THE ITERATED INTEGRALS OF $\ln(1 + x^n)$

TEWODROS AMDEBERHAN, CHRISTOPH KOUTSCHAN, VICTOR H. MOLL,
AND ERIC S. ROWLAND

Abstract. For a polynomial P, we consider the sequence of iterated integrals of $\ln P(x)$. This sequence is expressed in terms of the zeros of $P(x)$. In the special case of $\ln(1 + x^2)$, arithmetic properties of certain coefficients arising are described. Similar observations are made for $\ln(1 + x^3)$.

1. Introduction

The evaluation of integrals, a subject that had an important role in the 19^{th} century, has been given a new life with the development of symbolic mathematics software such as *Mathematica* or *Maple*. The question of indefinite integrals was provided with an algorithmic approach beginning with work of J. Liouville [8] discussed in detail in Chapter IX of Lutzen [9]. A more modern treatment can be found in Ritt [21], R. H. Risch [19, 20], and M. Bronstein [3].

The absence of a complete algorithmic solution to the problem of evaluation of definite integrals justifies the validity of tables of integrals such as [1, 4, 18]. These collections have not been superseded, yet, by the software mentioned above.

The point of view illustrated in this paper is that the quest for evaluation of definite integrals may take the reader to unexpected parts of mathematics. This has been described by one of the authors in [14, 15]. The goal here is to consider the sequence of *iterated integrals* of a function $f_0(x)$, defined by

$$f_n(x) = \int_0^x f_{n-1}(t)\, dt \quad \text{if } n \geq 1. \tag{1.1}$$

This formula carries the implicit normalization $f_n(0) = 0$ for $n \geq 1$.

A classical formula for the iterated integrals is given by

$$f_n(x) = \frac{d^{-n}}{dx^{-n}} f(x) = \frac{1}{(n-1)!} \int_0^x f_0(t)\,(x-t)^{n-1}\, dt. \tag{1.2}$$

Expanding the kernel $(x - t)^{n-1}$ gives f_n in terms of the moments

$$M_j(x) = \int_0^x t^j f_0(t)\, dt \tag{1.3}$$

as

$$f_n(x) = \sum_{j=0}^{n-1} (-1)^j \frac{x^{n-1-j}}{j!\,(n-1-j)!} M_j(x). \tag{1.4}$$

Date: June 8, 2011.

2010 *Mathematics Subject Classification.* Primary 26A09, Secondary 11A25.

Key words and phrases. Iterated integrals, harmonic numbers, recurrences, valuations, hypergeometric functions.

The work presented here deals with the sequence starting at $f_0(x) = \ln(1+x^N)$. The main observation is that the closed-form expression of the iterated integrals contains a pure polynomial term and a linear combination of transcendental functions with polynomial coefficients. Some arithmetical properties of the pure polynomial term are described.

2. The iterated integral of $\ln(1+x)$

The iterated integral of $f_0(x) = \ln(1+x)$ was described in [13]. This sequence has the form

$$f_n(x) = A_{n,1}(x) + B_{n,1}(x)\ln(1+x) \tag{2.1}$$

where

$$A_{n,1}(x) = -\frac{1}{n!}\sum_{k=1}^{n}\binom{n}{k}(H_n - H_{n-k})\,x^k = -\frac{1}{n!}\sum_{k=1}^{n}\frac{x^k(x+1)^{n-k}}{k}, \tag{2.2}$$

$$B_{n,1}(x) = \frac{1}{n!}(1+x)^n,$$

where $H_n = 1 + \frac{1}{2} + \cdots + \frac{1}{n}$ is the nth harmonic number.

The expression for $B_{n,1}(x)$ is easily guessed from the symbolic computation of the first few values. The corresponding closed form for $A_{n,1}(x)$ was more difficult to find experimentally. Its study began with the analysis of its denominators, denoted here by $\alpha_{n,1}$. The fact that the ratio

$$\beta_{n,1} := \frac{\alpha_{n,1}}{n\,\alpha_{n-1,1}} \tag{2.3}$$

satisfies

$$\beta_{n,1} = \begin{cases} p & \text{if } n \text{ is a power of the prime } p \\ 1 & \text{otherwise} \end{cases} \tag{2.4}$$

was the critical observation in obtaining the closed form $A_{n,1}(x)$ given in (2.2). We recognize $\beta_{n,1}$ as $e^{\Lambda(n)}$, where

$$\Lambda(n) = \begin{cases} \ln p & \text{if } n \text{ is a power of the prime } p \\ 0 & \text{otherwise} \end{cases} \tag{2.5}$$

is the von Mangoldt function. This yields

$$\alpha_{n,1} = n!\prod_{j=2}^{n}\beta_{j,1} = n!\prod_{j=2}^{n}e^{\Lambda(j)},$$

and the relation

$$e^{\Lambda(n)} = \frac{\operatorname{lcm}(1,\ldots,n)}{\operatorname{lcm}(1,\ldots,n-1)} \tag{2.6}$$

shows that

$$\alpha_{n,1} = n!\operatorname{lcm}(1,\ldots,n). \tag{2.7}$$

Note 2.1. The harmonic number H_n appearing in (2.2) has challenging arithmetical properties. Written in reduced form as

$$H_n = \frac{N_n}{D_n}, \tag{2.8}$$

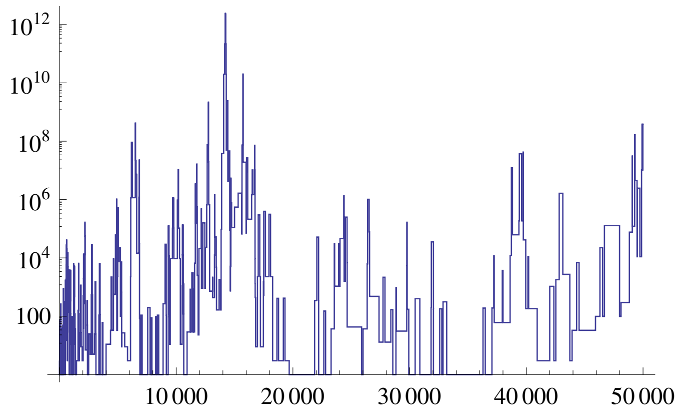

FIGURE 1. Logarithmic plot of the ratio L_n/D_n.

the denominator D_n divides the least common multiple $L_n := \text{lcm}(1, 2, \ldots, n)$. The complexity of the ratio L_n/D_n can be seen in Figure 1. It has been conjectured [5, page 304] that $D_n = L_n$ for infinitely many values of n.

The expressions for $A_{n,1}(x)$ and $B_{n,1}(x)$ can also be derived from (1.4). Letting $f_0(x) = \ln(1+x)$ yields

$$f_n(x) = \sum_{j=0}^{n-1} (-1)^j \frac{x^{n-1-j}}{j!(n-1-j)!} \int_0^x t^j \ln(1+t)\, dt. \tag{2.9}$$

Integration by parts gives

$$\int_0^x t^j \ln(1+t)\, dt = \frac{x^{j+1}\ \ln(1+x)}{j+1} - \frac{1}{j+1} \int_0^x \frac{t^{j+1}\, dt}{1+t}. \tag{2.10}$$

Replacing in (2.9) shows that the contribution of the first term reduces simply to $x^n \ln(1+x)$. Therefore

$$f_n(x) = \frac{1}{n!} x^n \ln(1+x) + \frac{1}{n!} \sum_{j=1}^{n} (-1)^j \binom{n}{j} x^{n-j} \int_0^x \frac{t^j\, dt}{1+t}. \tag{2.11}$$

It remains to provide a closed form for the integrals

$$I_j := \int_0^x \frac{t^j}{1+t}\, dt. \tag{2.12}$$

These can be produced by elementary methods by writing

$$\frac{t^j}{1+t} = \frac{t^j - (-1)^j}{1+t} + \frac{(-1)^j}{1+t}. \tag{2.13}$$

Replacing in (2.11) gives

$$\begin{aligned} f_n(x) &= \frac{1}{n!} x^n \ln(1+x) \\ &+ \frac{1}{n!} \sum_{j=1}^{n} (-1)^j \binom{n}{j} x^{n-j} \int_0^x \frac{t^j - (-1)^j}{t+1} \, dt \\ &+ \frac{1}{n!} \sum_{j=1}^{n} \binom{n}{j} x^{n-j} \int_0^x \frac{dt}{1+t}. \end{aligned}$$

