

SOME IDENTITIES INVOLVING HARMONIC NUMBERS

JÜRGEN SPIEB

ABSTRACT. Let H_n denote the n th harmonic number. Explicit formulas for sums of the form $\sum a_k H_k$ or $\sum a_k H_k H_{n-k}$ are derived, where the a_k are simple functions of k . These identities are generalized in a natural way by means of generating functions.

1. INTRODUCTION AND OVERVIEW

In the analysis of algorithms, harmonic numbers frequently occur. Quite often, sums of the form $\sum k^p H_k$ or more complicated expressions have to be manipulated. Knuth [4] dedicates one section to the study of these numbers and gives some basic identities. Riordan [6] gives some identities but does not collect these results for easy reference. Lafon [5] and Karr [2] mention the existence of such summation formulas without giving any of them explicitly.

Typical situations where the ability to sum harmonic numbers and their generalizations is needed can be found in Kemp [3], Sedgewick [8] or Spieß [9].

The author felt that a more systematic investigation of identities involving harmonic numbers would be helpful. The summation formulas presented here are stated in very general terms. But to see that they are useful for summing harmonic numbers, many examples are given. We present only the simplest relations. Because identities with binomial coefficients can be transformed in many different ways—e.g., by the use of inverse relations—it is impossible to be exhaustive. Asymptotic representations of some of the quantities we deal with are given by Knuth [4] and Roes [7].

Section 2 contains the basic definitions and introduces notation. Section 3 contains the main general summation formulas derived from a generating function. Section 4 contains recursion formulas which may be used to derive new identities and which may simplify calculations using a symbol manipulation system. Section 5 contains structure theorems. They are useful to give a quick orientation of the outcome of the summation. In the appendix special summation formulas are assembled for ease of use.

Received January 6, 1989.

1980 *Mathematics Subject Classification* (1985 Revision). Primary 05A19, 11B73, 68R05.

Key words and phrases. Harmonic numbers, symbolic computation, computational analysis, combinatorial identities.

© 1990 American Mathematical Society
0025-5718/90 \$1.00 + \$.25 per page

2. DEFINITIONS AND KNOWN RESULTS

Let \mathbf{Z} denote the integers and \mathbf{N} denote the natural numbers, with $0 \in \mathbf{N}$. The generalized harmonic numbers are defined by

$$H_n^{(r)} = \sum_{k=1}^n k^{-r}, \quad n \geq 1, r \in \mathbf{Z}.$$

It is convenient to set $H_n^{(r)} = 0$ for $n \leq 0$. As usual, we write H_n instead of $H_n^{(1)}$: $H_n = 1 + 1/2 + 1/3 + \dots + 1/n$. For $r \leq 0$, $H_n^{(r)}$ denotes a sum of powers: $H_n^{(0)} = n$, $H_n^{(-1)} = n(n+1)/2, \dots$.

Zave [10] gives the following series expansion: for all $m, r \in \mathbf{N}$

$$(1) \frac{1}{(1-z)^{m+1}} \left(\log \frac{1}{1-z} \right)^r = \sum_{k=m}^{\infty} P_r(H_k^{(1)} - H_m^{(1)}, \dots, H_k^{(r)} - H_m^{(r)}) \binom{k}{m} z^{k-m},$$

where the polynomial $P_r(x_1, \dots, x_r)$ is defined by $P_0 = 1$ and

$$P_r(x_1, \dots, x_r) = (-1)^r Y_r(-0!x_1, -1!x_2, \dots, -(r-1)!x_r),$$

where Y_r is the familiar Bell polynomial [6].

Some of the polynomials are

$$P_1(x_1) = x_1,$$

$$P_2(x_1, x_2) = x_1^2 - x_2,$$

$$P_3(x_1, x_2, x_3) = x_1^3 - 3x_1x_2 + 2x_3,$$

$$P_4(x_1, x_2, x_3, x_4) = x_1^4 - 6x_1^2x_2 + 8x_1x_3 + 3x_2^2 - 6x_4.$$

This expansion obviously contains as a special case ($r = 1, m = 0$) the generating function of the harmonic numbers,

$$\frac{-\log(1-z)}{(1-z)} = \sum_{k \geq 0} H_k z^k.$$

This, and the next more general form, is given by Knuth [4]:

$$\frac{-\log(1-z)}{(1-z)^{m+1}} = \sum_{k \geq m} (H_k - H_m) \binom{k}{m} z^{k-m}, \quad m \in \mathbf{N}.$$

The special case $r = 0$ of Zave's result is the binomial theorem:

$$\frac{1}{(1-z)^{m+1}} = \sum_{k \geq m} \binom{k}{m} z^{k-m}, \quad m \in \mathbf{N}.$$

Let X be an indeterminate; then, following Riordan [6], we write

$$(X)_p = X(X-1)(X-2) \cdots (X-p+1), \quad p \in \mathbf{N},$$

for the descending factorial function, with $(X)_0 = 1$.

It is convenient to extend this notation to negative-index values:

$$(2) \quad (X)_{-p} = \frac{1}{(X+1)(X+2)\cdots(X+p)}.$$

Then it is easy to see that

$$(3a) \quad \text{for all } p, q \in \mathbf{Z}: (X+p)_p \cdot (X)_q = (X+p)_{p+q},$$

$$(3b) \quad \text{for all } p \in \mathbf{Z}: (X)_p + p \cdot (X)_{p-1} = (X+1)_p.$$

If X stands for an element of \mathbf{Z} , then these equations are interpreted as

- (i) if both sides are defined, then they are equal, and
- (ii) if the left-hand side is defined, so is the right-hand side.

Equation (3b) is a useful generalization of the basic identity of the binomial coefficients: for $p > 0$, (3b) is equivalent to $\binom{X}{p} + \binom{X}{p-1} = \binom{X+1}{p}$. With this notation, the Stirling numbers of the first kind, $s(p, k)$, are defined by

$$(4a) \quad (X)_p = \sum_{k=0}^p s(p, k) X^k, \quad p \in \mathbf{N}.$$

The Stirling numbers of the second kind, $S(p, k)$, are defined by

$$(4b) \quad X^p = \sum_{k=0}^p S(p, k) (X)_k, \quad p \in \mathbf{N}.$$

These numbers may also be defined recursively by

$$s(p+1, k) = s(p, k-1) - p \cdot s(p, k),$$

$$S(p+1, k) = S(p, k-1) + k \cdot S(p, k),$$

with boundary values

$$\text{for all } n \in \mathbf{N}: s(n, 0) = s(0, n) = S(n, 0) = S(0, n) = \delta_{n0},$$

where δ_{pq} is the Kronecker delta, $\delta_{pp} = 1$, $\delta_{pq} = 0$ for $p \neq q$. Tables of the Stirling numbers can be found in Abramowitz and Stegun [1].

A very useful means to obtain new identities, starting from some known ones, is the application of inverse relations. Riordan [6] gives many inverse pairs which may be helpful. Here are the two simplest pairs:

$$(5a) \quad \text{if } a_n = \sum_{k=0}^n \binom{n}{k} b_k, \quad \text{then } b_n = \sum_{k=0}^n \binom{n}{k} (-1)^{k+n} a_k;$$

$$(5b) \quad \text{if } a_n = \sum_{k=n}^m \binom{k}{n} b_k, \quad \text{then } b_n = \sum_{k=n}^m \binom{k}{n} (-1)^{k+n} a_k.$$

3. GENERAL SUMMATION FORMULAS

Zave's result (1) can be exploited by standard combinatorial techniques. To simplify notation, we introduce the symbol

$$(6) \quad P(r, k, m) = P_r(H_k^{(1)} - H_m^{(1)}, \dots, H_k^{(r)} - H_m^{(r)}),$$

so that Zave's result can be restated as

$$(7) \quad \frac{(-\log(1-z))^r}{(1-z)^{m+1}} = \sum_{k \geq m} \binom{k}{m} P(r, k, m) z^{k-m}.$$

This expansion shows that for all k with $r \leq k < m+r$ we have $P(r, k, m) = 0$.

