TAYLOR SERIES FOR THE APÉRY NUMBERS

ERIC ROWLAND AND REEM YASSAWI

Abstract. The sequence of Apéry numbers can be interpolated to $\mathbb{C}$ by an entire function. We give a formula for the Taylor coefficients of this function, centered at the origin, as a $\mathbb{Z}$-linear combination of multiple zeta values. We also show that for integers n whose base-p digits belong to a certain set, the Apéry numbers satisfy a Lucas congruence modulo p^2.

1. Introduction

For each integer $n \geq 0$, the nth *Apéry number* is defined by

$$A(n) := \sum_{k \geq 0} \binom{n}{k}^2 \binom{n+k}{k}^2.$$

These numbers arose in Apéry's proof of the irrationality of $\zeta(3)$. This sum is finite, since $\binom{n}{k} = 0$ when $k > n$. The sequence $A(n)_{n \geq 0}$ is

$$1, 5, 73, 1445, 33001, 819005, 21460825, 584307365, \dots.$$

The Apéry numbers satisfy the recurrence

$$n^3 A(n) - (34n^3 - 51n^2 + 27n - 5)A(n-1) + (n-1)^3 A(n-2) = 0 \tag{1}$$

for all integers $n \geq 2$.

Exceptional properties of the Apéry sequence have been observed in many settings [14]. Gessel [2, Theorem 1] showed that the Apéry numbers satisfy the congruence

$$A(d + pn) \equiv A(d)A(n) \mod p$$

for all $d \in \{0, 1, \dots, p-1\}$ and $n \geq 0$. Further, he extended this to a congruence modulo p^2 as follows. Define the sequence $A'(n)_{n \geq 0}$ by

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

where $H_k = 1 + \frac{1}{2} + \cdots + \frac{1}{k}$ is the kth harmonic number. The sequence $(A'(n))_{n \geq 0}$ is

$$0, 12, 210, 4438, 104825, \frac{13276637}{5}, 70543291, \frac{67890874657}{35}, \dots.$$

Then

$$A(d + pn) \equiv (A(d) + pnA'(d))\, A(n) \mod p^2 \tag{3}$$

for all $d \in \{0, 1, \dots, p-1\}$ and for all $n \geq 0$ [2, Theorem 4].

Gessel remarks that if $A(n)$ can be extended to a differentiable function $A(x)$ defined for $x \in \mathbb{R}_{\geq 0}$ such that $A(x)$ satisfies Recurrence (1), then $A'(n) = \left(\frac{d}{dx} A(x)\right)|_{x=n}$.

Date: May 10, 2020.

As shown by Zagier [14, Proposition 1] and proved in an automated way by Osburn and Straub [8, Remark 2.5], $A(n)$ can be extended to an entire function $A(z)$ satisfying

$$z^3 A(z) - (34z^3 - 51z^2 + 27z - 5)A(z-1) + (z-1)^3 A(z-2) = \frac{8}{\pi^2}(2z-1)(\sin(\pi z))^2 \tag{4}$$

for all $z \in \mathbb{C}$. Since $\frac{8}{\pi^2}(2z-1)(\sin(\pi z))^2$ and its derivative vanish at integer values of z, it follows that $A'(n) = \left(\frac{d}{dz}A(z)\right)|_{z=n}$, hence the notation $A'(n)$. Therefore the extension $A(z)$ confirms Gessel's intuition.

In this article we write the coefficients in the Taylor series of $A(z) = \sum_{m\geq 0} a_m z^m$ at $z = 0$ as an explicit $\mathbb{Z}$-linear combination of multiple zeta values. A striking fact is that the coefficient of each multiple zeta value is a signed power of 2. Let $s_1, s_2, \dots, s_j$ be positive integers with $s_1 \geq 2$. The *multiple zeta value* $\zeta(s_1, s_2, \dots, s_j)$ is defined as

$$\zeta(s_1, s_2, \dots, s_j) := \sum_{n_1 > n_2 > \cdots > n_j > 0} \frac{1}{n_1^{s_1} n_2^{s_2} \cdots n_j^{s_j}}.$$

The *weight* of $\zeta(s_1, s_2, \dots, s_j)$ is $s_1 + s_2 + \cdots + s_j$.

Let $\chi(m)$ be the characteristic function of the set of odd numbers. That is, $\chi(m) = 0$ if m is even and $\chi(m) = 1$ if m is odd. For a tuple $\mathbf{s} = (s_1, s_2, \dots, s_j)$, let $e(\mathbf{s}) = |\{i : 2 \leq i \leq j \text{ and } s_i = 2\}|$.

Theorem 1. *Let $A(z) = \sum_{m\geq 0} a_m z^m$ be the Apéry function. For each $m \geq 1$,*

$$a_m = \sum_{\mathbf{s}} (-1)^{\frac{m-s_1}{2}} 2^{e(\mathbf{s})+\chi(m)} \zeta(s_1, s_2, \dots, s_j),$$

where the sum is over all tuples $\mathbf{s} = (s_1, s_2, \dots, s_j)$, with $j \geq 1$, of non-negative integers satisfying

- $s_1 + s_2 + \cdots + s_j = m$,
- $s_1 = 3$ *if m is odd and $s_1 \in \{2, 4\}$ if m is even, and*
- $s_i \in \{2, 4\}$ *for all $i \in \{2, \dots, j\}$.*

The first several coefficients are

$$\begin{aligned}
a_0 &= 1 \\
a_1 &= 0 \\
a_2 &= \zeta(2) \\
a_3 &= 2\zeta(3) \\
a_4 &= \zeta(4) - 2\zeta(2,2) \\
a_5 &= -4\zeta(3,2) \\
a_6 &= \zeta(2,4) - 2\zeta(4,2) + 4\zeta(2,2,2) \\
a_7 &= 2\zeta(3,4) + 8\zeta(3,2,2) \\
a_8 &= \zeta(4,4) - 2\zeta(2,2,4) - 2\zeta(2,4,2) + 4\zeta(4,2,2) - 8\zeta(2,2,2,2) \\
a_9 &= -4\zeta(3,2,4) - 4\zeta(3,4,2) - 16\zeta(3,2,2,2).
\end{aligned}$$

Let $F(m)$ be the mth Fibonacci number. Since the number of integer compositions of m using parts 1 and 2 is $F(m+1)$, Theorem 1 expresses a_m as a linear combination of $F(\frac{m}{2}+1)$ multiple zeta values if m is even and $F(\frac{m-1}{2})$ multiple zeta values if m is odd.

Let $P(m)$ be the number of integer compositions of $m-3$ using parts 2 and 3. Then $P(m)$ is the mth Padovan number and satisfies the recurrence $P(m) = P(m-2)+P(m-3)$ with initial conditions $P(3)=1$, $P(4)=0$, $P(5)=1$. Let d_m be the dimension of the $\mathbb{Q}$-vector space spanned by the weight-m multiple zeta values. Recent progress by Brown [1] and Zagier [13] shows that $d_m \le P(m+3)$. For $m \ge 13$, the representation of a_m in Theorem 1 uses fewer than $P(m+3)$ multiple zeta values. Since $F(\frac{m}{2}+1) > P(m+3)$ for $m \in \{4,6,8,10,12\}$, this implies that $a_4, a_6, a_8, a_{10}, a_{12}$ can be written as $\mathbb{Q}$-linear combinations of fewer multiple zeta values than Theorem 1 provides. Namely,

$$\begin{aligned}
a_4 &= -\tfrac{1}{2}\zeta(4)\\
a_6 &= \tfrac{3}{2}\zeta(6) - 3\zeta(4,2)\\
a_8 &= -\tfrac{13}{24}\zeta(8) + 6\zeta(4,2,2)\\
a_{10} &= \tfrac{7}{8}\zeta(10) + 3\zeta(2,4,4) - 12\zeta(4,2,2,2)\\
a_{12} &= -\tfrac{915}{22112}\zeta(12) + 6\zeta(4,2,2,4) + 6\zeta(4,2,4,2) + 6\zeta(4,4,2,2) + 24\zeta(4,2,2,2,2).
\end{aligned}$$

In Section 2, we define the extension of $A(n)$ to the complex numbers. In Section 3 we prove that the Taylor coefficients of $A(z)$, centered at $z=0$, are $\mathbb{Q}$-linear combinations of multiple zeta values. We strengthen this in Section 4 to prove Theorem 1.

Returning to Gessel's congruence (3) in Section 5, we study base-p digits d such that $A(d+pn) \equiv A(d)A(n) \mod p^2$ for all $n \in \mathbb{Z}$. We show that d has this property if and only if $A(d) \equiv A(p-1-d) \mod p^2$. In particular, if all digits in the standard base-p representation $n_\ell \cdots n_1 n_0$ of a non-negative integer n satisfy $A(n_i) \equiv A(p-1-n_i) \mod p^2$, then

$$A(n) \equiv A(n_0)A(n_1)\cdots A(n_\ell) \mod p^2.$$

2. Extending $A(n)$ to the complex numbers

Let $\mathbb{N} = \{0,1,2,\dots\}$. Recall that the gamma function $\Gamma(z)$ is a meromorphic function satisfying

$$\Gamma(1) = 1 \text{ and } \Gamma(z+1) = z\Gamma(z)$$

for $z \notin -\mathbb{N}$. The gamma function has simple poles at the non-positive integers. For an exposition of the properties of $\Gamma(z)$, see [6].

