MULTIVARIATE APÉRY NUMBERS AND SUPERCONGRUENCES OF RATIONAL FUNCTIONS

ARMIN STRAUB

Department of Mathematics, University of Illinois at Urbana-Champaign
&
Max-Planck-Institut für Mathematik

Abstract. One of the many remarkable properties of the Apéry numbers $A(n)$, introduced in Apéry's proof of the irrationality of $\zeta(3)$, is that they satisfy the two-term supercongruences

$$A(p^r m) \equiv A(p^{r-1} m) \pmod{p^{3r}}$$

for primes $p \geqslant 5$. Similar congruences are conjectured to hold for all Apéry-like sequences. We provide a fresh perspective on the supercongruences satisfied by the Apéry numbers by showing that they extend to all Taylor coefficients $A(n_1, n_2, n_3, n_4)$ of the rational function

$$\frac{1}{(1 - x_1 - x_2)(1 - x_3 - x_4) - x_1 x_2 x_3 x_4}.$$

The Apéry numbers are the diagonal coefficients of this function, which is simpler than previously known rational functions with this property.

Our main result offers analogous results for an infinite family of sequences, indexed by partitions λ, which also includes the Franel and Yang–Zudilin numbers as well as the Apéry numbers corresponding to $\zeta(2)$. Using the example of the Almkvist–Zudilin numbers, we further indicate evidence of multivariate supercongruences for other Apéry-like sequences.

1. Introduction

The *Apéry numbers*

$$A(n) = \sum_{k=0}^{n} \binom{n}{k}^2 \binom{n+k}{k}^2 \tag{1}$$

played a crucial role in R. Apéry's proof [Apé79], [Poo79] of the irrationality of $\zeta(3)$ and have inspired much further work. Among many other interesting properties, they satisfy congruences with surprisingly large moduli, referred to as *supercongruences*, a term coined by F. Beukers [Beu85]. For instance, for all primes $p \geqslant 5$ and all positive integers r,

$$A(p^r m) \equiv A(p^{r-1} m) \pmod{p^{3r}}. \tag{2}$$

E-mail address: astraub@illinois.edu.

Date: October 22, 2014.

2010 *Mathematics Subject Classification.* Primary 11A07, 11B83; Secondary 11B37.

Key words and phrases. Apéry numbers, supercongruences, diagonals of rational functions, Almkvist–Zudilin numbers.

The special case $m = 1$, $r = 1$ was conjectured by S. Chowla, J. Cowles and M. Cowles [CCC80], who established the corresponding congruence modulo p^2. The case $r = 1$ was subsequently shown by I. Gessel [Ges82] and Y. Mimura [Mim83], while the general case has been proved by M. Coster [Cos88]. The proof is an adaption of F. Beukers' proof [Beu85] of the related congruence

$$A(p^r m - 1) \equiv A(p^{r-1} m - 1) \pmod{p^{3r}}, \tag{3}$$

again valid for all primes $p \geqslant 5$ and all positive integers r. That congruence (3) can be interpreted as an extension of (2) to negative integers is explained in Remark 1.3. For further congruence properties of the Apéry numbers we refer to [Cow80], [Beu87], [AO00], [Kil06].

Given a series

$$F(x_1, \ldots, x_d) = \sum_{n_1, \ldots, n_d \geqslant 0} a(n_1, \ldots, n_d) x_1^{n_1} \cdots x_d^{n_d}, \tag{4}$$

its *diagonal coefficients* are the coefficients $a(n, \ldots, n)$ and the *diagonal* is the ordinary generating function of the diagonal coefficients. For our purposes, F will always be a rational function. It is well-known, see for instance [LvdP90, Theorem 5.2], that the diagonal of a rational function satisfies a Picard–Fuchs linear differential equation and as such "comes from geometry". In particular, the diagonal coefficients satisfy a linear recurrence with polynomial coefficients.

Many sequences of number-theoretic interest can be represented as the diagonal coefficients of rational functions. In particular, it is known [Chr84], [LvdP90] that the Apéry numbers are the diagonal coefficients of the rational function

$$\frac{1}{(1 - x_1)[(1 - x_2)(1 - x_3)(1 - x_4)(1 - x_5) - x_1 x_2 x_3]}. \tag{5}$$

Several other rational functions of which the Apéry numbers are the diagonal coefficients are given in [BBC+13], where it is also discussed how these can be obtained from the representation of the Apéry numbers as the binomial sum (1). However, all of these rational function involve at least five variables and, in each case, the polynomial in the denominator factors. Our first result shows that, in fact, the Apéry numbers are the diagonal coefficients of a simpler rational function in only four variables.

Theorem 1.1. *The Apéry numbers $A(n)$, defined in (1), are the diagonal coefficients of*

$$\frac{1}{(1 - x_1 - x_2)(1 - x_3 - x_4) - x_1 x_2 x_3 x_4}. \tag{6}$$

Representing a sequence as the diagonal of a rational function has certain benefits. For instance, asymptotic results can be obtained directly and explicitly from this rational function. This is the subject of *multivariate asymptotics*, as developed in [PW02]. For details and a host of worked examples we refer to [PW08]. As a second example, the rational generating function provides a means to compute the sequence modulo a fixed prime power. Indeed, the diagonal of a rational function with integral Taylor coefficients, such as (6), is algebraic modulo p^α for any α [LvdP90]. A recent demonstration that this can be done very constructively is given in [RY13], where the values modulo p^α of sequences such as the Apéry numbers are, equivalently, encoded as finite automata.

We note that a statement such as Theorem 1.1 is more or less automatic to prove once discovered. For instance, given a rational function, we can always repeatedly employ a binomial series expansion to represent the Taylor coefficients as a nested sum of hypergeometric terms. In principle, creative telescoping [PWZ96] will then obtain a linear recurrence satisfied by the diagonal coefficients, in which case it suffices to check that the alternative expression satisfies the same recurrence and agrees for sufficiently many initial values.

For the rational function $F(\boldsymbol{x})$ given in (6), we can gain considerably more insight. Indeed, for all the Taylor coefficients $A(\boldsymbol{n})$, defined by

$$F(x_1, x_2, x_3, x_4) = \sum_{n_1, n_2, n_3, n_4 \geqslant 0} A(n_1, n_2, n_2, n_4) x_1^{n_1} x_2^{n_2} x_3^{n_3} x_4^{n_4}, \tag{7}$$

we find, for instance by applying MacMahon's Master Theorem [Mac15] as detailed in Section 4, the explicit formula

$$A(\boldsymbol{n}) = \sum_{k \in \mathbb{Z}} \binom{n_1}{k} \binom{n_3}{k} \binom{n_1 + n_2 - k}{n_1} \binom{n_3 + n_4 - k}{n_3}, \tag{8}$$

of which Theorem 1.1 is an immediate consequence.

An instance of our main result is the observation that the supercongruence (2) for the Apéry numbers generalizes to all coefficients (8) of the rational function (6) in the following sense.

Theorem 1.2. *Let $\boldsymbol{n} = (n_1, n_2, n_3, n_4) \in \mathbb{Z}^4$. The coefficients $A(\boldsymbol{n})$, defined in (7) and extended to negative integers by (8), satisfy, for primes $p \geqslant 5$ and positive integers r, the supercongruences*

$$A(p^r \boldsymbol{n}) \equiv A(p^{r-1} \boldsymbol{n}) \pmod{p^{3r}}. \tag{9}$$

Note that the Apéry numbers are $A(n) = A(n, n, n, n)$ so that (9) indeed generalizes (2). Our reason for allowing negative entries in $\boldsymbol{n}$ is that by doing so, we also generalize Beukers' supercongruence (3). Indeed, as explained in Remark 1.3 below, $A(n-1) = A(-n, -n, -n, -n)$. Theorem 1.2 is a special case of our main result, Theorem 3.2, in which we prove such supercongruences for an infinite family of sequences. This family includes other Apéry-like sequences such as the Franel and Yang–Zudilin numbers as well as the Apéry numbers corresponding to $\zeta(2)$.

We therefore review Apéry-like sequences in Section 2. Though no uniform reason is known, each Apéry-like sequence appears to satisfy a supercongruence of the form (2), some of which have been proved [Beu85], [Cos88], [CCS10], [OS11], [OS13], [OSS14] while others remain open [OSS14]. A major motivation for this note is to work towards an understanding of this observation. Our contribution to this question is the insight that, at least for several Apéry-like sequences, these supercongruences generalize to all coefficients of a rational function. Our main result, which includes the case of the Apéry numbers outlined in this introduction, is given in Section 3. In that section, we also record two further conjectural instances of this phenomenon. Finally, we provide proofs for our results in Sections 4 and 5.

