# Gessel–Lucas congruences for sporadic sequences

Armin Straub*
Department of Mathematics and Statistics
University of South Alabama

January 28, 2023

**Abstract**

For each of the 15 known sporadic Apéry-like sequences, we prove congruences modulo $p^2$ that are natural extensions of the Lucas congruences modulo $p$. This extends a result of Gessel for the numbers used by Apéry in his proof of the irrationality of $\zeta(3)$. Moreover, we show that each of these sequences satisfies two-term supercongruences modulo $p^{2r}$. Using special constant term representations recently discovered by Gorodetsky, we prove these supercongruences in the two cases that remained previously open.

## 1 Introduction

Sequences $A(n)$ that are integer solutions of either the three-term recurrence

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

or the three-term recurrence

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

subject to the initial conditions $A(-1) = 0$, $A(0) = 1$, are known as *Apéry-like sequences*. For instance, if $(a, b, c, d) = (17, 5, 1, 0)$, then the solution to (2) is the integer sequence

$$A(n) = \sum_{k=0}^{n} \binom{n}{k}^2 \binom{n+k}{k}^2, \tag{3}$$

which formed the basis of Apéry's proof [Apé79], [Poo79] of the irrationality of $\zeta(3)$. Systematic searches for Apéry-like sequences have been conducted by Zagier [Zag09] in

* *Email:* `straub@southalabama.edu`

the case of (1), and by Almkvist–Zudilin [AZ06] and Cooper [Coo12] in the case of (2). After normalizing, and apart from degenerate cases as well four hypergeometric and four Legendrian solutions in each case, only a small number of *sporadic* solutions have been found. Namely, Zagier [Zag09] found six sporadic solutions to (1), labeled $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}, \boldsymbol{D}, \boldsymbol{E}, \boldsymbol{F}$. Almkvist–Zudilin [AZ06] found six corresponding sporadic solutions to (1), labeled $(\alpha), (\gamma), (\delta), (\epsilon), (\eta), (\zeta)$, and Cooper [Coo12] found three additional sporadic solutions to (1) with $d \neq 0$, labeled $s_7, s_{10}, s_{18}$. Explicit formulas for the sequences $(\gamma)$, $\boldsymbol{F}$, $(\delta)$, $\boldsymbol{D}$ can be found in (3), (10), (15), (18), respectively. Tables of all $6 + 6 + 3 = 15$ sequences, including known representations as binomial sums such as (3) can be found, for instance, in [MS16], [OSS16] or [Gor21].

One reason that Apéry-like sequences have received attention in the literature is that they share (or, are believed to share) various remarkable arithmetic properties. For instance, they are connected to modular forms in different ways [SB85], [Beu87], [AO00], [OS19] and they satisfy unusually strong congruences [Beu85], [Cos88], [CCS10], [OS11], [OS13], [Str14], [OSS16], [AT16], [Gor19], [Gor21] that were coined *supercongruences* by Beukers. As we review in Section 3, these have been proven by various authors for 13 of the 15 sequences. We prove the supercongruences for one of the previously conjectural cases, namely the sequence labeled $\boldsymbol{F}$. For the other missing sequence, labeled $(\delta)$ and known as the Almkvist–Zudilin numbers, we prove a weaker version of the congruences which, however, is sufficient for our present purposes. Combined with the previously known cases, this results in the following uniform result.

**Theorem 1.1.** *Let $A(n)$ be one of the $6 + 6 + 3$ known sporadic Apéry-like sequences. Then, for all primes $p \geq 3$ and all positive integers $n, r$,*

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

As indicated above, Theorem 1.1 was previously known except in the two cases $\boldsymbol{F}$ and $(\delta)$. We prove these two new cases as Theorems 3.2 and 3.3 in Section 3. We note that the restriction to primes $p \geq 5$, or $p \geq 3$ in the case of Theorem 1.1, is natural for supercongruences due to Lemma 2.1. Numerical evidence suggests that Theorem 1.1 also holds for $p = 2$ except for the three sequences labeled $\boldsymbol{B}$, $(\delta)$ and $(\eta)$. We will not attempt to discuss the case $p = 2$ in further detail here. We note that, in the literature, the case $p = 3$ of Theorem 1.1 in the previously known cases is not always discussed in detail. However, the arguments, which for primes $p \geq 5$ yield slightly stronger congruences due to Lemma 2.1, still apply for $p = 3$ and the resulting congruences are sufficient for Theorem 1.1.

As another instance of a special arithmetic property, Gessel showed [Ges82, Theorem 1] that the Apéry numbers (3) satisfy the congruences

$$A(n) \equiv A(n_0)A(n_1) \cdots A(n_r) \pmod{p}, \tag{4}$$

where $n = n_0 + n_1 p + \cdots + n_r p^r$ is the $p$-adic expansion of $n$. The congruences (4) are known as *Lucas congruences* because they are of the same kind as the congruences that Lucas

[Luc78] showed for the binomial coefficients. Various sequences and families of sequences have since been shown [McI92], [Gra97], [SvS15], [RY15], [MS16], [Del18], [ABD19], [Gor21] to satisfy the Lucas congruences. For a historical survey, we further refer to [Me4].

In particular, it turns out that, as shown by Malik and the author [MS16] (see also [Gor21]), each sporadic Apéry-like sequence $A(n)$ satisfies the Lucas congruences.

**Theorem 1.2** ([MS16])**.** *Let $A(n)$ be one of the $6+6+3$ known sporadic Apéry-like sequences. Then, for all primes $p$ and all integers $n, k$ with $0 \leq k < n$,*

$$A(pn+k) \equiv A(k)A(n) \pmod{p}. \tag{5}$$

By iterating, one sees that the congruences (5) are equivalent to the the Lucas congruences (4). Gessel [Ges82] further proved an extension of the Lucas congruences for the Apéry numbers (3) modulo $p^2$. The second main result of this paper is to prove corresponding congruences for all known sporadic Apéry-like sequences. This results in the following extension of Theorem 1.2 modulo $p^2$.