For $n \ge 0$, we can write $A(n)$ as

$$\begin{aligned}
A(n) &= \sum_{k\ge 0} \binom{n}{k}^2 \binom{n+k}{k}^2\\
&= \sum_{k\ge 0} \frac{\Gamma(n+k+1)^2}{\Gamma(n-k+1)^2\Gamma(k+1)^4}.
\end{aligned}$$

We extend $A(n)$ to complex values by defining

$$A(z) = \sum_{k\ge 0} \frac{\Gamma(z+k+1)^2}{\Gamma(z-k+1)^2\Gamma(k+1)^4}.$$

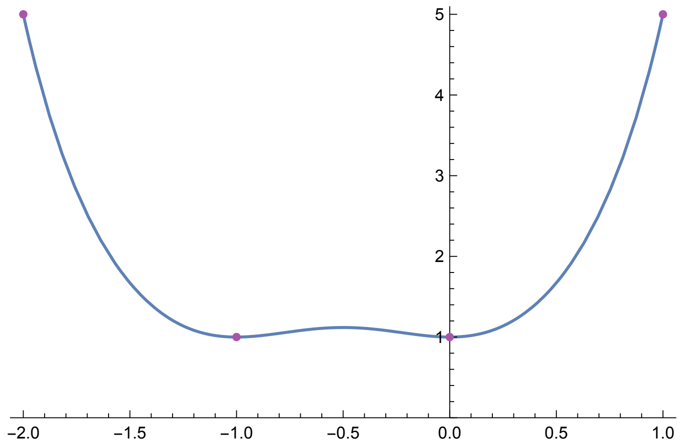

FIGURE 1. A plot of $A(z)$ for real z in the interval $-2 \le z \le 1$, showing the reflection symmetry $A(-1-z) = A(z)$.

Note that for each $k \in \mathbb{N}$, the function $\frac{\Gamma(z+k+1)^2}{\Gamma(z-k+1)^2\Gamma(k+1)^4}$ is entire. Furthermore for each $z \in \mathbb{C}$, the series $\sum_{k\ge 0} \frac{\Gamma(z+k+1)^2}{\Gamma(z-k+1)^2\Gamma(k+1)^4}$ is locally uniformly convergent. Thus $A(z)$ is an entire function, which we call the *Apéry function*. We remark that $A(z)$ can be written using the hypergeometric function ${}_4F_3$. Let $(z)_k := z(z+1)(z+2)\cdots(z+k-1)$ be the Pochhammer symbol (rising factorial). Writing $\frac{\Gamma(z+k+1)^2}{\Gamma(z-k+1)^2} = (-z)_k^2(z+1)_k^2$ shows that

$$A(z) = \sum_{k\ge 0} \frac{(-z)_k(-z)_k(z+1)_k(z+1)_k}{k!^4} = {}_4F_3(-z,-z,z+1,z+1;1,1,1;1). \tag{5}$$

Straub [11, Remark 1.3] proved the reflection formula $A(-1-n) = A(n)$ for all $n \in \mathbb{Z}$. Equation (5) shows that this formula also holds for non-integers, since the hypergeometric series is invariant under replacing z with $-1-z$.

Proposition 2. *For all $z \in \mathbb{C}$, we have $A(-1-z) = A(z)$.*

Figure 1 shows this symmetry on the real line. In light of Proposition 2, Theorem 1 also gives us the Taylor expansion of $A(z)$ at $z=-1$ for free.

3. TAYLOR COEFFICIENTS OF THE APÉRY FUNCTION

Let

$$A(z) = \sum_{k\ge 0} \frac{\Gamma(z+k+1)^2}{\Gamma(z-k+1)^2\Gamma(k+1)^4} = \sum_{m\ge 0} a_m z^m$$

be the Taylor series expansion of the Apéry function centered at the origin. In this section we prove that a_m is a $\mathbb{Q}$-linear combination of weight-m multiple zeta values. We compute a_m by directly evaluating the mth derivative $A^{(m)}(z)$ at $z=0$.

Consider the summand $\frac{\Gamma(z+k+1)^2}{\Gamma(z-k+1)^2 k!^4}$. The factor $\frac{1}{k!^4}$ is independent of z, so we factor it out and study the derivatives of $\frac{\Gamma(z+k+1)^2}{\Gamma(z-k+1)^2}$. These derivatives involve the digamma function $\psi(z) := \frac{\Gamma'(z)}{\Gamma(z)}$, that is, the logarithmic derivative of $\Gamma(z)$. For example, symbolically the first derivative is

$$\frac{d}{dz}\frac{\Gamma(z+k+1)^2}{\Gamma(z-k+1)^2} = \frac{\Gamma(z+k+1)^2}{\Gamma(z-k+1)^2}\left(2\psi(z+k+1) - 2\psi(z-k+1)\right), \tag{6}$$

which we can evaluate for all $z \in \mathbb{C} \setminus \{k-1, k-2, \dots\}$. In particular, this agrees with the expression for $A'(n)$ in Equation (2). Similarly, the second derivative is

$$\begin{aligned}\frac{d^2}{dz^2}\frac{\Gamma(z+k+1)^2}{\Gamma(z-k+1)^2} = \frac{\Gamma(z+k+1)^2}{\Gamma(z-k+1)^2}\big(&4\psi(z+k+1)^2 + 2\psi'(z+k+1)\\ &- 8\psi(z+k+1)\psi(z-k+1) + 4\psi(z-k+1)^2 - 2\psi'(z-k+1)\big).\end{aligned} \tag{7}$$

Remark 3. The singularities of $\frac{\Gamma(z+k+1)^2}{\Gamma(z-k+1)^2}$ are removable, so we interpret $\frac{\Gamma(z+k+1)^2}{\Gamma(z-k+1)^2}$ and all its derivatives to be the entire functions obtained by removing these singularities. The function $2\psi(z+k+1) - 2\psi(z-k+1)$ in Equation (6) is meromorphic with simple poles at $z \in \{k-1, k-2, \dots\}$. Therefore

$$\frac{\Gamma(z+k+1)^2}{\Gamma(z-k+1)^2}\left(2\psi(z+k+1) - 2\psi(z-k+1)\right)$$

is holomorphic on $\mathbb{C} \setminus \{k-1, k-2, \dots\}$. By the identity theorem for holomorphic functions, the two functions we obtain by removing the singularities on both sides of Equation (6) are equal for all $z \in \mathbb{C}$. The same argument applies to Equation (7) and further equations involving higher derivatives. Thus we interpret all expressions obtained from symbolically computing $\frac{d^m}{dz^m}\frac{\Gamma(z+k+1)^2}{\Gamma(z-k+1)^2}$ to be the entire functions obtained by removing their singularities.

In Lemma 4 we will be interested in the behavior of $\frac{\Gamma(z+k+1)^2}{\Gamma(z-k+1)^2}$ near $z = 0$. Iterating the recurrence $\Gamma(z+1) = z\Gamma(z)$ enables us to write

$$\Gamma(z-k+1) = \frac{\Gamma(z+1)}{z(z-1)(z-2)\cdots(z-k+1)}$$

for positive integers k and $z \notin \{\dots, k-2, k-1\}$. This implies that

$$\frac{\Gamma(z+k+1)^2}{\Gamma(z-k+1)^2} = \frac{\Gamma(z+k+1)^2(z(z-1)\cdots(z-k+1))^2}{\Gamma(z+1)^2} = O(z^2) \text{ as } z \to 0 \tag{8}$$

for positive k.

Since $\psi(z-k+1)$ arises when we differentiate $\frac{\Gamma(z+k+1)^2}{\Gamma(z-k+1)^2}$, we will also be interested in the behavior of $\psi^{(m)}(z-k+1)$ near $z = 0$. The recurrence for $\Gamma(z)$ also implies that the digamma function satisfies

$$\psi(z+1) = \psi(z) + \frac{1}{z} \tag{9}$$

for $z \notin -\mathbb{N}$. From this we can deduce that

$$\psi(z-k+1) = \psi(z+1) - \frac{1}{z} - \sum_{\ell=1}^{k-1}\frac{1}{z-\ell} \text{ for } 0 < |z| < 1 \tag{10}$$

for any positive integer k (where, if $k = 1$, the term $\sum_{\ell=1}^{k-1}\frac{1}{z-\ell}$ is the empty sum and therefore 0), so that $\psi^{(m)}(z-k+1)$ has Laurent, i.e. principal, part $\frac{(-1)^{m+1}m!}{z^{m+1}}$, and

$$\psi^{(m)}(z-k+1) = O\big(\frac{1}{z^{m+1}}\big) \tag{11}$$

for small z in a punctured disc about 0, for each $m \in \mathbb{N}$.

Notation. For each $m \geq 0$, define the expression $P_{m,k}(z)$ by

$$\frac{d^m}{dz^m} \frac{\Gamma(z+k+1)^2}{\Gamma(z-k+1)^2} = \frac{\Gamma(z+k+1)^2}{\Gamma(z-k+1)^2} P_{m,k}(z). \tag{12}$$

For $m = 0$, we have $P_{0,k}(z) = 1$. The expression for $P_{1,k}(z)$ appears in Equation (6), and $P_{2,k}(z)$ appears in Equation (7). Differentiating Equation (12) shows that $P_{m,k}(z)$ satisfies the recurrence

$$P_{m+1,k}(z) = P_{1,k}(z) P_{m,k}(z) + \tfrac{d}{dz} P_{m,k}(z). \tag{13}$$

The computation of the first Taylor coefficients of $A(z)$ in Example 5 hinges on the following fact, on which we will base the rest of our analysis of the Taylor expansion of the Apéry function.

Let $[z^i]f(z)$ denote the coefficient of z^i in the Laurent series of $f(z)$ at $z = 0$.

Lemma 4. *For each $m \geq 1$, we have $A^{(m)}(z)|_{z=0} = \sum_{k \geq 1} \frac{1}{k^2} [z^{-2}] P_{m,k}(z)$.*

Proof. By Equation (12), Equation (8), and Remark 3, we have, for $k \geq 1$, that

$$\begin{aligned}
\tfrac{d^m}{dz^m} \tfrac{\Gamma(z+k+1)^2}{\Gamma(z-k+1)^2}\big|_{z=0} &= \lim_{z \to 0} \left(\tfrac{\Gamma(z+k+1)^2 (z(z-1)\cdots(z-k+1))^2}{\Gamma(z+1)^2} P_{m,k}(z) \right) \\
&= k!^2 (k-1)!^2 \lim_{z \to 0} z^2 P_{m,k}(z) \\
&= k!^2 (k-1)!^2 [z^{-2}] P_{m,k}(z).
\end{aligned}$$

If $k = 0$, then $\frac{\Gamma(z+k+1)^2}{\Gamma(z-k+1)^2} = 1$, so its mth derivative is 0 for all $m \geq 1$. Therefore

$$A^{(m)}(z) = \sum_{k \geq 1} \frac{k!^2 (k-1)!^2 [z^{-2}] P_{m,k}(z)}{k!^4} = \sum_{k \geq 1} \frac{1}{k^2} [z^{-2}] P_{m,k}(z). \qquad \square$$

Example 5. Using (6) and (7), we can compute $A'(0)$ and $A''(0)$. Equations (6) and (11) imply that $P_{1,k}(z)$ has a simple pole at 0 for each k and so by Lemma 4, $A'(0) = 0$.