Remark 1.3. Let us indicate that congruence (3) can be interpreted as the natural extension of (2) to the case of negative integers m. To see this, generalize the definition (1) of the Apéry numbers $A(n)$ to all integers n by setting

$$A(n) = \sum_{k \in \mathbb{Z}} \binom{n}{k}^2 \binom{n+k}{k}^2. \tag{10}$$

Here, we assume the values of the binomial coefficients to be defined as the (limiting) values of the corresponding quotient of gamma functions, that is,

$$\binom{n}{k} = \lim_{z\to 0} \frac{\Gamma(z+n+1)}{\Gamma(z+k+1)\Gamma(z+n-k+1)}.$$

Since $\Gamma(z+1)$ has no zeros, and poles only at negative integers z, one observes that the binomial coefficient $\binom{n}{k}$ is finite for all integers n and k. Moreover, the binomial coefficient with integer entries is nonzero only if either $k \geqslant 0$ and $n-k \geqslant 0$, or if $n < 0$ and $k \geqslant 0$, or if $n < 0$ and $n-k \geqslant 0$. Note that in each of these cases $k \geqslant 0$ or $n-k \geqslant 0$, so that the symmetry $\binom{n}{k} = \binom{n}{n-k}$ allows us to compute these binomial coefficients in the obvious way. For instance, $\binom{-3}{-5} = \binom{-3}{2} = \frac{(-3)(-4)}{2!} = 6$. As carefully shown in [Spr08], for all integers n and k, we have the negation rule

$$\binom{n}{k} = \operatorname{sgn}(k)(-1)^k \binom{-n+k-1}{k}, \tag{11}$$

where $\operatorname{sgn}(k) = 1$ for $k \geqslant 0$ and $\operatorname{sgn}(k) = -1$ for $k < 0$. Applying (11) to the sum (10), we find that

$$A(-n) = A(n-1).$$

In particular, the congruence (3) is equivalent to (2) with $-m$ in place of m.

Remark 1.4. The proof of formula (8) in Section 4 shows that the coefficients can be expressed as

$$A(n_1,n_2,n_3,n_4) = \operatorname{ct} \frac{(x_1+x_2+x_3)^{n_1}(x_1+x_2)^{n_2}(x_3+x_4)^{n_3}(x_2+x_3+x_4)^{n_4}}{x_1^{n_1}x_2^{n_2}x_3^{n_3}x_4^{n_4}},$$

representing them as the constant terms of Laurent polynomials. In particular, the Apéry numbers (1) are the constant term of powers of a Laurent polynomial. Namely,

$$A(n) = \operatorname{ct} \left[\frac{(x_1+x_2)(x_3+1)(x_1+x_2+x_3)(x_2+x_3+1)}{x_1x_2x_3} \right]^n.$$

Since the Newton polyhedron of this Laurent polynomial has the origin as its only interior integral point, the results of [SvS09], [MV13] apply to show that $A(n)$ satisfies the Dwork congruences

$$A(p^r m + n)A(\lfloor n/p \rfloor) \equiv A(p^{r-1}m + \lfloor n/p \rfloor)A(n) \pmod{p^r}$$

for all primes p and all integers $m, n \geqslant 0$, $r \geqslant 1$. In particular,

$$A(p^r m) \equiv A(p^{r-1}m) \pmod{p^r}, \tag{12}$$

which is a weaker version of (2) that holds for the large class of sequences represented as the constant term of powers of a Laurent polynomial, subject only to the condition on the Newton polyhedron. This gives another indication why congruence (2) is referred to as a supercongruence. It would be of considerable interest to find similarly well-defined classes of sequences for which supercongruences, of the form (12) but modulo p^{kr} for $k > 1$, hold. Let us note that the case $r = 1$ of the Dwork congruences implies the Lucas congruences

$$A(n) \equiv A(n_0)A(n_1)\cdots A(n_\ell) \pmod{p},$$

where $n_0, \ldots, n_\ell \in \{0, 1, \ldots, p-1\}$ are the p-adic digits of $n = n_0 + n_1 p + \ldots + n_\ell p^\ell$. It is shown in [RY13] that Lucas congruences hold for all Taylor coefficients of certain rational functions. Additional divisibility properties in this direction are obtained in [Del13] for Apéry-like numbers as well as for constant terms of powers of certain Laurent polynomials. Finally, we note that an extension of Dwork congruences to the multivariate setting has been considered in [KR11]. In contrast to our approach, where, for instance, the Apéry numbers appear as the diagonal (multivariate) Taylor coefficients of a multivariate function $F(\boldsymbol{x})$, the theory developed in [KR11] is concerned with functions $G(\boldsymbol{x}) = G(x_1, \ldots, x_d)$ for which, say, the Apéry numbers are the (univariate) Taylor coefficients of the specialization $G(x, \ldots, x)$.

2. Review of Apéry-like numbers

The Apéry numbers $A(n)$ are characterized by the 3-term recurrence

$$(n+1)^3 u_{n+1} = (2n+1)(an^2 + an + b)u_n - n(cn^2 + d)u_{n-1}, \tag{13}$$

where $(a, b, c, d) = (17, 5, 1, 0)$, together with the initial conditions

$$u_{-1} = 0, \quad u_0 = 1. \tag{14}$$

As explained in [Beu02], the fact, that in the recursion (13) we divide by $(n+1)^3$ at each step, means that we should expect the denominator of u_n to grow like $(n!)^3$. While this is what happens for generic choice of the parameters (a, b, c, d), the Apéry numbers have the, from this perspective, exceptional property of being integral. Initiated by Beukers [Beu02], systematic searches have therefore been conducted for recurrences of this kind, which share the property of having an integer solution with initial conditions (14). This was done by D. Zagier [Zag09] for recurrences of the form

$$(n+1)^2 u_{n+1} = (an^2 + an + b)u_n - cn^2 u_{n-1}, \tag{15}$$

by G. Almkvist and W. Zudilin [AZ06] for recurrences of the form (13) with $d = 0$ and, more recently, by S. Cooper [Coo12] for recurrences of the form (13). In each case, apart from degenerate cases, only finitely many sequences have been discovered. For details and a possibly complete list of the sequences, we refer to [Zag09], [AZ06], [AvSZ11], [Coo12].

Remarkably, and still rather mysteriously, all of these sequences, often referred to as *Apéry-like*, share some of the interesting properties of the Apéry numbers. For instance, they all are the coefficients of modular forms expanded in terms of a corresponding modular function. In the case of the Apéry numbers $A(n)$, for instance, it was shown by Beukers [Beu87] that

$$\sum_{n \geqslant 0} A(n) \left(\frac{\eta(\tau)\eta(6\tau)}{\eta(2\tau)\eta(3\tau)} \right)^{12n} = \frac{\eta^7(2\tau)\eta^7(3\tau)}{\eta^5(\tau)\eta^5(6\tau)}, \tag{16}$$

where $\eta(\tau)$ is the Dedekind eta function $\eta(\tau) = e^{\pi i \tau/12} \prod_{n \geqslant 1}(1 - e^{2\pi i n \tau})$. The modular function and the modular form appearing in (16) are modular with respect to the congruence subgroup $\Gamma_0(6)$ of level 6 (in fact, they are modular with respect to a slightly larger group). While this relation with modular forms can be proven in each individual case, no conceptual explanation is available in the sense that if an additional Apéry-like sequence was found we would not know *a priori* that its generating function has a modular parametrization such as (16).

As a second example, it is conjectured and in some cases proven [OSS14] that each Apéry-like sequence satisfies a supercongruence of the form (2). Again, no uniform explanation is available and, the known proofs [Ges82], [Mim83], [Beu85], [Cos88] of the supercongruences (2) and (3) all rely on the explicit binomial representation (1) of the Apéry numbers. However, not all Apéry-like sequences have a comparably effective binomial representation so that, for instance, for the *Almkvist–Zudilin numbers* [AvSZ11, sequence (4.12) (δ)], [CZ10], [CCS10]

$$Z(n) = \sum_{k=0}^{n} (-3)^{n-3k} \binom{n}{3k} \binom{n+k}{n} \frac{(3k)!}{k!^3}, \tag{17}$$

which solve (13) with $(a, b, c, d) = (-7, -3, 81, 0)$, the supercongruence

$$Z(p^r m) \equiv Z(p^{r-1} m) \pmod{p^{3r}} \tag{18}$$

for primes $p \geqslant 3$ is conjectural only.

It would therefore be of particular interest to find alternative approaches to proving supercongruences. In this paper, we provide a new perspective on supercongruences of the form (18) by showing that they hold, at least for several Apéry-like sequences, for all coefficients $C(\boldsymbol{n})$ of a corresponding rational function, which has the sequence of interest as its diagonal coefficients. In such a case, one may then hope to use properties of the rational function to prove, for some $k > 1$, the supercongruence $C(p^r \boldsymbol{n}) \equiv C(p^{r-1} \boldsymbol{n})$ modulo p^{kr}. For instance, for fixed p^r, these congruences can be proved, at least in principle, by computing the multivariate generating functions of both $C(p^r \boldsymbol{n})$ and $C(p^{r-1} \boldsymbol{n})$, which are rational functions because they are multisections of a rational function, and comparing them modulo p^{kr}.