**Theorem 1.3.** *Let $A(n)$ be one of the $6+6+3$ known sporadic Apéry-like sequences. Then, for all primes $p \geq 3$ and all integers $n, k$ with $0 \leq k < n$,*

$$A(pn+k) \equiv A(k)A(n) + pnA'(k)A(n) \pmod{p^2}. \tag{6}$$

We will refer to the congruences (6) as *Gessel–Lucas congruences* modulo $p^2$. For a precise definition of the formal derivative $A'(n)$, we refer to Section 4. However, for certain of the sequences, $A'(n)$ can be obtained as the actual derivative $\frac{\mathrm{d}}{\mathrm{d}n}A(n)$ of a natural interpolation of $A(n)$. We prove Theorem 1.3 in Section 5 by extending the proof of Gessel [Ges82], who proved the congruences (6) for the Apéry numbers (3). To our knowledge, the Gessel–Lucas congruences (6) had not been previously observed for the other sequences covered by Theorem 1.3. A crucial ingredient for the proof is Theorem 1.1, the supercongruences satisfied by each sporadic sequence.

The reductions modulo prime powers $p^r$ of certain sequences $A(n)$, such as diagonals of rational functions and constant terms, can be described using the notion of *linear p-schemes* that was recently introduced by Rowland and Zeilberger [RZ14]. Using the language found in [Beu22], a linear $p$-scheme modulo $p^r$ for $A(n)$ with $s$ states consists of a vector of sequences $\boldsymbol{A}(n) = (A_1(n), \ldots, A_s(n))$, with $A_1(n) = A(n)$, as well as matrices $M(k)$ for $k \in \{0, 1, \ldots, p-1\}$ such that

$$\boldsymbol{A}(pn+k) \equiv M(k)\boldsymbol{A}(n) \pmod{p^r}$$

for all integers $n, k$ with $0 \leq k < n$. Based on earlier work of Rowland and Yassawi [RY15], Rowland and Zeilberger describe in [RZ14] algorithms to compute linear $p$-schemes for the values modulo $p^r$ of a sequence $A(n) = \mathrm{ct}[P(\boldsymbol{x})^n Q(\boldsymbol{x})]$ of constant terms, where $P, Q \in \mathbb{Z}[\boldsymbol{x}^{\pm 1}]$ are Laurent polynomials in $\boldsymbol{x} = (x_1, \ldots, x_d)$. As done in [RY15] and [Str22], one can derive upper bounds for the number of states for the $p$-schemes produced by these algorithms. However, as pointed out by Beukers [Beu22], these bounds appear inefficient except in small cases.

On the other hand, it appears to be an interesting and fruitful question to ask for general results to describe sequences whose reductions modulo $p^r$ can be expressed using linear $p$-schemes with few states. For instance, as observed in [HS22], a sequence $A(n)$ satisfies the Lucas congruences (4) modulo $p$ if and only if its modulo $p$ reductions can be encoded by a linear $p$-scheme with a single state. One way to interpret Theorem 1.3 is that it provides explicit two-state linear $p$-schemes for all sporadic Apéry-like sequences modulo $p^2$. It would be of considerable interest to better understand which other sequences share this property generalizing the classical Lucas congruences.

## 2 Notation and basic congruences

Suppose that $\boldsymbol{k} = (k_1, \ldots, k_\ell)$ is a tuple of nonnegative integers adding up to $n$, that is, $|\boldsymbol{k}| = k_1 + k_2 + \cdots + k_\ell = n$. Then we denote the corresponding multinomial coefficient as

$$\binom{n}{\boldsymbol{k}} = \binom{n}{k_1, \ldots, k_\ell} = \frac{n!}{k_1! k_2! \cdots k_\ell!}.$$

Throughout the paper, we use typical notation and write, for instance, $\lambda\boldsymbol{k}$ as short for $(\lambda k_1, \lambda k_2, \ldots, \lambda k_\ell)$.

The following version of Jacobsthal's binomial congruence is proved in [Ges83] and [Gra97] (and is extended in [Str14] to binomial coefficients which are allowed to have negative entries).

**Lemma 2.1.** *For primes $p \geq 5$, and integers $n, k$ and $r, s \geq 1$,*

$$\binom{p^r n}{p^s k} \Big/ \binom{p^{r-1} n}{p^{s-1} k} \equiv 1 \pmod{p^{r+s+\min(r,s)}}. \tag{7}$$

*For primes $p = 3$, respectively $p = 2$, the congruence (7) holds modulo $p^{r+s+\min(r,s)-1}$, respectively $p^{r+s+\min(r,s)-2}$.*

Also note that, if $p \nmid k$ and $s \leq r$, then we have the much simpler congruence

$$\binom{p^r n}{p^s k} = p^{r-s} \frac{n}{k} \binom{p^r n - 1}{p^s k - 1} \equiv 0 \pmod{p^{r-s}}. \tag{8}$$

Finally, for the proof of our main result, we record the following observation concerning Cooper's sporadic sequences $s_7, s_{10}, s_{18}$.

**Lemma 2.2.** *If $A(n)$ is one of the 3 known sporadic Apéry-like sequences that satisfy (2) with $d \neq 0$ then, for all primes $p$,*

$$A(p-1) \equiv 0 \pmod{p}.$$

*Proof.* This is a special case of [MS16, Theorem 6.6] and the discussion preceding it. □

## 3 Supercongruences

It was observed by Beukers [Beu85] that the Apéry numbers satisfy congruences of a certain type that are stronger than what is predicted by formal group theory. Such *supercongruences* were further studied by Coster [Cos88] who showed that the Apéry numbers satisfy

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

for all primes $p \geq 5$. The case $r = 1$ of these congruences had previously been observed by Gessel [Ges82]. Since then supercongruences have been established for other Apéry-like sequences [OS11], [OS13], [OSS16], [Gor19], [Gor21]. In fact, numerical evidence suggests that all Apéry-like sequences satisfy supercongruences. More precisely, the following conjecture appears in [OSS16].