Theorem 1. For all $n, p, q, u, v \in \mathbb{N}$

$$(8) \quad \sum_{k=p}^{n-q} \binom{k}{p} \binom{n-k}{q} P(u, k, p) P(v, n-k, q) = \binom{n+1}{p+q+1} P(u+v, n+1, p+q+1).$$

Proof. Obviously,

$$\frac{(-\log(1-z))^u}{(1-z)^{p+1}} \times \frac{(-\log(1-z))^v}{(1-z)^{q+1}} = \frac{(-\log(1-z))^{u+v}}{(1-z)^{p+q+1+1}}.$$

Use (1) on both sides. Equating coefficients gives the result. \square

In order to show the power of this theorem, we consider some special cases.

Example 1. $u = v = 0$ gives the well-known identity

$$\sum_{k=p}^{n-q} \binom{k}{p} \binom{n-k}{q} = \binom{n+1}{p+q+1}.$$

This formula can be rewritten in a more symmetric form by introducing a new integer parameter s . Replace k by $k+s$. Then $n-k = n+s-(s+k)$. Hence,

$$(9) \quad \sum_{k=p-s}^{n-q} \binom{s+k}{p} \binom{n-k}{q} = \binom{n+s+1}{p+q+1}, \quad s \in \mathbb{Z}.$$

Example 2. $u = 1, v = 0$ gives

Corollary 2. For all $n, p, q \in \mathbb{N}$ and $s \in \mathbb{Z}$

$$(10) \quad \sum_{k=p-s}^{n-q} \binom{s+k}{p} \binom{n-k}{q} H_{s+k} = \binom{n+s+1}{p+q+1} (H_{n+s+1} - H_{p+q+1} + H_p).$$

Proof. The parameter s has been introduced as above. The formula then is an immediate consequence of (8) and (9). \square

Example 3. $u = 1, v = 1$ gives

Corollary 3. For all $n, p, q \in \mathbb{N}$ and $s \in \mathbb{Z}$

$$(11) \quad \sum_{k=p-s}^{n-q} \binom{s+k}{p} \binom{n-k}{q} H_{s+k} H_{n-k} = \binom{n+s+1}{p+q+1} \left[(D + H_p)(D + H_q) - (H_{n+s+1}^{(2)} - H_t^{(2)}) \right],$$

where $t = p + q + 1$ and $D = H_{n+s+1} - H_t$.

Proof. Theorem 1 gives

$$\begin{aligned} \sum_{k=p}^{n-q} \binom{k}{p} \binom{n-k}{q} (H_k - H_p)(H_{n-k} - H_q) \\ = \binom{n+1}{p+q+1} [(H_{n+1} - H_t)^2 - (H_{n+1}^{(2)} - H_t^{(2)})]. \end{aligned}$$

Isolation of the term $H_k H_{n-k}$ on the left side of this equation, using (10), (9), and introduction of s gives the result. \square

The next theorem is also a simple consequence of (1).

Theorem 4. For all $m, n, p, r \in \mathbf{N}$ with $m \geq p$

$$(12) \quad \sum_{k \geq 0} (-1)^k \binom{p}{k} \binom{n-k}{m} P(r, n-k, m) = \binom{n-p}{m-p} P(r, n-p, m-p).$$

Proof. Using (1), we can write

$$\frac{(-\log(1-z))^r}{(1-z)^{m-p+1}} = \sum_{k \geq m-p} \binom{k}{m-p} P(r, k, m-p) z^{k-(m-p)};$$

but this is the same as

$$(1-z)^p \cdot \frac{(-\log(1-z))^r}{(1-z)^{m+1}} = \sum_{k=0}^p \sum_{j \geq m} (-1)^k \binom{p}{k} \binom{j}{m} P(r, j, m) z^{k+j-m}.$$

Rearranging this sum and equating coefficients gives the result. \square

Remark. For $k > \min(n-m, p)$ the terms in the sum vanish. So this theorem is only interesting for $n \geq m$.

We give two applications of Theorem 4:

Example 1. For $r = 0$ we obtain the well-known identity

$$\sum_{k \geq 0} (-1)^k \binom{p}{k} \binom{n-k}{m} = \binom{n-p}{m-p}, \quad m, n, p \in \mathbf{N}, \quad m \geq p.$$

Example 2. For $r = 1$ we obtain: for all $m, n, p \in \mathbf{N}, m \geq p,$

$$\sum_{k \geq 0} (-1)^k \binom{p}{k} \binom{n-k}{m} (H_{n-k} - H_m) = \binom{n-p}{m-p} (H_{n-p} - H_{m-p}).$$

This can be brought into a form which is easier to apply:

Corollary 5. For all $m, n, p \in \mathbf{N}$ with $m \geq p$

$$(13) \quad \sum_{k=0}^n (-1)^k \binom{p}{n-k} \binom{k}{m} H_k = (-1)^n \binom{n-p}{m-p} (H_{n-p} - H_{m-p} + H_m).$$

We now investigate recurrence relations for the functions $P(r, k, m)$.

Lemma 6. For all $r, k, m \in \mathbf{N}$ with $k \geq m$ and $r \geq 1$

$$(14) \quad P(r, k, m) = \frac{r}{m+1}P(r-1, k, m+1) + P(r, k, m+1).$$

Proof. Zave derives his result from the fact

$$\left(1 + \frac{z}{m+1}\right) \cdots \left(1 + \frac{z}{m+k}\right) = \sum_{r \geq 0} \frac{1}{r!} P(r, m+k, m) z^r,$$

which means

$$\frac{1}{r!} P(r, m+k, m) = \sum_{m+1 \leq k_1 < k_2 < \cdots < k_r \leq m+k} \frac{1}{k_1 k_2 \cdots k_r}.$$

But such a sum can be evaluated by splitting it into two parts: $k_1 = m+1$ and $k_1 > m+1$. Thus,

$$\begin{aligned} \frac{1}{r!} P(r, m+k, m) &= \frac{1}{m+1} \cdot \frac{1}{(r-1)!} P(r-1, m+k, m+1) \\ &\quad + \frac{1}{r!} P(r, m+k, m+1), \end{aligned}$$

and substitution of $m+k$ by k gives the result. \square

An immediate consequence is

Theorem 7. For all $m, n, r \in \mathbf{N}$ with $r \geq 1$

$$(15) \quad \sum_{k=1}^n \frac{r}{k} \cdot P(r-1, m, k) = P(r, m, 0) - P(r, m, n).$$

Proof. Using (14), we have

$$\begin{aligned} \sum_{k=0}^{n-1} \frac{r}{k+1} P(r-1, m, k+1) &= \sum_{k=0}^{n-1} [P(r, m, k) - P(r, m, k+1)] \\ &= P(r, m, 0) - P(r, m, n). \end{aligned}$$

Changing the summation index in the left sum yields the result. \square

We consider two special cases:

Example 1. $m = 0, r = 1$ gives $\sum_{k=1}^n \frac{1}{k} = H_n$.

Example 2. $m = 0, r = 2$ gives $\sum_{k=1}^n \frac{1}{k} H_k = \frac{1}{2}(H_n^2 + H_n^{(2)})$.

This formula is also given by Knuth [4].

Lemma 8. For all $k, r, m \in \mathbf{N}$ with $r \geq 1$ and $k \geq m$

$$(16) \quad k \cdot (P(r, k, m) - P(r, k-1, m)) = r \cdot P(r-1, k-1, m).$$

Proof. Apply the same technique as in the proof of Lemma 6, but this time splitting into the two parts $k_r = m+k$ and $k_r < m+k$. \square

The next theorem follows from Lemma 8.