Let us note that, in Example 3.9 below, we give a characterization of the Almkvist–Zudilin numbers (17) as the diagonal of a surprisingly simple rational function and conjecture that the supercongruences (18), which themselves have not been proved yet, again extend to all coefficients of this rational function. We hope that the simplicity of the rational function might help inspire a proof of these supercongruences.

3. Main result and examples

We now generalize what we have illustrated in the introduction for the Apéry numbers $A(n)$ to an infinite family of sequences $A_{\lambda,\varepsilon}(n)$, indexed by partitions λ and $\varepsilon \in \{-1, 1\}$, which includes other Apéry-like numbers such as the Franel and Yang–Zudilin numbers as well as the sequence used by Apéry in relation with $\zeta(2)$. Our main theorem is Theorem 3.2, in which we prove (multivariate) supercongruences for this family of sequences, thus unifying and extending a number of known supercongruences. To begin with, the sequences we are concerned with are introduced by the following extension of formula (8). Here, $\boldsymbol{x}^{\boldsymbol{n}}$ is short for $x_1^{n_1} x_2^{n_2} \cdots x_d^{n_d}$.

Theorem 3.1. *Let $\alpha \in \mathbb{C}$ and $\lambda = (\lambda_1, \ldots, \lambda_\ell) \in \mathbb{Z}_{>0}^{\ell}$ with $d = \lambda_1 + \ldots + \lambda_\ell$, and set $s(j) = \lambda_1 + \ldots + \lambda_{j-1}$. Then the Taylor coefficients of the rational function*

$$\left(\prod_{j=1}^{\ell} \left[1 - \sum_{r=1}^{\lambda_j} x_{s(j)+r} \right] - \alpha x_1 x_2 \cdots x_d \right)^{-1} = \sum_{\boldsymbol{n} \in \mathbb{Z}_{\geqslant 0}^d} A_{\lambda,\alpha}(\boldsymbol{n}) \boldsymbol{x}^{\boldsymbol{n}} \tag{19}$$

are given by

$$A_{\lambda,\alpha}(\boldsymbol{n}) = \sum_{k\in\mathbb{Z}} \alpha^k \prod_{j=1}^{\ell} \binom{n_{s(j)+1} + \ldots + n_{s(j)+\lambda_j} - (\lambda_j - 1)k}{n_{s(j)+1} - k, \ldots, n_{s(j)+\lambda_j} - k, k}. \tag{20}$$

The proof of this elementary but crucial result will be given in Section 4. Observe that the multivariate Apéry numbers $A(\boldsymbol{n})$, defined in (8), are the special case $A_{(2,2),1}(\boldsymbol{n})$.

Our main result, of which Theorem 1.2 is the special case $\lambda = (2,2)$ and $\varepsilon = 1$, follows next. Note that, if $\boldsymbol{n} \in \mathbb{Z}^d_{\geqslant 0}$, then the sum (20) defining $A_{\lambda,\alpha}(\boldsymbol{n})$ is finite and runs over $k = 0, 1, \ldots, \min(n_1, \ldots, n_d)$. On the other hand, if $\max(\lambda_1, \ldots, \lambda_\ell) \geqslant 2$, then $A_{\lambda,\alpha}(\boldsymbol{n})$ is finite for any $\boldsymbol{n} \in \mathbb{Z}^d$.

Theorem 3.2. *Let $\varepsilon \in \{-1, 1\}$, $\lambda = (\lambda_1, \ldots, \lambda_\ell) \in \mathbb{Z}^\ell_{>0}$ and assume that $\boldsymbol{n} \in \mathbb{Z}^d$, $d = \lambda_1 + \ldots + \lambda_\ell$, is such that $A_{\lambda,\varepsilon}(\boldsymbol{n})$, as defined in (20), is finite.*

(a) If $\ell \geqslant 2$, then, for all primes $p \geqslant 3$ and integers $r \geqslant 1$,

$$A_{\lambda,\varepsilon}(p^r\boldsymbol{n}) \equiv A_{\lambda,\varepsilon}(p^{r-1}\boldsymbol{n}) \pmod{p^{2r}}. \tag{21}$$

If $\varepsilon = 1$, then these congruences also hold for $p = 2$.

(b) If $\ell \geqslant 2$ and $\max(\lambda_1, \ldots, \lambda_\ell) \leqslant 2$, then, for primes $p \geqslant 5$ and integers $r \geqslant 1$,

$$A_{\lambda,\varepsilon}(p^r\boldsymbol{n}) \equiv A_{\lambda,\varepsilon}(p^{r-1}\boldsymbol{n}) \pmod{p^{3r}}. \tag{22}$$

A proof of Theorem 3.2 is given in Section 5. One of the novel features of the proof, which is based on the approach of Gessel [Ges82] and Beukers [Beu85], is that it proceeds in a uniform fashion for all $\boldsymbol{n} \in \mathbb{Z}^d$. As outlined in Remark 1.3, this allows us to also conclude, and to a certain extent explain, the shifted supercongruences (3), which, among Apéry-like numbers, are special to the Apéry numbers as well as their version (23) related to $\zeta(2)$. In cases where $\boldsymbol{n}$ has negative entries, the summation (20), while still finite, may include negative values for k (see Remark 1.3). We therefore extend classical results, such as Jacobsthal's binomial congruences, to the case of binomial coefficients with negative entries.

Example 3.3. For $\lambda = (2)$, the numbers (20) specialize to the *Delannoy numbers*

$$A_{(2),1}(\boldsymbol{n}) = \sum_{k\in\mathbb{Z}} \binom{n_1}{k}\binom{n_1 + n_2 - k}{n_1},$$

which, for $n_1, n_2 \geqslant 0$, count the number of lattice paths from $(0,0)$ to (n_1, n_2) with steps $(1,0)$, $(0,1)$ and $(1,1)$. The Delannoy numbers do not satisfy (21) or (22), thus demonstrating the necessity of the condition $\ell \geqslant 2$ in Theorem 3.2. They do satisfy (21) modulo p^r, by virtue of Remark 1.4.

Example 3.4. The Apéry-like sequence

$$B(n) = \sum_{k\in\mathbb{Z}} \binom{n}{k}^2 \binom{n+k}{k}, \tag{23}$$

which satisfies recurrence (15) with $(a,b,c) = (11,3,-1)$, was introduced by Apéry [Apé79], [Poo79] along with (1) and used to (re)prove the irrationality of $\zeta(2)$.

By Theorem 3.1 with $\lambda = (2,1)$ and $\varepsilon = 1$, the numbers $B(n)$ are the diagonal coefficients of the rational function

$$\frac{1}{(1-x_1-x_2)(1-x_3)-x_1x_2x_3} = \sum_{\boldsymbol{n}\in\mathbb{Z}^3_{\geqslant 0}} B(\boldsymbol{n})\boldsymbol{x}^{\boldsymbol{n}}. \tag{24}$$

In addition to the binomial sum for $B(\boldsymbol{n})$ provided by Theorem 3.1, MacMahon's Master Theorem 4.1 shows that $B(n_1,n_2,n_3)$ is the coefficient of $x_1^{n_1}x_2^{n_2}x_3^{n_3}$ in the product $(x_1+x_2+x_3)^{n_1}(x_1+x_2)^{n_2}(x_2+x_3)^{n_3}$. An application of Theorem 3.2 shows that, for $\boldsymbol{n}\in\mathbb{Z}^3$ and integers $r\geqslant 1$, the supercongruences

$$B(p^r\boldsymbol{n}) \equiv B(p^{r-1}\boldsymbol{n}) \pmod{p^{3r}} \tag{25}$$

hold for all primes $p\geqslant 5$. In the diagonal case $n_1=n_2=n_3$, this result was first proved by Coster [Cos88].

Proceeding as in Remark 1.3, and using the curious identity

$$\sum_{k=0}^{n}\binom{n}{k}^2\binom{n+k}{k} = \sum_{k=0}^{n}(-1)^{n+k}\binom{n}{k}\binom{n+k}{k}^2, \tag{26}$$

we find that $B(-n) = (-1)^{n-1}B(n-1)$ for $n>0$. Consequently, (25) implies the shifted supercongruences $B(p^rm-1)\equiv B(p^{r-1}m-1)$, which hold modulo p^{3r} for all primes $p\geqslant 5$ and were first proved in [Beu85], along with (3). We observe that, among the known Apéry-like numbers, the sequence $B(n)$ and the Apéry numbers (1) are the only ones to satisfy shifted supercongruences of the form (3) in addition to the supercongruences of the form (2).

