only if" part.

3 A Determinant Expansion

The following result is crucial for the "if" part of the main theorem in the next section. Let *coeff* denote the function that produces the row vector of coefficients of a polynomial or the matrix of coefficients of a list of polynomials. Thus $\mathrm{coeff}(\sum_{i=0}^{m} c_i x^i) = (c_i)_{i=0}^{m}$. Let $*$ denote convolution of sequences; thus $\mathrm{coeff}(p(x)q(x)) = \mathrm{coeff}(p(x)) * \mathrm{coeff}(q(x))$. Also, for a matrix N, let N° denote the column vector obtained by taking the Hadamard (entrywise) product of the columns in N. For example, for $N = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, $N^\circ = \begin{pmatrix} 2 \\ 12 \end{pmatrix}$.

Theorem 2 *Let m be a positive integer and let a_j $(1 \leq j \leq m)$, b_{ji} $(1 \leq i \leq j \leq m)$, c, e, x be indeterminates. Let N be the $m+1$ by m matrix with rows indexed $[0, m]$ and columns indexed $[1, m]$, and (i, j) entry*

$$\begin{cases} cx + a_j & \text{if } 0 \leq i < j \leq m, \\ ex + b_{ij} & \text{if } 1 \leq j \leq i \leq m. \end{cases}$$

Let M be the $m+1$ by $m+1$ matrix coeff(N°). For example, when $m = 2$,

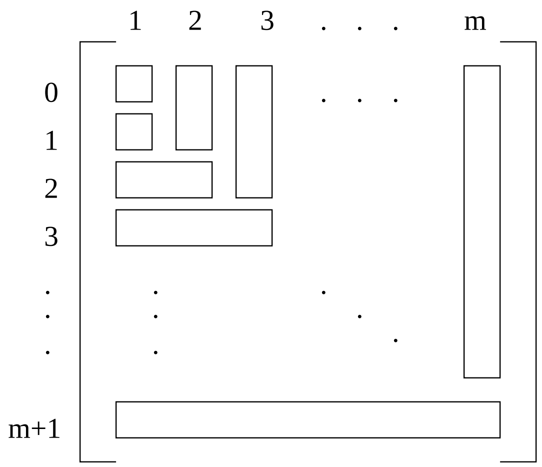
$$N = \begin{pmatrix} cx + a_1 & cx + a_2 \\ ex + b_{11} & cx + a_2 \\ ex + b_{21} & ex + b_{22} \end{pmatrix} \quad \text{and} \quad M = \begin{pmatrix} a_1 a_2 & (a_1 + a_2)c & c^2 \\ b_{11} a_2 & b_{11} c + a_2 e & ce \\ b_{21} b_{22} & (b_{21} + b_{22})e & e^2 \end{pmatrix}$$

Then $\det M = \prod_{1 \leq i \leq j \leq m} (ea_j - cb_{ji})$.

Proof We first show for $1 \leq i \leq j \leq m$ that $ea_j - cb_{ji}$ divides $\det M$ in the polynomial ring $\mathbb{Q}(e, c)[a\text{'s}, b\text{'s}]$. To do so, suppose $ea_j = cb_{ji}$ for some i, j. Let N_j denote the submatrix of

N consisting of rows 0 through j. Then $p_j := \prod_{j \le i \le m}(cx + a_i)$ is a factor in each entry of N_j°; we may write $N_j^\circ = (r_i)_{0 \le i \le j}\, p_j$ with $\deg r_i = j - 1$ $(0 \le i \le j)$. Now rows 0 through j of M constitute the submatrix $M_j = \text{coeff}(N_j^\circ) = \big(\text{coeff}(r_i)\big)_{0 \le i \le j} * \text{coeff}(p_j)$ (convolution of each $\text{coeff}(r_i)$ with $\text{coeff}(p_j)$). Since $R_j := \big(\text{coeff}(r_i)\big)_{0 \le i \le j}$ is a $j+1$ by j matrix, its rows are linearly dependent (over $\mathbb{Q}(e, c, a\text{'s}, b\text{'s})$) and there exists a nonzero vector $\mathbf{u} = (u_i)_{0 \le i \le j}$ such that $\mathbf{u}R_j = \mathbf{0}$. Hence $\mathbf{u}M_j = \mathbf{u}\big(R_j * \text{coeff}(p_j)\big) = (\mathbf{u}R_j) * \text{coeff}(p_j) = \mathbf{0} * \text{coeff}(p_j) = \mathbf{0}$ and M is singular. Thus $ea_j - cb_{ji}$ is a factor of $\det M$. Since each $ea_j - cb_{ji}$ is obviously prime in $\mathbb{Q}(e, c)[a\text{'s}, b\text{'s}]$, their product also divides $\det M$ and Theorem 2 follows by confirming the degrees agree and the coefficients of any one term agree.