Theorem 9. For all $r \in \mathbb{N}$ with $r \geq 1$

$$(17) \quad (-\log(1 - z))^r = \sum_{k \geq r} \frac{r}{k} \cdot P(r - 1, k - 1, 0) z^k.$$

Proof. We have

$$\begin{aligned} (-\log(1 - z))^r &= (1 - z) \cdot \frac{(-\log(1 - z))^r}{(1 - z)} \\ &= (1 - z) \cdot \sum_{k \geq 0} P(r, k, 0) z^k \quad (\text{by (1)}) \\ &= \sum_{k > 0} (P(r, k, 0) - P(r, k - 1, 0)) z^k \quad (\text{and using (15)}) \\ &= \sum_{k \geq r} \frac{r}{k} \cdot P(r - 1, k - 1, 0) z^k. \quad \square \end{aligned}$$

The special cases $r = 2$ and $r = 3$ may be of interest:

$$\begin{aligned} (-\log(1 - z))^2 &= \sum_{k \geq 0} \frac{2}{k + 1} H_k z^{k+1}, \\ (-\log(1 - z))^3 &= \sum_{k \geq 0} \frac{3}{k + 1} (H_k^2 - H_k^{(2)}) z^{k+1}. \end{aligned}$$

Theorem 9 can be used to get new identities:

Theorem 10. For all $m, n, r \in \mathbb{N}$ with $r \geq 1$ and $n \geq 1$

$$(18) \quad \sum_{k \geq 0} (-1)^k \binom{m + 1}{n - k} \binom{m + k}{m} P(r, m + k, m) = (-1)^n \frac{r}{n} P(r - 1, n - 1, 0).$$

Proof. This is a direct consequence of the obvious identity

$$(-\log(1 - z))^r = (1 - z)^{m+1} \cdot \frac{(-\log(1 - z))^r}{(1 - z)^{m+1}}.$$

Use of (17) and (1) gives the result. \square

By index transformation we get a simpler form equivalent to (18):

$$(19) \quad \sum_{k \geq 0} (-1)^k \binom{m + 1}{k} \binom{n - k}{m} P(r, n - k, m) = \frac{r}{n - m} P(r - 1, n - 1 - m, 0).$$

We consider two examples:

Example 1. $r = 1$ gives

$$\sum_{k \geq 0} (-1)^k \binom{m + 1}{n - k} \binom{m + k}{m} (H_{m+k} - H_m) = \frac{(-1)^n}{n}.$$

Then, using the identity $\sum_{k \geq 0} (-1)^k \binom{m+1}{n-k} \binom{m+k}{m} = 0$ for all $m, n \in \mathbb{N}$ with $n > 0$, which is derived from $(1-z)^{m+1} \cdot 1/(1-z)^{m+1} = 1$, to simplify this sum, we get:

Corollary 11. For all $m, n \in \mathbb{N}$ with $n \geq 1$

$$(20) \quad \sum_{k \geq 0} (-1)^k \binom{m+1}{n-k} \binom{m+k}{m} H_{m+k} = \frac{(-1)^n}{n}.$$

This may be recast into a simpler form:

$$(20a) \quad \sum_{k \geq 0} (-1)^k \binom{m+1}{k} \binom{n-k}{m} H_{n-k} = \frac{1}{n-m} \quad \text{for } n > m.$$

Example 2. $r = 2$ gives similarly

Corollary 12. For all $m, n \in \mathbb{N}$ with $n \geq 1$

$$(21) \quad \sum_{k \geq 0} (-1)^k \binom{m+1}{n-k} \binom{m+k}{m} (H_{m+k}^2 - H_{m+k}^{(2)}) = (-1)^n \frac{2}{n} (H_{n-1} + H_m).$$

But we can use Theorem 9 in other ways to obtain useful identities:

Theorem 13. For all $n, r, s \in \mathbb{N}$ with $r \geq 1, s \geq 1, n \geq r + s - 1$

$$(22) \quad \sum_{k=r}^{n-s+1} \frac{r}{k} \cdot \frac{s}{n-k+1} P(r-1, k-1, 0) P(s-1, n-k, 0) = \frac{r+s}{n+1} P(r+s-1, n, 0).$$

Proof. Apply Theorem 9 to both sides of

$$(-\log(1-z))^r \cdot (-\log(1-z))^s = (-\log(1-z))^{r+s}$$

and compare the coefficients of the power series. \square

Example 1. $r = 1, s = 1$ gives

$$(23) \quad \sum_{k=1}^n \frac{1}{k} \cdot \frac{1}{n-k+1} = \frac{2}{n+1} H_n.$$

Example 2. $r = 2, s = 1$ gives

$$(24) \quad \sum_{k=2}^n \frac{2}{k} \cdot \frac{1}{n-k+1} H_{k-1} = \frac{3}{n+1} (H_n^2 - H_n^{(2)}).$$

Example 3. $r = 2, s = 2$ gives

$$(25) \quad \sum_{k=2}^n \frac{2}{k} \cdot \frac{2}{n-k+1} H_{k-1} H_{n-k} = \frac{4}{n+1} (H_n^3 - 3H_n H_n^{(2)} + 2H_n^{(3)}).$$

With the same proof technique we can combine (1) and (17) to obtain:

Theorem 14. For all $n, m, r, s \in \mathbb{N}$ with $s \geq 1$ and $n \geq m + r + s$

$$(26) \quad \sum_{k=m+r}^{n-s} \binom{k}{m} \frac{s}{n-k} P(r, k, m) \cdot P(s-1, n-k-1, 0) = \binom{n}{m} P(r+s, n, m).$$

Proof. Use the identity

$$\frac{(-\log(1-z))^r}{(1-z)^{m+1}} (-\log(1-z))^s = \frac{(-\log(1-z))^{r+s}}{(1-z)^{m+1}}$$

and (1) and (17). \square

We give some examples:

Example 1. $r = 0, s = 1$ gives

$$(27) \quad \sum_{k=m}^{n-1} \binom{k}{m} \frac{1}{n-k} = \binom{n}{m} (H_n - H_m).$$

Using the inverse relation (5b) we get

$$\frac{1}{n-m} = \sum_{k=m}^{n-1} (-1)^{k+m} \binom{k}{m} \binom{n}{k} (H_n - H_k) = \sum_{k=m}^n (-1)^{k+m} \binom{k}{m} \binom{n}{k} (H_n - H_k),$$

but since $\sum_{k=m}^n (-1)^{k+m} \binom{k}{m} \binom{n}{k} = 0$ for $n > m$, we get

$$(28) \quad \sum_{k=m}^n (-1)^k \binom{k}{m} \binom{n}{k} H_k = \frac{(-1)^{m+1}}{n-m}, \quad n > m.$$

Using $\binom{k}{m} \binom{n}{k} = \binom{n}{m} \binom{n-m}{k-m}$, we get

$$(29) \quad \sum_{k=m}^n (-1)^k \binom{n-m}{k-m} H_k = \binom{n}{m}^{-1} \cdot \frac{(-1)^{m+1}}{n-m}.$$

Finally, if we replace $n - m$ by n and $k - m$ by k , we obtain

$$(30) \quad \sum_{k=0}^n (-1)^k \binom{n}{k} H_{k+m} = \frac{-1}{n \binom{n+m}{n}}, \quad n > m \geq 0.$$

It is not hard to see that (30) holds also for $n = m$. The case $m = 0$ can be found in Riordan [6].

Example 2. $r = 1, s = 1$, after a simplification using (27), gives

$$(31) \quad \sum_{k=m}^{n-1} \binom{k}{m} \frac{1}{n-k} H_k = \binom{n}{m} [H_n (H_n - H_m) - H_n^{(2)} + H_m^{(2)}].$$

Example 3. $r = 0, s = 2$ gives

$$(32) \quad \sum_{k=m}^{n-1} \binom{k}{m} \frac{2}{n-k} H_{n-k-1} = \binom{n}{m} [(H_n - H_m)^2 - H_n^{(2)} + H_m^{(2)}].$$

Example 4. $r = 1, s = 2$ gives

$$(33) \sum_{k=m}^{n-1} \binom{k}{m} \frac{2}{n-k} (H_k - H_m) H_{n-k-1} \\ = \binom{n}{m} [(H_n - H_m)^3 - 3(H_n - H_m)(H_n^{(2)} - H_m^{(2)}) + 2(H_n^{(2)} - H_m^{(2)})].$$

Using (32), one could derive a formula for $\sum_{k=m}^{n-1} \binom{k}{m} (2/(n-k)) H_k H_{n-k-1}$ which is slightly simpler than (33).

By repeated integration of a power series $\sum a_k z^k$ one gets a series of the form $\sum_k \binom{k}{-q} a_k z^k$. And from $(1-z)^{-1} \sum_k \binom{k}{-q} a_k z^k = \sum_n (\sum_k \binom{k}{-q} a_k) z^n$ one can obtain interesting summation identities. The next theorem is one of this type:

Theorem 15. For all $q, r, n \in \mathbb{N}, q > 1$

$$\sum_{k=0}^n \frac{P(r, k, 0)}{(k+q) \cdots (k+1)} = \frac{r!}{(q-1)!(q-1)^{r+1}} \\ - \frac{1}{(n+q) \cdots (n+2)} \sum_{j=0}^r \frac{\binom{r}{r-j}}{(q-1)^{r+1-j}} P(j, n+1, 0).$$

Proof. We show this by induction on n .

