#A68 INTEGERS 26 (2026)

COMBINATORIAL IDENTITIES INSPIRED BY KNUTH'S OLD SUM

Akalu Tefera
Department of Mathematics, Grand Valley State University, Allendale, Michigan
`teferaa@gvsu.edu`

Aklilu Zeleke
Lyman Briggs College and Department of Statistics and Probability, Michigan State University, East Lansing, Michigan
`zeleke@msu.edu`

Received: 3/16/26, Accepted: 5/10/26, Published: 6/8/26

Abstract

We present new combinatorial proofs for a family of alternating binomial identities arising in the study of generalized forms of Knuth's old sum. Our approach uses sign-reversing involutions and equivalence relations on words over a finite set of alphabets, yielding new combinatorial identities. These results also give rise to new integer sequences corresponding to special cases of the sums studied here.

1. Introduction

Alternating binomial sums involving powers of two arise naturally in many areas of combinatorics, number theory, and computer science [6, 3, 1, 4]. A classical prototype is Knuth's old sum, a binomial identity that was studied by several authors [7, 8, 9, 10]. In this work, we investigate the sequence

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

from which interesting combinatorial identities with rich combinatorial structures are generated.

Initial computational explorations suggest that the values of $S(n,j)$, for fixed j, define new integer sequences not previously cataloged. For example, for $j=0$ and $j=1$, the sequences $\{S(n,0)\}_{n\geq 0}$ and $\{S(n,1)\}_{n\geq 0}$ appear in the On-Line Encyclopedia of Integer Sequences as OEIS A390529 and OEIS A390530, respectively [5].

DOI: 10.5281/zenodo.20597521

The present paper develops a uniform combinatorial framework that explains these identities and proves them rigorously. Our approach utilizes direct enumeration of equivalence classes and sign-reversing involutions (reminiscent of the DIE method [2]) on words over finite alphabets. Both approaches lead to closed forms involving alternating sum expressions.

The resulting identities generate new integer sequences and point toward broader families of combinatorial and hypergeometric-type identities. We conclude with a brief section indicating how the main results can also be verified by the Wilf–Zeilberger (WZ) method.

2. Main Results

Theorem 1. *For nonnegative integers n and j with $N = n + 2j + 1$, we have*

$$\sum_{k=0}^{n}(2^{n-k}+1)2^{n-k-1}\binom{N}{n-k}(-1)^k = \binom{N}{n}\sum_{k=0}^{n}(2^{n-k}+1)2^{n-k-1}\frac{\binom{n}{k}}{\binom{2j+1+k}{2j+1}}(-1)^k. \quad (1)$$

Theorem 2. *For nonnegative integers n and j with $N = n + 2j + 1$, we have*

$$\sum_{k=0}^{n}(2^{n-k}+1)2^{n-k-1}\binom{N}{n-k}(-1)^k = \sum_{k=0}^{n}\binom{2j+k}{k}2^{k-1}(1+2^k3^{n-k}). \quad (2)$$

3. Proofs of Main Results

Lemma 1. *Let S be the set of all words of length n over the alphabet $\Sigma = \{a, b, c, d\}$. Let E be the set of words in S that contain an even number of a's. Then*

$$|E| = \frac{4^n + 2^n}{2} = 2^{n-1}(2^n + 1). \quad (3)$$

Proof. Note that $|S| = 4^n$. We partition S into two disjoint subsets E and O, where O is the set of words in S with an odd number of a's. We consider two cases.

Case 1: The word w has no a or b. Let

$$T := \{w \in S : w \text{ uses only letters from } \{c, d\}\}.$$

Then $|T| = 2^n$. Every word in T has zero occurrences of a, and hence belongs to E. In particular, $T \subseteq E$.

Case 2: The word w has at least one a or b. Let $R := S \setminus T$ be the set of words that contain at least one a or b. Define an involution f on R as follows: read w from left to right until you encounter the first a or b. If it is an a, change it to b; if it is a b, change it to a. For example,

$$f(cbda) = cada, \qquad f(cada) = cbda.$$

The map f is an involution that toggles the number of a's by ±1. This swaps the parity of the count of a's (even to odd or odd to even). Consequently, R is partitioned into disjoint 2-cycles, each containing one word from E and one from O. Thus, exactly half of the words in R lie in E.