To compute $A''(0)$, note that $\psi(z + k + 1)$ is analytic at $z = 0$ for each $k \geq 0$. In particular, any term involving only $\psi(z + k + 1)$ and its derivatives has Laurent part 0. By Equation (10), $\psi(z - k + 1)$ has only a simple pole at $z = 0$. Therefore we can ignore the term $-8\psi(z + k + 1)\psi(z - k + 1)$. Therefore Equations (7) and (10) imply that

$$\begin{aligned}
[z^{-2}] P_{2,k}(z) &= [z^{-2}] \left(4\psi(z-k+1)^2 - 2\psi'(z-k+1) \right) \\
&= [z^{-2}] \left(4 \left(\psi(z+1) - \frac{1}{z} - \sum_{\ell=1}^{k-1} \frac{1}{z-\ell} \right)^2 - 2 \left(\psi'(z+1) + \frac{1}{z^2} + \sum_{\ell=1}^{k-1} \frac{1}{(z-\ell)^2} \right) \right) \\
&= \frac{2}{k^2},
\end{aligned}$$

so that $A''(0) = \sum_{k \geq 1} \frac{2}{k^2} = 2\zeta(2)$, and $a_2 = \frac{A''(0)}{2!} = \zeta(2)$.

The next proposition gives the series coefficients of $P_{1,k}(z)$. Let $H_{k-1}^{(s)} = \sum_{\ell=1}^{k-1} \frac{1}{\ell^s}$ be the generalized harmonic number, where $H_{k-1}^{(s)} = 0$ if $k = 1$.

Proposition 6. *Let k be a positive integer. The Laurent series of the function $P_{1,k}(z) = 2(\psi(z+k+1) - \psi(z-k+1))$ is*

$$\frac{2}{z} + 2\sum_{s=0}^{\infty} \left((-1)^s H_k^{(s+1)} - H_{k-1}^{(s+1)}\right) z^s,$$

and this series converges in $0 < |z| < 1$.

Proof. We use the series expansions of $\psi(z+k+1)$ and $\psi(z-k+1)$. Let γ be Euler's constant. Using Equation (9) and the Taylor series expansion of $\psi(z+1)$ [6, Equation (48)], we have

$$\begin{aligned}\psi(z+k+1) &= \psi(z+1) + \sum_{\ell=1}^{k} \frac{1}{z+\ell} \\ &= -\gamma + \sum_{s=1}^{\infty} (-1)^{s+1}\zeta(s+1) z^s + \sum_{\ell=1}^{k} \frac{1}{z+\ell}\end{aligned}$$

for $|z| < 1$. Expanding each term in powers of z, we get

$$\begin{aligned}\psi(z+k+1) &= -\gamma + \sum_{s=1}^{\infty} (-1)^{s+1}\zeta(s+1) z^s + \sum_{\ell=1}^{k} \frac{1}{\ell} \sum_{s=0}^{\infty} \left(\frac{-z}{\ell}\right)^s \\ &= -\gamma + H_k + \sum_{s=1}^{\infty} \left((-1)^{s+1}\zeta(s+1) + (-1)^s H_k^{(s+1)}\right) z^s.\end{aligned}$$

By Equation (10), for $0 < |z| < 1$ we have

$$\begin{aligned}\psi(z-k+1) &= -\frac{1}{z} + \psi(z+1) - \sum_{\ell=1}^{k-1} \frac{1}{z-\ell} \\ &= -\frac{1}{z} - \gamma + \sum_{s=1}^{\infty} (-1)^{s+1}\zeta(s+1) z^s + \sum_{\ell=1}^{k-1} \frac{1}{\ell} \sum_{s=0}^{\infty} \left(\frac{z}{\ell}\right)^s \\ &= -\frac{1}{z} - \gamma + H_{k-1} + \sum_{s=1}^{\infty} \left((-1)^{s+1}\zeta(s+1) + H_{k-1}^{(s+1)}\right) z^s.\end{aligned}$$

The series for $\psi(z+k+1)$ and $\psi(z-k+1)$ give the series for $2(\psi(z+k+1) - \psi(z-k+1))$. □

More generally, the Laurent coefficients of $P_{m,k}(z)$ are of a form given by the following lemma. For fixed m and ℓ, note that $[z^\ell]P_{m,k}(z)$ is a symbolic expression in k, and so the coefficients in the $\mathbb{Q}$-linear combination guaranteed by the lemma do not depend on k.

Lemma 7. *For all positive integers m and all $\ell \in \mathbb{Z}$, the coefficient $[z^\ell]P_{m,k}(z)$ is a $\mathbb{Q}$-linear combination of expressions of the form*

$$\frac{1}{k^{r_1}} H_{k-1}^{(r_2)} \cdots H_{k-1}^{(r_j)},$$

where r_1 is a non-negative integer, $r_2, \dots, r_j$ are positive even integers, and $r_1 + r_2 + \cdots + r_j = \ell + m$.

Proof. If $m = 1$ then Proposition 6 implies $[z^\ell]P_{1,k}(z) = \frac{2}{k^{\ell+1}}$ if $\ell \geq 0$ is even and $[z^\ell]P_{1,k}(z) = -4H_{k-1}^{(\ell+1)} - 2\frac{1}{k^{\ell+1}}$ if $\ell \geq 0$ is odd, as desired. If $\ell = -1$ then $[z^\ell]P_{1,k}(z) = 2$, and if $\ell \leq -2$ then $[z^\ell]P_{1,k}(z) = 0$ is the empty sum. Hence the statement of the lemma is true for $m = 1$.

By induction using Equation (13), $z^2 P_{m,k}(z)$ is analytic at $z = 0$ for each $m \geq 1$. From Equation (13), if $\ell \geq -2$ we have

$$[z^\ell]P_{m,k}(z) = \left(\sum_{i=-1}^{\ell+2} [z^i]P_{1,k}(z)\,[z^{\ell-i}]P_{m-1,k}(z)\right) + (\ell+1)\,[z^{\ell+1}]P_{m-1,k}(z). \tag{14}$$

Now a second induction on m using Proposition 6 shows that $[z^\ell]P_{m,k}(z)$ is a $\mathbb{Q}$-linear combination of expressions of the form

$$\frac{1}{k^{r_1}} H_{k-1}^{(r_2)} \cdots H_{k-1}^{(r_j)}$$

where r_1 is non-negative integer, $r_2, \dots, r_j$ are positive even integers, and $r_1 + r_2 + \cdots + r_j = \ell + m$. □

In light of Lemma 4, we are interested in $\frac{1}{m!k^2}[z^{-2}]P_{m,k}(z)$. The following table shows the first several values, computed recursively, where we have replaced each $H_k^{(r)}$ that arises with $H_{k-1}^{(r)} + \frac{1}{k^r}$ to obtain the form in Lemma 7.

m	$\frac{1}{m!k^2}[z^{-2}]P_{m,k}(z)$
1	0
2	$\frac{1}{k^2}$
3	$\frac{2}{k^3}$
4	$\frac{1}{k^4} - \frac{2}{k^2}H_{k-1}^{(2)}$
5	$-\frac{4}{k^3}H_{k-1}^{(2)}$
6	$-\frac{2}{k^4}H_{k-1}^{(2)} + \frac{2}{k^2}H_{k-1}^{(2)}H_{k-1}^{(2)} - \frac{1}{k^2}H_{k-1}^{(4)}$
7	$\frac{4}{k^3}H_{k-1}^{(2)}H_{k-1}^{(2)} - \frac{2}{k^3}H_{k-1}^{(4)}$
8	$\frac{2}{k^4}H_{k-1}^{(2)}H_{k-1}^{(2)} - \frac{4}{3k^2}H_{k-1}^{(2)}H_{k-1}^{(2)}H_{k-1}^{(2)} - \frac{1}{k^4}H_{k-1}^{(4)} + \frac{2}{k^2}H_{k-1}^{(2)}H_{k-1}^{(4)} - \frac{2}{3k^2}H_{k-1}^{(6)}$
9	$-\frac{8}{3k^3}H_{k-1}^{(2)}H_{k-1}^{(2)}H_{k-1}^{(2)} + \frac{4}{k^3}H_{k-1}^{(2)}H_{k-1}^{(4)} - \frac{4}{3k^3}H_{k-1}^{(6)}$

We can now deduce, as a simple corollary of Lemma 7, that all Taylor coefficients of the Apéry function are $\mathbb{Q}$-linear combinations of multiple zeta values. The idea is to partition sums of $H_{k-1}^{(r_2)} \cdots H_{k-1}^{(r_j)}$ into sums over regions that yield multiple zeta values, as was done by Hoffman [4] to express the product of two multiple zeta values as a linear combination of multiple zeta values. This product gives one of the relations that Ihara, Kaneko, and Zagier [5] refer to as *double shuffle relations*. We illustrate with an example.