1. $n = 0$: For all $r \in \mathbb{N}$ we have $P(r, 0, 0) = \delta_{r0}$. For $r = 0$ the left-hand side is $(q!)^{-1}$, and the right-hand side is

$$\frac{1}{(q-1)!(q-1)} - \frac{1}{q!} \cdot \frac{1}{q-1} = \frac{1}{q!}.$$

For $r > 0$ the left-hand side is zero as is the right-hand side.

2. Assume the result is true for some n . To simplify notation, we introduce

$$a = \frac{r!}{(q-1)!(q-1)^{r+1}}, \quad b_{-1} = 0, \quad b_j = \frac{\binom{r}{r-j}}{(q-1)^{r+1-j}}$$

and

$$P_j = P(j, n+1, 0), \quad j \in [0, r].$$

Then $j \cdot b_j = (q-1) \cdot b_{j-1}$, and with the notation introduced in (2) we may restate the assertion as

$$S_n = \sum_{k=0}^n \binom{k}{-q} P(r, k, 0) = a - (n+1)_{1-q} \sum_{j=0}^r b_j P_j.$$

Then

$$\begin{aligned}
 S_{n+1} &= S_n + (n+1)_{-q} P_r = a - (n+1)_{1-q} \sum_{j=0}^r b_j P_j + (n+1)_{-q} P_r \\
 &= a - (n+1)_{1-q} \sum_{j=0}^r b_j P_j + (n+1)_{-q} \sum_{j=0}^r (q-1) b_j P_j \\
 &\quad - (n+1)_{-q} \sum_{j=0}^{r-1} (q-1) b_j P_j \\
 &= a - (n+2)_{1-q} \sum_{j=0}^r b_j P_j - (n+1)_{-q} \sum_{j=1}^r (q-1) b_{j-1} P_{j-1} \quad (\text{using (3b)}) \\
 &= a - (n+2)_{1-q} \sum_{j=0}^r b_j P_j - (n+1)_{-q} \sum_{j=0}^r j b_j P_{j-1};
 \end{aligned}$$

but $jP_{j-1} = (n+2)(P(j, n+2, 0) - P_j)$ by Lemma 6, so that

$$S_{n+1} = a - (n+2)_{1-q} \sum_{j=0}^r b_j P_j - (n+1)_{-q} (n+2) \sum_{j=0}^r b_j (P(j, n+2, 0) - P_j).$$

By (3a) we have $(n+1)_{-q} (n+2) = (n+2)_{1-q}$. So we finally get

$$S_{n+1} = a - (n+2)_{1-q} \sum_{j=0}^r b_j P(j, n+2, 0). \quad \square$$

To illustrate this result, we consider some examples:

Example 1. $r = 0$ gives

$$\sum_{k=0}^n \frac{1}{(k+q) \cdots (k+1)} = \frac{1}{q-1} \left(\frac{1}{(q-1)!} - \frac{1}{(n+q) \cdots (n+2)} \right).$$

Example 2. $r = 1$ gives

$$(34) \quad \sum_{k=0}^n \frac{H_k}{(k+q) \cdots (k+1)} = \frac{1}{(q-1)^2} \left(\frac{1}{(q-1)!} - \frac{1+(q-1)H_{n+1}}{(n+q) \cdots (n+2)} \right).$$

Example 3. $r = 2$ gives

$$\begin{aligned}
 &\sum_{k=0}^n \frac{H_k^2 - H_k^{(2)}}{(k+q) \cdots (k+1)} \\
 (35) \quad &= \frac{1}{(q-1)^3} \left(\frac{2}{(q-1)!} - \frac{2+2(q-1)H_{n+1} + (q-1)^2(H_{n+1}^2 - H_{n+1}^{(2)})}{(n+q) \cdots (n+2)} \right).
 \end{aligned}$$

The case $q = 1$ had to be excluded from Theorem 15. But from Theorem 9 we obtain:

Theorem 16. For all $n, r \in \mathbb{N}$

$$(36) \quad \sum_{k=0}^n \frac{P(r, k, 0)}{k+1} = \frac{P(r+1, n+1, 0)}{r+1}.$$

Proof. Using (7) and Theorem 9, we have

$$\begin{aligned} \frac{1}{r+1} \sum_{n \geq 0} P(r+1, n+1, 0) z^{n+1} &= \frac{1}{r+1} \frac{(-\log(1-z))^{r+1}}{(1-z)} \\ &= \frac{1}{1-z} \sum_{k \geq 0} \frac{P(r, k, 0)}{k+1} z^{k+1} = \sum_{n \geq 0} \left(\sum_{k=0}^n \frac{P(r, k, 0)}{k+1} \right) z^{n+1}. \quad \square \end{aligned}$$

Example. If $r = 1$, then

$$\sum_{k=0}^n \frac{H_k}{k+1} = \frac{1}{2} (H_{n+1}^2 - H_{n+1}^{(2)}).$$

One would like to generalize these results to sums of the form

$$\sum_{k=0}^n \frac{P(r, k, 0)}{(k+p)_q}$$

with $p \neq q$. But even for the simplest case $\sum_{k=0}^n P(r, k, 0)/k$, a closed form cannot be found. This is a consequence of the fact that $(-\log(1-z))^r/z$ cannot be integrated in closed form.

Theorem 7 looks promising but is of help only for $r = 0$ and $r = 1$. For all $k \in \mathbb{N}$ we have $P(0, k, 0) = P(0, 0, k)$ and $P(1, k, 0) = -P(1, 0, k)$. Repeated application of (3b) in the form $p \cdot (k-1)_{-p-1} = (k-1)_{-p} - (k)_{-p}$ gives

$$\frac{1}{(k+p)_{p+1}} = \frac{1}{p!} \frac{1}{k} - \sum_{j=1}^p \frac{1}{(p)_{p+1-j}} \cdot \frac{1}{(k+j)_j},$$

so that

$$(37) \quad \sum_{k=0}^n \frac{H_k}{(k+p)_{p+1}} = \frac{H_{n+1}^2 - H_{n+1}^{(2)}}{2p!} - \sum_{j=1}^p \frac{1}{(p)_{p+1-j}} \cdot \sum_{k=0}^n \frac{H_k}{(k+j)_j}.$$

The sums on the right-hand side can be evaluated by Theorem 15.

For an extension of this result, see Theorem 32.

The key for more general summation formulas is contained in the next two lemmas.

Lemma 17. For $n, p \in \mathbb{N}$ let

$$U_p = \sum_{k=1}^n \frac{H_k}{k+p} \quad \text{and} \quad V_p = \sum_{k=p+1}^n \frac{H_k}{k-p};$$

then

$$\text{for } p > 0: \quad U_{p+1} = U_p - \frac{H_p}{p} + \frac{H_{n+p} - H_n}{p} + \frac{H_n}{n+p+1},$$

$$\text{for } n > p: \quad V_{p+1} = V_p + \frac{H_{p+1}}{p+1} - \frac{H_n - H_{n-p-1}}{p+1} - \frac{H_n}{n-p}.$$

Proof. We prove only the relation for U_p :

$$\begin{aligned} U_p - U_{p+1} &= \sum_{k=1}^n \frac{H_k - H_{k-1}}{k+p} - \frac{H_n}{n+p+1} = \sum_{k=1}^n \frac{1}{(k+p) \cdot k} - \frac{H_n}{n+p+1} \\ &= \frac{1}{p} \left(\frac{1}{1} - \frac{1}{p+1} + \frac{1}{2} - \frac{1}{p+2} + \dots + \frac{1}{n} - \frac{1}{p+n} \right) - \frac{H_n}{n+p+1} \\ &= \frac{1}{p} (H_n - (H_{n+p} - H_p)) - \frac{H_n}{n+p+1} \\ &= \frac{H_p}{p} - \frac{H_{n+p} - H_n}{p} - \frac{H_n}{n+p+1}. \quad \square \end{aligned}$$

The solution of these recurrence relations gives immediately:

Lemma 18. For $n, p \in \mathbb{N}$

$$\begin{aligned} \sum_{k=1}^n \frac{H_k}{k+p+1} &= \frac{1}{2} (H_n^2 - H_n^{(2)}) - \frac{1}{2} (H_p^2 + H_p^{(2)}) \\ &\quad + \sum_{k=1}^p \frac{H_p - H_{k-1}}{n+k} + \sum_{k=1}^{p+1} \frac{1}{n+k} H_n, \end{aligned}$$

and for $p \leq n$

$$\begin{aligned} \sum_{k=p+1}^n \frac{H_k}{k-p} &= \frac{1}{2} (H_n^2 + H_n^{(2)}) + \frac{1}{2} (H_p^2 + H_p^{(2)}) \\ &\quad - \sum_{k=1}^p \frac{H_p - H_{k-1}}{n+1-k} - \sum_{k=1}^p \frac{1}{n+1-k} H_n. \end{aligned}$$

Another line of attack to get new identities and to rediscover old ones is the next lemma which can be found in [4].