**Conjecture 3.1.** *Let $A(n)$ be one of the $6 + 6 + 3$ known sporadic Apéry-like sequences. Then, for all primes $p \geq 5$ and all positive integers $n, r$,*

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

*where $\lambda = 3$ except in the five cases $\boldsymbol{B}$, $\boldsymbol{C}$, $\boldsymbol{E}$, $\boldsymbol{F}$ and $s_{18}$ in which case $\lambda = 2$.*

The four cases $\boldsymbol{A}$, $\boldsymbol{D}$, $(\gamma)$, $s_{10}$ follow from Coster's work [Cos88]. Osburn and Sahu proved the case $\boldsymbol{C}$ in [OS11] as well as the cases $\boldsymbol{E}$ and $(\alpha)$ in [OS13]; together with the author, they further established the cases $(\epsilon)$, $(\eta)$, $s_7$ and $s_{18}$ in [OSS16]. More recently, Gorodetsky proved the case $(\zeta)$ in [Gor19] as well as the case $\boldsymbol{B}$ [Gor21].

In Theorem 3.2, we prove the previously open case $\boldsymbol{F}$ of Conjecture 3.1. For the final missing case $(\delta)$, we prove in Theorem 3.3 a weaker version of Conjecture 3.1 where the conjectured exponent $\lambda = 3$ is replaced with $\lambda = 2$. The crucial ingredient for these proofs are suitable constant terms expressions that were recently obtained by Gorodetsky [Gor21]. Combined with the previously known cases, this proves Theorem 1.1, which is the weaker version of Conjecture 3.1 where $\lambda = 2$ for all sequences. While the case $(\delta)$ of Conjecture 3.1 remains open in general, we note that Amdeberhan and Tauraso [AT16] prove the case $r = 1$ of the corresponding conjectured supercongruences (9).

We recall that the sporadic Apéry-like sequence labeled $\boldsymbol{F}$ by Zagier [Zag09] is the sequence

$$A_{\boldsymbol{F}}(n) = \sum_{k=0}^{n} (-1)^k 8^{n-k} \binom{n}{k} \sum_{l=0}^{k} \binom{k}{l}^3, \tag{10}$$

which solves the three-term recurrence (1) with $(a, b, c) = (17, 6, 72)$.

**Theorem 3.2.** *For all primes $p \geq 3$ and all positive integers $n, r$,*

$$A_{\boldsymbol{F}}(p^r n) \equiv A_{\boldsymbol{F}}(p^{r-1} n) \pmod{p^{2r}}.$$

*Proof.* By expressing $A_{\boldsymbol{F}}(n)$ as the constant term of $\Lambda(x,y)^n$ where

$$\Lambda(x,y) = \frac{(x-y+1)(y-x+1)(x+y-1)(x+y+1)(x^2+y^2+1)}{(xy)^2},$$

Gorodetsky [Gor21] obtained the alternative representation

$$A_{\boldsymbol{F}}(n) = \sum_{(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c},\boldsymbol{d},\boldsymbol{e})\in U(n)} (-1)^{a_1+b_2+c_3} \binom{n}{\boldsymbol{a}}\binom{n}{\boldsymbol{b}}\binom{n}{\boldsymbol{c}}\binom{n}{\boldsymbol{d}}\binom{n}{\boldsymbol{e}}$$

as a multiple binomial sum. Here, the sum is over the set

$$U(n) = \left\{ (\boldsymbol{a},\boldsymbol{b},\boldsymbol{c},\boldsymbol{d},\boldsymbol{e}) \in \mathbb{Z}_{\geq 0}^{15} \ : \ \begin{array}{rl} a_1+a_2+a_3 &= n \\ b_1+b_2+b_3 &= n \\ c_1+c_2+c_3 &= n\,, \\ d_1+d_2+d_3 &= n \\ e_1+e_2+e_3 &= n \end{array} \quad \begin{array}{c} a_i+b_i+c_i+d_i+2e_i \;=\; 2n \\ \text{for each } i\in\{1,2,3\} \end{array} \right\}.$$

In the sequel, we use the notation $\boldsymbol{k} = (\boldsymbol{a},\boldsymbol{b},\boldsymbol{c},\boldsymbol{d},\boldsymbol{e}) \in \mathbb{Z}_{\geq 0}^{15}$. As usual, we say that $p$ divides $\boldsymbol{k}$ (and write $p|\boldsymbol{k}$) if $p$ divides each component of $\boldsymbol{k}$. If $\boldsymbol{k}\in U(n)$, then we write

$$B(\boldsymbol{k}) := (-1)^{a_1+b_2+c_3} \binom{n}{\boldsymbol{a}}\binom{n}{\boldsymbol{b}}\binom{n}{\boldsymbol{c}}\binom{n}{\boldsymbol{d}}\binom{n}{\boldsymbol{e}}.$$

Using this notation, we have

$$A_{\boldsymbol{F}}(p^r n) = \sum_{\boldsymbol{k}\in U(p^r n)} B(\boldsymbol{k}) = \sum_{\substack{\boldsymbol{k}\in U(p^r n)\\ p|\boldsymbol{k}}} B(\boldsymbol{k}) + \sum_{\substack{\boldsymbol{k}\in U(p^r n)\\ p\nmid\boldsymbol{k}}} B(\boldsymbol{k}). \tag{11}$$

We claim (and will show below) that, for all primes $p\geq 3$ and all $\boldsymbol{k}\in U(p^r n)$, we have

$$B(\boldsymbol{k}) \equiv B(\boldsymbol{k}/p) \pmod{p^{2r}} \tag{12}$$

where the right-hand side is to be interpreted as 0 if $p\nmid\boldsymbol{k}$. Combining (11) and (12), we conclude that

$$A_{\boldsymbol{F}}(p^r n) \equiv \sum_{\substack{\boldsymbol{k}\in U(p^r n)\\ p|\boldsymbol{k}}} B(\boldsymbol{k}/p) = \sum_{\boldsymbol{k}\in U(p^{r-1} n)} B(\boldsymbol{k}) = A_{\boldsymbol{F}}(p^{r-1}n) \pmod{p^2},$$

which is what we set out to prove. Note that the middle equality uses that $\boldsymbol{k}\in U(p^{r-1}n)$ if and only if $p\boldsymbol{k}\in U(p^r n)$.