The first and last line add up to $(x+1)^n \ln(1+x)/n!$, which yields the closed-form expression for $B_{n,1}(x)$. Expanding the quotient in the second line produces

$$\frac{1}{n!} \sum_{j=1}^{n} (-1)^j \binom{n}{j} x^{n-j} \sum_{r=0}^{j-1} \frac{(-1)^r}{j-r} x^{j-r} = \frac{1}{n!} \sum_{j=0}^{n-1} \binom{n}{j} x^j \sum_{r=1}^{n-j} \frac{(-1)^r}{r} x^r. \tag{2.14}$$

The double sum can be written as

$$\frac{1}{n!} \sum_{j=0}^{n} \sum_{r=1}^{n-j} \binom{n}{j} \frac{(-1)^r}{r} x^{j+r} = \frac{1}{n!} \sum_{a=1}^{n} \left[\sum_{r=1}^{a} \binom{n}{a-r} \frac{(-1)^r}{r} \right] x^a. \tag{2.15}$$

The expression for $A_{n,1}(x)$ now follows from the identity

$$\sum_{r=1}^{a} \binom{n}{a-r} \frac{(-1)^r}{r} = -\binom{n}{a} \left[H_n - H_{n-a} \right]. \tag{2.16}$$

An equivalent form, with $m = n - a$, is given by

$$U(a) := \sum_{r=1}^{a} \frac{(-1)^{r-1} \binom{a}{r}}{r \binom{m+r}{r}} = H_{m+a} - H_m. \tag{2.17}$$

To establish this identity, we employ the WZ method [16]. Define the pair of functions

$$F(r,a) = \frac{(-1)^{r-1} \binom{a}{r}}{r \binom{m+r}{r}} \quad \text{and} \quad G(r,a) = \frac{(-1)^r \binom{a}{r-1}}{(m+a+1) \binom{m+r-1}{r-1}}. \tag{2.18}$$

It can be easily checked that

$$F(r, a+1) - F(r,a) = G(r+1, a) - G(r,a). \tag{2.19}$$

Summing both sides of this equation over r, from 1 to $a+1$, leads to

$$U(a+1) - U(a) = \frac{1}{m+a+1}. \tag{2.20}$$

Now sum this identity over a, from 1 to $k-1$, to obtain

$$U(k) - U(1) = \sum_{a=1}^{k-1} \frac{1}{m+a+1} = \sum_{a=m+2}^{m+k} \frac{1}{r} = H_{m+k} - H_{m+1}. \tag{2.21}$$

Combining this with the initial condition $U(1) = \frac{1}{m+1}$ gives the result.

3. The method of roots

The iterated integrals of the function $f_0(x) = \ln P(x)$ for a general polynomial

$$P(x) = \prod_{j=1}^{m} (x + z_j) \tag{3.1}$$

are now expressed in terms of the roots z_j using an explicit expression for the iterated integrals of $f_0(x) = \ln(x+a)$.

Theorem 3.1. The iterated integral of $f_0(x) = \ln(x+a)$ is given by

$$f_n(x) = -\frac{1}{n!}\sum_{k=1}^{n} \frac{x^k (x+a)^{n-k}}{k} - \frac{(x+a)^n - x^n}{n!} \ln a + \frac{(x+a)^n}{n!} \ln(x+a). \tag{3.2}$$

Proof. A symbolic calculation of the first few values suggests the ansatz $f_n(x) = S_n(x) + T_n(x)\ln(x+a)$ for some polynomials S_n, T_n. The relation $f_n' = f_{n-1}$ and the form of S_n, T_n given in (3.2) give the result by induction. $\square$

In the special case $P(x) = 1 + x^N$, the previous result can be made more explicit.

Theorem 3.2. Let $a = u + iv$ be a root of $1 + x^N = 0$. Then the contribution of a and $\bar{a} = u - iv$ to the iterated integral of $\ln(1+x^N)$ is given by

$$\begin{aligned} -\frac{1}{n!}\sum_{k=1}^{n} \frac{x^k}{k} &\left[(x+a)^{n-k} + (x+\bar{a})^{n-k}\right] \\ &+ \frac{1}{in!}\left[(x+a)^n - (x+\bar{a})^n\right] \arctan\left(\frac{vx}{1+ux}\right) \\ &+ \frac{(x+a)^n + (x+\bar{a})^n}{2n!} \ln[(1+ux)^2 + v^2x^2]. \end{aligned}$$

Proof. First observe that $\ln(x+a) - \ln a = \ln(\bar{a}x + 1)$; hence for $f_0(x) = \ln(x+a)$ Theorem 3.1 takes the form

$$f_n(x) = -\frac{1}{n!}\sum_{k=1}^{n} \frac{x^k (x+a)^{n-k}}{k} + \frac{x^n}{n!} \ln a + \frac{(x+a)^n}{n!} \ln(\bar{a}x + 1). \tag{3.3}$$

Since

$$\begin{aligned} \ln(ax+1) &= \ln|ax+1| + i\,\mathrm{Arg}(ax+1), \\ \ln(\bar{a}x+1) &= \ln|ax+1| - i\,\mathrm{Arg}(ax+1), \end{aligned}$$

and $\ln a + \ln \bar{a} = 2\ln|a| = 0$, it follows that the total contribution of a and $\bar{a}$ is given by

$$\begin{aligned} -\frac{1}{n!}\sum_{k=1}^{n} \frac{x^k}{k} &\left[(x+a)^{n-k} + (x+\bar{a})^{n-k}\right] + \frac{[(x+a)^n \ln(\bar{a}x+1) + (x+\bar{a})^n \ln(ax+1)]}{n!} \\ = -\frac{1}{n!}\sum_{k=1}^{n} \frac{x^k}{k} &\left[(x+a)^{n-k} + (x+\bar{a})^{n-k}\right] + \frac{[(x+a)^n + (x+\bar{a})^n]}{n!} \ln|ax+1| \\ &- \frac{i\,[(x+a)^n - (x+\bar{a})^n]}{n!} \mathrm{Arg}(ax+1). \end{aligned}$$

The stated result comes from expressing the logarithmic terms in their real and imaginary parts. $\square$

Corollary 3.3. Let $n \in \mathbb{N}$. Then

$$\sum_{k=1}^{n} \frac{1}{k} \int_0^x t^k (t+a)^{n-k} \, dt = \frac{1}{n+1} \sum_{k=1}^{n} \frac{x^k (x+a)^{n+1-k}}{k} + \frac{\left[x^{n+1} - (x+a)^{n+1} + a^{n+1}\right]}{(n+1)^2}.$$

Proof. Integrate both sides of the identity in Theorem 3.1 and use the relation $f_{n-1}' = f_n$ to obtain the result inductively. □

Note 3.4. The identity in Corollary 3.3 can be expressed in terms of the function

$$\Phi_n(x,a) := \sum_{k=1}^{n} \frac{1}{k} x^k (x+a)^{n-k} \tag{3.4}$$

in the form

$$\int_0^x \Phi_n(t,a) \, dt = \frac{x+a}{n+1} \Phi_n(x,a) + \frac{1}{(n+1)^2} \left[x^{n+1} + a^{n+1} - (x+a)^{n+1}\right]. \tag{3.5}$$

The function $\Phi_n(x,a)$ admits the hypergeometric representation

$$\Phi_n(x,a) = -\frac{x^{n+1}}{(n+1)(x+a)} \, {}_2F_1\left(\begin{matrix} 1, 1+n \\ 2+n \end{matrix} ; \frac{x}{x+a} \right) - (x+a)^n \ln\left(\frac{a}{x+a} \right).$$

With this representation, the identity in Corollary 3.3 now becomes

$$\int_0^x \left(\frac{t}{1-t} \right)^{n+1} {}_2F_1\left(\begin{matrix} 1, 1+n \\ 2+n \end{matrix} ; t \right) \frac{dt}{1-t} = \frac{1}{n+1} \left(\frac{x}{1-x} \right)^{n+1} \left[{}_2F_1\left(\begin{matrix} 1, 1+n \\ 2+n \end{matrix} ; x \right) - 1 \right].$$

4. The iterated integral of $\ln(1+x^2)$

In this section we consider the iterated integral of $f_0(x) = \ln(1+x^2)$ defined by

$$f_n(x) = \int_0^x f_{n-1}(t) \, dt. \tag{4.1}$$

The first few examples, given by

$$\begin{aligned}
f_1(x) &= -2x + 2 \arctan x + x \, \ln(1+x^2) \\
f_2(x) &= -\tfrac{3}{2} x^2 + 2x \, \arctan x + \tfrac{1}{2}(x^2-1) \ln(1+x^2) \\
f_3(x) &= -\tfrac{11}{18} x^3 + \tfrac{1}{3} x + (x^2 - \tfrac{1}{3}) \, \arctan x + \left(\tfrac{1}{6} x^3 - \tfrac{1}{2} x \right) \ln(1+x^2),
\end{aligned}$$

suggest the form

$$f_n(x) = A_{n,2}(x) + B_{n,2}(x) \arctan x + C_{n,2}(x) \ln(1+x^2) \tag{4.2}$$

for some polynomials $A_{n,2}, B_{n,2}, C_{n,2}$. Theorem 3.2 can be employed to obtain a closed form for these polynomials. It follows that $f_n(x)$ satisfies

$$\begin{aligned}
n! f_n(x) = &- \sum_{k=1}^{n} \frac{x^k}{k} \left[(x+i)^{n-k} + (x-i)^{n-k} \right] \\
&- i \left[(x+i)^n - (x-i)^n \right] \arctan x + \frac{1}{2} \left[(x+i)^n + (x-i)^n \right] \ln(1+x^2).
\end{aligned} \tag{4.3}$$

The expressions for $A_{n,2}$, $B_{n,2}$, $C_{n,2}$ may be read from here.