Example 8. The sum

$$\sum_{k_1 \geq 1} \frac{1}{k_1^{r_1}} H_{k_1-1}^{(r_2)} H_{k_1-1}^{(r_3)} = \sum_{k_1 \geq 1} \sum_{k_2=1}^{k_1-1} \sum_{k_3=1}^{k_1-1} \frac{1}{k_1^{r_1} k_2^{r_2} k_3^{r_3}}$$

occurs in the expression $\frac{1}{k^2}[z^{-2}]P_{m,k}(z)$ for $m = 6$ with $r_1 = r_2 = r_3 = 2$. We partition this sum into 3 sums that yield multiple zeta values:

$$\sum_{k_1>k_2=k_3>0} \frac{1}{k_1^{r_1} k_2^{r_2} k_3^{r_3}} + \sum_{k_1>k_2>k_3>0} \frac{1}{k_1^{r_1} k_2^{r_2} k_3^{r_3}} + \sum_{k_1>k_3>k_2>0} \frac{1}{k_1^{r_1} k_2^{r_2} k_3^{r_3}}$$
$$= \zeta(r_1, r_2 + r_3) + \zeta(r_1, r_2, r_3) + \zeta(r_1, r_3, r_2).$$

In general, the sum

$$\sum_{k_1 \geq 1} \frac{1}{k_1^{r_1}} H_{k_1-1}^{(r_2)} \cdots H_{k_1-1}^{(r_j)} = \sum_{k_1 \geq 1} \sum_{k_2=1}^{k_1-1} \cdots \sum_{k_j=1}^{k_1-1} \frac{1}{k_1^{r_1} k_2^{r_2} \cdots k_j^{r_j}} \tag{15}$$

can be partitioned into sums corresponding to weak orderings of the $j-1$ summation variables $k_2, \ldots, k_j$. For example, the regions of summation for $j = 4$ are

$$\begin{array}{lll}
k_1 > k_2 = k_3 = k_4 > 0, & & \\
k_1 > k_2 = k_3 > k_4 > 0, & k_1 > k_2 = k_4 > k_3 > 0, & k_1 > k_3 = k_4 > k_2 > 0, \\
k_1 > k_4 > k_2 = k_3 > 0, & k_1 > k_3 > k_2 = k_4 > 0, & k_1 > k_2 > k_3 = k_4 > 0, \\
k_1 > k_2 > k_3 > k_4 > 0, & k_1 > k_2 > k_4 > k_3 > 0, & k_1 > k_3 > k_2 > k_4 > 0, \\
k_1 > k_3 > k_4 > k_2 > 0, & k_1 > k_4 > k_2 > k_3 > 0, & k_1 > k_4 > k_3 > k_2 > 0.
\end{array}$$

The number of such orderings is the $(j-1)$st ordered Bell number [10, A000670]. Furthermore, each of these regions of summation yields a multiple zeta value, the sum of all of which is equal to the expression in Equation (15). For example, one of the multiple zeta values is

$$\sum_{k_1>k_3>k_2=k_4>0} \frac{1}{k_1^{r_1} k_2^{r_2} k_3^{r_3} k_4^{r_4}} = \zeta(r_1, r_3, r_2 + r_4).$$

Corollary 9. *Let $A(z) = \sum_{m \geq 0} a_m z^m$ be the Apéry function. For each $m \geq 1$, the coefficient a_m is a $\mathbb{Q}$-linear combination of weight-m multiple zeta values.*

Proof. By Lemmas 4 and 7,

$$a_m = \frac{1}{m!} \sum_{k \geq 1} \frac{1}{k^2} [z^{-2}] P_{m,k}(z)$$

is a $\mathbb{Q}$-linear combination of numbers of the form

$$\sum_{k \geq 1} \frac{1}{k^{2+r_1}} H_{k-1}^{(r_2)} \cdots H_{k-1}^{(r_j)}.$$

As in Example 8, the cube $\{1, \ldots, k-1\}^{j-1}$ can be partitioned into a disjoint finite union of sets S_ω indexed by weak orderings ω of $j-1$ objects. Each weak ordering ω yields the sum $\sum_{k_1 \geq 1} \frac{1}{k_1^{2+r_1}} \sum_{(k_2,\ldots,k_j) \in S_\omega} \frac{1}{k_2^{r_2} \cdots k_j^{r_j}}$, which equals a multiple zeta value whose first argument is $2 + r_1$ and whose remaining arguments are sums of $r_2, \ldots, r_j$ determined by ω. Each multiple zeta value converges because the first argument $2 + r_1$ is at least 2. $\square$

4. Proof of Theorem 1

Corollary 9 shows that a_m is a $\mathbb{Q}$-linear combination of multiple zeta values. In this section we express a_m as an explicit $\mathbb{Z}$-linear combination of multiple zeta values, as given by Theorem 1.

Equation (14) expresses the coefficients of $P_{m,k}(z)$ in terms of coefficients of $P_{1,k}(z)$ and $P_{m-1,k}(z)$. In Lemma 10 we write $[z^{-2}]P_{m,k}(z)$ in terms of coefficients of $P_{m-\ell,k}(z)$, and in Corollary 11, we express these coefficients directly in terms of coefficients of $P_{1,k}(z)$, which are given by Proposition 6. To make the following proof visually easier, we use the notation

$$\alpha_{i,k} := [z^i]P_{1,k}(z) = \begin{cases} -4H_{k-1}^{(i+1)} - \frac{2}{k^{i+1}} & \text{if } i \geq 0 \text{ is odd} \\ \frac{2}{k^{i+1}} & \text{if } i \geq 0 \text{ is even} \\ 2 & \text{if } i = -1 \\ 0 & \text{if } i \leq -2. \end{cases}$$

We define $q_{m,k}$ for all $m \geq 0$ as follows. Let $q_{0,k} = 1$, and recursively define

$$q_{m,k} := \frac{1}{m} \sum_{i=0}^{m-1} q_{i,k}\, \alpha_{m-1-i,k}. \tag{16}$$

The first few expressions are

$$\begin{aligned} q_{0,k} &= 1 \\ q_{1,k} &= \tfrac{2}{k} \\ q_{2,k} &= \tfrac{1}{k^2} - 2H_{k-1}^{(2)} \\ q_{3,k} &= -\tfrac{4}{k} H_{k-1}^{(2)} \\ q_{4,k} &= -H_{k-1}^{(4)} - \tfrac{2}{k^2} H_{k-1}^{(2)} + 2H_{k-1}^{(2)} H_{k-1}^{(2)} \\ q_{5,k} &= -\tfrac{2}{k} H_{k-1}^{(4)} + \tfrac{4}{k} H_{k-1}^{(2)} H_{k-1}^{(2)}. \end{aligned}$$

Lemma 10. *Let $m \in \mathbb{N}$, and let $1 \leq \ell \leq m-1$. Then*

$$[z^{-2}]P_{m,k}(z) = \ell! \sum_{i=0}^{\ell} q_{i,k}\, [z^{\ell-i-2}]P_{m-\ell,k}(z). \tag{17}$$

Proof. Our proof technique is to fix arbitrary m and to verify the statement by induction for $\ell = 1, \ldots, m-1$. If $\ell = 1$, since $P_{m,k}(z) = P_{1,k}(z)P_{m-1,k}(z) + P'_{m-1,k}(z)$ by Equation (13), we have

$$\begin{aligned} [z^{-2}]P_{m,k}(z) &= \alpha_{-1,k}\, [z^{-1}]P_{m-1,k}(z) + \alpha_{0,k}\, [z^{-2}]P_{m-1,k}(z) - [z^{-1}]P_{m-1,k}(z) \\ &= [z^{-1}]P_{m-1,k}(z) + \alpha_{0,k}\, [z^{-2}]P_{m-1,k}(z), \end{aligned}$$

and Equation (17) is verified for $\ell = 1$.

Now suppose that (17) is true; we will show that

$$[z^{-2}]P_{m,k}(z) = (\ell+1)! \sum_{i=0}^{\ell+1} q_{i,k}\, [z^{\ell-i-1}]P_{m-\ell-1,k}(z). \tag{18}$$

Equation (14) gives

$$\begin{aligned}[z^{\ell-2-j}]P_{m-\ell,k}(z) &= \sum_{i=-1}^{\ell-j} \alpha_{i,k}\,[z^{\ell-2-j-i}]P_{m-\ell-1,k}(z) + (\ell-j-1)\,[z^{\ell-j-1}]P_{m-\ell-1,k}(z) \\ &= \sum_{i=0}^{\ell-j} \alpha_{i,k}\,[z^{\ell-2-j-i}]P_{m-\ell-1,k}(z) + (\ell-j+1)\,[z^{\ell-j-1}]P_{m-\ell-1,k}(z)\end{aligned}$$

for $j = 0, \ldots, \ell$. Substituting this into (17), we have

$$\begin{aligned}&\frac{1}{\ell!}[z^{-2}]P_{m,k}(z) \\ &\quad= \sum_{j=0}^{\ell} q_{j,k}\,[z^{\ell-2-j}]P_{m-\ell,k}(z) \\ &\quad= \sum_{j=0}^{\ell}\sum_{i=0}^{\ell-j} q_{j,k}\alpha_{i,k}\,[z^{\ell-2-j-i}]P_{m-\ell-1,k}(z) + \sum_{j=0}^{\ell}(\ell-j+1)q_{j,k}\,[z^{\ell-j-1}]P_{m-\ell-1,k}(z).\end{aligned}$$

To collect like coefficients of $P_{m-\ell-1,k}(z)$, we make the substitution $i = h-1-j$ in the first sum and $j = h$ in the second. This gives

$$\begin{aligned}&\frac{1}{\ell!}[z^{-2}]P_{m,k}(z) \\ &\quad= \sum_{j=0}^{\ell}\sum_{h=j+1}^{\ell+1} q_{j,k}\alpha_{h-1-j,k}\,[z^{\ell-h-1}]P_{m-\ell-1,k}(z) + \sum_{h=0}^{\ell}(\ell-h+1)q_{h,k}\,[z^{\ell-h-1}]P_{m-\ell-1,k}(z) \\ &\quad= \sum_{h=1}^{\ell+1}\left(\sum_{j=0}^{h-1} q_{j,k}\alpha_{h-1-j,k}\right)[z^{\ell-h-1}]P_{m-\ell-1,k}(z) + \sum_{h=0}^{\ell}(\ell-h+1)q_{h,k}\,[z^{\ell-h-1}]P_{m-\ell-1,k}(z) \\ &\quad= \sum_{h=1}^{\ell+1} hq_{h,k}\,[z^{\ell-h-1}]P_{m-\ell-1,k}(z) + \sum_{h=0}^{\ell}(\ell-h+1)q_{h,k}\,[z^{\ell-h-1}]P_{m-\ell-1,k}(z)\end{aligned}$$

after applying Equation (16). Combining the sums gives

$$\frac{1}{\ell!}[z^{-2}]P_{m,k}(z) = (\ell+1)\sum_{h=0}^{\ell+1} q_{h,k}\,[z^{\ell-h-1}]P_{m-\ell-1,k}(z),$$

which is equivalent to Equation (18). □

Corollary 11. *For all $m \geq 2$,*

$$[z^{-2}]P_{m,k}(z) = (m-1)!\sum_{i=0}^{m-2} q_{i,k}\,\alpha_{m-3-i,k}. \tag{19}$$