Corollary 3 *Let N be an $m+1$ by m matrix with linear polynomials in one indeterminate as entries. Partition N into offset row and column segments as indicated.* (*Each vertical column segment sits atop the last position in the corresponding row segment.*)



Suppose for $1 \le j \le m$ that all entries in column segment j are equal and this common entry does not divide any of the entries in row segment j.

Then the $m+1$ by $m+1$ matrix $M = \text{coeff}(N^\circ)$ is invertible.

Proof The matrix N is of the form in Theorem 2. Clearly, a factor $ea_j - cb_{ji}$ $(1 \le i \le j \le m)$ in $\det M$ is 0 if and only if $cx + a_j$ is proportional to $ex + b_{ji}$, that is, divides $ex + b_{ji}$. But these polynomials lie in corresponding row and column segments and thus the hypothesis ensures that one does not divide the other. Hence $\det M \ne 0$ and M is invertible.

4 Main Theorem: Proof of "If"

We seek an expression for $\frac{1}{P(n)}\binom{an+b}{cn+d}$ as a rational-coefficient linear combination of binomial coefficients. Due to the basic identity $\binom{n}{m} = \binom{n-1}{m} + \binom{n-1}{m-1}$, we can always reduce an upper parameter at the expense of increasing the number of terms in the linear combination. Thus we look for a combination in which all the upper parameters are the same. It will turn out that a suitable upper parameter is determined by the factors in $P(n)$ that appear in U (the upper range in the numerator set). Specifically, it is $an+b-u$ where u is the location of the last term in U that appears in $P(n)$ (and $u=0$ if there is no such term).

By hypothesis, each (linear) factor of $P(n)$ appears in U or V or possibly both. Let $(an+b+1-i)_{i\in I} \cup (en+f+j)_{j\in J}$ be a complete listing of these appearances where I and J are finite subsets (one of them may be empty) of the positive integers. Set $u = \max I$ and $v = \max J$ (with $\max \emptyset := 0$). Let $r_i = an+b+1-i,\ s_i = en+f+v+1-i$ and $t_i = cn+d-u-v+i$ so that

$$\binom{an+b}{cn+d} = \frac{\overbrace{r_1 r_2 \dots r_u}\dots\dots\overbrace{s_1 s_2 \dots s_v}}{\underbrace{t_{u+v} t_{u+v-1} \dots t_1}\dots\dots} = \frac{\prod_{i=1}^{u} r_i \prod_{j=1}^{v} s_j}{\prod_{i=1}^{u+v} t_i}\binom{an+b-u}{cn+d-(u+v)}$$

We claim that all appearances of $P(n)$'s factors in $N \cup D$ occur within the three groupings displayed in the middle expression. This is true for the numerator N by definition of u and v. And if a $P(n)$ factor $gn+h$ appears in D, then by hypothesis it appears in both U and V, say in the ith position in U and the jth position in V. As noted earlier, a simple calculation then shows that the position in D at which $gn+h$ appears is $i+j$. Since $i \le u$ and $j \le v$, it follows that $i+j \le u+v$ and so the $(i+j)$ term in D is one of the displayed t's. Hence the claim.

Next we have to determine appropriate lower parameters for the binomial coefficients in the desired linear combination. This turns out to be a little tricky; rather than being consecutive as one might expect, they turn out to form an interval with a hole in it. To this end, define $L = \{i \in [1, u+v] : t_i | s_j \text{ in the ring } \mathbb{Q}[n] \text{ for some } j \text{ with } 1 \le j \le i\}$. Since the j here is necessarily unique, we get a map $\phi : L \to [1, u+v]$ satisfying $t_i \,|\, s_{\phi(i)}$ and $\phi(i) \le i,\ i \in L$. Also, it is easy to check that L is either empty or an interval of integers. (The reader might like to look ahead to the illustrative example at the end of this section.) Suitable lower parameters are determined by removing L from the set $[1, u+v]$ and adjoining 0. Thus we set $K := [1, u+v] \backslash L$ and the rest of the proof is devoted to showing that there exist (unique) rational numbers $(c_i)_{i\in K\cup\{0\}}$, such that

$$\sum_{i\in K\cup\{0\}} c_i \binom{an+b-u}{cn+d-(u+v)+i} = \frac{1}{P(n)}\binom{an+b}{cn+d}. \tag{2}$$

Factoring out $\binom{an+b-u}{cn+d-(u+v)} / \prod_{j\in K} t_j$ from each side, (2) is equivalent to

$$c_0 \prod_{j\in K} t_j + \sum_{i\in K} c_i \frac{s_1 \dots s_i t_{i+1} \dots t_{u+v}}{\prod_{j\in L} t_j} = \frac{\prod_{i=1}^{u} r_i \prod_{j=1}^{v} s_j}{P(n) \prod_{j\in L} t_j}. \tag{3}$$

We will show that (i) both sides of (3) are polynomials in n, and (ii) equating coefficients of like powers of n in these polynomials yields a system of linear equations for the c_i's with a coefficient matrix to which the Corollary to Theorem 2 applies (and which is therefore invertible).