Example 3.5. As a consequence of Theorem 3.1 with $\lambda=(3,1)$ and $\varepsilon=1$, the numbers

$$C(n) = \sum_{k=0}^{n}\binom{n}{k}^2\binom{n+k}{k}\binom{n+2k}{k}$$

are the diagonal coefficients of the rational function $1/((1-x_1-x_2-x_3)(1-x_4)-x_1x_2x_3x_4)$. By Theorem 3.2, it follows that $C(p^rn)\equiv C(p^{r-1}n)$ modulo p^{2r}, for all primes p. We note that this congruence does not, in general, hold modulo a larger power of p, as is illustrated by $C(5) = 4,009,657 \not\equiv 7 = C(1)$ modulo 5^3. This demonstrates that in Theorem 3.2(a) the modulus p^{2r} of the congruences cannot, in general, be replaced with p^{3r}, even for $p\geqslant 5$.

Example 3.6. Next, we consider the sequences

$$Y_d(n) = \sum_{k=0}^{n}\binom{n}{k}^d. \tag{27}$$

The numbers $Y_3(n)$ satisfy the recurrence (15) with $(a,b,c)=(7,2,-8)$ and are known as *Franel numbers* [Fra94], while the numbers $Y_4(n)$, corresponding to $(a,b,c,d)=(6,2,-64,4)$ in (13), are sometimes referred to as *Yang–Zudilin numbers* [CCS10]. It follows from Theorem 3.1 with $\lambda=(1,1,\ldots,1)$ and $\varepsilon=1$, that

$$\frac{1}{(1-x_1)(1-x_2)\cdots(1-x_d)-x_1x_2\cdots x_d} = \sum_{\boldsymbol{n}\in\mathbb{Z}^d_{\geqslant 0}} Y_d(\boldsymbol{n})\boldsymbol{x}^{\boldsymbol{n}}, \tag{28}$$

where

$$Y_d(\boldsymbol{n}) = \sum_{k\geqslant 0} \binom{n_1}{k}\binom{n_2}{k}\cdots\binom{n_d}{k}. \tag{29}$$

It is proved in [CCS10] that $Y_d(pn) \equiv Y_d(n)$ modulo p^3 for primes $p \geqslant 5$ if $d \geqslant 2$. These congruences are generalized to the multivariate setting by Theorem 3.2, which shows that, if $d \geqslant 2$, then, for $\boldsymbol{n} \in \mathbb{Z}^d_{\geqslant 0}$ and integers $r \geqslant 1$,

$$Y_d(p^r\boldsymbol{n}) \equiv Y_d(p^{r-1}\boldsymbol{n}) \pmod{p^{3r}} \tag{30}$$

for primes $p \geqslant 5$. Note that

$$Y_2(\boldsymbol{n}) = \sum_{k\in\mathbb{Z}} \binom{n_1}{k}\binom{n_2}{k} = \binom{n_1+n_2}{n_1}.$$

Hence, congruence (30) includes, in particular, the appealing binomial congruence

$$\binom{pa}{pb} \equiv \binom{a}{b} \pmod{p^3},$$

which is attributed to W. Ljunggren [Gra97] and which generalizes the classical congruences by C. Babbage, J. Wolstenholme and J. W. L. Glaisher. It is further refined by E. Jacobsthal's binomial congruence, which we review in Lemma 5.1 and which the proof of Theorem 3.2 crucially depends on.

Let us conclude this section with two conjectural examples, which suggest that our results are not an isolated phenomenon.

Example 3.7. As noted in the introduction for the Apéry numbers, there is no unique rational function of which a given sequence is the diagonal. For instance, the Franel numbers $Y_3(n)$ are also the diagonal coefficients of the rational function

$$\frac{1}{1-(x_1+x_2+x_3)+4x_1x_2x_3}. \tag{31}$$

A rational function $F(\boldsymbol{x})$ is said to be *positive* if its Taylor coefficients (4) are all positive. The Askey–Gasper rational function (31), whose positivity is proved in [AG77] and [GRZ83], is an interesting instance of a rational function on the boundary of positivity (if the 4 is replaced by $4+\varepsilon$, for any $\varepsilon > 0$, then the resulting rational function is not positive). The present work was, in part, motivated by the observation [SZ14] that for several of the rational functions, which have been shown or conjectured to be on the boundary of positivity, the diagonal coefficients are arithmetically interesting sequences with links to modular forms. Note that the Askey–Gasper rational function (31) corresponds to the choice $\lambda = (3)$ and $\alpha = -4$ in Theorem 3.1, which makes its Taylor coefficients $G(\boldsymbol{n}) = A_{(3),-4}(\boldsymbol{n})$ explicit. We also note that an application of MacMahon's Master Theorem 4.1 shows that $G(n_1, n_2, n_3)$ is the coefficient of $x_1^{n_1}x_2^{n_2}x_3^{n_3}$ in the product $(x_1 - x_2 - x_3)^{n_1}(x_2 - x_1 - x_3)^{n_2}(x_3 - x_1 - x_2)^{n_3}$. Although it is unclear how one might adjust the proof of Theorem 3.2, numerical evidence suggests that the coefficients $G(\boldsymbol{n})$ satisfy supercongruences modulo p^{3r} as well.

Conjecture 3.8. *The coefficients $G(\boldsymbol{n})$ of the rational function (31) satisfy, for primes $p \geqslant 5$ and integers $r \geqslant 1$,*

$$G(p^r\boldsymbol{n}) \equiv G(p^{r-1}\boldsymbol{n}) \pmod{p^{3r}}.$$

Example 3.9. Remarkably, the previous example has a four-variable analog, which involves the Almkvist–Zudilin numbers $Z(n)$, introduced in (17). Namely, the numbers $Z(n)$ are the diagonal coefficients of the unexpectedly simple rational function

$$\frac{1}{1-(x_1+x_2+x_3+x_4)+27x_1x_2x_3x_4}, \tag{32}$$

as can be deduced from Theorem 3.1 with $\lambda=(4)$ and $\alpha=-27$. Again, numerical evidence suggests that the coefficients $Z(\boldsymbol{n})$ of (32) satisfy supercongruences modulo p^{3r}. This is particularly interesting, since even the univariate congruences (18) are conjectural at this time.

Conjecture 3.10. *The coefficients $Z(\boldsymbol{n})$ of the rational function (32) satisfy, for primes $p \geqslant 5$ and integers $r \geqslant 1$,*

$$Z(p^r\boldsymbol{n}) \equiv Z(p^{r-1}\boldsymbol{n}) \pmod{p^{3r}}.$$

Remark 3.11. The rational functions (31) and (32) involved in the previous examples make it natural to wonder whether supercongruences might similarly exist for the family of rational functions given by

$$\frac{1}{1-(x_1+x_2+\ldots+x_d)+(d-1)^{d-1}x_1x_2\cdots x_d}.$$

This does not, however, appear to be the case for $d \geqslant 5$. In fact, no value $b \neq 0$ in

$$\frac{1}{1-(x_1+x_2+\ldots+x_d)+bx_1x_2\cdots x_d}$$

appears to give rise to supercongruences (by computing coefficients, we have ruled out supercongruences modulo p^{2r} for integers $|b|<100,000$ and $d \leqslant 25$).

4. The Taylor coefficients

This section is devoted to proving Theorem 3.1. Before we give a general proof, we offer an alternative approach based on MacMahon's Master Theorem, to which we refer at several occasions in this note and which offers additional insight into the Taylor coefficients by expressing them as coefficients of certain polynomials (see also Remark 1.4). This approach, which we apply here to prove formula (8), is based on the following result of P. MacMahon [Mac15], coined by himself "a master theorem in the Theory of Permutations". Here, $[\boldsymbol{x}^{\boldsymbol{m}}]$ denotes the coefficient of $x_1^{m_1}\cdots x_n^{m_n}$ in the expansion of what follows.

Theorem 4.1. *For $\boldsymbol{x}=(x_1,\ldots,x_n)$, matrices $A \in \mathbb{C}^{n\times n}$ and $\boldsymbol{m}=(m_1,\ldots,m_n) \in \mathbb{Z}_{\geqslant 0}^n$,*

$$[\boldsymbol{x}^{\boldsymbol{m}}]\prod_{i=1}^n\left(\sum_{j=1}^n A_{i,j}x_j\right)^{m_i} = [\boldsymbol{x}^{\boldsymbol{m}}]\frac{1}{\det(I_n-AX)},$$

where X is the diagonal $n\times n$ matrix with entries $x_1,\ldots,x_n$.