In particular,

$$\frac{1}{m!}[z^{-2}]P_{m,k}(z) = q_{m-2,k}. \tag{20}$$

Proof. Letting $\ell = m-1$ in Equation (17) gives

$$[z^{-2}]P_{m,k}(z) = (m-1)!\sum_{i=0}^{m-1} q_{i,k}\,\alpha_{m-3-i,k}.$$

Since $\alpha_{-2,k} = [z^{-2}]P_{1,k}(z) = 0$, we can omit the summand for $i = m-1$. To see the last assertion, we use Equations (19) and (16), obtaining

$$\begin{aligned}\frac{1}{m!}[z^{-2}]P_{m,k}(z) &= \frac{(m-1)!}{m!}\sum_{i=0}^{m-2} q_{i,k}\,\alpha_{m-3-i,k} \\ &= \frac{1}{m}\left(\sum_{i=0}^{m-3} q_{i,k}\,\alpha_{m-3-i,k} + 2q_{m-2,k}\right) \\ &= \frac{1}{m}\left((m-2)q_{m-2,k} + 2q_{m-2,k}\right) \\ &= q_{m-2,k}. \qquad \square\end{aligned}$$

Next we define a finite analogue of the multiple zeta value $\zeta(s_2, \dots, s_j)$. These terms will appear in the proof of Theorem 13, from which Theorem 1 follows.

Definition. Define the *multiple harmonic number*

$$H_{k-1}^{(s_2,\dots,s_j)} := \sum_{k-1>k_2>\cdots>k_j>0} \frac{1}{k_2^{s_2}\cdots k_j^{s_j}}.$$

Note that $H_{k-1}^{()}$ is a sum over the one tuple of length 0, for which the summand is the empty product; therefore $H_{k-1}^{()} = 1$.

Remark 12. Thanks to Equations (16) and (20), we only ever need to multiply a multiple harmonic number of the form $H_k^{(s_2,\dots,s_j)}$ by a multiple harmonic number of the form $H_k^{(i)}$. In this case, the product is

$$\begin{aligned}&H_{k-1}^{(s_2,s_3,\dots,s_j)} \cdot H_{k-1}^{(i)} \\ &\quad = H_{k-1}^{(i,s_2,s_3,\dots,s_j)} + H_{k-1}^{(s_2,i,s_3,\dots,s_j)} + \cdots + H_{k-1}^{(s_2,s_3,\dots,s_j,i)} \qquad (21)\\ &\quad\quad + H_{k-1}^{(s_2+i,s_3,\dots,s_j)} + \cdots + H_{k-1}^{(s_2,s_3,\dots,s_j+i)}. \qquad (22)\end{aligned}$$

We call the summands in (21) *insertions* and the summands in (22) *additions.* Using this relation to convert products of multiple harmonic numbers into linear combinations of multiple harmonic numbers, we can write the first few terms of $(q_{m,k})_{m\geq 0}$ as follows.

$$\begin{aligned}q_{0,k} &= 1 \\ q_{1,k} &= \tfrac{2}{k} \\ q_{2,k} &= \tfrac{1}{k^2} - 2H_{k-1}^{(2)} \\ q_{3,k} &= -\tfrac{4}{k}H_{k-1}^{(2)} \\ q_{4,k} &= H_{k-1}^{(4)} - \tfrac{2}{k^2}H_{k-1}^{(2)} + 4H_{k-1}^{(2,2)} \\ q_{5,k} &= \tfrac{2}{k}H_{k-1}^{(4)} + \tfrac{8}{k}H_{k-1}^{(2,2)} \\ q_{6,k} &= \tfrac{1}{k^2}H_{k-1}^{(4)} - 2H_{k-1}^{(2,4)} - 2H_{k-1}^{(4,2)} + \tfrac{4}{k^2}H_{k-1}^{(2,2)} - 8H_{k-1}^{(2,2,2)} \\ q_{7,k} &= -\tfrac{4}{k}H_{k-1}^{(2,4)} - \tfrac{4}{k}H_{k-1}^{(4,2)} - \tfrac{16}{k}H_{k-1}^{(2,2,2)}\end{aligned}$$

Comparing $q_{m-2,k}$ to the Taylor coefficient a_m listed in Section 1, we see that they agree with the last statement of Corollary 11. Indeed, we convert from $q_{m-2,k}$ to

a_m by adding 2 to s_1 in each term $\frac{1}{k^{s_1}} H_k^{(s_2,\ldots,s_j)}$ to account for the factor $\frac{1}{k^2}$, and then summing over $k \geq 1$ to obtain a multiple zeta value $\zeta(2+s_1, s_2, \ldots, s_j)$.

The following theorem shows that the coefficient of $\frac{1}{k^{s_1}} H_{k-1}^{(s_2,\ldots,s_j)}$ in $q_{m-2,k}$ is equal to the integer that Theorem 1 claims is the coefficient of $\zeta(2+s_1, s_2, \ldots, s_j)$ in a_m. Recall that $\chi(m) = 0$ if m is even and $\chi(m) = 1$ if m is odd.

Theorem 13. *For each $m \geq 0$, let $S(m)$ be the set tuples $\mathbf{s} = (s_1, s_2, \ldots, s_j)$, with $j \geq 1$, of non-negative integers satisfying*

- $s_1 + s_2 + \cdots + s_j = m$,
- $s_1 = 1$ *if m is odd and* $s_1 \in \{0, 2\}$ *if m is even, and*
- $s_i \in \{2, 4\}$ *for all* $i \in \{2, \ldots, j\}$.

Let $e(\mathbf{s}) = |\{i : 2 \leq i \leq j$ and $s_i = 2\}|$. For each $m \geq 0$,

$$q_{m,k} = \sum_{\mathbf{s} \in S(m)} (-1)^{\frac{m-s_1}{2}} 2^{e(\mathbf{s})+\chi(m)} \frac{1}{k^{s_1}} H_{k-1}^{(s_2,\ldots,s_j)}. \tag{23}$$

We give some intuition for the proof of Theorem 13. The multiple harmonic numbers that do not appear in the expression for $q_{m,k}$ do arise recursively from products $H_{k-1}^{(s_2,s_3,\ldots,s_j)} \cdot H_{k-1}^{(i)}$ in (21)–(22), but they do so in three ways, and the three corresponding coefficients sum to zero. The multiple harmonic numbers that do appear in $q_{m,k}$ also arise in at most three ways, but here, somewhat mysteriously, the coefficients accumulate in the stated way.

Proof of Theorem 13. We show that both sides of Equation (23) satisfy the same recurrence and initial condition. The initial condition is straightforward. Namely, for $m = 0$ we have $q_{0,k} = 1$. Since $S(0) = \{(0)\}$,

$$\sum_{\mathbf{s} \in S(0)} (-1)^{\frac{0-s_1}{2}} 2^{e(\mathbf{s})+\chi(0)} \frac{1}{k^{s_1}} H_{k-1}^{(s_2,\ldots,s_j)} = (-1)^{\frac{0-0}{2}} 2^{0+0} \frac{1}{k^0} H_{k-1}^{()} = 1.$$

Therefore Equation (23) holds for $m = 0$.

By definition, $q_{m,k}$ satisfies the recurrence given by Equation (16), that is,

$$q_{m,k} = \frac{1}{m} \sum_{i=0}^{m-1} q_{i,k} \left(-4\chi(m-1-i) H_{k-1}^{(m-i)} + (-1)^{m-1-i} \frac{2}{k^{m-i}} \right).$$

It suffices to show, for all $m \geq 1$, that the right side of Equation (23), namely

$$\sum_{\mathbf{s} \in S(m)} (-1)^{\frac{m-s_1}{2}} 2^{e(\mathbf{s})+\chi(m)} \frac{1}{k^{s_1}} H_{k-1}^{(s_2,\ldots,s_j)}, \tag{24}$$

also satisfies this recurrence. For each tuple $\mathbf{t} = (t_1, t_2, \ldots, t_\ell)$ of non-negative integers with $t_1 + t_2 + \cdots + t_\ell = m$, we compare the coefficient of $\frac{1}{k^{t_1}} H_{k-1}^{(t_2,\ldots,t_\ell)}$ in (24) to the coefficient of $\frac{1}{k^{t_1}} H_{k-1}^{(t_2,\ldots,t_\ell)}$ in

$$\frac{1}{m} \sum_{i=0}^{m-1} \left(\sum_{\mathbf{s} \in S(i)} (-1)^{\frac{i-s_1}{2}} 2^{e(\mathbf{s})+\chi(i)} \frac{1}{k^{s_1}} H_{k-1}^{(s_2,\ldots,s_j)} \right) \left(-4\chi(m-1-i) H_{k-1}^{(m-i)} + (-1)^{m-1-i} \frac{2}{k^{m-i}} \right) \tag{25}$$

and show that they are equal. In the latter, for odd $m-1-i$ we rewrite the product $H_{k-1}^{(s_2,\ldots,s_j)} H_{k-1}^{(m-i)}$ as a sum of multiple harmonic numbers of the form (21) and (22). There are two cases. For a tuple $\mathbf{t}$, we will use the set of positions $P_a(\mathbf{t}) = \{j \in \{2, \dots, \ell\} : t_j = a\}$.