Combining Cases 1 and 2, we get

$$\begin{aligned} |E| &= |T| + \tfrac{1}{2}|R| \\ &= 2^n + \tfrac{1}{2}(4^n - 2^n) \\ &= \frac{4^n + 2^n}{2}, \end{aligned}$$

which is precisely Equation (3). $\square$

Proof of Theorem 1. Let S be the set of words of length N over the alphabet

$$\Gamma := \{a, b, c, d, A, *\}$$

satisfying the following conditions: (i) exactly $2j + 1$ positions contain the symbol "$*$"; (ii) the remaining n positions contain letters from $\{a, b, c, d, A\}$; (iii) among the lowercase letters $\{a, b, c, d\}$, the number of a's is even.

For any $w \in S$, let $A(w)$ denote the number of A's in w. Assign to w the weight

$$\mathrm{wt}(w) = (-1)^{A(w)}.$$

Define a relation $\sim$ on S as follows. For $w, w' \in S$, say that $w \sim w'$ if and only if (i) w and w' have the same set L of lowercase positions, and (ii) at each position of L, the lowercase letters in w and w' agree. One can verify that $\sim$ is an equivalence relation on S. Let $\mathcal{S}$ denote the set of equivalence classes under this relation.

Since n and j are fixed, once the set L of lowercase positions and the lowercase letters occupying those positions are specified, the positions outside L must contain exactly $(2j + 1)\,*$'s. All the remaining positions must be occupied by A's. In particular, every word u in the class $[w]$ has the same number of A's. Thus, $A(u)$ is constant on $[w]$, and hence every $u \in [w]$ has the same weight $\mathrm{wt}(u) = (-1)^{A(u)}$. Therefore, the weight of a class is well defined.

Set $\mathrm{wt}([w]) := \mathrm{wt}(w)$. The total weight

$$\sum_{[w]\in\mathcal{S}} \mathrm{wt}([w])$$

will be computed in two different ways. First, group the equivalence classes according to the number of uppercase letters. Let $[w] \in \mathcal{S}$ be fixed, and let $k = A(w)$ be the number of A's in any representative of $[w]$. Set $m = n - k$; this is the number of lowercase letters.

An equivalence class with parameter k is determined by two choices: a choice of the m lowercase positions $L \subseteq \{1, \ldots, N\}$, and a choice of the lowercase word on L having an even number of a's. The first choice can be made in $\binom{N}{m}$ ways, and the second, by Lemma 1, in

$$(2^m + 1)2^{m-1} = (2^{n-k} + 1)2^{n-k-1}$$

ways. Since every such class has weight $(-1)^k$, the total contribution from all classes with parameter k is

$$(2^{n-k} + 1)2^{n-k-1}\binom{N}{n-k}(-1)^k.$$

Summing over k gives

$$\sum_{[w]\in\mathcal{S}} \mathrm{wt}([w]) = \sum_{k=0}^{n}(2^{n-k} + 1)2^{n-k-1}\binom{N}{n-k}(-1)^k. \tag{4}$$

The same total weight can be computed by first choosing the set of positions that contain letters and then correcting for the multiplicity with which each equivalence class is represented. To see the source of this multiplicity, consider the case $n = 2$, $j = 1$, and $k = 1$, so that $N = 2j + 1 + n = 5$. Fix the lowercase position set $L = \{3\}$. The letter in position 3 is lowercase, and the remaining four positions lie outside L. Since $k = 1$, exactly one of these four positions contains the letter A, producing four configurations, as shown in Figure 1.

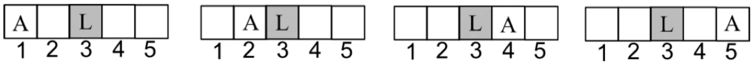

Figure 1: Fixing the lowercase position set $L = \{3\}$ (shaded box) produces four representatives of a single equivalence class, corresponding to the different placements of A outside L.