Proof of formula (8). We note that

$$\frac{1}{(1-x_1-x_2)(1-x_3-x_4)-x_1x_2x_3x_4} = \frac{1}{\det(I_4-MX)},$$

where M and X are the matrices

$$M = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{pmatrix}, \quad X = \begin{pmatrix} x_1 & & & \\ & x_2 & & \\ & & x_3 & \\ & & & x_4 \end{pmatrix}.$$

An application of MacMahon's Master Theorem 4.1 therefore shows that the coefficients $A(\boldsymbol{n})$, with $\boldsymbol{n} = (n_1, n_2, n_3, n_4)$, are given by

$$A(\boldsymbol{n}) = [\boldsymbol{x}^{\boldsymbol{n}}](x_1 + x_2 + x_3)^{n_1}(x_1 + x_2)^{n_2}(x_3 + x_4)^{n_3}(x_2 + x_3 + x_4)^{n_4}.$$

In order to extract the requisite coefficient, we expand the right-hand side as

$$\begin{aligned} & (x_1 + x_2 + x_3)^{n_1}(x_1 + x_2)^{n_2}(x_3 + x_4)^{n_3}(x_2 + x_3 + x_4)^{n_4} \\ = & \sum_{k_1, k_4} \binom{n_1}{k_1}\binom{n_4}{k_4} x_2^{n_4 - k_4} x_3^{n_1 - k_1}(x_1 + x_2)^{k_1 + n_2}(x_3 + x_4)^{n_3 + k_4} \\ = & \sum_{k_1, k_2, k_3, k_4} \binom{n_1}{k_1}\binom{n_4}{k_4}\binom{k_1 + n_2}{k_2}\binom{n_3 + k_4}{k_3} x_1^{k_1 + n_2 - k_2} x_2^{n_4 - k_4 + k_2} x_3^{n_1 - k_1 + k_3} x_4^{n_3 + k_4 - k_3}. \end{aligned}$$

The summand contributes to $x_1^{n_1}x_2^{n_2}x_3^{n_3}x_4^{n_4}$ if and only if $n_i - k_i = n_j - k_j$ for all $i, j = 1, \ldots, 4$. Writing $k = n_i - k_i$ for the common value, we obtain

$$A(n_1, n_2, n_3, n_4) = \sum_{k \in \mathbb{Z}} \binom{n_1}{k}\binom{n_4}{k}\binom{n_1 - k + n_2}{n_2 - k}\binom{n_3 + n_4 - k}{n_3 - k},$$

which is equivalent to the claimed (8). □

Proof of Theorem 3.1. Recall the elementary formula

$$\frac{1}{(1 - x)^{k+1}} = \sum_{n \geqslant 0} \binom{n + k}{k} x^n,$$

for integers $k \geqslant 0$. Combined with an application of the multinomial theorem, it implies that

$$\frac{1}{(1 - x_1 - \ldots - x_\rho)^{k+1}} = \sum_{n_1 \geqslant 0} \cdots \sum_{n_\rho \geqslant 0} \binom{n_1 + \ldots + n_\rho + k}{n_1, \ldots, n_\rho, k} x_1^{n_1} \cdots x_\rho^{n_\rho},$$

and hence

(33)

$$\frac{(x_1 \cdots x_\rho)^k}{(1 - x_1 - \ldots - x_\rho)^{k+1}} = \sum_{n_1 \geqslant 0} \cdots \sum_{n_\rho \geqslant 0} \binom{n_1 + \ldots + n_\rho - (\rho - 1)k}{n_1 - k, \ldots, n_\rho - k, k} x_1^{n_1} \cdots x_\rho^{n_\rho}.$$

Here, we used that the multinomial coefficient vanishes if $k > \min(n_1, \ldots, n_\rho)$. Geometrically expanding the left-hand side of (19), we find that

$$\left(\prod_{j=1}^{\ell}\left[1 - \sum_{r=1}^{\lambda_j} x_{s(j)+r}\right] - \alpha x_1 x_2 \cdots x_d\right)^{-1} = \sum_{k \geqslant 0} \alpha^k \prod_{j=1}^{\ell} \frac{(x_{s(j)+1} \cdots x_{s(j)+\lambda_j})^k}{\left[1 - \sum_{r=1}^{\lambda_j} x_{s(j)+r}\right]^{k+1}},$$

which we further expand using (33) to get

$$\sum_{k \geqslant 0} \alpha^k \sum_{\boldsymbol{n} \in \mathbb{Z}_{\geqslant 0}^d} \boldsymbol{x}^{\boldsymbol{n}} \prod_{j=1}^{\ell} \binom{n_{s(j)+1} + \ldots + n_{s(j)+\lambda_j} - (\lambda_j - 1)k}{n_{s(j)+1} - k, \ldots, n_{s(j)+\lambda_j} - k, k} = \sum_{\boldsymbol{n} \in \mathbb{Z}_{\geqslant 0}^d} A_{\lambda,\alpha}(\boldsymbol{n}) \boldsymbol{x}^{\boldsymbol{n}},$$

with $A_{\lambda,\alpha}(\boldsymbol{n})$ as in (20). □

5. The supercongruences

Our proof of Theorem 3.2, which generalizes the supercongruence in Theorem 1.2, builds upon the respective proofs in [Ges82] and [Beu85].

We need a number of lemmas in preparation. To begin with, we prove the following extension of Jacobsthal's binomial congruence [Ges83], [Gra97] to binomial coefficients which are allowed to have negative entries (see Remark 1.3).

Lemma 5.1. *For all primes p and all integers a, b,*

$$\binom{ap}{bp} \Big/ \binom{a}{b} \equiv \varepsilon \pmod{p^q}, \tag{34}$$

where q is the power of p dividing $p^3ab(a-b)/12$ and where $\varepsilon = 1$, unless $p = 2$ and $(a, b) \equiv (0, 1)$ modulo 2 in which case $\varepsilon = -1$.

Proof. Congruence (34), for nonnegative a, b, is proved in [Ges83] (alternatively, a proof for $p \geqslant 5$ is given in [Gra97]). We therefore only indicate how to extend (34) to negative values of a or b. Note that, for all $a, b \in \mathbb{Z}$ with $b \neq 0$,

$$\binom{a}{b} = \frac{a}{b}\binom{a-1}{b-1},$$

and hence

$$\binom{ap}{bp} \Big/ \binom{a}{b} = \binom{ap-1}{bp-1} \Big/ \binom{a-1}{b-1}.$$

We claim that the extension of (34) to the case $a < 0$ and $b < 0$ therefore follows from

$$\binom{a}{b} = \binom{-b-1}{-a-1}(-1)^{a-b}\operatorname{sgn}(a-b), \tag{35}$$

where sgn is defined as in Remark 1.3. This is clear for $p \geqslant 3$. Write $\varepsilon(a, b) = -1$ if $(a, b) \equiv (0, 1)$ modulo 2 and $\varepsilon(a, b) = 1$ otherwise. It is straightforward to check that

$$(-1)^{a-b}\varepsilon(-b, -a) = \varepsilon(a, b),$$

which shows the case $p = 2$.

Similarly, if $a < 0$ and $b > 0$, then we may apply

$$\binom{a}{b} = \binom{b-a-1}{-a-1}(-1)^{b+1}\operatorname{sgn}(a-b)\operatorname{sgn}(-a-1)$$

as well as

$$(-1)^{b}\varepsilon(b-a, -a) = \varepsilon(a, b).$$

A derivation of the above binomial identities, which are valid for all $a, b \in \mathbb{Z}$, may be found in [Spr08]. □

Much simpler and well-known is the following congruence.

Lemma 5.2. *Let $p \geqslant 5$ be a prime, and $\varepsilon \in \{-1, 1\}$. Then, for all integers $r \geqslant 0$,*

$$\sum_{k=1, p\nmid k}^{p^r-1} \frac{\varepsilon^k}{k^2} \equiv 0 \pmod{p^r}. \tag{36}$$

Proof. Let α be an odd integer, not divisible by p, such that $\alpha^2 \not\equiv 1$ modulo p (take, for instance, $\alpha = 3$). Then,

$$\frac{1}{\alpha^2}\sum_{k=1,p\nmid k}^{p^r-1}\frac{\varepsilon^k}{k^2} = \sum_{k=1,p\nmid k}^{p^r-1}\frac{\varepsilon^k}{(\alpha k)^2} \equiv \sum_{k=1,p\nmid k}^{p^r-1}\frac{\varepsilon^k}{k^2} \pmod{p^r},$$

since the second and third sum run over the same residues modulo p^r (note that $\varepsilon^{\alpha k} = \varepsilon^k$ since α is odd). As α^2 is not divisible by p, the congruence (36) follows. □

The next lemmas establish properties of the summands of the numbers $A_{\lambda,\varepsilon}(\boldsymbol{n})$ as introduced in (20), which will be needed in our proof of Theorem 3.2. Throughout this section, we fix the notation of Theorem 3.2, letting $\lambda = (\lambda_1, \ldots, \lambda_\ell) \in \mathbb{Z}_{>0}^\ell$ with $d = \lambda_1 + \ldots + \lambda_\ell$ and setting $s(j) = \lambda_1 + \ldots + \lambda_{j-1}$.