First assume $t_2, \dots, t_\ell \in \{2, 4\}$. The term $\frac{1}{k^{t_1}} H_{k-1}^{(t_2,\dots,t_\ell)}$ arises in the sum (25) in several possible ways: from insertions by inserting $m - i = 2$ or $m - i = 4$, from additions by adding $m - i = 2$ to an existing 2, and as $\frac{1}{k^{s_1}} H_{k-1}^{(t_2,\dots,t_\ell)} \cdot \frac{1}{k^{m-i}}$ for $m - i = t_1 - s_1$. This final product gives rise to $\frac{1}{k^{t_1}} H_{k-1}^{(t_2,\dots,t_\ell)}$ for each $s_1 \leq \min(t_1 - 1, 2)$; here $s_1 = t_1$ is excluded since the sum does not include $m - i = 0$, and $s_1 \leq 2$ by definition of $S(m)$. Therefore, if $t_1 \in \{0, 1, 2\}$ then the coefficient of $\frac{1}{k^{t_1}} H_{k-1}^{(t_2,\dots,t_\ell)}$ in (25) is

$$\frac{1}{m}\Bigg(\sum_{h \in P_2(\mathbf{t})} (-1)^{\frac{(m-2)-t_1}{2}} 2^{e(\mathbf{t})-1+\chi(m-2)} \cdot (-4) + \sum_{h \in P_4(\mathbf{t})} (-1)^{\frac{(m-4)-t_1}{2}} 2^{e(\mathbf{t})+\chi(m-4)} \cdot (-4)$$
$$+ \sum_{h \in P_4(\mathbf{t})} (-1)^{\frac{(m-2)-t_1}{2}} 2^{e(\mathbf{t})+1+\chi(m-2)} \cdot (-4)$$
$$+ \sum_{s_1=0}^{\min(t_1-1,2)} (-1)^{\frac{(m-t_1+s_1)-s_1}{2}} 2^{e(\mathbf{t})+\chi(m-t_1+s_1)} \cdot (-1)^{m-1-(m-t_1+s_1)} 2 \Bigg).$$

Factoring out common terms in this expression gives

$$(-1)^{\frac{m-t_1}{2}} \frac{2^{e(\mathbf{t})+\chi(m)}}{m} \Bigg(\sum_{h \in P_2(\mathbf{t})} 2 - \sum_{h \in P_4(\mathbf{t})} 4 + \sum_{h \in P_4(\mathbf{t})} 8$$
$$+ \sum_{s_1=0}^{\min(t_1-1,2)} (-1)^{t_1-s_1-1} 2^{\chi(m-t_1+s_1)-\chi(m)+1} \Bigg).$$

We show that the four sums in the parentheses add up to m. For the first three, we have

$$\begin{aligned}
\sum_{h \in P_2(\mathbf{t})} 2 - \sum_{h \in P_4(\mathbf{t})} 4 + \sum_{h \in P_4(\mathbf{t})} 8 &= \sum_{h \in P_2(\mathbf{t})} 2 + \sum_{h \in P_4(\mathbf{t})} 4 \\
&= t_2 + \cdots + t_\ell \\
&= m - t_1.
\end{aligned}$$

Since $m = t_1 + t_2 + \cdots + t_\ell \equiv t_1 \mod 2$, the final sum can be written

$$\sum_{s_1=0}^{\min(t_1-1,2)} (-2)^{\chi(s_1)-\chi(t_1)+1} = \begin{cases} 0 & \text{if } t_1 = 0 \\ 1 & \text{if } t_1 = 1 \\ 2 & \text{if } t_1 = 2 \\ 0 & \text{if } t_1 \geq 3. \end{cases} \tag{26}$$

Therefore it follows for $t_1 \in \{0, 1, 2\}$ that the coefficient of $\frac{1}{k^{t_1}} H_{k-1}^{(t_2,\dots,t_\ell)}$ in (25) is $(-1)^{\frac{m-t_1}{2}} 2^{e(\mathbf{t})+\chi(m)}$, which equals the coefficient of $\frac{1}{k^{t_1}} H_{k-1}^{(t_2,\dots,t_\ell)}$ in (24), as desired. If $t_1 \geq 3$, then the coefficient of $\frac{1}{k^{t_1}} H_{k-1}^{(t_2,\dots,t_\ell)}$ in (24) is 0. In (25), this term only arises as $\frac{1}{k^{s_1}} H_{k-1}^{(t_2,\dots,t_\ell)} \cdot \frac{1}{k^{m-i}}$, by the definition of $S(i)$, so Equation (26) shows that the coefficient of $\frac{1}{k^{t_1}} H_{k-1}^{(t_2,\dots,t_\ell)}$ in (25) is also 0.

Second, assume $t_h \notin \{2,4\}$ for some $h \geq 2$. Since $\frac{1}{k^{t_1}} H_{k-1}^{(t_2,\ldots,t_\ell)}$ does not appear in (24), we show that its coefficient in (25) is 0. This term only arises in (25) from insertions and additions, and there is exactly one index h such that $t_h \notin \{2,4\}$, since insertions and additions only modify one entry in a tuple. Moreover, t_h must be even. Since insertions and additions do not result in entries $t_h = 0$, we have $t_h \geq 6$. The term $\frac{1}{k^{t_1}} H_{k-1}^{(t_2,\ldots,t_\ell)}$ arises in (25) from insertions by inserting $m-i = t_h$ and from additions by adding $m-i = t_h - 2$ to an existing 2 or $m-i = t_h - 4$ to an existing 4. Therefore, if $t_1 \in \{0,1,2\}$ then the coefficient of $\frac{1}{k^{t_1}} H_{k-1}^{(t_2,\ldots,t_\ell)}$ in (25) is

$$\frac{1}{m}\Bigg((-1)^{\frac{(m-t_h)-t_1}{2}} 2^{e(\mathbf{t})+\chi(m-t_h)} \cdot (-4) + (-1)^{\frac{(m-t_h+2)-t_1}{2}} 2^{e(\mathbf{t})+1+\chi(m-t_h+2)} \cdot (-4) + (-1)^{\frac{(m-t_h+4)-t_1}{2}} 2^{e(\mathbf{t})+\chi(m-t_h+4)} \cdot (-4)\Bigg),$$

which simplifies to

$$(-1)^{\frac{m-t_h-t_1}{2}} \frac{2^{e(\mathbf{t})+\chi(m-t_h)}}{m} (-4+8-4) = 0.$$

If $t_1 \geq 3$, then $\frac{1}{k^{t_1}} H_{k-1}^{(t_2,\ldots,t_\ell)}$ does not appear in (25) since $s_1 = t_1$ for insertions and additions. $\square$

We are now ready to prove Theorem 1.

Proof of Theorem 1. The case $m = 1$ follows from Example 5, where we showed $a_1 = 0$, and the fact that the sum in Theorem 1 is empty since there are no tuples $(s_1, s_2, \ldots, s_j)$ with $s_1 = 3$ whose entries sum to 1.

Let $m \geq 2$. By Lemma 4, $a_m = \frac{1}{m!} \sum_{k \geq 1} \frac{1}{k^2} [z^{-2}] P_{m,k}(z)$. Equation (20) implies

$$a_m = \sum_{k \geq 1} \frac{1}{k^2} q_{m-2,k}.$$

By Theorem 13,

$$\begin{aligned} a_m &= \sum_{k \geq 1} \frac{1}{k^2} \sum_{\mathbf{s} \in S(m-2)} (-1)^{\frac{m-2-s_1}{2}} 2^{e(\mathbf{s})+\chi(m-2)} \frac{1}{k^{s_1}} H_{k-1}^{(s_2,\ldots,s_j)} \\ &= \sum_{\mathbf{s} \in S(m-2)} (-1)^{\frac{m-(2+s_1)}{2}} 2^{e(\mathbf{s})+\chi(m)} \zeta(2+s_1, s_2, \ldots, s_j). \end{aligned}$$

Theorem 1 follows, since the tuples in $S(m-2)$ have first entry $s_1 \in \{0,1,2\}$ and the tuples in Theorem 1 have first entry in $\{2,3,4\}$. $\square$

As discussed in Section 1, the coefficients $a_4, a_6, a_8, a_{10}, a_{12}$ can be written as $\mathbb{Q}$-linear combinations of fewer multiple zeta values than given by Theorem 1. The strategy given in the following example can be used to reduce a_m for all even $m \geq 4$.

Example 14. For $m = 10$, Theorem 1 gives

$$\begin{aligned} a_{10} = \zeta(2,4,4) - 2\zeta(4,2,4) - 2\zeta(4,4,2) \\ + 4\zeta(2,2,2,4) + 4\zeta(2,2,4,2) + 4\zeta(2,4,2,2) - 8\zeta(4,2,2,2) \\ + 16\zeta(2,2,2,2,2). \end{aligned}$$

We use insertions and additions analogous to (21) and (22) to rewrite several products $\zeta(s_1, s_2, \dots, s_j)\zeta(i)$. We first express $-2\zeta(4,2,4) - 2\zeta(4,4,2)$ in terms of $\zeta(2,4,4)$ and $\zeta(10)$. We have

$$\zeta(4,4)\zeta(2) = \zeta(2,4,4) + \zeta(4,2,4) + \zeta(4,4,2) + \zeta(6,4) + \zeta(4,6).$$

The relations $\zeta(4)\zeta(4) = 2\zeta(4,4) + \zeta(8)$ and $\zeta(4)\zeta(6) = \zeta(4,6) + \zeta(6,4) + \zeta(10)$ allow us to write

$$\begin{aligned} -2\zeta(4,2,4) - 2\zeta(4,4,2) &= 2\zeta(2,4,4) + 2\zeta(4)\zeta(6) - 2\zeta(10) - \zeta(4)^2\zeta(2) + \zeta(8)\zeta(2) \\ &= 2\zeta(2,4,4) - \tfrac{3}{40}\zeta(10) \end{aligned}$$

using $\zeta(2) = \frac{\pi^2}{6}$, $\zeta(4) = \frac{\pi^4}{90}$, $\zeta(6) = \frac{\pi^6}{945}$, $\zeta(8) = \frac{\pi^8}{9450}$, and $\zeta(10) = \frac{\pi^{10}}{93555}$. Next we rewrite

$$4\zeta(2,2,2,4) + 4\zeta(2,2,4,2) + 4\zeta(2,4,2,2).$$

For this we use

$$\begin{aligned} \zeta(2,2,2)\zeta(4) - \zeta(2,2,2,4) - \zeta(2,2,4,2) - \zeta(2,4,2,2) - \zeta(4,2,2,2) \\ = \zeta(2,2,6) + \zeta(2,6,2) + \zeta(6,2,2) \\ = \zeta(2,2)\zeta(6) - (\zeta(8,2) + \zeta(2,8)) \\ = \zeta(2,2)\zeta(6) - (\zeta(2)\zeta(8) - \zeta(10)). \end{aligned}$$