4.1. **Recurrences.** The polynomials $A_{n,2}, B_{n,2}, C_{n,2}$ can also be found as solutions to certain recurrences. Differentiation of (4.1) yields $f_n'(x) = f_{n-1}(x)$. It is easy to check that this relation, with the initial conditions $f_n(0) = 0$ and $f_0(x) = \ln(1+x^2)$, is equivalent to (4.1). Replacing the ansatz (4.2) produces

$$\begin{aligned} A_{n,2}'(x) + B_{n,2}'(x)\arctan x + \frac{B_{n,2}(x)}{1+x^2} + C_{n,2}'(x)\ln(1+x^2) + C_{n,2}(x)\frac{2x}{1+x^2} \\ = A_{n-1,2}(x) + B_{n-1,2}(x)\arctan x + C_{n-1,2}(x)\ln(1+x^2). \end{aligned}$$

A natural linear independence assumption yields the system of recurrences

$$\begin{aligned} &(4.4) & B_{n,2}'(x) &= B_{n-1,2}(x) \\ & & B_{0,2}(x) &= 0 \\ &(4.5) & C_{n,2}'(x) &= C_{n-1,2}(x) \\ & & C_{0,2}(x) &= 1 \\ &(4.6) & A_{n,2}'(x) &= A_{n-1,2}(x) - \frac{B_{n,2}(x) + 2xC_{n,2}(x)}{1+x^2} \\ & & A_{0,2}(x) &= 0. \end{aligned}$$

Note 4.1. The definition (4.1) determines completely the function $f_n(x)$. In particular, given the form (4.2), the polynomials $A_{n,2}, B_{n,2}$ and $C_{n,2}$ are uniquely specified. Observe however that the recurrence (4.4) does not determine $B_{n,2}(x)$ uniquely. At each step, there is a constant of integration to be determined. In order to address this ambiguity, the first few values of $B_{n,2}(0)$ are determined *empirically*, and the condition

$$B_{n,2}(0) = \begin{cases} 2(-1)^{\frac{n-1}{2}}/n! & \text{if } n \text{ is odd} \\ 0 & \text{if } n \text{ is even} \end{cases} \tag{4.7}$$

is added to the recurrence (4.4). The polynomials $B_{n,2}(x)$ are now uniquely determined. Similarly, the initial condition

$$C_{n,2}(0) = \begin{cases} (-1)^{\frac{n}{2}}/n! & \text{if } n \text{ is even} \\ 0 & \text{if } n \text{ is odd} \end{cases} \tag{4.8}$$

adjoined to (4.5), determines $C_{n,2}$. The initial condition imposed on $A_{n,2}$ is simply $A_{n,2}(0) = 0$.

The recurrence (4.4) is then employed to produce a list of the first few values of $B_{n,2}(x)$. These are then used to guess the closed-form expression for this family. The same is true for $C_{n,2}(x)$.

Proposition 4.2. The recurrence (4.4) and the (heuristic) initial condition (4.7) yield

$$\begin{aligned} B_{n,2}(x) &= \frac{2}{n!}\sum_{j=0}^{\frac{n-1}{2}} (-1)^j \binom{n}{2j+1} x^{n-2j-1} \\ &= \frac{1}{i\,n!}\left[(x+i)^n - (x-i)^n\right]. \end{aligned} \tag{4.9}$$

Similarly, the polynomial $C_{n,2}$ is given by

$$C_{n,2}(x) = \frac{1}{n!}\sum_{j=0}^{\lfloor \frac{n}{2} \rfloor} (-1)^j \binom{n}{2j} x^{n-2j} \tag{4.10}$$
$$= \frac{1}{2\,n!}\left[(x+i)^n + (x-i)^n\right].$$

In particular, the degree of $B_{n,2}$ is $n-1$, and the degree of $C_{n,2}$ is n.

Proof. This follows directly from the recurrences (4.4) and (4.5). □

Corollary 4.3. The recurrence for $A_{n,2}$ can be written as

$$A'_{n,2}(x) = A_{n-1,2}(x) - \frac{1}{n!}\left[(x+i)^{n-1} + (x-i)^{n-1}\right]. \tag{4.11}$$

In particular, the degree of $A_{n,2}$ is n.

Proof. Simply replace the explicit expressions for $B_{n,2}$ and $C_{n,2}$ in the recurrence (4.6). □

4.2. **Trigonometric forms.** A trigonometric form of the polynomials $B_{n,2}$ and $C_{n,2}$ is established next.

Proposition 4.4. The polynomials $B_{n,2}$ and $C_{n,2}$ are given by

$$B_{n,2}(x) = \frac{2}{n!}(x^2+1)^{n/2}\sin(n\,\mathrm{arccot}\,x)$$
$$C_{n,2}(x) = \frac{1}{n!}(x^2+1)^{n/2}\cos(n\,\mathrm{arccot}\,x).$$

In particular,

$$\frac{C_{n,2}(x)}{B_{n,2}(x)} = \frac{1}{2}\cot(n\,\mathrm{arccot}\,x). \tag{4.12}$$

Proof. The polar form

$$x + i = \sqrt{x^2+1}\,\left[\cos(\mathrm{arccot}\,x) + i\sin(\mathrm{arccot}\,x)\right] \tag{4.13}$$

produces

$$(x+i)^n = (x^2+1)^{n/2}\left[\cos(n\,\mathrm{arccot}\,x) + i\sin(n\,\mathrm{arccot}\,x)\right]. \tag{4.14}$$

A similar expression for $(x-i)^n$ gives the result. □

Proof. A second proof follows from the Taylor series

$$\frac{\sin(z\arctan t)}{(1+t^2)^{z/2}} = \sum_{k=0}^{\infty} \frac{(-1)^k (z)_{2k+1}}{(2k+1)!} t^{2k+1} \tag{4.15}$$

and

$$\frac{\cos(z\arctan t)}{(1+t^2)^{z/2}} = \sum_{k=0}^{\infty} \frac{(-1)^k (z)_{2k}}{(2k)!} t^{2k} \tag{4.16}$$

where $(z)_n$ denotes the Pochhammer symbol. These series were established in [2] in the context of integrals related to the Hurwitz zeta function.

Indeed, the formula for $B_{n,2}(x)$ comes from (4.15) replacing t by $1/x$ and z by $-n$ to obtain

$$\sin(n \operatorname{arccot} x)\,(x^2+1)^{n/2} = -x^n \sum_{k=0}^{\infty} \frac{(-1)^k(-n)_{2k+1}}{(2k+1)!} x^{-2k-1}. \tag{4.17}$$

The result (4.9) now follows from the identity

$$(-n)_{2k+1} = \begin{cases} -n!/(n-2k-1)! & \text{if } 2k+1 \leq n \\ 0 & \text{otherwise.} \end{cases} \tag{4.18}$$

A similar argument gives the form of $C_{n,2}(x)$ in (4.10). □

Note 4.5. The rational function R_n that gives

$$\cot(n\theta) = R_n(\cot\theta) \tag{4.19}$$

appears in (4.12) in the form

$$R_n(x) = \frac{2C_{n,2}(x)}{B_{n,2}(x)}. \tag{4.20}$$

This rational function plays a crucial role in the development of *rational Landen transformations* [10]. These are transformations of the coefficients of a rational integrand that preserve the value of a definite integral. For example, the map

$$\begin{aligned} a &\mapsto a\left((a+3c)^2-3b^2\right)/\Delta \\ b &\mapsto b\left(3(a-c)^2-b^2\right)/\Delta \\ c &\mapsto c\left((3a+c)^2-3b^2\right)/\Delta, \end{aligned}$$

where $\Delta = (3a+c)(a+3c)-b^2$, preserves the value of

$$\int_{-\infty}^{\infty} \frac{dx}{ax^2+bx+c} = \frac{2\pi}{\sqrt{4ac-b^2}}. \tag{4.21}$$

The reader will find in [12] a survey of this type of transformation and [11] the example given above. The reason for the appearance of $R_n(x)$ in the current context remains to be clarified.