Lemma 5.3. *Let $\boldsymbol{n} \in \mathbb{Z}^d$, $k \in \mathbb{Z}$, and define*

$$A_\lambda(\boldsymbol{n}; k) = \prod_{j=1}^{\ell} \binom{n_{s(j)+1} + \ldots + n_{s(j)+\lambda_j} - (\lambda_j - 1)k}{n_{s(j)+1} - k, \ldots, n_{s(j)+\lambda_j} - k, k}. \tag{37}$$

(a) If $\ell \geqslant 2$, then, for all primes p and integers $r \geqslant 1$,

$$A_\lambda(p^r\boldsymbol{n}; pk) \equiv A_\lambda(p^{r-1}\boldsymbol{n}; k) \pmod{p^{2r}}. \tag{38}$$

(b) If $\ell \geqslant 2$ and $\max(\lambda_1, \ldots, \lambda_\ell) \leqslant 2$, then, for primes $p \geqslant 5$ and integers $r \geqslant 1$,

$$A_\lambda(p^r\boldsymbol{n}; pk) \equiv A_\lambda(p^{r-1}\boldsymbol{n}; k) \pmod{p^{3r}}. \tag{39}$$

Proof. We show (38) and (39) by proving that for integers $r, s \geqslant 1$ and k such that $p \nmid k$,

$$A_\lambda(p^r\boldsymbol{n}; p^s k) \equiv A_\lambda(p^{r-1}\boldsymbol{n}; p^{s-1}k) \pmod{p^{\alpha r}}, \tag{40}$$

where $\alpha = 2$ or $\alpha = 3$ depending on whether $\max(\lambda_1, \ldots, \lambda_\ell) \leqslant 2$.

Let us first consider the case $\ell \geqslant 2$ and $\max(\lambda_1, \ldots, \lambda_\ell) \leqslant 2$. Then each factor of (37) is a single binomial, if $\lambda_j = 1$, or of the form

$$\binom{m_1}{k}\binom{m_1 + m_2 - k}{m_1},$$

if $\lambda_j = 2$. Let p be a prime such that $p \geqslant 5$. It follows from Jacobsthal's congruence (34) that

$$\binom{p^r m_1}{p^s k} \Big/ \binom{p^{r-1} m_1}{p^{s-1} k} \equiv 1 \pmod{p^{r+s+\min(r,s)}}$$

as well as

$$\binom{p^r(m_1 + m_2) - p^s k}{p^r m_1} \Big/ \binom{p^{r-1}(m_1 + m_2) - p^{s-1}k}{p^{r-1} m_1} \equiv 1 \pmod{p^{r+2\min(r,s)}}.$$

Consequently,

$$A_\lambda(p^r\boldsymbol{n}; p^s k) = c A_\lambda(p^{r-1}\boldsymbol{n}; p^{s-1}k) \tag{41}$$

with $c \equiv 1$ modulo $p^{r+2\min(r,s)}$. If $s \geqslant r$, this proves congruence (40) with $\alpha = 3$. On the other hand, suppose $s \leqslant r$. Since $p \nmid k$, we have

$$\binom{p^r n}{p^s k} = p^{r-s}\frac{n}{k}\binom{p^r n - 1}{p^s k - 1} \equiv 0 \pmod{p^{r-s}}.$$

Since $\ell \geqslant 2$, it follows that $p^{2(r-s)}$ divides $A_\lambda(p^r\boldsymbol{n}; p^s k)$. Since $(r+2s)+2(r-s) = 3r$, the congruence (40), with $\alpha = 3$, now follows from (41). This shows (b).

Let us now turn to the proof of (a). Assume that $\ell \geqslant 2$. Note that, for any positive integer ρ,

$$\binom{m_1 + \ldots + m_\rho - (\rho-1)k}{m_1 - k, \ldots, m_\rho - k, k} = \binom{m_1}{k}\binom{m_1 + (m_2 - k) + \ldots + (m_\rho - k)}{m_1, m_2 - k, \ldots, m_\rho - k},$$

so that, as in the previous case, $p^{\ell(r-s)}$ divides $A_\lambda(p^r\boldsymbol{n}; p^s k)$ if $r \geqslant s$.

Initially, assume that $p \geqslant 3$. By further unravelling the multinomial coefficient as a product of binomial coefficients and applying Jacobsthal's congruence (34) as above, we find that

$$A_\lambda(p^r\boldsymbol{n}; p^s k) = cA_\lambda(p^{r-1}\boldsymbol{n}; p^{s-1}k)$$

with $c \equiv 1$ modulo $p^{3\min(r,s)-\delta}$ and $\delta = 0$, if $p \geqslant 5$, and $\delta = 1$, if $p = 3$. In light of $p^{2(r-s)}$ dividing $A_\lambda(p^r\boldsymbol{n}; p^s k)$ if $r \geqslant s$, we conclude congruence (40) with $\alpha = 2$.

Now, consider $p = 2$. If $r \geqslant 2$ and $s \geqslant 2$, then the sign ε in Jacobsthal's congruence (34) is always $+1$ when applying the above approach, and we again find that (40) holds with $\alpha = 2$. On the other hand, if $r = 1$, then it suffices to use the (combinatorial) congruence

$$\binom{pa}{pb} \equiv \binom{a}{b} \pmod{p^2},$$

which holds for all primes p. It remains to consider the case $r \geqslant 2$ and $s = 1$. Applying the approach employed for $p \geqslant 3$, we find that

$$A_\lambda(p^r\boldsymbol{n}; p^s k) = cA_\lambda(p^{r-1}\boldsymbol{n}; p^{s-1}k), \tag{42}$$

where $c \equiv \pm 1$ modulo $p^{3\min(r,s)-2} = 2$. If $\max(\lambda_1, \ldots, \lambda_\ell) \leqslant 2$, then we, in fact, have $c \equiv (-1)^\ell$ modulo $p^{r+2\min(r,s)-2} = 2^r$. Since $A_\lambda(p^r\boldsymbol{n}; p^s k)$ is divisible by $p^{\ell(r-1)}$, congruence (40) trivially holds with $\alpha = 2$ if $\ell \geqslant 3$. Hence, we may assume that $\ell = 2$. If $\max(\lambda_1, \lambda_2) \leqslant 2$, then $c \equiv 1$ modulo 2^r in (42) and, since both sides of (42) are divisible by 2^{2r-2}, congruence (40) with $\alpha = 2$ again follows. Finally, suppose that there is j such that $\lambda_j \geqslant 3$. Then the factor corresponding to j in (37) is of the form

$$\binom{m_1}{k}\binom{m_1 + m_2 - k}{m_1}\binom{m_1 + m_2 + m_3 - 2k}{m_3 - k}\binom{m_1 + \ldots + m_\rho - (\rho-1)k}{m_1 + m_2 + m_3 - 2k, m_4 - k, \ldots}.$$

Note that, for even m_1, m_2, m_3 and odd k, the third binomial in this product is even. Hence, $A_\lambda(p^r\boldsymbol{n}; p^s k)$ is divisible by $2^{2(r-1)+1} = 2^{2r-1}$. In light of (42), this proves congruence (40) with $\alpha = 2$. □

The next congruence, with $k \geqslant 0$, has been used in [Beu85]. For our present purpose, we extend it to the case of negative k.

Lemma 5.4. *For primes p, integers m, k and integers $r \geqslant 1$,*

$$\binom{p^r m - 1}{k}(-1)^k \equiv \binom{p^{r-1}m - 1}{[k/p]}(-1)^{[k/p]} \pmod{p^r}. \tag{43}$$

Proof. First, assume that $k \geqslant 0$. Following [Beu85, Lemma 2], we split the defining product of the binomial coefficient, according to whether the index is divisible by

p or not, to obtain

$$\begin{aligned}\binom{p^r m-1}{k} &= \prod_{j=1}^{k} \frac{p^r m-j}{j}\\ &= \prod_{j=1,p\nmid j}^{k} \frac{p^r m-j}{j} \prod_{\lambda=1}^{[k/p]} \frac{p^{r-1}m-\lambda}{\lambda}\\ &= \binom{p^{r-1}m-1}{[k/p]} \prod_{j=1,p\nmid j}^{k} \frac{p^r m-j}{j}.\end{aligned}$$

Congruence (43), with $k \geqslant 0$, follows upon reducing modulo p^r.