By definition of $\sim$, the equivalence class is determined solely by the set L and the lowercase letter in position 3. Thus, varying the locations of the A's among the remaining positions does not change the class, and all four configurations represent the same equivalence class. Hence, the number of representatives of this class is

$$\binom{2j+1+k}{k} = \binom{4}{1} = 4.$$

The general case is similar. Let $N = 2j+1+n$. First choose the set B of the n positions that contain letters. This can be done in $\binom{N}{n}$ ways. Next choose a subset $U \subseteq B$ of size k consisting of the positions occupied by A. This can be done in $\binom{n}{k}$ ways. The lowercase positions are then $L = B \setminus U$, so that $|L| = n-k$. Fill the positions in L with a lowercase word containing an even number of a's. By Lemma 1, this can be done in

$$(2^{n-k}+1)2^{n-k-1}$$

ways. Now fix L and its lowercase word. The k occurrences of A occupy a k-subset of the $N-(n-k) = 2j+1+k$ positions outside L. Hence, each equivalence class is represented

$$\binom{2j+1+k}{k} = \binom{2j+1+k}{2j+1}$$

times. Each such choice determines an equivalence class of weight $(-1)^k$. After dividing by this multiplicity factor, we obtain

$$\sum_{[w]\in\mathcal{S}} \mathrm{wt}([w]) = \binom{N}{n}\sum_{k=0}^{n}(2^{n-k}+1)2^{n-k-1}\frac{\binom{n}{k}}{\binom{2j+1+k}{2j+1}}(-1)^k. \tag{5}$$

Hence, from Equations (4) and (5), Equation (1) follows. □

Proof of Theorem 2. Set $N = n+2j+1$ and $r = 2j+1$. Let $\mathcal{U}$ consist of words of length N over the alphabet $\{*, b, A, C\}$ containing exactly r $*$'s, such that every occurrence of C lies to the right of the last $*$. For $u \in \mathcal{U}$, define

$$\mathrm{wt}(u) = \frac{1}{2}(-1)^{C(u)},$$

where $C(u)$ denotes the number of C's in u.

Let $\mathcal{V}$ consist of words of length N over the alphabet $\{*, b, c, d, A, D\}$ containing exactly r $*$'s, such that every occurrence of D lies to the right of the last $*$. For $v \in \mathcal{V}$, define

$$\mathrm{wt}(v) = \frac{1}{2}(-1)^{D(v)},$$

where $D(v)$ denotes the number of D's in v.

Let $\mathcal{S}$ be the disjoint union of the two families $\mathcal{U}$ and $\mathcal{V}$; that is, an element of $\mathcal{S}$ is a word together with the specification of whether it is viewed as an element of $\mathcal{U}$ or of $\mathcal{V}$. The total weight

$$\sum_{w\in\mathcal{S}} \mathrm{wt}(w)$$

will be computed in two ways.

For the first computation, group the words by the number of letters. Consider the contribution from $\mathcal{U}$. Fix $m \in \{0,1,\ldots,n\}$, and consider words in $\mathcal{U}$ having exactly m letters from $\{b, A\}$. Choose their positions in $\binom{N}{m}$ ways and fill them in 2^m ways. The remaining $N-m$ positions can be filled by only $*$'s and C's. Since the word contains exactly r $*$'s and every C must lie to the right of the last $*$, the leftmost r of these positions are $*$'s and the remaining $n-m$ positions are C's. Thus, every such word has exactly $n-m$ occurrences of C, and hence the total contribution is

$$\frac{1}{2}\binom{N}{m}2^m(-1)^{n-m}.$$

Summing over m gives

$$\sum_{u\in\mathcal{U}} \mathrm{wt}(u) = \frac{1}{2}\sum_{m=0}^{n}\binom{N}{m}2^m(-1)^{n-m}.$$

The contribution from $\mathcal{V}$ is computed similarly. Fix $m \in \{0,1,\ldots,n\}$, and consider words in $\mathcal{V}$ having exactly m letters from $\{b,c,d,A\}$. Choose their positions in $\binom{N}{m}$ ways and fill them in 4^m ways. The remaining $N-m$ positions can contain only $*$'s and D's. By the same reasoning, the leftmost r of these positions are $*$'s and the remaining $n-m$ positions are D's. Hence, every such word has exactly $n-m$ occurrences of D, and their total contribution is

$$\frac{1}{2}\binom{N}{m}4^m(-1)^{n-m}.$$

Therefore,

$$\sum_{v\in\mathcal{V}} \mathrm{wt}(v) = \frac{1}{2}\sum_{m=0}^{n}\binom{N}{m}4^m(-1)^{n-m}.$$

Combining these two cases, the sum of the weights of words in $\mathcal{S}$ equals

$$\sum_{w\in\mathcal{S}} \mathrm{wt}(w) = \frac{1}{2}\sum_{m=0}^{n}\binom{N}{m}(2^m+4^m)(-1)^{n-m}. \tag{6}$$