4.3. **An automatic derivation of a recurrence for $A_{n,2}$.** The formula (1.2) for for the iterated integral can be used in the context of computer algebra methods. In the case discussed here, the integral

$$I_n(x) = \frac{1}{(n-1)!}\int_0^x (x-t)^{n-1}\ln(1+t^2)\,dt \tag{4.22}$$

gives the desired iterated integrals of $\ln(1+x^2)$ for $n \geq 1$.

A standard application of the holonomic systems approach, as implemented in the *Mathematica* package `HolonomicFunctions` [6], yields a recurrence in n for (4.22). The reader will find in [7] a description of the use of this package in the evaluation of definite integrals. The recurrence

$$\begin{aligned} n^2(n-1)I_n(x) &= x(3n-2)(n-1)\,I_{n-1}(x) - \left(3nx^2-4x^2+n\right)I_{n-2}(x) \\ &\qquad + x\left(x^2+1\right)I_{n-3}(x) \end{aligned} \tag{4.23}$$

is delivered immediately by the package. Using the linear independence of $\arctan x$ and $\ln(1+x^2)$, it follows that each of the sequences $A_{n,2}$, $B_{n,2}$, and $C_{n,2}$ must also satisfy the recurrence (4.23). Symbolic methods for solving recurrences are employed next to produce the explicit expressions for $A_{n,2}$, $B_{n,2}$, and $C_{n,2}$ given above.

Petkovšek's algorithm Hyper [17] (as implemented in the *Mathematica* package `Hyper`, for example) computes a basis of hypergeometric solutions of a linear recurrence with polynomial coefficients. Given (4.23) as input, it outputs the two solutions $(x+i)^n/n!$ and $(x-i)^n/n!$. The initial values are used to obtain the correct linear combinations of these solutions. This produces the expressions for $B_{n,2}(x)$ and $C_{n,2}(x)$ given in Proposition 4.2.

However, the third solution is not hypergeometric and it will give the polynomials $A_{n,2}(x)$. It can be found by Schneider's *Mathematica* package `Sigma` [22]:

$$A_{n,2}(x) = \frac{i}{n!}\left(x\left((x+i)^n-(x-i)^n\right)+\sum_{k=2}^{n}\frac{x^k\left((x-i)^{n-k+1}-(x+i)^{n-k+1}\right)}{(k-1)k}\right),$$

with the initial values

$$A_{0,2}(x) = 0, \qquad A_{1,2}(x) = -2x, \qquad A_{2,2}(x) = -\tfrac{3}{2}x^2.$$

In summary:

Theorem 4.6. Define $a_k = k(k-1)$ for $k \geq 2$ and $a_1 = -1$. The polynomial $A_{n,2}(x)$ introduced in (4.2) is given by

$$A_{n,2}(x) = \frac{1}{i\,n!}\sum_{k=1}^{n}\frac{x^k}{a_k}\left[(x+i)^{n-k+1}-(x-i)^{n-k+1}\right]. \tag{4.24}$$

This can be written as

$$A_{n,2}(x) = \frac{1}{n!}\sum_{k=1}^{n}\frac{(n-k+1)!}{a_k}x^kB_{n-k+1,2}(x). \tag{4.25}$$

Note that the expression for $A_{n,2}$ given before is equivalent to the forms appearing in Theorem 4.6.

Note 4.7. Similar procedures applied to the case of $\ln(1+x)$ yield the evaluation given in (2.2).

5. Arithmetical properties

In this section we discuss arithmetical properties of the polynomials $B_{n,2}$ and $A_{n,2}$. The explicit formula for $B_{n,2}$ produces some elementary results.

Proposition 5.1. *Let $m, n \in \mathbb{N}$ such that m divides n. Then $B_{m,2}(x)$ divides $B_{n,2}(x)$ as polynomials in $\mathbb{Q}[x]$.*

Proof. This follows directly from (4.9) and the divisibility of $a^n - b^n$ by $a^m - b^m$. □

For odd n, the quotient of $B_{2n,2}(x)$ by $B_{n,2}(x)$ admits a simple expression.

Proposition 5.2. *Let $n \in \mathbb{N}$. Define*

$$B^*_{n,2}(x) = x^{\deg B_{n,2}}B_{n,2}(1/x). \tag{5.1}$$

Then, for n odd,

$$\binom{2n}{n} B_{2n,2}(x) = (-1)^{\frac{n-1}{2}} x B_{n,2}(x) B^*_{n,2}(x). \tag{5.2}$$

In particular, the sequence of coefficients in $B_{2n}(x)$ is palindromic.

Proof. The proof is elementary. Observe that

$$\begin{aligned} B^*_{n,2}(x) &= \frac{x^{n-1}}{in!}\left[\left(\frac{1}{x}+i\right)^n - \left(\frac{1}{x}-i\right)^n\right] \\ &= \frac{1}{ixn!}\left[(1+ix)^n - (1-ix)^n\right] \\ &= \frac{i^{n-1}}{n!x}\left[(x-i)^n - (-1)^n(x+i)^n\right]. \end{aligned}$$

It follows that

$$B^*_{n,2}(x) = \frac{(-1)^{\frac{n-1}{2}}}{xn!}\left[(x+i)^n + (x-i)^n\right], \tag{5.3}$$

and the result now follows directly. □

The explicit expression (4.24) for the polynomial $A_{n,2}$ can be written in terms of the polynomials

$$\varphi_m(x) = (x+i)^m - (x-i)^m \tag{5.4}$$

as

$$A_{n,2}(x) = \frac{i}{n!}\left[x\varphi_n(x) - \sum_{k=2}^{n} \frac{x^k \varphi_{n-k+1}(x)}{k(k-1)}\right]. \tag{5.5}$$

The polynomial $A_{n,2}$ is of degree n and has rational coefficients.

By analogy with the properties of denominators of $A_{n,1}(x)$ mentioned in Section 2 and discussed at greater length in [13], we now study the denominators $A_{n,2}(x)$ from an arithmetic point of view. The first result is elementary.

Proposition 5.3. Let

$$\alpha_{n,2} := \text{denominator of } A_{n,2}(x). \tag{5.6}$$

Then $\alpha_{n,2}$ divides $n!\,\mathrm{lcm}(1,2,\ldots,n)$.

Proof. The result follows from (5.5) and the fact that the polynomials $\varphi_m(x)$ have integer coefficients. □

As in (2.3), it is useful to consider the ratio

$$\beta_{n,2} := \frac{\alpha_{n,2}}{n\,\alpha_{n-1,2}}. \tag{5.7}$$

Symbolic computations suggest the following.

Conjecture 5.4. The sequence $\beta_{n,2}$ is given by

$$\beta_{n,2} = \begin{cases} p & \text{if } n = p^r \text{ for some prime } p \text{ and } r \in \mathbb{N} \text{ and } n \neq 2\cdot 3^m + 1 \\ \frac{1}{3} & \text{if } n = 2\cdot 3^m \text{ for some } m \in \mathbb{N} \\ 3p & \text{if } n = 2\cdot 3^m + 1 \text{ and } n = p^r \text{ for some } m, r \in \mathbb{N} \\ 3 & \text{if } n = 2\cdot 3^m + 1 \text{ for some } m \in \mathbb{N} \text{ and } n \neq p^r \\ 1 & \text{otherwise.} \end{cases} \tag{5.8}$$

The formulation of this conjecture directly in terms of the denominators of $A_{n,2}(x)$ is as follows.

Conjecture 5.5. The denominator $\alpha_{n,2}$ of $A_{n,2}(x)$ is given by

$$\alpha_{n,2} = \begin{cases} 1 & \text{if } n = 1 \\ n!\,\operatorname{lcm}(1,2,\ldots,n)/6 & \text{if } n = 2\cdot 3^m \text{ for some } m \geq 1 \\ n!\,\operatorname{lcm}(1,2,\ldots,n)/2 & \text{otherwise.} \end{cases} \tag{5.9}$$

This conjecture shows that the cancellations produced by the polynomials $\varphi_m(x)$ in (5.5) have an arithmetical nature.