It therefore only remains to prove (12). First, consider the case $\boldsymbol{k}\in U(p^r n)$ with $p\nmid\boldsymbol{k}$. Without loss of generality, we may assume that $p\nmid a_i$ for some $i\in\{1,2,3\}$ (otherwise, the same argument applies with $a_i$ replaced by one of $b_i, c_i, d_i, e_i$). It follows from the constraint $a_i + b_i + c_i + d_i + 2e_i = 2p^r n$ that at least one of $b_i, c_i, d_i, e_i$ is also not divisible by $p$ (note that we excluded the case $p = 2$ so that the factors of 2 in the constraint can be ignored). Without loss, we assume that $p \nmid b_i$. Note that we have the divisibility

$$\binom{p^r n}{a_i}\binom{p^r n}{b_i} \quad \Big| \quad B(\boldsymbol{k}), \tag{13}$$

and that, by (8), the two binomial coefficients on the left-hand side are each divisible by $p^r$. It follows that $B(\boldsymbol{k})$ is divisible by $p^{2r}$, which proves congruence (12) in the case $p \nmid \boldsymbol{k}$.

Finally, consider (12) in the case $p|\boldsymbol{k}$. Write $s = \min(\nu_p(a_1), r)$ and $t = \min(\nu_p(a_2), r)$ and suppose that $s \geq t$ (this is no loss of generality; otherwise, we can swap $a_1$ and $a_2$ throughout the proof). Consequently, $a_3 = p^r n - a_1 - a_2$ is divisible by $p^t$ as well. It follows from Lemma 2.1 applied to each binomial coefficient on the right-hand side of

$$\binom{p^r n}{\boldsymbol{a}} = \binom{p^r n}{a_1}\binom{p^r n - a_1}{a_2}$$

(note that $p^s$ divides $p^r n - a_1$), together with $p|\boldsymbol{a}$, that

$$\binom{p^r n}{\boldsymbol{a}} \Big/ \binom{p^{r-1} n}{\boldsymbol{a}/p} \equiv 1 \pmod{p^{s+2t-\varepsilon}},$$

where $\varepsilon = 0$ if $p \geq 5$ and $\varepsilon = 1$ if $p = 3$. The same argument applies with $\boldsymbol{b}, \boldsymbol{c}, \boldsymbol{d}, \boldsymbol{e}$ in place of $\boldsymbol{a}$. Suppose that the value of the quantity $s + 2t$ is smallest for $\boldsymbol{a}$ compared to the corresponding values for $\boldsymbol{b}, \boldsymbol{c}, \boldsymbol{d}, \boldsymbol{e}$ (this is no loss of generality; otherwise, we can swap $\boldsymbol{a}$ for one of $\boldsymbol{b}, \boldsymbol{c}, \boldsymbol{d}, \boldsymbol{e}$ in the remaining argument). We therefore have

$$\frac{B(\boldsymbol{k})}{B(\boldsymbol{k}/p)} \equiv 1 \pmod{p^{s+2t-\varepsilon}}. \tag{14}$$

If $t = r$, then the claim follows and we are done. Otherwise, $\nu_p(a_2) = t < r$. It follows from the constraint $a_i + b_i + c_i + d_i + 2e_i = 2p^r n$ that at least one of $b_2, c_2, d_2, e_2$ is not divisible by $p^{t+1}$. Without loss of generality, suppose that $\nu_p(b_2) \leq t$. Note that we have the divisibility (13) with $i = 2$ and that, by (8), the two binomial coefficients on the left-hand side of (13) are each divisible by $p^{r-t}$. This shows that $B(\boldsymbol{k})$ is divisible by $p^{2r-2t}$. Combining this with (14), we conclude that

$$B(\boldsymbol{k}) \equiv B(\boldsymbol{k}/p) \quad \left(\operatorname{mod} p^{2r+s-\varepsilon}\right)$$

for all $\boldsymbol{k} \in U(p^r n)$ with $p|\boldsymbol{k}$. Since $s \geq 1$ and $\varepsilon \in \{0, 1\}$, this is slightly stronger, in the case $p|\boldsymbol{k}$, than the claimed congruences (12). □

The same argument we used to prove Theorem 3.2 applies similarly to the sporadic Apéry-like sequence labeled ($\delta$) by Almkvist–Zudilin [AZ06]. This sequence has the binomial sum representation

$$A_\delta(n) = \sum_{k=0}^{n} (-1)^k 3^{n-3k} \binom{n}{3k} \binom{n+k}{n} \frac{(3k)!}{k!^3}, \tag{15}$$

solves the three-term recurrence (2) with $(a,b,c,d) = (7,3,81,0)$, and is also known as the Almkvist–Zudilin numbers. Combined with Theorem 3.2, the following proves Theorem 1.1.