Therefore $4\zeta(2,2,2,4)+4\zeta(2,2,4,2)+4\zeta(2,4,2,2)$ can be written using $\zeta(2,2)\zeta(6)$, $\zeta(2,2,2)\zeta(4)$, $\zeta(4,2,2,2)$, and $\zeta(10)$. Finally, we use

$$\zeta(\underbrace{2,\dots,2}_{j}) = \frac{\pi^{2j}}{(2j+1)!}$$

(see for example [4]) to write $\zeta(2,2)$, $\zeta(2,2,2)$, and $\zeta(2,2,2,2,2)$. Consolidating these results, we obtain

$$a_{10} = \tfrac{7}{8}\zeta(10) + 3\zeta(2,4,4) - 12\zeta(4,2,2,2).$$

5. Lucas congruences modulo p^2

Gessel [2] proved three theorems on congruences for $A(n)$ where $n \geq 0$. In this section we generalize these theorems to $n \in \mathbb{Z}$, making substantial use of the reflection formula $A(-1-z) = A(z)$ from Proposition 2. We simplify one of the arguments by using the fact that we can differentiate $A(z)$. As an application, we show that some Apéry numbers satisfy a Lucas congruence modulo p^2.

First we generalize Gessel's result that the Apéry numbers satisfy a Lucas congruence modulo p [2, Theorem 1].

Theorem 15. *Let p be a prime. For all $d \in \{0, 1, \dots, p-1\}$ and for all $n \in \mathbb{Z}$, we have $A(d + pn) \equiv A(d)A(n) \mod p$.*

Proof. Gessel proved the statement for $n \geq 0$. Let $n \leq -1$. By Proposition 2,

$$\begin{aligned} A(d+pn) &= A(-1-(d+pn)) \\ &= A((p-1-d)+p\,(-1-n)) \\ &\equiv A(p-1-d)A(-1-n) \mod p \\ &= A(p-1-d)A(n). \end{aligned}$$

Malik & Straub [7, Lemma 6.2] proved that $A(p-1-d) \equiv A(d) \mod p$, which completes the proof. □

Next we generalize Gessel's congruence for $A(pn)$ modulo p^3 for $p \geq 5$ and variants for $p=2$ and $p=3$ [2, Theorem 3].

Theorem 16. *For all $n \in \mathbb{Z}$,*

- $A(n) \equiv 5^n \mod 8$ *for all $n \geq 0$ and $A(n) \equiv 5^{n+1} \mod 8$ for all $n \leq -1$,*
- $A(d+3n) \equiv A(d)A(n) \mod 9$ *for all $d \in \{0,1,2\}$, and*
- $A(pn) \equiv A(n) \equiv A(pn+p-1) \mod p^3$ *for all primes $p \geq 5$.*

A special case of a theorem of Straub [11, Theorem 1.2] shows that $A(pn) \equiv A(n) \mod p^3$ for all $n \in \mathbb{Z}$ and all primes $p \geq 5$. We prove this result another way, using an approach similar to Gessel's.

Proof of Theorem 16. Gessel proved $A(n) \equiv 5^n \mod 8$ for all $n \geq 0$. For $n \leq -1$, we use Proposition 2 to write

$$\begin{aligned} A(n) = A(-1-n) &\equiv 5^{-1-n} \mod 8 \\ &\equiv 5^{1+n} \mod 8 \end{aligned}$$

since $5^{-1} \equiv 5 \mod 8$.

For $p=3$, the proof is similar to the proof of Theorem 15. Gessel proved the statement for $n \geq 0$, so for $n \leq -1$ we have

$$\begin{aligned} A(d+3n) &= A(-1-(d+3n)) \\ &= A((2-d)+3(-1-n)) \\ &\equiv A(2-d)A(-1-n) \mod 9 \\ &\equiv A(d)A(n) \mod 9 \end{aligned}$$

since one checks that $A(2-d) \equiv A(d) \mod 9$.

Let $p \geq 5$. Gessel proved $A(pn) \equiv A(n) \mod p^3$ for all $n \geq 0$. We show $A(pn+p-1) \equiv A(n) \mod p^3$ for all $n \geq 0$. We write

$$\begin{aligned} A(pn+p-1) &= \sum_{k=0}^{pn+p-1} \binom{pn+p-1}{k}^2 \binom{pn+p-1+k}{k}^2 \\ &= \sum_{d=0}^{p-1} \sum_{m=0}^{n} \binom{pn+p-1}{pm+d}^2 \binom{p\,(n+m+1)+d-1}{pm+d}^2 \\ &= \sum_{d=0}^{p-1} \sum_{m=0}^{n} \binom{pn+p-1}{pm+d}^2 \frac{p^2(n+1)^2}{(p\,(n+m+1)+d)^2} \binom{p\,(n+m+1)+d}{pm+d}^2 \\ &= S_0 + S_1 \end{aligned}$$

where

$$S_0 = \sum_{m=0}^{n} \binom{pn+p-1}{pm}^2 \frac{(n+1)^2}{(n+m+1)^2} \binom{p\,(n+m+1)}{pm}^2$$

is the summand for $d = 0$, and

$$S_1 = \sum_{d=1}^{p-1} \sum_{m=0}^{n} \binom{pn+p-1}{pm+d}^2 \frac{p^2(n+1)^2}{(p\,(n+m+1)+d)^2} \binom{p\,(n+m+1)+d}{pm+d}^2.$$

For S_0, we have

$$\begin{aligned} S_0 &= \sum_{m=0}^{n} \frac{(pn+p-pm)^2}{(pn+p)^2} \binom{pn+p}{pm}^2 \frac{(n+1)^2}{(n+m+1)^2} \binom{p\,(n+m+1)}{pm}^2 \\ &\equiv \sum_{m=0}^{n} \frac{(n-m+1)^2}{(n+m+1)^2} \binom{n+1}{m}^2 \binom{n+m+1}{m}^2 \mod p^3 \\ &= \sum_{m=0}^{n} \binom{n}{m}^2 \binom{n+m}{m}^2 \\ &= A(n) \end{aligned}$$

by Ljunggren's congruence $\binom{pa}{pb} \equiv \binom{a}{b} \mod p^3$, which holds for all primes $p \geq 5$ [3]. For S_1, we have

$$\begin{aligned} S_1 &\equiv p^2 \sum_{d=1}^{p-1} \sum_{m=0}^{n} \binom{pn+p-1}{pm+d}^2 \frac{(n+1)^2}{d^2} \binom{p\,(n+m+1)+d}{pm+d}^2 \mod p^3 \\ &\equiv p^2 \sum_{d=1}^{p-1} \sum_{m=0}^{n} \binom{p-1}{d}^2 \binom{n}{m}^2 \frac{(n+1)^2}{d^2} \binom{d}{d}^2 \binom{n+m+1}{m}^2 \mod p^3 \end{aligned}$$

by the Lucas congruence for binomial coefficients modulo p. Since

$$\binom{p-1}{d} = \frac{(p-1)(p-2)\cdots(p-d)}{1\cdot 2\cdots d} \equiv \frac{(-1)(-2)\cdots(-d)}{1\cdot 2\cdots d} \equiv (-1)^d \mod p,$$

we obtain

$$\begin{aligned} S_1 &\equiv p^2 \left(\sum_{d=1}^{p-1} \frac{1}{d^2}\right) \sum_{m=0}^{n} \binom{n}{m}^2 (n+1)^2 \binom{n+m+1}{m}^2 \mod p^3 \\ &\equiv 0 \mod p^3 \end{aligned}$$

since $H_{p-1}^{(2)} = \sum_{d=1}^{p-1} \frac{1}{d^2} \equiv 0 \mod p$, as established by Wolstenholme [12]. Therefore $A(pn+p-1) = S_0 + S_1 \equiv A(n) \mod p^3$.

Now for $n \leq -1$ we have

$$\begin{aligned} A(pn) &= A(-1-pn) \\ &= A((p-1) + p\,(-1-n)) \\ &\equiv A(-1-n) \mod p^3 \\ &= A(n) \end{aligned}$$

and

$$\begin{aligned} A(pn+p-1) &= A(-1-(pn+p-1)) \\ &= A(p\,(-1-n)) \\ &\equiv A(-1-n) \mod p^3 \\ &= A(n). \end{aligned}$$ □

Finally, we generalize Gessel's congruence for $A(d+pn)$ modulo p^2 [2, Theorem 4]. Recall that $A'(n)$ is given by Equation (2). Since $A'(n) \in \mathbb{Q}$ for every $n \geq 0$, it follows that if the denominator of $A'(n)$ is not divisible by p then we can interpret $A'(n)$ modulo p^2.

Theorem 17. *Let p be a prime, and let $d \in \{0, 1, \ldots, p-1\}$. The denominator of $A'(d)$ is not divisible by p. Moreover, for all $n \in \mathbb{Z}$,*

$$A(d+pn) \equiv (A(d) + pnA'(d))\, A(n) \mod p^2 \tag{27}$$

Proof. Gessel proved the statement for $n \geq 0$. The same approach allows us to prove the general case.

Fix $n \in \mathbb{Z}$. For each $d \in \{0, 1, \ldots, p-1\}$, define $c_d \in \{0, 1, \ldots, p-1\}$ such that $A(d+pn) \equiv A(d)A(n) + pc_d \mod p^2$; this can be done by Theorem 15. Let $c_{-1} = 0$. (The value of c_{-1} does not actually matter, since it will be multiplied by 0.) We show that $(c_d)_{0 \leq d \leq p-1}$ and $(nA'(d)A(n))_{0 \leq d \leq p-1}$ satisfy the same recurrence and initial conditions modulo p; this will imply $c_d \equiv nA'(d)A(n) \mod p$. Theorem 16 implies that $A(pn) \equiv A(n) \mod p^2$, so $c_0 = 0$. Since $A'(0) = 0$, the initial conditions are equal.