Proof that Conjecture 5.5 implies Conjecture 5.4. Assume that (5.9) holds for $n \geq 1$. If $n = 2\cdot 3^m$, then $\alpha_{n,2}$ contains one fewer power of 3 than $n\,\alpha_{n-1,2}$. If $n = 2\cdot 3^m + 1$, then $\alpha_{n,2}$ contains one more power of 3 than $n\,\alpha_{n-1,2}$. If $n = p^r$ is a prime power, then $\alpha_{n,2}$ contains one more power of p than $n\,\alpha_{n-1,2}$. Otherwise each prime appears the same number of times in $\alpha_{n,2}$ and $n\,\alpha_{n-1,2}$. ☐

The first reduction is obtained by expanding the inner sum in (5.5). Define

$$G_n(x) = -2i \sum_{k=0}^{\lfloor n/2\rfloor-1} (-1)^k \left[\sum_{j=2k+1}^{n-1} \frac{1}{(n-j)(n-j+1)} \binom{j}{2k+1} \right] x^{n-2k}. \tag{5.10}$$

Proposition 5.6. We have

$$A_{n,2}(x) = \frac{i}{n!}\left[x\left((x+i)^n - (x-i)^n\right) + G_n(x)\right]. \tag{5.11}$$

Proof. Expanding the terms $(x+i)^{n-k+1}$ and $(x-i)^{n-k+1}$ in the expression for $A_{n,2}(x)$ yields the sum

$$\sum_{j=1}^{n-1} \frac{x^{n+1-j}}{(n-j)(n-j+1)} \sum_{k=0}^{j} \binom{j}{k} x^{j-k} i^k \left((-1)^k - 1\right) \tag{5.12}$$

so only odd k contribute to it. Reversing the order of summation gives the result. ☐

The next result compares the denominator $\alpha_{n,2}$ of $A_{n,2}(x)$ and the denominator of G_n, denoted by γ_n.

Corollary 5.7. For $n \geq 1$, the denominators $\alpha_{n,2}$ and γ_n satisfy

$$\alpha_{n,2} = n!\,\gamma_n. \tag{5.13}$$

We now rephrase Conjecture 5.5 as the following.

Conjecture 5.8. For $n \geq 2$,

$$\gamma_n = \begin{cases} \operatorname{lcm}(1,2,\ldots,n)/6 & \text{if } n = 2\cdot 3^m \text{ for some } m \geq 1 \\ \operatorname{lcm}(1,2,\ldots,n)/2 & \text{otherwise.} \end{cases} \tag{5.14}$$

The next theorem establishes part of this conjecture, namely the exceptional role that the prime $p = 3$ plays. The proof employs the notation

$$g_{n,k}(j) = \frac{1}{(n-j)(n-j+1)} \binom{j}{2k+1} \tag{5.15}$$

so that

$$G_n(x) = -2i \sum_{k=0}^{\lfloor n/2 \rfloor -1} (-1)^k h_{n,k} x^{n-2k} \tag{5.16}$$

with

$$h_{n,k} := \sum_{j=2k+1}^{n-1} g_{n,k}(j) = \sum_{\ell=1}^{n-1-2k} \frac{1}{\ell(\ell+1)} \binom{n-\ell}{2k+1}. \tag{5.17}$$

Therefore, for $n \geq 2$,

$$\gamma_n = \frac{1}{2} \cdot \text{lcm} \left\{ \text{denominator of } h_{n,k} : \, 0 \leq k \leq \lfloor n/2 \rfloor - 1 \right\}. \tag{5.18}$$

Let $\nu_p(n)$ be the exponent of the highest power of p dividing n — the *p-adic valuation of n*. The denominators in the terms forming the sum $h_{n,k}$ are consecutive integers bounded by n. Therefore

$$\nu_3(\gamma_n) \leq \nu_3(\text{lcm}(1, 2, \ldots, n)). \tag{5.19}$$

In fact we can establish $\nu_3(\gamma_n)$ precisely.

Theorem 5.9. The 3-adic valuation of γ_n is given by

$$\nu_3(\gamma_n) = \begin{cases} \nu_3(\text{lcm}(1, 2, \ldots, n)) - 1 & \text{if } n = 2 \cdot 3^m \text{ for some } m \geq 1 \\ \nu_3(\text{lcm}(1, 2, \ldots, n)) & \text{otherwise.} \end{cases}$$

Proof. The analysis is divided into two cases.

Case 1. Assume that $n = 2 \cdot 3^m$. We show that $\nu_3(\gamma_n) = m - 1$.

The bound (5.19) shows that $\nu_3(\gamma_n) \leq m$.

Claim: $\nu_3(\gamma_n) \neq m$. To prove this, the coefficient

$$h_{n,k} = \sum_{\ell=1}^{n-1-2k} \frac{1}{\ell(\ell+1)} \binom{n-\ell}{2k+1} \tag{5.20}$$

is written as

$$h_{n,k} = S_1(n,k) + S_2(n,k) \tag{5.21}$$

where $S_1(n,k)$ is the sum of all the terms in $h_{n,k}$ with a denominator divisible by 3^m and $S_2(n,k)$ contains the remaining terms. This is the highest possible power of 3 that appears in the denominator of $h_{n,k}$.

It is now shown that the denominator of the sum $S_1(n,k)$ is never divisible by 3^m.

Step 1. The sum $S_1(n,k)$ contains at most two terms.

Proof. The index ℓ satisfies $\ell \leq 2 \cdot 3^m - 1 - 2k < 2 \cdot 3^m$. The only choices of ℓ that produce denominators divisible by 3^m are $\ell = 3^m, 3^m - 1$ and $\ell = 2 \cdot 3^m - 1$. The term corresponding to this last choice is $\frac{1}{(2 \cdot 3^m - 1) \cdot 2 \cdot 3^m} \binom{1}{2k+1}$, so it only occurs for $k = 0$. In this situation, the term corresponding to $\ell = 3^m$ is $1/(3^m + 1)$ and it does not contribute to S_1. $\square$

Step 2. If $\frac{1}{2}(3^m - 1) < k \leq 3^m - 1$, then $S_1(n,k)$ is the empty sum. Therefore the denominator of $h_{n,k}$ is not divisible by 3^m.

Proof. The index ℓ in the sum defining $h_{n,k}$ satisfies $1 \leq \ell \leq 2 \cdot 3^m - 1 - 2k$. The assumption on k guarantees that neither $\ell = 3^m$ nor $\ell = 3^m - 1$ appear in this range. □

Step 3. If $k = \frac{1}{2}(3^m - 1)$, then the denominator of $S_1(n,k)$ is not divisible by 3^m.

Proof. In this case the sum $S_1(n,k)$ is

$$\frac{1}{3^m(3^m+1)} + \frac{3^m+1}{(3^m-1)3^m} = \frac{3^m+3}{3^{2m}-1}.$$ □

Step 4. If $0 < k < \frac{1}{2}(3^m - 1)$, then the denominator of $S_1(n,k)$ is not divisible by 3^m.

Proof. The proof of this step employs a theorem of Kummer stating that $\nu_p(\binom{a}{b})$ is equal to the number of borrows involved in subtracting b from a in base p. By Kummer's theorem, $\binom{3^m}{2k+1}$ and $\binom{3^m+1}{2k+1}$ are divisible by 3, so neither of the two terms in $S_1(n,k)$ has denominator divisible by 3^m. □

Step 5. If $k = 0$, then the denominator of $S_1(n,k)$ is not divisible by 3^m.

Proof. For $k = 0$ we have

$$h_{n,0} = \sum_{\ell=1}^{n-1} \frac{n-\ell}{\ell(\ell+1)} = \sum_{\ell=1}^{n-1} \left(\frac{n-\ell}{\ell} - \frac{n-(\ell+1)}{\ell+1} - \frac{1}{\ell+1} \right) = n - H_n, \tag{5.22}$$

and the two terms in H_n whose denominators are divisible by 3^m add up to

$$\frac{1}{3^m} + \frac{1}{2 \cdot 3^m} = \frac{1}{2 \cdot 3^{m-1}} \tag{5.23}$$

with denominator not divisible by 3^m. □

It follows that, for $n = 2 \cdot 3^m$, the denominator of the term $h_{n,k}$ is not divisible by 3^m. Thus, $\nu_3(\gamma_n) \leq m - 1$.