Lemma 19. Let

$$F(w, z) = \sum_{r \geq 0} \sum_{n \geq 0} \frac{P(r, n, 0)}{r!} w^r z^n;$$

then $F(w, z) = 1/(1-z)^{w+1}$.

Proof. By (7) we have

$$\begin{aligned} \sum_{r \geq 0} \left(\sum_{n \geq 0} \frac{P(r, n, 0)}{r!} z^n \right) w^r &= \frac{1}{(1-z)} \sum_{r \geq 0} \frac{(-\log(1-z))^r}{r!} w^r \\ &= \frac{1}{(1-z)} \exp(-w \cdot \log(1-z)) = \frac{1}{(1-z)^{w+1}}. \quad \square \end{aligned}$$

Now we can derive many identities:

Example 1. $F(w, z) \cdot F(w, -z) = 1/(1 - z^2)^{w+1} = F(w, z^2)$, which can be written as an identity for the power series

$$(38) \quad \sum_{j=0}^r \sum_{k=0}^{2n} (-1)^k \frac{P(r-j, 2n-k, 0)}{(r-j)!} \cdot \frac{P(j, k, 0)}{j!} = P(r, n, 0)$$

and

$$\sum_{j=0}^r \sum_{k=0}^{2n+1} (-1)^k \frac{P(r-j, 2n+1-k, 0)}{(r-j)!} \cdot \frac{P(j, k, 0)}{j!} = 0.$$

The last equation is trivial because the terms in the sum cancel pairwise.

From the first equation we obtain, for $r = 1$,

$$(39) \quad \sum_{k=0}^{2n} (-1)^k H_k = \frac{1}{2} H_n.$$

Example 2. $F(w, z) \cdot F(-w, z) = 1/(1 - z)^2 = F(0, z)/(1 - z)$, which is equivalent to

$$(40) \quad \sum_{j=0}^r (-1)^j \sum_{k=0}^n \frac{P(r-j, n-k, 0)}{(r-j)!} \cdot \frac{P(j, k, 0)}{j!} = \begin{cases} n+1 & \text{for } r = 0, \\ 0 & \text{for } r > 0. \end{cases}$$

For $r = 2$, this means

$$\sum_{k=0}^n (H_k^2 - H_k^{(2)}) = \sum_{k=0}^n H_k H_{n-k}.$$

Example 3. $F(1, z) = 1/(1 - z)^2 = \sum_{n \geq 0} (n+1)z^n$, which means

$$(41) \quad \sum_{r \geq 0} \frac{P(r, n, 0)}{r!} = n+1 \quad \text{for all } n \in \mathbf{N}.$$

Only a finite number of terms in this sum do not vanish.

4. RECURSION FORMULAS

In this section we give recursion formulas for some sums. These are especially useful when working with a formula manipulating system like REDUCE, which allows recursive procedures.

In practical computations one often needs the values of $T(r, k)$, which are defined by

$$T(r, k) = \frac{1}{r!} P_r(H_k^{(1)}, H_k^{(2)}, \dots, H_k^{(r)}) = \frac{1}{r!} P(r, k, 0).$$

Then, for all $k, r \in \mathbf{N}$,

$$T(r, k+1) = T(r, k) + \frac{T(r-1, k)}{k+1}$$

with boundary conditions $T(0, k) = 1$ and $T(r, 0) = 0$ ($r > 0$).

This is a simple consequence of Lemma 6.

The next lemma is helpful to obtain recursion formulas.

Lemma 20. For all $n, p \in \mathbb{N}$, let $X(p) = \sum_{k=1}^n k^p a_k$ and $b_k = a_k - a_{k-1}$, $1 \leq k \leq n$; then

$$(p + 1)X(p) = (n + 1)^{p+1} a_n - a_0 - \sum_{j=0}^{p-1} \binom{p + 1}{j} X(j) - \sum_{k=1}^n k^{p+1} b_k.$$

Proof. We have

$$\begin{aligned} X(p + 1) + (p + 1)X(p) + \sum_{j=0}^{p-1} \binom{p + 1}{j} X(j) &= \sum_{j=0}^{p+1} \binom{p + 1}{j} X(j) = \sum_{k=1}^n a_k \sum_{j=0}^{p+1} \binom{p + 1}{j} k^j \\ &= \sum_{k=1}^n (k + 1)^{p+1} a_k = (n + 1)^{p+1} a_n + \sum_{k=1}^{n-1} (k + 1)^{p+1} a_k \\ &= (n + 1)^{p+1} a_n + \sum_{k=1}^n k^{p+1} a_{k-1} - a_0 \\ &= (n + 1)^{p+1} a_n + X(p + 1) - \sum_{k=1}^n k^{p+1} b_k - a_0. \end{aligned}$$

Comparing the first and last lines gives the result. \square

An immediate consequence of this lemma is:

Theorem 21. For all $n, p \in \mathbb{N}$, $r \in \mathbb{Z}$, let $Q(p, r) = \sum_{k=1}^n k^p H_k^{(r)}$; then

$$(42) \quad (p + 1)Q(p, r) = (n + 1)^{p+1} H_n^{(r)} - H_n^{(r-p-1)} - \sum_{j=0}^{p-1} \binom{p + 1}{j} Q(j, r).$$

Proof. Set $a_k = H_k^{(r)}$; then $a_0 = 0$ and $b_k = H_k^{(r)} - H_{k-1}^{(r)} = k^{-r}$. With $X(p) = Q(p, r)$, Lemma 20 gives the result. \square

Remarks. 1. The case $r = 0$ gives a well-known identity for sums of powers:

$$(p + 2)Q(p, 0) = (n + 1)^{p+1} n - \sum_{j=0}^{p-1} \binom{p + 1}{j} Q(j, 0).$$

2. The case $p = 0$ reduces to

$$(43) \quad \sum_{k=1}^n H_k^{(r)} = (n + 1)H_n^{(r)} - H_n^{(r-1)}.$$

One can also sum the q th powers of H_k .

Theorem 22. For all $n, q \in \mathbb{N}$, $p \in \mathbb{Z}$, let $R(p, q) = \sum_{k=1}^n k^p H_k^q$; then for all $p \geq -1$

$$(p + 1)R(p, q) = (n + 1)^{p+1} H_n^q - \sum_{j=0}^{p-1} \binom{p + 1}{j} R(j, q) + \sum_{j=1}^q (-1)^j \binom{q}{j} R(p + 1 - j, q - j) - \delta_{0q},$$

where δ_{0q} is the Kronecker delta.

Proof. Set $a_k = H_k^q$; then $a_0 = \delta_{0q}$ and

$$b_k = H_k^q - H_{k-1}^q = H_k^q - \left(H_k - \frac{1}{k}\right)^q = - \sum_{j=1}^q (-1)^j \binom{q}{j} H_k^{q-j} k^{-j}.$$

With $X(p) = R(p, r)$, Lemma 20 gives the result. \square

Remarks. 1. The case $q = 0$ gives a well-known recursion for the sum of powers.

2. For $q > 0$ the terms $R(p + 1 - j, q - j)$ in the second sum may produce a negative first argument. But we know only $R(-r, 0) = H_n^{(r)}$ for $r > 0$ and $R(-1, 1) = \frac{1}{2}(H_n^2 + H_n^{(2)})$. So the theorem seems to be useful only for $q \leq p + 3$.