**Theorem 3.3.** *For all primes $p \geq 3$ and all positive integers $n, r$,*

$$A_\delta(p^r n) \equiv A_\delta(p^{r-1} n) \pmod{p^{2r}}.$$

*Proof.* By expressing $A_\delta(n)$ as the constant term of $\Lambda(x,y,z)^n$ where

$$\Lambda(x,y,z) = \frac{(x+y-1)(x+z+1)(y-x+z)(y-z+1)}{xyz},$$

Gorodetsky [Gor21] obtained the alternative representation

$$A_\delta(n) = \sum_{(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c},\boldsymbol{d}) \in U(n)} (-1)^{a_2+b_1+d_3} \binom{n}{\boldsymbol{a}} \binom{n}{\boldsymbol{b}} \binom{n}{\boldsymbol{c}} \binom{n}{\boldsymbol{d}}$$

as a multiple binomial sum. Here, the sum is over the set

$$U(n) = \left\{ (\boldsymbol{a},\boldsymbol{b},\boldsymbol{c},\boldsymbol{d}) \in \mathbb{Z}_{\geq 0}^{12} : \begin{array}{rl} a_1+a_2+a_3 &= n \\ b_1+b_2+b_3 &= n \\ c_1+c_2+c_3 &= n \\ d_1+d_2+d_3 &= n \end{array}, \quad \begin{array}{rl} b_1+c_1+d_1 &= n \\ a_1+b_2+d_2 &= n \\ a_2+b_3+c_2 &= n \end{array} \right\}.$$

As in the proof of Theorem 3.2, we consider

$$B(\boldsymbol{k}) := (-1)^{a_2+b_1+d_3} \binom{n}{\boldsymbol{a}} \binom{n}{\boldsymbol{b}} \binom{n}{\boldsymbol{c}} \binom{n}{\boldsymbol{d}}$$

for $\boldsymbol{k} = (\boldsymbol{a},\boldsymbol{b},\boldsymbol{c},\boldsymbol{d}) \in U(n)$. The same argument as for the congruences (12) applies and allows us to show that, again, $B(\boldsymbol{k}) \equiv B(\boldsymbol{k}/p)$ modulo $p^{2r}$, with the understanding that $B(\boldsymbol{k}/p) = 0$ if $p \nmid \boldsymbol{k}$. As a consequence, we once more conclude that

$$\begin{aligned} A_\delta(p^r n) &= \sum_{\substack{\boldsymbol{k} \in U(p^r n) \\ p \mid \boldsymbol{k}}} B(\boldsymbol{k}) + \sum_{\substack{\boldsymbol{k} \in U(p^r n) \\ p \nmid \boldsymbol{k}}} B(\boldsymbol{k}) \\ &\equiv \sum_{\substack{\boldsymbol{k} \in U(p^r n) \\ p \mid \boldsymbol{k}}} B(\boldsymbol{k}/p) = \sum_{\boldsymbol{k} \in U(p^{r-1} n)} B(\boldsymbol{k}) = A_\delta(p^{r-1} n) \pmod{p^2}, \end{aligned}$$

as claimed. ∎

## 4 The formal derivative of recurrence sequences

Suppose that $c_0(n), \ldots, c_r(n) \in \mathbb{Z}[n]$ are polynomials with $c_0(0) = 0$ or, equivalently, $c_0(n) \in n\mathbb{Z}[n]$ and $c_0(n) \neq 0$ for all $n \in \mathbb{Z}_{>0}$. Then there exists a unique sequence $A(n)$ which satisfies the linear recurrence recurrence

$$\sum_{j=0}^{r} c_j(n) A(n-j) = 0 \tag{16}$$

for all $n \geq 0$, subject to the initial conditions $A(0) = 1$ and $A(j) = 0$ for $j < 0$. In the sequel, our interest will be limited to the cases where recursion (16) is either (1) or (2). In particular, for our purposes we have $r = 2$. Note that, for (1) or (2) only the initial condition $A(0) = 1$ is significant because in (16) with $n = 1$ the term involving $A(-1)$ vanishes due to $c_2(1) = 0$.

Suppose further that $c_0(n) \in n^2\mathbb{Z}[n]$ (which is satisfied for the recurrences (1) and (2)). Then we can introduce the formal derivative $A'(n)$ of $A(n)$ as the unique sequence satisfying

$$\sum_{j=0}^{r} c_j(n) A'(n-j) + \sum_{j=0}^{r} c_j'(n) A(n-j) = 0 \tag{17}$$

(where $c_j'(n) = \frac{\mathrm{d}}{\mathrm{d}n} c_j(n)$ is the ordinary derivative of the polynomial $c_j(n)$), subject to the initial conditions $A'(j) = 0$ for $j \leq 0$.

We note that, in some cases, the formal derivative $A'(n)$ can be obtained as a usual derivative of a natural interpolation of $A(n)$. Namely, suppose that the sequence $A(n)$ can be extended to a smooth function $A(n)$ that is defined for all real $n$, or a suitable subset of the reals, in such a way that (16) holds for all such $n$. By differentiating (16), it then follows that the usual derivative $\frac{\mathrm{d}}{\mathrm{d}n} A(n)$ satisfies the recursion (17). Therefore, provided that the initial conditions line up as well, the usual derivative agrees with the formal derivative defined above. This is illustrated in Example 4.1. We then indicate in Example 4.2 that the same conclusion can still be drawn if, instead of (16), the extension of $A(n)$ to real $n$ satisfies an inhomogeneous recurrence with an additional term that (together with its derivative) vanishes when $n$ is an integer.

**Example 4.1.** The sporadic sequence $\boldsymbol{D}$ (which is connected to $\zeta(2)$ in the same way that the Apéry numbers (3) are connected to $\zeta(3)$) is given by

$$A_{\boldsymbol{D}}(n) = \sum_{k=0}^{n} \binom{n}{k}^2 \binom{n+k}{k} \tag{18}$$

and solves (1) with $(a, b, c) = (11, 3, -1)$. The sequence $A_{\boldsymbol{D}}'(n)$ is therefore characterized by solving the recurrence

$$\begin{aligned} (n+1)^2 A_{\boldsymbol{D}}'(n+1) \quad = \quad & (11n^2 + 11n + 3) A_{\boldsymbol{D}}'(n) + n^2 A_{\boldsymbol{D}}'(n-1) \\ & -2(n+1) A_{\boldsymbol{D}}(n+1) + 11(2n+1) A_{\boldsymbol{D}}(n) + 2n A_{\boldsymbol{D}}(n-1) \end{aligned} \tag{19}$$

with $A'_{\boldsymbol{D}}(0) = 0$ (note that we don't actually need an additional initial value for $A'_{\boldsymbol{D}}(-1)$). The resulting initial values for $A'_{\boldsymbol{D}}(n)$ for $n = 1, 2, \ldots$ are

$$5, \quad \frac{75}{2}, \quad \frac{1855}{6}, \quad \frac{10875}{4}, \quad \frac{299387}{12}, \quad \frac{943397}{4}, \quad \frac{63801107}{28}, \quad \frac{1253432797}{56}, \quad \ldots$$