Claim: $\nu_3(\gamma_n) \geq m - 1$. This is established by checking that 3^{m-1} divides the denominator of $h_{n,0}$. Indeed, there are six terms in $h_{n,0} = n - H_n$ whose denominators are divisible by 3^{m-1}, and their sum is

$$\sum_{\ell=1}^{6} \frac{1}{\ell \cdot 3^{m-1}} = \frac{H_6}{3^{m-1}} = \frac{49}{20 \cdot 3^{m-1}}. \tag{5.24}$$

Therefore 3^{m-1} divides the denominator of $h_{n,0}$. This completes Case 1.

Case 2. Assume now that n is not of the form $2 \cdot 3^m$. This states that the base 3 representation of n is not of the form $200 \cdots 00_3$.

Let $r = \lfloor \log_3 n \rfloor$, so that 3^r is the largest power of 3 less than or equal to n. We show that $\nu_3(\gamma_n) = r$ by exhibiting a value of the index k so that the denominator of $h_{n,k}$ is divisible by 3^r.

Step 1. Assume first that the base 3 representation of n begins with 1. Then choose $k = 0$. As before, $h_{n,0} = n - H_n$. Observe that each term in the sum

$$\text{lcm}(1,2,\ldots,n) \cdot H_n = \sum_{\ell=1}^{n} \frac{\text{lcm}(1,2,\ldots,n)}{\ell} \tag{5.25}$$

is an integer. The condition on the base 3 representation of n guarantees that only one of these integers, namely the one corresponding to $\ell = 3^r$, is not divisible by

3. Thus there is no extra cancellation of powers of 3 in H_n, and as a result the denominator of H_n is divisible by 3^r.

Step 2. Assume now that the base 3 representation of n begins with 2. Choose $k = \frac{1}{2}(3^r + 3^{\nu_3(n)})$. As in the discussion in Case 1, there are at most two terms in the sum

$$h_{n,k} = \sum_{\ell=1}^{n-1-2k} \frac{1}{\ell(\ell+1)} \binom{n-\ell}{2k+1} \tag{5.26}$$

with denominator divisible by 3^r. The sum of these terms is

$$\frac{1}{3^r(3^r+1)} \binom{n-3^r}{2k+1} + \frac{1}{(3^r-1)3^r} \binom{n-3^r+1}{2k+1}. \tag{5.27}$$

If $n \equiv 2 \mod 3$, then Kummer's theorem shows that 3 divides the second binomial coefficient but not the first; otherwise, 3 divides the first binomial coefficient but not the second. Therefore $h_{n,k}$ has precisely one term with denominator divisible by 3^r. The argument is complete. $\square$

Corollary 5.10. The 3-adic valuation of the denominator $\alpha_{n,2}$ of $A_{n,2}(x)$ is

$$\nu_3(\alpha_{n,2}) = \begin{cases} \nu_3(n!\ \mathrm{lcm}(1,2,\ldots,n)) - 1 & \text{if } n = 2\cdot 3^m \text{ for some } m \geq 1 \\ \nu_3(n!\ \mathrm{lcm}(1,2,\ldots,n)) & \text{otherwise.} \end{cases}$$

Note 5.11. It easily follows that the 2-adic valuation of γ_n is

$$\nu_2(\gamma_n) = \nu_2(\mathrm{lcm}(1,2,\ldots,n)) - 1. \tag{5.28}$$

Indeed, from (5.18) one has $\nu_2(\gamma_n) \leq \nu_2(\mathrm{lcm}(1,2,\ldots,n)) - 1$. On the other hand, the denominator of $h_{n,0} = n - H_n$ is divisible by the highest power of 2, i.e., by $2^{\lfloor \log_2 n \rfloor}$, which implies (5.28).

The proof of Conjecture 5.5 has been reduced to the identity

$$\nu_p(\gamma_n) = \nu_p(\mathrm{lcm}(1,2,\ldots,n)) \tag{5.29}$$

for all primes $p > 3$.

The sequence $2\cdot 3^m$ appearing in the previous discussion also appears in relation with the denominators of the harmonic numbers H_n. As before, write

$$H_n = \frac{N_n}{D_n} \tag{5.30}$$

in reduced form. The next result considers a special case of the quotient D_{n-1}/D_n of denominators of consecutive harmonic numbers. The general case will be described elsewhere [23].

Theorem 5.12. Let $n \in \mathbb{N}$. Then $D_{2\cdot 3^n - 1} = 3D_{2\cdot 3^n}$.

Proof. An elementary argument shows that $\nu_2(D_n) = \lfloor \log_2 n \rfloor$. Therefore N_n is odd and D_n is even.

Observe that

$$\begin{aligned} \frac{N_{2\cdot 3^n}}{D_{2\cdot 3^n}} &= \frac{N_{2\cdot 3^n - 1}}{D_{2\cdot 3^n - 1}} + \frac{1}{2\cdot 3^n} \\ &= \frac{2\cdot 3^n N_{2\cdot 3^n - 1} + D_{2\cdot 3^n - 1}}{2\cdot 3^n D_{2\cdot 3^n - 1}}. \end{aligned} \tag{5.31}$$

Therefore the denominator $D_{2\cdot 3^n}$ is obtained from $2\cdot 3^n D_{2\cdot 3^n-1}$ by canceling the factor

$$w = \gcd\left(2\cdot 3^n \cdot N_{2\cdot 3^n-1} + D_{2\cdot 3^n-1}, 2\cdot 3^n\cdot D_{2\cdot 3^n-1}\right). \tag{5.32}$$

That is,

$$2\cdot 3^n\cdot D_{2\cdot 3^n-1} = w\cdot D_{2\cdot 3^n}. \tag{5.33}$$

Lemma 5.13. The number w has the form $2^\alpha\cdot 3^\beta$, for some $\alpha,\beta\geq 0$.

Proof. Any prime factor p of w divides

$$2\cdot 3^n\cdot(2\cdot 3^n\cdot N_{2\cdot 3^n-1}+D_{2\cdot 3^n-1}) - 2\cdot 3^n\cdot D_{2\cdot 3^n-1} = 2^2\cdot 3^{2n}\cdot N_{2\cdot 3^n-1}.$$

Then p is a common divisor of $2\cdot 3^n\cdot N_{2\cdot 3^n-1}$ and $2\cdot 3^n\cdot D_{2\cdot 3^n-1}$. The harmonic numbers are in reduced form, so p must be 2 or 3. □

The relation (5.33) becomes $2\cdot 3^n D_{2\cdot 3^n-1} = 2^\alpha\cdot 3^\beta D_{2\cdot 3^n}$, and replacing this in (5.31) yields

$$2^\alpha\cdot 3^\beta N_{2\cdot 3^n} = 2\cdot 3^n N_{2\cdot 3^n-1} + D_{2\cdot 3^n-1}. \tag{5.34}$$

Define $t = \lfloor \log_2(2\cdot 3^n-1)\rfloor > 1$ and write $D_{2\cdot 3^n-1} = 2^t C_{2\cdot 3^n-1}$ with $C_{2\cdot 3^n-1}$ an odd integer. Then (5.34) becomes

$$2^{\alpha-1}\cdot 3^\beta N_{2\cdot 3^n} - 2^{t-1}C_{2\cdot 3^n-1} = 3^n N_{2\cdot 3^n-1}. \tag{5.35}$$

A simple analysis of the parity of each term in (5.35) shows that the only possibility is $\alpha = 1$.

The relation (5.33) now becomes

$$3^n\cdot D_{2\cdot 3^n-1} = 3^\beta\cdot D_{2\cdot 3^n}. \tag{5.36}$$

In the computation of the denominator $D_{2\cdot 3^n}$ we have the sum

$$1+\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{3^n}+\cdots+\frac{1}{2\cdot 3^n-1}+\frac{1}{2\cdot 3^n} \tag{5.37}$$

so that the maximum power of 3 that appears in a denominator forming the sum (5.37) is 3^n. Simply observe that $3^{n+1} > 2\cdot 3^n-1$. The combination of all the fractions in the sum (5.37) with denominator 3^n is

$$\frac{1}{3^n}+\frac{1}{2\cdot 3^n} = \frac{2+1}{2\cdot 3^n} = \frac{1}{2\cdot 3^{n-1}}. \tag{5.38}$$

It follows that the maximum power of 3 in (5.37) is at most 3^{n-1}.