Let $d \in \{1, 2, \ldots, p-1\}$. Write Equation (1) as

$$\sum_{i=0}^{2} r_i(n)A(n-i) = 0, \tag{28}$$

where each $r_i(n)$ is a polynomial in n with integer coefficients. Note that Equation (28) holds for all $n \in \mathbb{Z}$. Substituting $d + pn$ for n in Equation (28) gives

$$\sum_{i=0}^{2} r_i(d+pn)A(d-i+pn) = 0.$$

If $d-i = -1$ then $r_i(d+pn) = r_2(1+pn) = (pn)^3 \equiv 0 \mod p^2$, hence the arbitrary value of c_{-1}. Therefore, using the Taylor expansion of $r_i(n)$, we have

$$\sum_{i=0}^{2} \big(r_i(d) + pnr_i'(d)\big)\big(A(d-i)A(n) + pc_{d-i}\big) \equiv 0 \mod p^2.$$

Since $\sum_{i=0}^{2} r_i(d)A(d-i) = 0$, expanding and dividing by p gives

$$\sum_{i=0}^{2} \big(r_i(d)c_{d-i} + nr_i'(d)A(d-i)A(n)\big) \equiv 0 \mod p.$$

This gives a recurrence satisfied by $(c_d)_{0 \leq d \leq p-1}$ that can be used to compute $c_1, c_2, \ldots, c_{p-1}$ since $r_0(d) = d^3 \not\equiv 0 \mod p$.

To obtain a recurrence for $(nA'(d)A(n))_{0\le d\le p-1}$, we differentiate Equation (28) to obtain

$$\sum_{i=0}^{2}\big(r_i(d)A'(d-i)+r_i'(d)A(d-i)\big)=0.$$

Since $A'(0)$ and $A'(1)$ are integers and $r_0(d)\not\equiv 0 \mod p$, the denominator of $A'(d)$ is not divisible by p. Multiplying by $nA(n)$ gives

$$\sum_{i=0}^{2}\big(r_i(d)nA'(d-i)A(n)+nr_i'(d)A(d-i)A(n)\big)=0.$$

Subtracting this from the recurrence for $(c_d)_{0\le d\le p-1}$ shows that

$$\sum_{i=0}^{2} r_i(d)\big(c_{d-i}-nA'(d-i)A(n)\big)\equiv 0 \mod p.$$

Since $r_0(d)\not\equiv 0 \mod p$, it follows that $c_d \equiv nA'(d)A(n) \mod p$ for all $d\in\{0,1,\dots,p-1\}$. □

In the case $p=3$, Theorem 17 gives a second proof of the congruence $A(d+3n)\equiv A(d)A(n) \mod 9$ from Theorem 16, since $A'(0)\equiv A'(1)\equiv A'(2)\equiv 0 \mod 3$. For larger primes, in general $A(d+pn)\not\equiv A(d)A(n) \mod p^2$. However, if we restrict to certain sets of base-p digits, then we do obtain congruences that hold modulo p^2. For example, if $d\in\{0,2,4\}$, then

$$A(d+5n)\equiv A(d)A(n) \mod 25.$$

This was proven by the authors [9] by computing an automaton for $A(n) \bmod 25$. Since $A(0)\equiv 1\equiv A(4) \mod 25$ and $A(2)\equiv 23 \mod 25$, this implies $A(n)\equiv 23^{e_2(n)} \mod 25$ for all $n\ge 0$ whose base-5 digits belong to $\{0,2,4\}$, where $e_2(n)$ is the number of 2s in the base-5 representation of n. The following theorem generalizes this result to other primes.

We say that the set $D\subseteq\{0,1,\dots,p-1\}$ supports a *Lucas congruence* for the sequence $s(n)_{n\in\mathbb{Z}}$ modulo p^α if $s(d+pn)\equiv s(d)s(n) \mod p^\alpha$ for all $d\in D$ and for all $n\in\mathbb{Z}$. As mentioned in the proof of Theorem 15, Malik and Straub [7, Lemma 6.2] proved that $A(d)\equiv A(p-1-d) \mod p$ for each $d\in\{0,1,\dots,p-1\}$. Let $D(p)$ be the set of base-p digits for which this congruence holds modulo p^2; that is,

$$D(p)=\left\{d\in\{0,1,\dots,p-1\} : A(d)\equiv A(p-1-d) \mod p^2\right\}.$$

Theorem 18. *The set $D(p)$ is the maximum set of digits that supports a Lucas congruence for the Apéry numbers modulo p^2.*

Proof. Let $d\in D(p)$, so that $A(d)\equiv A(p-1-d) \mod p^2$. Letting $n=-1$ in Theorem 17 gives $A(d-p)\equiv A(d)-pA'(d) \mod p^2$. Applying Proposition 2, we find

$$\begin{aligned} pA'(d) &\equiv A(d)-A(d-p) \mod p^2\\ &= A(d)-A(p-1-d)\\ &\equiv 0 \mod p^2.\end{aligned}$$

Therefore it follows from Theorem 17 that, for all $n\in\mathbb{Z}$,

$$\begin{aligned} A(d+pn) &\equiv \left(A(d)+pnA'(d)\right)A(n) \mod p^2\\ &\equiv A(d)A(n) \mod p^2.\end{aligned}$$

Therefore $D(p)$ supports a Lucas congruence for the Apéry numbers modulo p^2.

To see that $D(p)$ is the maximum such set, assume $A(d+pn) \equiv A(d)A(n) \mod p^2$ for all $n \in \mathbb{Z}$. Then

$$\begin{aligned}(A(d) + pnA'(d))\, A(n) &\equiv A(d+pn) \mod p^2\\ &\equiv A(d)A(n) \mod p^2,\end{aligned}$$

and it follows that $pnA'(d)A(n) \equiv 0 \mod p^2$ for all $n \in \mathbb{Z}$. Therefore $A(d) - A(p-1-d) = A(d) - A(d-p) \equiv pA'(d) \equiv 0 \mod p^2$. □

As a special case, we obtain the following congruence, since $\{0, p-1\} \subseteq D(p)$ by Theorem 16, and $A(0) = 1 \equiv A(p-1) \mod p^2$.

Corollary 19. *Let $p \neq 2$ and $n \geq 0$. If the base-p digits of n all belong to $\{0, \frac{p-1}{2}, p-1\}$, then $A(n) \equiv A(\frac{p-1}{2})^{e(n)} \mod p^2$ where $e(n)$ is the number of occurrences of the digit $\frac{p-1}{2}$.*

These are the first several primes with digit sets $D(p)$ containing at least 4 digits:

p	$D(p)$
7	$\{0, 2, 3, 4, 6\}$
23	$\{0, 7, 11, 15, 22\}$
43	$\{0, 5, 18, 21, 24, 37, 42\}$
59	$\{0, 6, 29, 52, 58\}$
79	$\{0, 18, 39, 60, 78\}$
103	$\{0, 17, 51, 85, 102\}$
107	$\{0, 14, 21, 47, 53, 59, 85, 92, 106\}$
127	$\{0, 17, 63, 109, 126\}$
131	$\{0, 62, 65, 68, 130\}$
139	$\{0, 68, 69, 70, 138\}$
151	$\{0, 19, 75, 131, 150\}$
167	$\{0, 35, 64, 83, 102, 131, 166\}$

A natural question, which we do not address here, is the following. How big can $|D(p)|$ be, as a function of p?

Theorem 16 implies the following Lucas congruence modulo p^3.

Theorem 20. *Let $p \geq 5$ and $n \geq 0$. If the base-p digits of n all belong to $\{0, p-1\}$, then $A(n) \equiv 1 \mod p^3$.*

Experiments do not suggest the existence of any additional Lucas congruences for the Apéry numbers modulo p^3. We leave this as open question.

References

[1] Francis Brown, Mixed Tate motives over $\mathbb{Z}$, *Annals of Mathematics* **175** (2012) 949–976.

[2] Ira Gessel, Some congruences for Apéry numbers, *Journal of Number Theory* **14** (1982) 362–368.

[3] Andrew Granville, Binomial coefficients modulo prime powers, *Canadian Mathematical Society Conference Proceedings* **20** (1997) 253–275.

[4] Michael E. Hoffman, Multiple harmonic series, *Pacific Journal of Mathematics* **152** (1992) 275–290.

[5] Kentaro Ihara, Masanobu Kaneko and Don Zagier, Derivation and double shuffle relations for multiple zeta values, *Compositio Mathematica* **142** (2006) 307–338.

[6] J. L. W. V. Jensen and T. H. Gronwall, An elementary exposition of the theory of the gamma function, *Annals of Mathematics* **17** (1916) 124–166.

[7] Amita Malik and Armin Straub, Divisibility properties of sporadic Apéry-like numbers, *Research in Number Theory* **2** (2016) Article 5.

[8] Robert Osburn and Armin Straub, Interpolated sequences and critical L-values of modular forms, Chapter 14 in *Elliptic Integrals, Elliptic Functions and Modular Forms in Quantum Field Theory* (2019), Texts & Monographs in Symbolic Computation, Springer.

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

[10] Neil Sloane et al., The On-Line Encyclopedia of Integer Sequences, `http://oeis.org`.

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

[12] Joseph Wolstenholme, On certain properties of prime numbers, *The Quarterly Journal of Pure and Applied Mathematics* **5** (1862) 35–39.

[13] Don Zagier, Evaluation of the multiple zeta values $\zeta(2,\ldots,2,3,2,\ldots,2)$, *Annals of Mathematics* **175** (2012) 977–1000.

[14] Don Zagier, Arithmetic and topology of differential equations, *Proceedings of the 2016 ECM* (2017).

Department of Mathematics, Hofstra University, Hempstead, NY, USA

School of Mathematics and Statistics, Open University, Milton Keynes, UK