Indeed, one can show (for instance, using creative telescoping) that the formal derivative has the explicit formula

$$A'_{\boldsymbol{D}}(n) = 5\sum_{k=0}^{n} \binom{n}{k}^2 \binom{n+k}{k} (H_n - H_k) \tag{20}$$

involving harmonic sums. On the other hand, the series

$$A_{\boldsymbol{D}}(x) = \sum_{k=0}^{\infty} \binom{x}{k}^2 \binom{x+k}{k}$$

converges for complex $x$ with $\operatorname{Re} x > -1$ and therefore defines an interpolation of the sporadic sequence $A_{\boldsymbol{D}}(n)$. As shown in [OS19], this interpolation satisfies the homogeneous functional equation

$$(x+1)^2 A_{\boldsymbol{D}}(x+1) - (11x^2 + 11x + 3)A_{\boldsymbol{D}}(x) - x^2 A_{\boldsymbol{D}}(x-1) = 0$$

for all complex $x$ with $\operatorname{Re} x > -1$ (this is (1) with $(a, b, c) = (11, 3, -1)$). We can then differentiate this equation to obtain (19). By verifying that the derivative of $A_{\boldsymbol{D}}(x)$ vanishes for $x = 0$, we conclude that, for positive integers $n$, the values $A'_{\boldsymbol{D}}(n)$ in (20) agree with the values of the actual derivative of the interpolation $A_{\boldsymbol{D}}(x)$.

**Example 4.2.** In a similar way, Zagier [Zag18, Section 7] interpolates the Apéry numbers $A(n)$, defined by the binomial sum (3), using the series

$$A(x) = \sum_{k=0}^{\infty} \binom{x}{k}^2 \binom{x+k}{k}^2,$$

which is well-defined for all complex $x$. Somewhat surprisingly, Zagier showed that, unlike the previous example, the interpolation $A(x)$ satisfies the functional equation (see [OS19] for an algorithmic derivation using creative telescoping)

$$(x+1)^3 A(x+1) - (2x+1)(17x^2 + 17x + 5)A(x) + x^3 A(x-1) = \frac{8}{\pi^2}(2x+1)\sin^2(\pi x)$$

which is an inhomogeneous version of the recurrence (2) with $(a, b, c, d) = (17, 5, 1, 0)$ satisfied the Apéry numbers. However, note that the inhomogeneous term and its derivative vanish for any integer $x$. We can therefore conclude as in the previous example that, for positive integers $n$, the values $A'(n)$ of the formal derivative agree with the derivative values of Zagier's interpolation. As recorded by Gessel [Ges82], these are given by

$$A'(n) = 2\sum_{k=0}^{n} \binom{n}{k}^2 \binom{n+k}{k}^2 (H_{n+k} - H_{n-k}).$$

**Remark 4.3.** Write $F(x) = \sum_{n\geq 0} A(n)x^n$. Then the recurrence (16) translates into the differential equation

$$\sum_{j=0}^{r} c_j(\theta)x^j F(x) = 0, \tag{21}$$

where $\theta = x\frac{\mathrm{d}}{\mathrm{d}x}$ is the Euler operator. Note that $c_0(n)$ is the indicial polynomial of this differential equation. Since we assumed that $c_0(n) \in n^2\mathbb{Z}[n]$ and $c_0(n) \neq 0$ for all $n \in \mathbb{Z}_{>0}$, it follows that, up to scaling, $F(x)$ is the unique power series solution of (21). Moreover, if we set $G(x) = \sum_{n\geq 0} A'(n)x^n$, then we claim that $\log(x)F(x)+G(x)$ is a second solution of (21). To see this, note that it follows inductively from $\theta \log(x)F(x) = \log(x)\theta F(x) + F(x)$ that $\theta^k \log(x)F(x) = \log(x)\theta^k F(x) + k\theta^{k-1}F(x)$. Consequently, we have

$$c(\theta)\log(x)F(x) = \log(x)c(\theta)F(x) + c'(\theta)F(x)$$

for any polynomial $c(n) \in \mathbb{C}[n]$. On the other hand, we readily verify that $c(\theta)x^j = x^j c(\theta + j)$. Applying these formulae, as well as using (21), we find

$$\begin{aligned}
\sum_{j=0}^{r} c_j(\theta)x^j \log(x)F(x) &= \sum_{j=0}^{r} x^j c_j(\theta+j)\log(x)F(x) \\
&= \sum_{j=0}^{r} x^j(\log(x)c_j(\theta+j)F(x) + c_j'(\theta+j)F(x)) \\
&= \sum_{j=0}^{r} x^j c_j'(\theta+j)F(x) \\
&= \sum_{j=0}^{r} c_j'(\theta)x^j F(x).
\end{aligned}$$

Hence,

$$\sum_{j=0}^{r} c_j(\theta)x^j(\log(x)F(x) + G(x)) = \sum_{j=0}^{r}[c_j(\theta)x^j G(x) + c_j'(\theta)x^j F(x)] = 0,$$

where the final equality follows from the equivalent recursion (17) for the coefficients.

## 5 Gessel–Lucas congruences modulo $p^2$

We are now in a convenient position to prove Theorem 1.3, which is restated below for convenience. This an extension of Gessel's result [Ges82, Theorem 4] for the Apéry numbers and our proof proceeds along the same lines, with an extra argument required for the three sporadic Apéry-like sequences that satisfy (2) with $d \neq 0$. Here, the formal derivative $A'(n)$ is as introduced in Section 4.