Example. We have

$$\sum_{k=1}^n H_k^3 = R(0, 3) = (n + 1)H_n^3 - 3R(0, 2) + 3R(-1, 1) - R(-2, 0);$$

but since $R(0, 2) = (n + 1)H_n^2 - 2R(0, 1) + R(-1, 0)$ and $R(0, 1) = (n + 1)H_n - R(0, 0)$, we get

$$\sum_{k=1}^n H_k^3 = (n + 1)H_n^3 - \frac{3}{2}(2n + 1)H_n^2 + 3(2n + 1)H_n - 6n + \frac{1}{2}H_n^{(2)}.$$

Similarly, one could sum the q th powers of $H_k^{(r)}$.

Theorem 23. For all $n, p \in \mathbb{N}$, $r \in \mathbb{Z}$, let $V(p, r) = \sum_{k=1}^n k^p H_{n-k}^{(r)}$; then

$$(p + 1)V(p, r) = ((n + 1)^{p+1} - 1)H_n^{(r)} - \sum_{j=0}^p (-1)^{j+p} \binom{p + 1}{j} (n + 1)^j H_n^{(r-p-1+j)} - \sum_{j=0}^{p-1} \binom{p + 1}{j} V(j, r).$$

Proof. Set $a_k = H_{n-k}^{(r)}$; then we have $a_0 = H_n^{(r)}$ and $b_k = H_{n-k}^{(r)} - H_{n-k+1}^{(r)} = -(n - k + 1)^{-r}$, so that

$$\begin{aligned} \sum_{k=1}^n k^{p+1} b_k &= (-1)^{p+1} \sum_{k=1}^n [n + 1 - k - (n + 1)]^{p+1} \cdot (n + 1 - k)^{-r} \\ &= (-1)^{p+1} \sum_{j=0}^{p+1} \binom{p + 1}{j} (-1)^j (n + 1)^j H_n^{(r-p-1+j)}. \end{aligned}$$

With $X(p) = R(p, r)$, Lemma 20 gives the result. \square

The polynomials $P(r, k, m)$ which have been introduced in §3 may be summed along the same line:

Theorem 24. For all $n, p, r \in \mathbf{N}$, $r \geq 1$, let $Z(p, r) = \sum_{k=1}^n k^p P(r, k, 0)$; then

$$(p+1)Z(p, r) = (n+1)^{p+1}P(r, n, 0) - \sum_{j=0}^{p-1} \binom{p+1}{j} Z(j, r) \\ + \sum_{j \geq 1} (-1)^{j-1} (r)_j Z(p+1-j, r-j).$$

Proof. Set $a_k = P(r, k, 0)$; then, as $r \geq 1$, we have $a_0 = 0$, and, by Lemma 8,

$$b_k = a_k - a_{k-1} = P(r, k, 0) - P(r, k-1, 0) = \frac{r}{k} P(r-1, k-1, 0).$$

Let $T(p, r) = \sum_{k=1}^n k^p b_k$; then

$$T(p, r) = r \cdot \sum_{k=1}^n k^p P(r-1, k-1, 0) \\ = r \cdot \sum_{k=1}^n k^p P(r-1, k, 0) \\ - r \cdot \sum_{k=1}^n k^p (P(r-1, k, 0) - P(r-1, k-1, 0)) \\ = r \cdot Z(p, r-1) - r \cdot T(p-1, r-1);$$

hence

$$T(p, r) = r \cdot Z(p, r-1) - r \cdot (r-1)Z(p-1, r-2) + \dots \\ = \sum_{j \geq 1} (-1)^{j-1} (r)_j Z(p+1-j, r-j).$$

Lemma 20 then gives the result. \square

This theorem is useful to compute $\sum_{k=1}^n k^p (H_k^2 - H_k^{(2)})$ recursively ($r = 2$).

5. STRUCTURE THEOREMS

Quite often it is easy to solve a problem if one knows that a solution exists. If one knows in advance what the solution of the problem will look like, one can decide whether the solution will be simple enough for it to be worthwhile working out the exact solution or just truncating it to obtain the dominant terms.

The results given here are a kind of existence theorems for closed form representations of sums involving generalized harmonic numbers. The closed form of a given sum can then be obtained either by interpolations—which means solving a system of linear equations—or better by application of the summation formulas given in this section. The proofs given are all constructive.

Example. Try to find a closed form representation of the sum

$$S(n) = \sum_{k=1}^n (2k + 3)H_k^2.$$

We have two obvious strategies to find a closed form representation of this sum:

(1) By Theorem 30 below, the sum is of the form

$$S(n) = (a_2 n^2 + a_1 n + a_0)H_n^2 + (b_2 n^2 + b_1 n + b_0)H_n + (c_2 n^2 + c_1 n + c_0).$$

Evaluate $S(0)$, $S(1)$, \dots , $S(8)$ and then solve this linear system for the unknown coefficients a_i , b_i , c_i ($i = 0, 1, 2$). This approach should be used only if exact arithmetic (with rational numbers) can be used. Otherwise, numerical difficulties will arise.

(2) Using Lemma 29 below, we write

$$S(n) = \sum_{k=1}^n \left(2 \binom{k}{1} + 3 \binom{k}{0} \right) H_k^2 = \dots,$$

and after some rearrangement we obtain

$$S(n) = (n + 1)(n + 3)H_n^2 - (n^2 + 5n + 2)H_n + n(n + 9)/2.$$

The next theorem and its proof technique is fundamental for structure theorems. This method of proof is also used by Knuth [4].

Theorem 25. For all $r \in \mathbf{Z}$, $n, p \in \mathbf{N}$

$$(44) \quad \sum_{k=1}^n \binom{k}{p} H_k^{(r)} = \binom{n+1}{p+1} H_n^{(r)} - \frac{1}{(p+1)!} \sum_{k=1}^{p+1} s(p+1, k) H_n^{(r-k)}.$$

Proof. By definition of the $H_k^{(r)}$ we have

$$\begin{aligned} \sum_{k=1}^n \binom{k}{p} H_k^{(r)} &= \sum_{k=1}^n \sum_{j=1}^k \binom{k}{p} j^{-r} = \sum_{j=1}^n \sum_{k=j}^n \binom{k}{p} j^{-r} \\ &= \sum_{j=1}^n j^{-r} \sum_{k=1}^n \binom{k}{p} - \sum_{j=1}^n j^{-r} \sum_{k=1}^{j-1} \binom{k}{p} \\ &= \binom{n+1}{p+1} H_n^{(r)} - \sum_{j=1}^n j^{-r} \binom{j}{p+1} \\ &= \binom{n+1}{p+1} H_n^{(r)} - \frac{1}{(p+1)!} \sum_{j=1}^n \sum_{i=1}^{p+1} s(p+1, i) j^{i-r}. \end{aligned}$$

Interchanging the last sums gives the result. \square

Example 1. $p = 0$ gives

$$\sum_{k=1}^n H_k^{(r)} = (n+1)H_n^{(r)} - H_n^{(r-1)}.$$

Example 2. $r = 2$ gives

$$\sum_{k=1}^n \binom{k}{p} H_k^{(2)} = \binom{n+1}{p+1} H_n^{(2)} - \frac{1}{(p+1)!} [s(p+1, 1)H_n + s(p+1, 2)H_n^{(0)} + \dots + s(p+1, p+1)H_n^{(1-p)}];$$

but $s(p+1, 1) = (-1)^p \cdot p!$ and the terms $H_n^{(0)}, \dots, H_n^{(1-p)}$ are all polynomials of degree at most p of the parameter n . So we have

Corollary 26. For all $n, p \in \mathbb{N}$

$$\sum_{k=1}^n \binom{k}{p} H_k^{(2)} = \binom{n+1}{p+1} H_n^{(2)} - \frac{(-1)^p}{p+1} H_n + A_p(n),$$

where $A_p(n)$ is a polynomial in n of degree p .

An immediate consequence of Theorem 25 is:

Theorem 27. For all $n \in \mathbb{N}$, $r \in \mathbb{Z}$, if $A(k)$ is a polynomial in k of degree q , then

$$\sum_{k=1}^n A(k)H_k^{(r)} = (n+1)B(n)H_n^{(r)} + \sum_{k=1}^{q+1} c_k H_n^{(r-k)},$$

where $B(n)$ is a polynomial in n of degree q . The constants c_1, \dots, c_{q+1} depend on the coefficients of $A(k)$.