On the other hand, assume $k < 0$. Since (43) is trivial if $m > 0$, we let $m \leqslant 0$. We use the basic symmetry relation

$$\binom{p^r m-1}{k} = \binom{p^r m-1}{p^r m-k-1}$$

and note that, since $k < 0$, the binomials are zero unless $p^r m - k - 1 \geqslant 0$. Observe that, for all integers k, m,

$$[(p^r m-k-1)/p] = p^{r-1}m + [-(k+1)/p] = p^{r-1}m - [k/p] - 1. \tag{44}$$

Thus, assuming $p^r m - 1 - k \geqslant 0$, we may apply (43) to find

$$\begin{aligned}\binom{p^r m-1}{k}(-1)^k &= \binom{p^r m-1}{p^r m-k-1}(-1)^k\\ &\equiv \binom{p^{r-1}m-1}{p^{r-1}m-[k/p]-1}(-1)^{[k/p]}(-1)^{p^r m+p^{r-1}m}\\ &= \binom{p^{r-1}m-1}{[k/p]}(-1)^{[k/p]}(-1)^{p^r m+p^{r-1}m} \pmod{p^r}.\end{aligned}$$

It only remains to note that $p^r m + p^{r-1}m = p^{r-1}(p+1)m$ is even unless $p = 2$ and $r = 1$. Hence, in all cases, $(-1)^{p^r m+p^{r-1}m} \equiv 1$ modulo p^r. ☐

Lemma 5.5. *For primes p, integers m_1, m_2, k and integers $r \geqslant 1$,*

$$\binom{p^r m_1 + p^r m_2 - k - 1}{p^r m_1} \equiv \binom{p^{r-1}m_1 + p^{r-1}m_2 - [k/p] - 1}{p^{r-1}m_1} \pmod{p^r}.$$

Proof. By an application of (11),

$$\binom{m_1+m_2-k-1}{m_1} = \operatorname{sgn}(m_2-k-1)(-1)^{m_2-k-1}\binom{-m_1-1}{m_2-k-1}.$$

Since, for all $a \in \mathbb{Z}$, $\operatorname{sgn}(a) = \operatorname{sgn}([a/p])$, the claimed congruence therefore follows from (44) and Lemma 5.4. ☐

The following generalizes [Beu85, Lemma 3] to our needs.

Lemma 5.6. *Let p be a prime and $\boldsymbol{n} \in \mathbb{Z}^d$.*

- *Let $a_k \in \mathbb{Z}_p$, with $k \in \mathbb{Z}$, be such that, for all $l, s \in \mathbb{Z}$ with $s \geqslant 0$,*

$$\sum_{[k/p^s]=l} a_k \equiv 0 \pmod{p^s}.$$

- *Let $C(\boldsymbol{n};k)$ be such that, for all $k,r\in\mathbb{Z}$ with $r\geqslant 0$,*

$$C(p^r\boldsymbol{n};k)\equiv C(p^{r-1}\boldsymbol{n};[k/p]) \pmod{p^r}. \tag{45}$$

Then, for all $r,l\in\mathbb{Z}$ with $r\geqslant 0$,

$$\sum_{[k/p^r]=l} a_k C(p^r\boldsymbol{n};k)\equiv 0 \pmod{p^r}. \tag{46}$$

Proof. The claim is trivial for $r=0$. Fix $r>0$ and assume, for the purpose of induction on r, that the congruence (46) holds for the exponent $r-1$ in place of r. By the assumption (45) on $C(\boldsymbol{n};k)$, we have that, modulo p^r,

$$\begin{aligned}\sum_{[k/p^r]=l} a_k C(p^r\boldsymbol{n};k) &\equiv \sum_{[k/p^r]=l} a_k C(p^{r-1}\boldsymbol{n};[k/p])\\ &= \sum_{[m/p^{r-1}]=l}\left(\sum_{[k/p]=m} a_k\right) C(p^{r-1}\boldsymbol{n};m)\\ &= p\sum_{[m/p^{r-1}]=l} b_m C(p^{r-1}\boldsymbol{n};m),\end{aligned}$$

where b_m is the sequence

$$b_m=\frac{1}{p}\sum_{[k/p]=m} a_k.$$

We note that, for all $s,l\in\mathbb{Z}$ with $s\geqslant 0$,

$$\sum_{[m/p^s]=l} b_m=\frac{1}{p}\sum_{[m/p^s]=l}\sum_{[k/p]=m} a_k=\frac{1}{p}\sum_{[k/p^{s+1}]=m} a_k\equiv 0 \pmod{p^s},$$

so that we may apply our induction hypothesis (46) with $r-1$ to conclude

$$\sum_{[k/p^r]=l} a_k C(p^r\boldsymbol{n};k)=p\sum_{[m/p^{r-1}]=l} b_m C(p^{r-1}\boldsymbol{n};m)\equiv 0 \pmod{p^r}.$$

The claim therefore follows by induction. □

We are now in a comfortable position to prove Theorem 3.2.

Proof of Theorem 3.2. In terms of the numbers $A_{\lambda,\varepsilon}(\boldsymbol{n};k)$, defined in (37), we have

$$A_{\lambda,\varepsilon}(\boldsymbol{n})=\sum_{k\geqslant 0}\varepsilon^k A_\lambda(\boldsymbol{n};k)=\sum_{s\geqslant 0} G_s(\boldsymbol{n}),$$

where

$$G_s(\boldsymbol{n})=\sum_{p\nmid k}\varepsilon^{p^s k} A_\lambda(\boldsymbol{n};p^s k).$$

Suppose that $\ell\geqslant 2$. Further, suppose that $p\geqslant 3$, or that $p=2$ and $\varepsilon=1$. Then $\varepsilon^{p^s k}=\varepsilon^{p^{s-1}k}$, and it follows from Lemma 5.3 that, for $s\geqslant 1$,

$$G_s(p^r\boldsymbol{n})\equiv G_{s-1}(p^{r-1}\boldsymbol{n}) \pmod{p^{2r}}.$$

In order to prove that $A_{\lambda,\varepsilon}(p^r\boldsymbol{n})\equiv A_{\lambda,\varepsilon}(p^{r-1}\boldsymbol{n})$ modulo p^{2r}, it therefore remains only to show that $G_0(p^r\boldsymbol{n})\equiv 0$ modulo p^{2r}. This, however, is immediate because, as observed in the proof of Lemma 5.3, $A_\lambda(p^r\boldsymbol{n};k)$, with $p\nmid k$, is divisible by $p^{\ell r}$. This proves congruence (21).

Now, suppose that $\ell \geqslant 2$ and $\max(\lambda_1, \ldots, \lambda_\ell) \leqslant 2$. Let p be a prime such that $p \geqslant 5$. It again follows from $\varepsilon^{p^s k} = \varepsilon^{p^{s-1} k}$ and Lemma 5.3 that, for $s \geqslant 1$,

$$G_s(p^r \boldsymbol{n}) \equiv G_{s-1}(p^{r-1}\boldsymbol{n}) \pmod{p^{3r}}.$$

To prove that $A_{\lambda,\varepsilon}(p^r \boldsymbol{n}) \equiv A_{\lambda,\varepsilon}(p^{r-1}\boldsymbol{n})$ modulo p^{3r}, we have to show that $G_0(p^r \boldsymbol{n}) \equiv 0$ modulo p^{3r}. As in the previous case, this is trivial if $\ell \geqslant 3$. We thus assume $\ell = 2$.

Note that, since $\max(\lambda_1, \ldots, \lambda_\ell) \leqslant 2$, each factor of $A_\lambda(\boldsymbol{n}; k)$ is of the form

$$\binom{m_1}{k}, \quad \text{or} \quad \binom{m_1}{k}\binom{m_1 + m_2 - k}{m_1}.$$

Using the basic identity

$$\binom{m_1}{k} = \frac{m_1}{k}\binom{m_1 - 1}{k - 1},$$

it is clear that the numbers

$$B_\lambda(\boldsymbol{n}; k) = \frac{k^2}{n_1 n_{1+\lambda_1}} A_\lambda(\boldsymbol{n}; k)$$

are integers. Moreover, it follows from Lemmas 5.4 and 5.5, and the fact that $\ell = 2$, that the integers $C_\lambda(\boldsymbol{n}; k) = B_\lambda(\boldsymbol{n}; k+1)$ satisfy, for all $k, r \in \mathbb{Z}$ with $r \geqslant 0$,

$$C(p^r \boldsymbol{n}; k) \equiv C(p^{r-1}\boldsymbol{n}; [k/p]) \pmod{p^r}.$$

If $p \nmid k$ then $[(k-1)/p] = [k/p]$ so that, in particular,

$$C(p^r \boldsymbol{n}, k-1) \equiv C(p^r \boldsymbol{n}, [k/p]) \equiv C(p^r \boldsymbol{n}; k) \pmod{p^r}.$$

By construction,

$$G_0(p^r \boldsymbol{n}) = p^{2r} n_1 n_{1+\lambda_1} \sum_{p \nmid k} \frac{\varepsilon^k}{k^2} C(p^r \boldsymbol{n}; k-1),$$

so that, in order to show that $G_0(p^r \boldsymbol{n}) \equiv 0$ modulo p^{3r}, it suffices to prove

$$\sum_{p \nmid k} \frac{\varepsilon^k}{k^2} C(p^r \boldsymbol{n}; k) \equiv 0 \pmod{p^r}. \tag{47}$$