**Theorem 5.1.** *Let $A(n)$ be one of the $6+6+3$ known sporadic Apéry-like sequences. Then, for all primes $p \geq 3$ and all integers $n, k$ with $0 \leq k < n$,*

$$A(pn+k) \equiv A(k)A(n) + pnA'(k)A(n) \pmod{p^2}. \tag{22}$$

*Proof.* By Theorem 1.2, each sporadic Apéry-like sequence $A(n)$ satisfies the Lucas congruences. That is, $A(pn+k) \equiv A(k)A(n)$ modulo $p$. We therefore have

$$A(pn+k) \equiv A(k)A(n) + pa(n,k) \pmod{p^2}$$

for some $a(n,k) \in \mathbb{Z}_p$. Our goal is to show that $a(n,k) \equiv nA'(k)A(n) \pmod{p}$.

By the case $r = 1$ of the supercongruences in Theorem 1.1, we know that each of the sequences $A(n)$ in question satisfies the congruences

$$A(pn) \equiv A(n) \pmod{p^2}.$$

This verifies the case $k = 0$ in (22) because, by our definition, $A'(0) = 0$. In the sequel, we therefore assume that $k \geq 1$.

Substituting $n$ by $pn+k$ in the recurrence (16) for $A(n)$, which is either of the form (1) or of the form (2), we have

$$\sum_{j=0}^{r} c_j(pn+k)A(pn+k-j) = 0$$

for all $n, k \in \mathbb{Z}_{\geq 0}$. If we normalize $c_0(n) = n^\lambda$ with $\lambda \in \{2,3\}$, then $c_2(n) = c(n-1)^2$ or $c_2(n) = (n-1)(c(n-1)^2+d)$ depending on whether (16) is of the form (1) or of the form (2).

We claim that the terms with $k - j < 0$ do not contribute modulo $p^2$. Note that $k - j < 0$ only if $k = 1, j = 2$, so that we need to show that $c_j(pn+k)A(pn+k-j) = c_2(pn+1)A(pn-1)$ is divisible by $p^2$. We note that $c_2(pn+1)$ is divisible by $p^2$ unless we have $c_2(n) = (n-1)(c(n-1)^2+d)$ with $d \neq 0$. In that latter case, which consists of the 3 known sporadic Apéry-like sequences that satisfy (2) with $d \neq 0$, $c_2(pn+1)$ is only divisible by $p$. However, in those cases, we can combine Lemma 2.2, by which $A(p-1)$ is divisible by $p$, with the Lucas congruences (5) to conclude that

$$A(pn-1) = A(p(n-1)+p-1) \equiv A(n-1)A(p-1) \equiv 0 \pmod{p}.$$

This shows that, in all cases, $c_2(pn+1)A(pn-1)$ is divisible by $p^2$.

Since the terms with $k-j<0$ do not contribute modulo $p^2$ (and because $A(j)=0$ for $j<0$), we can apply the Lucas congruences to obtain

$$\sum_{j=0}^{r} c_j(pn+k)(A(k-j)A(n)+pa(n,k-j)) \equiv 0 \pmod{p^2},$$

with the understanding that $a(n,j)=0$ if $j<0$. Using the Taylor expansion of the polynomials $c_j(n)$, this becomes

$$\sum_{j=0}^{r}(c_j(k)+pnc_j'(k))(A(k-j)A(n)+pa(n,k-j)) \equiv 0 \pmod{p^2}.$$

Expanding, followed by applying the recurrence (16), this is equivalent to

$$\sum_{j=0}^{r}(c_j(k)a(n,k-j)+nc_j'(k)A(k-j)A(n)) \equiv 0 \pmod{p}. \tag{23}$$

Since $c_0(n)=n^\lambda$ with $\lambda\in\{2,3\}$, we have $c_0(k)\not\equiv 0 \pmod{p}$ for all $k\in\{1,2,\dots,p-1\}$. Therefore, the congruence (23) together with $a(n,0)=0$ characterizes the values modulo $p$ of $a(n,k)$ for $k\in\{1,2,\dots,p-1\}$.

On the other hand, replacing $n$ by $k$ in (17), and multiplying with $nA(n)$, we find

$$\sum_{j=0}^{r}(c_j(k)nA'(k-j)A(n)+nc_j'(k)A(k-j)A(n))=0.$$

Comparison with (23) shows that $a(n,k)\equiv nA'(k)A(n) \pmod{p}$, which is what we set out to show. □

# 6 Conclusions

Samol and van Straten [SvS15] (see also [MV16]) showed that, if the Newton polytope of a Laurent polynomial $P(\boldsymbol{x})\in\mathbb{Z}[\boldsymbol{x}^{\pm1}]$, with $\boldsymbol{x}=(x_1,\dots,x_d)$, has the origin as its only interior integral point, then $A(n)=\operatorname{ct}[P(\boldsymbol{x})^n]$, the sequence formed by the constant terms of powers of $P(\boldsymbol{x})$, satisfies the *Dwork congruences*

$$A(p^rm+n)A(\lfloor n/p\rfloor) \equiv A(p^{r-1}m+\lfloor n/p\rfloor)A(n) \pmod{p^r} \tag{24}$$

for all primes $p$ and all integers $m,n\ge 0$, $r\ge 1$. The case $r=1$ of these congruences is equivalent to the Lucas congruences (4). In a similar spirit, is there a natural extension of the Gessel–Lucas congruences that we prove in Theorem 1.3 to modulus $p^{2r}$? Presumably, such an extension should also contain the supercongruences of Theorem 1.1 as a special case.

In another direction, it is natural to investigate an extension of Theorem 1.3 modulo $p^3$. More generally, it would be of interest to understand the minimal number of states required for a linear $p$-scheme describing sporadic Apéry-like sequences modulo $p^r$. We refer to recent work of Beukers [Beu22] for recent promising progress in this regard.