The terms in (5.37) that contain exactly 3^{n-1} in the denominator are

$$\frac{1}{3^{n-1}}, \frac{1}{2\cdot 3^{n-1}}, \frac{1}{4\cdot 3^{n-1}}, \frac{1}{5\cdot 3^{n-1}}, \tag{5.39}$$

and these combine with the two terms with denominator exactly divisible by 3^n to produce

$$\left(1+\frac{1}{2}+\frac{1}{4}+\frac{1}{5}\right)\cdot\frac{1}{3^{n-1}}+\frac{1}{2\cdot 3^{n-1}} = \frac{49}{20\cdot 3^{n-1}}. \tag{5.40}$$

The rest of the terms in (5.37) have at most a power of 3^{n-2} in the denominator. The total sum can be written as

$$\frac{49}{20\cdot 3^{n-1}}+\frac{x_n}{y_n\cdot 3^{n-2}} = \frac{49y_n+60x_n}{20y_n\cdot 3^{n-1}} \tag{5.41}$$

and no cancellation occurs. Therefore 3^{n-1} is the 3-adic valuation of $D_{2\cdot 3^n}$. Write $D_{2\cdot 3^n} = 3^{n-1}\cdot E_{2\cdot 3^n}$, where $E_{2\cdot 3^n}$ is not divisible by 3.

Now consider the denominator $D_{2\cdot 3^n-1}$. Observe that

$$1+\frac{1}{2}+\cdots+\frac{1}{2\cdot 3^n-1} = \frac{1}{3^n}+\frac{x_n}{y_n\cdot 3^{n-1}} = \frac{y_n+3x_n}{y_n\cdot 3^n}, \tag{5.42}$$

with y_n not divisible by 3. Therefore 3^n is the 3-adic valuation of $D_{2\cdot 3^n-1}$. Write $D_{2\cdot 3^n-1} = 3^n\cdot E_{2\cdot 3^n-1}$ where $E_{2\cdot 3^n-1}$ is not divisible by 3.

The relation (5.36) now reads $3^{2n}E_{2\cdot 3^n-1} = 3^{\beta+n-1}E_{2\cdot 3^n}$ and this gives $\beta = n+1$. Replacing in (5.36) produces $D_{2\cdot 3^n-1} = 3D_{2\cdot 3^n}$, as claimed. □

6. The iterated integral of $\ln(1+x^3)$

In this final section we consider the iterated integral of $\ln(1+x^3)$. The first value is

$$\begin{aligned} f_1(x) = {} & \frac{1}{6}\left(\sqrt{3}\pi-18x\right) \\ & -\sqrt{3}\arctan\left(\frac{1-2x}{\sqrt{3}}\right)+(x+1)\ln(x+1)+\frac{1}{2}(2x-1)\ln(x^2-x+1). \end{aligned}$$

This and additional values suggest the ansatz

$$f_n(x) = A_{n,3}(x)+B_{n,3}(x)u+C_{n,3}(x)v+D_{n,3}(x)w, \tag{6.1}$$

where

$$\begin{aligned} u &= \sqrt{3}\arctan\left(\frac{1-2x}{\sqrt{3}}\right) \\ v &= \ln(x+1) \\ w &= \ln(x^2-x+1) \end{aligned}$$

and where $A_{n,3}$ is a polynomial in $\mathbb{Q}[\sqrt{3}\pi, x]$ and $B_{n,3}$, $C_{n,3}$, and $D_{n,3}$ are polynomials in $\mathbb{Q}[x]$.

The method of roots described in Section 3 shows that $f_n(x)$ can be expressed in terms of $\ln(x+1)$, $\ln(x+\omega)$, and $\ln(x+\bar{\omega})$, where $\omega = e^{2\pi i/3} = \frac{1}{2}(-1+i\sqrt{3})$ satisfies $\omega^3 = 1$. Using the relation

$$\ln(-2i(x+\omega)) = \frac{1}{2}\ln(x^2-x+1)+i\arctan\left(\frac{1-2x}{\sqrt{3}}\right)+\ln(2), \tag{6.2}$$

we can convert between (6.1) and expressions in terms of $\ln(x+\omega)$ and $\ln(x+\bar{\omega})$.

As was the case for the iterated integrals of $\ln(1+x)$ and $\ln(1+x^2)$, it is easy to conjecture closed forms for all but one of these polynomials.

Theorem 6.1. Define

$$\chi_3(k) = \begin{cases} 0 & \text{if } k\equiv 0 \mod 3 \\ 1 & \text{if } k\equiv 1 \mod 3 \\ -1 & \text{if } k\equiv 2 \mod 3 \end{cases}$$

and

$$\lambda(k) = \begin{cases} 1 & \text{if } k\equiv 0 \mod 3 \\ 0 & \text{if } k\not\equiv 0 \mod 3. \end{cases}$$

Then

$$
\begin{aligned}
B_{n,3}(x) &= -\frac{1}{n!}\sum_{k=0}^{n}\chi_3(n-k)\binom{n}{k}x^k \\
C_{n,3}(x) &= \frac{1}{n!}(x+1)^n = \frac{1}{n!}\sum_{k=0}^{n}\binom{n}{k}x^k \\
D_{n,3}(x) &= \frac{1}{2n!}\sum_{k=0}^{n}(3\lambda(n-k)-1)\binom{n}{k}x^k.
\end{aligned}
$$

Proof. The method of roots developed in Section 3 shows that the iterated integral can be expressed in the form (6.1). The polynomials $A_{n,3}, B_{n,3}, C_{n,3}, D_{n,3}$ will be linear combinations of the powers $(x+1)^n$, $(x+\omega)^n$, and $(x+\bar{\omega})^n$. Comparing initial values, it is found that

$$
\begin{aligned}
B_{n,3}(x) &= \frac{i}{\sqrt{3}n!}\big((x+\omega)^n-(x+\bar{\omega})^n\big) \\
C_{n,3}(x) &= \frac{1}{n!}(x+1)^n \\
D_{n,3}(x) &= \frac{1}{2n!}\big((x+\omega)^n+(x+\bar{\omega})^n\big).
\end{aligned}
$$

Note that the above expressions can also be automatically found as solutions of the fourth-order recurrence that `HolonomicFunctions` derives in this case:

$$
\begin{aligned}
(6.3)\quad &(n-2)(n-1)n^2F_n \\
&= (n-2)(n-1)(4n-3)xF_{n-1} - 3(n-2)(2n-3)x^2F_{n-2} \\
&\qquad + \left[(4n-9)x^3+n\right]F_{n-3} - x(x^3+1)F_{n-4}.
\end{aligned}
$$

The above closed forms for $B_{n,3}$ and $D_{n,3}$ can be used to derive explicit expressions for their coefficients:

$$
B_{n,3}(x) = \frac{i}{\sqrt{3}n!}\sum_{k=0}^{n}\binom{n}{k}x^k\left(\omega^{n-k}-\bar{\omega}^{n-k}\right).
$$

The value of the last parenthesis can be found by case distinction using the fact that $\omega^3=\bar{\omega}^3=1$:

$$
\begin{aligned}
n-k &\equiv 0 \mod 3: & 1-1 &= 0, \\
n-k &\equiv 1 \mod 3: & \omega-\bar{\omega} &= i\sqrt{3}, \\
n-k &\equiv 2 \mod 3: & \omega^2-\bar{\omega}^2 &= -i\sqrt{3}.
\end{aligned}
$$

It follows that

$$
B_{n,3}(x) = -\frac{1}{n!}\sum_{k=0}^{n}\chi_3(n-k)\binom{n}{k}x^k.
$$

The similar computation for $D_{n,3}(x)$ is left to the reader. $\square$

Note 6.2. In order to obtain these functions from a purely symbolic approach, consider a brute force evaluation of $f_n(x)$ by using *Mathematica* to evaluate (1.4). The results are expressed in terms of the functions

$$
(6.4)\qquad h_1(x) = {}_2F_1\left(\begin{matrix}1/3,\,1\\4/3\end{matrix};-x^3\right) \quad\text{and}\quad h_2(x) = {}_2F_1\left(\begin{matrix}2/3,\,1\\5/3\end{matrix};-x^3\right),
$$

where

$$ {}_2F_1\left(\begin{matrix} a,\, b \\ c \end{matrix} ; z \right) = \sum_{k=0}^{\infty} \frac{(a)_k\,(b)_k}{(c)_k\,k!} z^k \tag{6.5} $$

is the classical hypergeometric series. The first few values are

$$ \begin{aligned} f_1(x) &= -3x + x\,\ln(1+x^3) + 3xh_1(x) \\ f_2(x) &= -\frac{9x^2}{4} + \frac{1}{2}x^2\ln(1+x^3) + 3x^2h_1(x) - \frac{3x^2}{4}h_2(x) \\ f_3(x) &= -\frac{11x^3}{12} + \frac{1}{6}(x^3+1)\ln(1+x^3) + \frac{3x^3}{2}h_1(x) - \frac{3x^3}{4}h_2(x). \end{aligned} $$

The hypergeometric series $h_1(x)$ and $h_2(x)$ can be expressed as

$$ \begin{aligned} h_1(x) &= \frac{1}{3x}\left(\ln(1+x) + \omega\ln(1+\bar{\omega}x) + \bar{\omega}\ln(1+\omega x)\right) \\ h_2(x) &= \frac{2}{3x^2}\left(-\ln(1+x) - \omega\ln(1+\omega x) - \bar{\omega}\ln(1+\bar{\omega}x)\right), \end{aligned} $$

with $\omega = e^{2\pi i/3}$ as before. These expressions can be transformed into the functions in (6.1).