Consider the right side of (3). All the factors in $P(n)$ appear in its numerator by definition of u and v. For $j \in L$, we have $t_j \,|\, s_{\phi(j)}$. If $\phi(j) \le v$, then $s_{\phi(j)}$ is present in the numerator. If on the other hand $\phi(j) > v$, we claim: t_j also divides some r_i with $1 \le i \le u$. In fact, $i = u+\phi(j)-j$ works. First, $i \ge 1$ since $i > u+v-j \ge 0$ and $i \le u$ since $\phi(j) \le j$. Second, $t_j \mid s_{\phi(j)}$ implies $t_j \,|\, t_j + s_{\phi(j)} = \big(cn+d-(u+v)+j\big) + \big(en+f+v+1-\phi(j)\big) = an+b+1-u+j-\phi(j) = an+b+1-i = r_i$. Hence the claim. Thus every factor in the denominator divides a factor in the numerator. And if a factor in $P(n)$ also appears among $\{t_j\}_{j\in L}$, then by hypothesis it appears twice in N and hence appears twice in the numerator. So the right side of (3) is indeed a polynomial $P_{\text{rhs}}(n)$ and its degree is $u+v-\deg P - |L| = |K| - \deg P$.

As for the left side of (3), it is clearly a polynomial if $L = \emptyset$. Otherwise, since $K = [1, u+v]\backslash L$ and L consists of consecutive integers in $[1, u+v]$, K may be written as a disjoint union of intervals $K_s \cup K_b$ (K_s for the smaller numbers, here one of K_s, K_b may be empty). For $i \in K_s$, summand i is $c_i(\prod_{j=1}^{i} s_j)(\prod_{k\in K, k>i} t_k)$. Now suppose $i \in K_b$. Since $t_j \mid s_{\phi(j)}$ for $j \in L$ and $\phi(j) \le j \le \max L < i$, each t in the denominator of summand i divides some s in the numerator, leaving a quotient $q := e/c$ (e and c being the coefficients of n in the s's and t's respectively) Hence the left side of (3) is the polynomial $P_{\text{lhs}}(n) =$

$$c_0 \prod_{j\in K} t_j + \sum_{i\in K_s} c_i \left(\prod_{j=1}^{i} s_j \prod_{k\in K, k>i} t_k \right) + \sum_{i\in K_b} c_i q^{|L|} \left(\prod_{j\in[1,i], j\notin \operatorname{rng}\phi} s_j \times t_{i+1}\dots t_{u+v} \right) \tag{4}$$

and its degree is $|K|$.

Equating coefficients of powers of n in these polynomials gives a linear system of equations for the linear combination coefficients c_i. To apply Corollary 3 to the coefficient matrix of this system, arrange the factors in the products occurring in $P_{\text{lhs}}(n)$ into a (block) matrix $N =$

$$\begin{array}{c c} & \begin{array}{cc} K_s & K_b \end{array} \\ \begin{array}{c} K_s \cup \{0\} \\ K_b \end{array} & \begin{pmatrix} N_1 & N_2 \\ N_3 & N_4 \end{pmatrix} \end{array}$$

with rows and columns indexed as indicated. For blocks N_1 and N_4, the ij entry is t_j if $i < j$ and s_j if $i \ge j$. For N_2, the ij entry is t_j for all i. For N_3, each row is $(s_j)_{j\in K_s\cup L, j\notin\phi(L)}$ (order immaterial). Thus, in matrix terms, $P_{\text{lhs}}(n) = \mathbf{c}N^\circ$ where $\mathbf{c} = \big(c_0, (c_i)_{i\in K_s}, (q^{|L|}c_i)_{i\in K_b}\big)$ incorporates the $q^{|L|}$ factors.

Now equate coefficients of powers of n in $P_{\text{lhs}}(n) = P_{\text{rhs}}(n)$, that is, in $\mathbf{c}N^\circ = P_{\text{rhs}}(n)$, by applying the coeff operator of §3, to obtain

$$\mathbf{c}\ \text{coeff}(N^\circ) = \text{coeff}(P_{\text{rhs}}(n)).$$

This is a linear system of $|K|+1$ equations in the $|K|+1$ unknowns $\mathbf{c}$. The coefficient matrix $M = \text{coeff}(N^\circ)$ is invertible because Corollary 3 applies to N. The hypothesis of the Corollary is met because, for all $j \in K = K_s \cup K_b$, all entries of N directly above position (j,j) are equal to t_j, and all entries at or to its left are of the form s_i with $i \le j$. And t_j does not divide any such s_i or else j would lie in L whereas by definition of K, j does not lie in L.