Define $a_k = \varepsilon^k / k^2$, if $p \nmid k$, and $a_k = 0$ otherwise. Since $p \geqslant 5$, it follows from Lemma 5.2 that, for all $l, s \in \mathbb{Z}$ with $s \geqslant 0$,

$$\sum_{[k/p^s]=l} a_k = \sum_{k=1, p \nmid k}^{p^s - 1} \frac{\varepsilon^{lp^s + k}}{(lp^s + k)^2} \equiv \varepsilon^l \sum_{k=1, p\nmid k}^{p^s - 1} \frac{\varepsilon^k}{k^2} \equiv 0 \pmod{p^s}.$$

Hence, the conditions of Lemma 5.6 are met, allowing us to conclude that

$$\sum_{p \nmid k} \frac{\varepsilon^k}{k^2} C(p^r \boldsymbol{n}; k) = \sum_l \sum_{[k/p^r]=l} a_k C(p^r \boldsymbol{n}; k) \equiv 0 \pmod{p^r}.$$

This shows (47) and completes our proof. □

Acknowledgements. I wish to thank Robert Osburn and Wadim Zudilin for interesting and motivating discussions on Apéry-like numbers and supercongruences. Much appreciated comments on an earlier version of this note have also been sent by Bruce Berndt and Brundaban Sahu. Moreover, I am very grateful to the referee for several suggestions that improved this paper. Finally, I thank the Max-Planck-Institute for Mathematics in Bonn, where most of this work was completed, for providing wonderful working conditions.

References

[AG77] R. Askey and G. Gasper. Convolution structures for Laguerre polynomials. *Journal d'Analyse Mathématique*, 31(1):48–68, December 1977.

[AO00] S. Ahlgren and K. Ono. A Gaussian hypergeometric series evaluation and Apéry number congruences. *Journal für die reine und angewandte Mathematik*, 2000(518), January 2000.

[Apé79] R. Apéry. Irrationalité de $\zeta(2)$ et $\zeta(3)$. *Astérisque*, 61:11–13, 1979.

[AvSZ11] G. Almkvist, D. van Straten, and W. Zudilin. Generalizations of Clausen's formula and algebraic transformations of Calabi–Yau differential equations. *Proc. Edinburgh Math. Soc.*, 54(2):273–295, 2011.

[AZ06] G. Almkvist and W. Zudilin. Differential equations, mirror maps and zeta values. In *Mirror symmetry. V*, volume 38 of *AMS/IP Stud. Adv. Math.*, pages 481–515. Amer. Math. Soc., Providence, RI, 2006.

[BBC+13] A. Bostan, S. Boukraa, G. Christol, S. Hassani, and J.-M. Maillard. Ising n-fold integrals as diagonals of rational functions and integrality of series expansions. *Journal of Physics A: Mathematical and Theoretical*, 46(18):185202, May 2013.

[Beu85] F. Beukers. Some congruences for the Apéry numbers. *Journal of Number Theory*, 21(2):141–155, October 1985.

[Beu87] F. Beukers. Another congruence for the Apéry numbers. *Journal of Number Theory*, 25(2):201–210, February 1987.

[Beu02] F. Beukers. On Dwork's accessory parameter problem. *Mathematische Zeitschrift*, 241(2):425–444, October 2002.

[CCC80] S. Chowla, J. Cowles, and M. Cowles. Congruence properties of Apéry numbers. *Journal of Number Theory*, 12(2):188–190, May 1980.

[CCS10] H. H. Chan, S. Cooper, and F. Sica. Congruences satisfied by Apéry-like numbers. *International Journal of Number Theory*, 06(01):89, 2010.

[Chr84] G. Christol. Diagonales de fractions rationnelles et equations différentielles. In *Study group on ultrametric analysis, 10th year: 1982/83, No. 2*, pages Exp. No. 18, 10. Inst. Henri Poincaré, Paris, 1984.

[Coo12] S. Cooper. Sporadic sequences, modular forms and new series for $1/\pi$. *The Ramanujan Journal*, 29(1–3):163–183, 2012.

[Cos88] M. J. Coster. *Supercongruences*. PhD thesis, Universiteit Leiden, 1988.

[Cow80] J. Cowles. Some congruence properties of three well-known sequences: Two notes. *Journal of Number Theory*, 12(1):84–86, February 1980.

[CZ10] H. H. Chan and W. Zudilin. New representations for Apéry-like sequences. *Mathematika*, 56:107–117, 2010.

[Del13] E. Delaygue. Arithmetic properties of Apéry-like numbers. *Preprint*, October 2013. Available at: http://arxiv.org/abs/1310.4131.

[Fra94] J. Franel. On a question of Laisant. *L'Intermédiaire des Mathématiciens*, 1:45–47, 1894.

[Ges82] I. M. Gessel. Some congruences for Apéry numbers. *Journal of Number Theory*, 14(3):362–368, June 1982.

[Ges83] I. M. Gessel. Some congruences for generalized Euler numbers. *Canad. J. Math*, 35(4):687–709, 1983.

[Gra97] A. Granville. Arithmetic properties of binomial coefficients I: Binomial coefficients modulo prime powers. *CMS Conf. Proc.*, 20:253–275, 1997.

[GRZ83] J. Gillis, B. Reznick, and D. Zeilberger. On elementary methods in positivity theory. *SIAM J. Math. Anal.*, 14(2):396–398, 1983.

[Kil06] T. Kilbourn. An extension of the Apéry number supercongruence. *Acta Arithmetica*, 123(4):335–348, 2006.

[KR11] C. Krattenthaler and T. Rivoal. Multivariate p-adic formal congruences and integrality of Taylor coefficients of mirror maps. In *Arithmetic and Galois theories of differential equations*, volume 23 of *Sémin. Congr.*, pages 241–269. Soc. Math. France, Paris, 2011.

[LvdP90] L. Lipshitz and A. J. van der Poorten. Rational functions, diagonals, automata and arithmetic. In *Number theory (Banff, AB, 1988)*, Berlin, 1990. de Gruyter.

[Mac15] P. A. MacMahon. *Combinatory Analysis*. Cambridge Univ. Press, 1915.

[Mim83] Y. Mimura. Congruence properties of Apéry numbers. *Journal of Number Theory*, 16(1):138–146, February 1983.

[MV13] A. Mellit and M. Vlasenko. Dwork's congruences for the constant terms of powers of a Laurent polynomial. *Preprint*, June 2013. Available at: http://arxiv.org/abs/1306.5811.

[OS11] R. Osburn and B. Sahu. Supercongruences for Apéry-like numbers. *Advances in Applied Mathematics*, 47(3):631–638, September 2011.

[OS13] R. Osburn and B. Sahu. A supercongruence for generalized Domb numbers. *Functiones et Approximatio Commentarii Mathematici*, 48(1):29–36, March 2013.

[OSS14] R. Osburn, B. Sahu, and A. Straub. Supercongruences for sporadic sequences. *Proceedings of the Edinburgh Mathematical Society*, 2014. To appear. Preprint available at: http://arxiv.org/abs/1312.2195.

[Poo79] A. v. d. Poorten. A proof that Euler missed ... Apéry's proof of the irrationality of $\zeta(3)$. *Math. Intelligencer*, 1(4):195–203, 1979.

[PW02] R. Pemantle and M. C. Wilson. Asymptotics of multivariate sequences: I. smooth points of the singular variety. *Journal of Combinatorial Theory, Series A*, 97(1):129–161, January 2002.

[PW08] R. Pemantle and M. C. Wilson. Twenty combinatorial examples of asymptotics derived from multivariate generating functions. *SIAM Review*, 50(2):199–272, January 2008.

[PWZ96] M. Petkovsek, H. Wilf, and D. Zeilberger. *A=B*. A. K. Peters, 1996.

[RY13] E. Rowland and R. Yassawi. Automatic congruences for diagonals of rational functions. *Preprint*, October 2013. Available at: http://arxiv.org/abs/1310.8635.

[Spr08] R. Sprugnoli. Negation of binomial coefficients. *Discrete Mathematics*, 308(22):5070–5077, November 2008.

[SvS09] K. Samol and D. van Straten. Dwork congruences and reflexive polytopes. *Preprint*, November 2009. Available at: http://arxiv.org/abs/0911.0797.

[SZ14] A. Straub and W. Zudilin. Positivity of rational functions and their diagonals. *Journal of Approximation Theory*, 2014. To appear. Preprint available at: http://arxiv.org/abs/1312.3732.

[Zag09] D. Zagier. Integral solutions of Apery-like recurrences. In J. Harnad and P. Winternitz, editors, *Groups and symmetries. From Neolithic to John McKay*, volume 47. American Mathematical Society, 2009.