As in previous sections, the closed-form expression for the *pure* polynomial part $A_{n,3}(x)$ is more elaborate. By writing

$$ f_n(x) = \tilde{A}_{n,3}(x) + \tilde{B}_{n,3}(x)\ln(x+1) + \tilde{C}_{n,3}(x)\ln(x+\omega) + \tilde{D}_{n,3}(x)\ln(x+\bar{\omega}), $$

we obtain a polynomial $\tilde{A}_{n,3}(x)$ with rational coefficients. The first values are given by

$$ \tilde{A}_{0,3}(x) = 0, \quad \tilde{A}_{1,3}(x) = -3x, \quad \tilde{A}_{2,3}(x) = -\tfrac{9}{4}x^2, \quad \tilde{A}_{3,3}(x) = -\tfrac{11}{12}x^3. \tag{6.6} $$

Schneider's *Mathematica* package `Sigma` can be used again to obtain, from the recurrence (6.3) and the initial conditions, the expression

$$ \tilde{A}_{n,3}(x) = -\frac{1}{n!}\sum_{k=1}^{n} \frac{x^k}{k}\left[(x+1)^{n-k} + (x+\omega)^{n-k} + (x+\bar{\omega})^{n-k}\right]. \tag{6.7} $$

6.1. **Arithmetical properties of $\tilde{A}_{n,3}$.** Define $\alpha_{n,3}$ to be the denominator of $\tilde{A}_{n,3}$ and $\beta_{n,3} = \alpha_{n,3}/(n\alpha_{n-1,3})$.

Conjecture 6.3. The sequence $\beta_{n,3}$ is given by

$$ \beta_{n,3} = \begin{cases} p & \text{if } n = p^m \neq 3 \text{ for some prime } p \text{ and } m \in \mathbb{N} \\ \frac{1}{11} & \text{if } n = 3\cdot 11^m \text{ for some } m \in \mathbb{N} \\ 11 & \text{if } n = 3\cdot 11^m + 1 \text{ for some } m \in \mathbb{N} \\ 1 & \text{otherwise.} \end{cases} \tag{6.8} $$

Observe that this expression for $\beta_{n,3}$ does not have the exceptional case where $3\cdot 11^m + 1$ is a prime power that appears in $\beta_{n,2}$ given in (5.8). This is ruled out by the following.

Lemma 6.4. Let $m \in \mathbb{N}$. Then $3\cdot 11^m + 1$ is not a prime power.

Proof. The number $3\cdot 11^m + 1$ is even, so only the prime 2 needs to be checked. We have $3\cdot 11^m + 1 \equiv 3^{m+1} + 1 \not\equiv 0 \mod 8$ since the powers of 3 are 1 or 3 modulo 8. Therefore $3\cdot 11^m + 1$ (since it is larger than 4) is not a power of 2. $\square$

Acknowledgements. The authors wish to thank Xinyu Sun for discussions on the paper, specially on Section 5. The authors also wish to thank the referee for a detailed reading of the paper and many corrections. The work of the third author was partially supported by NSF DMS 0070567. The work of the second author was partially supported by the same grant as a postdoctoral fellow at Tulane University and by the Austrian Science Fund (FWF):P20162-N18. The work of the last author was partially supported by Tulane VIGRE Grant 0239996.

References

[1] A. Apelblat. *Tables of Integrals and Series.* Verlag Harry Deutsch, Thun; Frankfurt am Main, 1996.
[2] G. Boros, O. Espinosa, and V. Moll. On some families of integrals solvable in terms of polygamma and negapolygamma functions. *Integrals Transforms and Special Functions*, 14:187–203, 2003.
[3] M. Bronstein. *Symbolic Integration I. Transcendental functions*, volume 1 of *Algorithms and Computation in Mathematics.* Springer-Verlag, 1997.
[4] I. S. Gradshteyn and I. M. Ryzhik. *Table of Integrals, Series, and Products.* Edited by A. Jeffrey and D. Zwillinger. Academic Press, New York, 7th edition, 2007.
[5] R. Graham, D. Knuth, and O. Patashnik. *Concrete Mathematics.* Addison Wesley, Boston, 2nd edition, 1994.
[6] C. Koutschan. *Advanced Applications of the Holonomic Systems Approach.* PhD thesis, RISC, Johannes Kepler University, Linz, Austria, 2009.
[7] C. Koutschan and V. Moll. The integrals in Gradshteyn and Ryzhik. Part 18: Some automatic proofs. *Scientia*, 20:93–111, 2011.
[8] J. Liouville. Sur les transcendentes elliptiques de premiére et de seconde espéce, considérées comme fonctions de leur amplitude. *Journ. Ec. Polyt.*, 14, (23. cahier):37–83, 1834.
[9] J. Lützen. *Joseph Liouville* $1809-1882$*. Master of Pure and Applied Mathematics*, volume 15 of *Studies in the History of Mathematics and Physical Sciences.* Springer-Verlag, New York, 1990.
[10] D. Manna and V. Moll. Rational Landen transformations on $\mathbb{R}$. *Math. Comp.*, 76:2023–2043, 2007.
[11] D. Manna and V. Moll. A simple example of a new class of Landen transformations. *Amer. Math. Monthly*, 114:232–241, 2007.
[12] D. Manna and V. Moll. Landen Survey. *MSRI Publications: Probability, Geometry and Integrable Systems. In honor of Henry McKean* 75*th birthday*, 55:201–233, 2008.
[13] L. Medina, V. Moll, and E. Rowland. Iterated primitives of logarithmic powers. *International Journal of Number Theory*, To appear, 2011.
[14] V. Moll. The evaluation of integrals: a personal story. *Notices of the AMS*, 49:311–317, 2002.
[15] V. Moll. Seized opportunities. *Notices of the AMS*, pages 476–484, 2010.
[16] M. Petkovšek, H. Wilf, and D. Zeilberger. *A=B.* A. K. Peters, Ltd., 1st edition, 1996.
[17] M. Petkovšek. Hypergeometric solutions of linear recurrences with polynomial coefficients. *Journal of Symbolic Computation*, 14(2/3):243–264, 1992.
[18] A. P. Prudnikov, Yu. A. Brychkov, and O. I. Marichev. *Integrals and Series.* Gordon and Breach Science Publishers, 1992.
[19] R. H. Risch. The problem of integration in finite terms. *Trans. Amer. Math. Soc.*, 139:167–189, 1969.
[20] R. H. Risch. The solution of the problem of integration in finite terms. *Bull. Amer. Math. Soc.*, 76:605–608, 1970.
[21] J. F. Ritt. *Integration in finite terms. Liouville's theory of elementary functions.* New York, 1948.
[22] C. Schneider. Symbolic summation assists Combinatorics. *Séminaire Lotharingien de Combinatoire*, 56:1–36, April 2007. Article B56b.
[23] X. Sun and V. Moll. Denominators of harmonic numbers. *In preparation*, 2011.

Department of Mathematics, Tulane University, New Orleans, LA 70118
E-mail address: tamdeber@math.tulane.edu

Research Institute for Symbolic Computation, Johannes Kepler University, 4040 Linz, Austria
E-mail address: ckoutsch@risc-uni-linz.ac.at

Department of Mathematics, Tulane University, New Orleans, LA 70118
E-mail address: vhm@math.tulane.edu

Department of Mathematics, Tulane University, New Orleans, LA 70118
E-mail address: erowland@tulane.edu