To illustrate, for $\frac{1}{(6n+14)(4n+13)}\binom{6n+15}{2n+8}$, we have $u = 2$, $v = 6$, $r_i = 6n+16-i$, $s_i = 4n+14-i$, $t_i = 2n+i$. Since $t_3 \mid s_8$, $t_4 \mid s_6$, $t_5 \mid s_4$ and $t_6 \mid s_2$, we have $L = \{5,6\}$ with $\phi(5) = 4$, $\phi(6) = 2$. This makes $K_s = [1,4]$ and $K_b = [7,8]$. The common factor in (2) is $\binom{6n+13}{2n}/((2n+1)(2n+2)(2n+3)(2n+4)(2n+7)(2n+8))$. After dividing this out, the polynomial remaining on the right side is $2^2(6n+15)(4n+8)(4n+9)(4n+11)$ while that on the left is $(c_0, c_1, c_2, c_3, c_4, 4c_7, 4c_8)N^\circ$ where $N =$

$$\begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 7 \\ 8 \end{array}
\begin{array}{c}
\begin{array}{cccccc} 1 & 2 & 3 & 4 & 7 & 8 \end{array} \\
\begin{pmatrix}
2n+1 & 2n+2 & 2n+3 & 2n+4 & 2n+7 & 2n+8 \\
4n+13 & 2n+2 & 2n+3 & 2n+4 & 2n+7 & 2n+8 \\
4n+13 & 4n+12 & 2n+3 & 2n+4 & 2n+7 & 2n+8 \\
4n+13 & 4n+12 & 4n+11 & 2n+4 & 2n+7 & 2n+8 \\
4n+13 & 4n+12 & 4n+11 & 4n+10 & 2n+7 & 2n+8 \\
4n+13 & 4n+8 & 4n+11 & 4n+9 & 4n+7 & 2n+8 \\
4n+13 & 4n+8 & 4n+11 & 4n+9 & 4n+7 & 4n+6
\end{pmatrix}
\end{array}$$

5 Concluding Remarks

Theorem 1 enables one to tell by inspection if a linear binomial sequence $\frac{1}{P(n)}\binom{an+b}{cn+d}$ has bounded denominators. The theorem readily extends to sequences of the form $\frac{Q(n)}{P(n)}\binom{an+b}{cn+d}$ where both P and Q have linear factors. Indeed, if $gn+h$ is a factor in $P(n)$ with g, h relatively prime, and $g'n+h'$ is a factor in $Q(n)$, then the prime values of $gn+h$ can divide $g'n+h'$ for only finitely many values of n unless $gn+h$ divides $g'n+h'$ (as polynomials in n over $\mathbb{Q}$), in which case they can be cancelled. Thus the criterion of Theorem 1 also applies to $\frac{Q(n)}{P(n)}\binom{an+b}{cn+d}$.

The algorithm of Theorem 1 often yields the "smallest" sequence of integers among all multiples of the original sequence that are integral. But it does not always do so. It does not necessarily even yield the smallest sequence expressible as an integral linear combination of binomials. For example, $\binom{5n}{2n}$ will be returned unchanged whereas $\frac{1}{5}\binom{5n}{2n} = \binom{5n-1}{2n} - \binom{5n-1}{2n-1}$. Here is another phenomenon: $\binom{4n}{2n-1}$ is also returned unchanged while

$$\frac{1}{8}\binom{4n}{2n-1} = n^3\binom{4n-1}{2n-1} - n^3\binom{4n-1}{2n-2} - (4n-1)(4n-3)\binom{4n-5}{2n-3}$$

is clearly a sequence of integers. We conjecture that every such rational multiple of a linear binomial that yields a sequence of integers is similarly expressible as a linear combination of binomial coefficients with polynomial coefficients in $\mathbb{Z}[n]$. It would be interesting to characterize those cases where the coefficients can be taken to be constants, to extend the algorithm of Theorem 1 to sums $\sum_i \frac{P_i(n)}{Q_i(n)}\binom{a_i n+b_i}{c_i n+d_i}$, and to sharpen it to yield "smallest" sequences.

Acknowledgment

I would like to thank George Gilbert for conjecturing Theorem 1 in response to a problem proposal I submitted to *Mathematics Magazine*.

References

[1] Tom M. Apostol, *Introduction to Analytic Number Theory*, Springer-Verlag, 1976

[2] Graham, Knuth, Patashnik, *Concrete Mathematics* (2nd edition), Addison-Wesley, 1989.

[3] N. J. A. Sloane and Simon Plouffe, *The Encyclopedia of Integer Sequences*, Academic Press, 1995.

AMS Classification Number: 11B65