Proof. $A(k)$ has a representation $A(k) = \sum_{p=0}^q a_p \binom{k}{p}$. Application of (44) to all terms of this sum gives

$$\begin{aligned} \sum_{k=1}^n A(k)H_k^{(r)} &= \sum_{p=0}^q a_p \sum_{k=1}^n \binom{k}{p} H_k^{(r)} \\ &= \sum_{p=0}^q a_p \left[\binom{n+1}{p+1} H_n^{(r)} - \sum_{k=1}^{p+1} \frac{s(p+1, k)}{(p+1)!} H_n^{(r-k)} \right]. \end{aligned}$$

With

$$B(n) = \sum_{p=0}^q \binom{n}{p} \frac{1}{p+1} a_p \quad \text{and} \quad c_k = - \sum_{p=0}^q a_p \frac{s(p+1, k)}{(p+1)!}$$

we obtain the stated result. \square

For $r = 1$ this theorem specializes to:

Corollary 28. If $A(k)$ is a polynomial of degree q , then

$$\sum_{k=1}^n A(k)H_k = (n+1)B(n)H_n + C(n),$$

where $B(n), C(n)$ are polynomials of degree q .

Even powers of harmonic numbers can be summed. We give just one result, but omit the proof.

Lemma 29. For all $n, p \in \mathbf{N}$

$$(45) \quad \sum_{k=1}^n \binom{k}{p} H_k^2 = \binom{n+1}{p+1} H_n^2 - \left[2 \binom{n}{p+1} + (-1)^p \right] \frac{H_n}{p+1} \\ + \frac{2}{(p+1)^2} \binom{n}{p+1} - \frac{1}{(p+1)!} \sum_{k=2}^{p+1} s(p+1, k) H_n^{(2-k)}.$$

One can obtain a similar, but more complicated, identity for $\sum_{k=1}^n \binom{k}{p} H_k^3$. We therefore have:

Theorem 30. For all $n, p \in \mathbf{N}$, $r \in \mathbf{Z}$, $m \in \{0, 1, 2, 3\}$, if $A_0(k), \dots, A_m(k)$ are polynomials in k of degree at most q , then

$$\sum_{k=1}^n \sum_{p=0}^m A_p(k) H_k^p = \sum_{p=0}^m B_p(n) H_n^p + C(n) H_n^{(2)},$$

where the $B_0(n), \dots, B_m(n)$, $C(n)$ are polynomials in n of degree at most $q+1$.

We conjecture that this theorem holds for all $m \in \mathbf{N}$.

From Theorem 1 we can derive, for $u, v, n, p \in \mathbf{N}$,

$$\sum_{k=0}^n \binom{k}{p} P(u, k, p) P(v, n-k, 0) = \binom{n+1}{p+1} P(u+v, n+1, p+1).$$

This may be restated as

Theorem 31. For all $u, v, n, p \in \mathbf{N}$, if $A_0(k), \dots, A_m(k)$ are polynomials in k of degree at most q , then

$$\sum_{k=1}^n \sum_{p=0}^m A_p(k) P(u, k, p) P(v, n-k, 0) = \sum_{p=0}^m B_p(n) P(u+v, n+1, p+1),$$

where the $B_0(n), \dots, B_m(n)$ are polynomials in n of degree at most $q+1$.

We shall now try to find the structure of sums of the form $\sum_{k=1}^n H_k / (k+p)_q$. By partial fraction expansion we have

$$\frac{1}{(k+p)_q} = \sum_{j=1}^q \frac{a_j}{k+p+1-j} \quad \text{with} \quad \sum_{j=1}^q a_j = 0.$$

Therefore,

$$\sum_{k=1}^n \frac{H_k}{(k+p)_q} = \sum_{j=1}^q a_j \sum_{k=1}^n \frac{H_k}{k+p+1-j}.$$

These sums can be evaluated by Lemma 18.

Because of $\sum_{j=1}^q a_j = 0$, many terms cancel, and we get:

Theorem 32. For all $n, p, q \in \mathbb{N}$

$$\sum_{k=1}^n \frac{H_k}{(k+p)_q} = c_1 + \sum_{k=p-q+2}^p \frac{d_k}{n+k} H_n + \sum_{k=1}^{p-1} \frac{b_k}{n+k} \quad \text{for } 0 < q \leq p,$$

$$\sum_{k=q-p}^n \frac{H_k}{(k+p)_q} = c_1 + \sum_{k=p-q+2}^p \frac{d_k}{n+k} H_n + \sum_{k=p-q+2}^{p-1} \frac{b_k}{n+k} + c_2 H_n^{(2)}$$

for $q \geq p + 1 > 1$,

$$\sum_{k=p+q}^n \frac{H_k}{(k-p)_q} = c_1 + \sum_{k=p}^{p+q-2} \frac{d_k}{n-k} H_n + \sum_{k=0}^{p+q-2} \frac{b_k}{n-k},$$

where c_k, d_k, b_k are constants depending on p and q .

6. OPEN QUESTIONS

1. It seems natural to consider the coefficients of

$$\left(\frac{-\log(1-z)}{1-z} \right)^{-1} = \frac{1}{z} + \sum_{k \geq 0} a_k z^k,$$

where $a_0 = -H_2$, $a_1 = H_2^2 - H_3$, $a_2 = -H_2^3 + 2H_2H_3 - H_4, \dots$. Obviously, many nice summation formulas can be derived for these quantities. Do these numbers occur in a natural way in applications?

2. Is the conjecture following Theorem 30 true?

3. Is it possible to introduce a few new functions so that any sum $\sum k^p H_k^q$ can be expressed by these functions for $p, q \in \mathbb{Z}$?

7. CONCLUSION

We have shown in this paper that

- (1) many complicated-looking summation problems with harmonic numbers can be solved;
- (2) there are two natural generalizations of the harmonic numbers, namely $H_n^{(r)}$ and the function $P(r, n, m)$, which share many properties with the harmonic numbers; and
- (3) sums of powers behave similarly to sums of harmonic numbers.

APPENDIX

In this appendix we give some special sums which might be helpful if one carries out symbolic computations ‘by hand’, in contrast to using a formula manipulation system.

All equations can be derived from the results of §3 with the aid of

Lemma 33. *Given summation formulas $\sum_{k=0}^n \binom{k}{p} a_k = F(n, p)$, $n, p \in \mathbf{N}$, one has*

$$(46) \quad \sum_{k=0}^n k^p a_k = \sum_{k=0}^p S(p, k) \cdot k! \cdot F(n, k).$$

Proof. Using (4b), we have

$$\begin{aligned} \sum_{k=0}^n k^p a_k &= \sum_{k=0}^n a_k \sum_{j=0}^p S(p, j) \cdot (k)_j = \sum_{j=0}^p S(p, j) \cdot j! \sum_{k=0}^n \binom{k}{j} a_k \\ &= \sum_{j=0}^p S(p, j) \cdot j! \cdot F(n, j). \quad \square \end{aligned}$$

From Corollary 2 we can derive as special cases:

$$\sum_{k=0}^n \binom{k}{p} H_k = \binom{n+1}{p+1} \left[H_{n+1} - \frac{1}{p+1} \right]$$

and

$$\sum_{k=0}^n \binom{k}{p} H_{n-k} = \binom{n+1}{p+1} [H_{n+1} - H_{p+1}].$$

From Corollary 3 we obtain

$$\begin{aligned} \sum_{k=0}^n \binom{k}{p} H_k H_{n-k} &= \binom{n+1}{p+1} \left[H_{n+1}^2 - \left(H_{p+1} + \frac{1}{p+1} \right) H_{n+1} \right. \\ &\quad \left. - H_{n+1}^{(2)} + \frac{1}{p+1} H_{p+1} + H_{p+1}^{(2)} \right]. \end{aligned}$$

From Corollary 26 we have

$$\sum_{k=0}^n \binom{k}{p} H_k^{(2)} = \binom{n+1}{p+1} H_n^{(2)} - \frac{(-1)^p}{p+1} H_n + \dots,$$

and Lemma 29 states

$$\sum_{k=0}^n \binom{k}{p} H_k^2 = \binom{n+1}{p+1} H_n^2 - [\dots] H_n + \dots.$$

To complete this list, we note

$$\sum_{k=0}^n \binom{k}{p} = \binom{n+1}{p+1}.$$

Using these equations and Lemma 33, we get:

$$\sum_{k=0}^n k^p = A(p, n),$$

$$\sum_{k=0}^n k^p H_k = A(p, n)H_{n+1} - B(p, n),$$

$$\sum_{k=0}^n k^p H_{n-k} = A(p, n)H_{n+1} - C(p, n),$$

$$\sum_{k=0}^n k^p H_k H_{n-k} = A(p, n)(H_{n+1}^2 - H_{n+1}^{(2)}) - (B(p, n) + C(p, n))H_{n+1} + D(p, n),$$

$$\sum_{k=0}^n k^p H_k^{(2)} = A(p, n)H_n^{(2)} - B_p H_n - E(p, n),$$

$$\sum_{k=0}^n k^p H_k^2 = A(p, n)H_n^2 - F(p, n)H_n + G(p, n),$$

where the $A(p, n), \dots, G(p, n)$ are polynomials in n and the B_p are the Bernoulli numbers. The $A(p, n)$ are well known (see, e.g., Riordan [6]).