Recently, through a careful and clever search, Gorodetsky was able to find new constant term representations [Gor21] for several of the sporadic Apéry-like sequences. Combined with the result of Samol and van Straten, these can be used to establish Theorem 1.2 for 14 of the 15 sporadic sequences. In the case of the sporadic sequence $(\eta)$, however, we presently only have the lengthy and technical proof given in [MS16]. It would of interest to also find a suitable constant term representation for the sporadic sequence $(\eta)$. More generally, it would be valuable to have general results in the spirit of [SvS15] that would allow us to prove Theorem 1.3 in the presence of suitable constant term expressions. However, while the result of Samol and van Straten [SvS15] shows that large families of sequences satisfy the Lucas congruences, the congruences in Theorem 1.3 are considerably more rare. In particular, we note that the congruences (6) imply the congruences $A(pn) \equiv A(n) \pmod{p^2}$, which are an instance of the supercongruences discussed in Section 3.

**Acknowledgements.** The author is grateful to Frits Beukers for interesting discussions, motivation, and for sharing early versions of the paper [Beu22].

## References

[ABD19] Boris Adamczewski, Jason P. Bell, and Éric Delaygue. Algebraic independence of $G$-functions and congruences "à la Lucas". *Annales Scientifiques de l'École Normale Supérieure*, 52(3):515–559, 2019.

[AO00] Scott Ahlgren and Ken Ono. A Gaussian hypergeometric series evaluation and Apéry number congruences. *Journal für die reine und angewandte Mathematik*, 2000(518):187–212, January 2000.

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

[AT16] Tewodros Amdeberhan and Roberto Tauraso. Supercongruences for the Almkvist–Zudilin numbers. *Acta Arithmetica*, 173:255–268, 2016.

[AZ06] Gert Almkvist and Wadim 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.

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

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

[Beu22] Frits Beukers. $p$-linear schemes for sequences modulo $p^r$. *Preprint*, November 2022. arXiv:2211.15240.

[CCS10] Heng Huat Chan, Shaun Cooper, and Francesco Sica. Congruences satisfied by Apéry-like numbers. *International Journal of Number Theory*, 6(1):89–97, 2010.

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

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

[Del18] Eric Delaygue. Arithmetic properties of Apéry-like numbers. *Compositio Mathematica*, 154(2):249–274, February 2018.

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

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

[Gor19] Ofir Gorodetsky. $q$-congruences, with applications to supercongruences and the cyclic sieving phenomenon. *International Journal of Number Theory*, 2019.

[Gor21] Ofir Gorodetsky. New representations for all sporadic Apéry-like sequences, with applications to congruences. *Experimental Mathematics*, 2021. DOI:10.1080/10586458.2021.1982080.

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

[HS22] Joel A. Henningsen and Armin Straub. Generalized Lucas congruences and linear $p$-schemes. *Advances in Applied Mathematics*, 141:1–20, #102409, 2022.

[Luc78] Edouard Lucas. Sur les congruences des nombres Eulériens et des coefficients différentiels des fonctions trigonométriques, suivant un module premier. *Bulletin de la Société Mathématique de France*, 6:49–54, 1878.

[McI92] Richard J. McIntosh. A generalization of a congruential property of Lucas. *American Mathematical Monthly*, 99(3):231–238, March 1992.

[Me4] Romeo Meštrović. Lucas' theorem: its generalizations, extensions and applications (1878–2014). *Preprint*, September 2014. arXiv:1409.3820.

[MS16] Amita Malik and Armin Straub. Divisibility properties of sporadic Apéry-like numbers. *Research in Number Theory*, 2(1):#5, 26 p., 2016.

[MV16] Anton Mellit and Masha Vlasenko. Dwork's congruences for the constant terms of powers of a Laurent polynomial. *International Journal of Number Theory*, 12(2):313–321, 2016.

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

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

[OS19] Robert Osburn and Armin Straub. Interpolated sequences and critical $L$-values of modular forms. In J. Blümlein, P. Paule, and C. Schneider, editors, *Elliptic Integrals, Elliptic Functions and Modular Forms in Quantum Field Theory*, pages 327–349. Springer, 2019.

[OSS16] Robert Osburn, Brundaban Sahu, and Armin Straub. Supercongruences for sporadic sequences. *Proceedings of the Edinburgh Mathematical Society*, 59(2):503–518, 2016.

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

[RY15] Eric Rowland and Reem Yassawi. Automatic congruences for diagonals of rational functions. *Journal de Théorie des Nombres de Bordeaux*, 27(1):245–288, 2015.

[RZ14] Eric Rowland and Doron Zeilberger. A case study in meta-automation: automatic generation of congruence automata for combinatorial sequences. *Journal of Difference Equations and Applications*, 20(7):973–988, 2014.

[SB85] Jan Stienstra and Frits Beukers. On the Picard-Fuchs equation and the formal Brauer group of certain elliptic K3-surfaces. *Mathematische Annalen*, 271(2):269–304, April 1985.

[Str14] Armin Straub. Multivariate Apéry numbers and supercongruences of rational functions. *Algebra & Number Theory*, 8(8):1985–2008, 2014.

[Str22] Armin Straub. On congruence schemes for constant terms and their applications. *Research in Number Theory*, 8(3):1–21, #42, 2022.

[SvS15] Kira Samol and Duco van Straten. Dwork congruences and reflexive polytopes. *Annales mathématiques du Québec*, 39(2):185–203, October 2015.

[Zag09] Don B. Zagier. Integral solutions of Apéry-like recurrence equations. In *Groups and symmetries*, volume 47 of *CRM Proc. Lecture Notes*, pages 349–366. Amer. Math. Soc., Providence, RI, 2009.

[Zag18] Don B. Zagier. The arithmetic and topology of differential equations. In Volker Mehrmann and Martin Skutella, editors, *Proceedings of the European Congress of Mathematics, Berlin, 18-22 July, 2016*, pages 717–776. European Mathematical Society, 2018.