Table 1 lists some values for small p of these polynomials.

TABLE 1

p	$A(p, n)$	$B(p, n)$	$C(p, n)$
0	$n + 1$	$n + 1$	$n + 1$
1	$\frac{n(n+1)}{2}$	$\frac{n(n+1)}{2} \cdot \frac{1}{2}$	$\frac{n(n+1)}{2} \cdot \frac{3}{2}$
2	$\frac{n(n+1)(2n+1)}{6}$	$\frac{n(n+1)}{3} \cdot \frac{4n+5}{12}$	$\frac{n(n+1)}{3} \cdot \frac{22n+5}{12}$
3	$\left(\frac{n(n+1)}{2}\right)^2$	$\frac{n(n+1)}{4} \cdot \frac{(n+2)(3n+1)}{12}$	$\frac{n(n+1)}{4} \cdot \frac{25n^2 + 13n - 2}{12}$

p	$D(p, n)$	$E(p, n)$	$F(p, n)$	$G(p, n)$
0	$2(n+1)$	0	$2n+1$	$2n$
1	$n(n+1)$	$\frac{n}{2}$	$\frac{n^2 - n - 1}{2}$	$\frac{n(n-3)}{4}$
2	$n(n+1) \cdot \frac{71n+37}{108}$	$\frac{n(n-2)}{6}$	$\frac{4n^3 - 3n^2 - n + 3}{18}$	$n \cdot \frac{8n^2 - 15n + 25}{108}$
3	$n^2(n+1) \cdot \frac{35n+37}{72}$	$n(n-1) \frac{2n-1}{24}$	$n(n^2-1) \frac{3n-2}{24}$	$n(n-1) \frac{9n^2 - 5n + 10}{288}$

It remains to prove that $\sum_{k=0}^n k^p H_k^{(2)} = A(p, n)H_n^{(2)} - B_p H_n - E(p, n)$. But this results from $\sum_{k=0}^n \binom{k}{p} H_k^{(2)} = \dots - (-1)^p H_n / (p+1) + \dots$ and the following lemma, which connects the Stirling numbers with the Bernoulli numbers.

Lemma 34. For all $p \in \mathbb{N}$

$$(47) \quad \sum_{k=0}^p (-1)^k \frac{S(p, k) \cdot k!}{k+1} = B_p.$$

Proof. The following identities are well known (see e.g. Riordan [6]):

$$(e^z - 1)^k = k! \cdot \sum_{p \geq 0} S(p, k) \frac{z^p}{p!} \quad \text{and} \quad \frac{z}{e^z - 1} = \sum_{p \geq 0} B_p \cdot \frac{z^p}{p!}.$$

Then,

$$\begin{aligned} \sum_{p \geq 0} \sum_{k=0}^p (-1)^k \frac{S(p, k) \cdot k!}{k+1} \frac{z^p}{p!} &= \sum_{k \geq 0} \frac{(-1)^k}{k+1} \cdot k! \cdot \sum_{p \geq 0} S(p, k) \frac{z^p}{p!} \\ &= \sum_{k \geq 0} \frac{(-1)^k}{k+1} (e^z - 1)^k = \frac{1}{e^z - 1} \cdot \log(1 - (1 - e^z)) = \frac{z}{e^z - 1}. \end{aligned}$$

Comparing coefficients gives the result. \square

The first few Bernoulli numbers are:

$$B_0 = 1, \quad B_1 = -1/2, \quad B_2 = 1/6, \quad B_3 = 0, \quad B_4 = -1/30.$$

The next formulas are immediate consequences of Lemma 18.

$$\begin{aligned} \sum_{k=1}^n \frac{H_k}{(k+2)(k+1)} &= 1 - \frac{1}{n+2} H_n - \frac{1}{n+1}, \\ \sum_{k=1}^n \frac{H_k}{(k+1)k} &= H_n^{(2)} - \frac{1}{n+1} H_n, \\ \sum_{k=2}^n \frac{H_k}{k(k-1)} &= 2 - \frac{1}{n} H_n - \frac{1}{n}, \\ \sum_{k=3}^n \frac{H_k}{(k-1)(k-2)} &= \frac{1}{2} \left(\frac{9}{2} - \frac{2}{n-1} H_n - \frac{1}{n-1} - \frac{1}{n} \right), \\ \sum_{k=1}^n \frac{H_k}{(k+2)(k+1)k} &= \frac{1}{2} \left[H_n^{(2)} - 1 - \left(\frac{1}{n+1} - \frac{1}{n+2} \right) H_n + \frac{1}{n+1} \right], \\ \sum_{k=2}^n \frac{H_k}{(k+1)k(k-1)} &= \frac{1}{2} \left[-H_n^{(2)} + \frac{5}{2} - \left(\frac{1}{n} - \frac{1}{n+1} \right) H_n - \frac{1}{n} \right], \end{aligned}$$

$$\sum_{k=3}^n \frac{H_k}{k(k-1)(k-2)} = \frac{1}{4} \left[2 - 2 \left(\frac{1}{n-1} - \frac{1}{n} \right) H_n - \frac{1}{n-1} + \frac{1}{n} \right],$$

$$\sum_{k=2}^n \frac{H_k}{(k+2)(k+1)k(k-1)} = \frac{1}{6} \left[-2H_n^{(2)} + \frac{23}{6} - \left(\frac{1}{n} - \frac{2}{n+1} + \frac{1}{n+2} \right) H_n - \frac{1}{n} - \frac{1}{n+1} \right],$$

$$\sum_{k=3}^n \frac{H_k}{(k+1)k(k-1)(k-2)} = \frac{1}{12} \left[2H_n^{(2)} - 2 - \left(\frac{2}{n-1} - \frac{4}{n} + \frac{2}{n+1} \right) H_n - \frac{1}{n-1} + \frac{3}{n} \right].$$

BIBLIOGRAPHY

1. M. Abramowitz and I. A. Stegun, *Handbook of mathematical functions*, Dover, New York, 1965.
2. M. Karr, *Summation in finite terms*, J. Assoc. Comput. Mach. **28** (1981), 305–350.
3. R. Kemp, *Fundamentals of the average case analysis of particular algorithms*, Wiley-Teubner Series in Computer Science, Stuttgart, 1984.
4. D. E. Knuth, *The art of computer programming*, Vols. 1–3, Addison-Wesley, Reading, Mass., 1968.
5. J. C. Lafon, *Summation in finite terms*, Computing Suppl. **4** (1982), 71–77.
6. J. Riordan, *Combinatorial identities*, R. E. Krieger, Huntington, N.Y., 1979.
7. P. B. M. Roes, *A note on linear programming design: A combinatorial problem*, Comm. ACM **9** (1966), 340–342.
8. R. Sedgewick, *The analysis of quicksort programs*, Acta Inform. **7** (1977), 327–355.
9. J. Spieß, *Mathematische Probleme der Analyse eines Algorithmus*, Z. Angew. Math. Mech. **63** (1983), T429–T431.
10. D. A. Zave, *A series expansion involving the harmonic numbers*, Inform. Process. Lett. **5** (1976), 75–77.

INSTITUT FÜR PROGRAMMIERSPRACHEN UND INFORMATIONSSYSTEME, TECHNISCHE UNIVERSITÄT
 BRAUNSCHWEIG, D-3300 BRAUNSCHWEIG, FEDERAL REPUBLIC OF GERMANY
 E-mail address: spies@dbisinf6.bitnet