Since

$$\frac{2^m+4^m}{2} = (2^m+1)2^{m-1},$$

Equation (6) becomes

$$\sum_{w\in\mathcal{S}} \mathrm{wt}(w) = \sum_{m=0}^{n} \binom{N}{m}(2^m+1)2^{m-1}(-1)^{n-m}.$$

Replacing m by $n-k$, we obtain

$$\sum_{w\in\mathcal{S}} \mathrm{wt}(w) = \sum_{k=0}^{n}(2^{n-k}+1)2^{n-k-1}\binom{N}{n-k}(-1)^k,$$

which is the left-hand side of Equation (2).

The same total weight can also be evaluated using sign-reversing involutions. For $u \in \mathcal{U}$, begin reading the word from the last occurrence of $*$ from left to right. If every such letter is b, leave u fixed. Otherwise, let p be the first position to the right of the last $*$ whose letter is not b. Then this letter must be either A or C. Replace A by C and C by A at position p.

This map is an involution on $\mathcal{U}$. For every non-fixed word u, it toggles exactly one letter, changing A to C or C to A. Therefore, the number $C(u)$ changes by $+1$ or by -1. Hence, the sign of $\mathrm{wt}(u)$ is reversed. It follows that all non-fixed words cancel in pairs, and the total contribution of $\mathcal{U}$ equals the total weight of its fixed points.

A fixed word in $\mathcal{U}$ has only b's to the right of the last $*$. If exactly k letters occur before the last $*$, then, before the last $*$, we have $(r-1)$ $*$'s and k letters from $\{b, A\}$. This can be arranged in $\binom{r-1+k}{k}$ ways, and the k letters can be chosen in 2^k ways. Thus,

$$\sum_{u\in\mathcal{U}} \mathrm{wt}(u) = \frac{1}{2}\sum_{k=0}^{n}\binom{r-1+k}{k}2^k.$$

Similarly, for $v \in \mathcal{V}$, begin reading the word from the last occurrence of $*$ from left to right. If every such letter lies in $\{b, c, d\}$, leave v fixed. Otherwise, let p be the first position to the right of the last $*$ whose letter is not in $\{b, c, d\}$. Then this letter must be either A or D. Replace A by D and D by A at position p.

Again, this map is an involution on $\mathcal{V}$. For every non-fixed word v, it toggles exactly one letter, changing A to D or D to A. Therefore, the number $D(v)$ changes by $+1$ or by -1. Hence, the sign of $\mathrm{wt}(v)$ is reversed. Thus, all non-fixed words cancel in pairs, and only fixed points contribute.

A fixed word in $\mathcal{V}$ has only letters from $\{b, c, d\}$ to the right of the last $*$. If exactly k letters occur before the last $*$, then, before the last $*$, we have $(r-1)$ $*$'s and k letters from $\{b, c, d, A\}$. This can be arranged in $\binom{r-1+k}{k}$ ways, and these k letters can be chosen in 4^k ways. The remaining $n-k$ positions to the right of the last $*$ can be filled in 3^{n-k} ways. Hence,

$$\sum_{v\in\mathcal{V}} \mathrm{wt}(v) = \frac{1}{2}\sum_{k=0}^{n}\binom{r-1+k}{k}4^k3^{n-k}.$$

Adding the weight contributions of $\mathcal{U}$ and $\mathcal{V}$ yields

$$\begin{aligned}\sum_{w\in\mathcal{S}} \mathrm{wt}(w) &= \frac{1}{2}\sum_{k=0}^{n}\binom{r-1+k}{k}(2^k+4^k3^{n-k})\\ &= \sum_{k=0}^{n}\binom{r-1+k}{k}2^{k-1}(1+2^k3^{n-k}).\end{aligned}$$

Since $r-1=2j$, we have $\binom{r-1+k}{k}=\binom{2j+k}{k}$, and therefore

$$\sum_{w\in\mathcal{S}} \mathrm{wt}(w) = \sum_{k=0}^{n}\binom{2j+k}{k}2^{k-1}(1+2^k3^{n-k}),$$

which is the right-hand side of Equation (2).

Both computations evaluate the same total weight $\sum_{w\in\mathcal{S}}\mathrm{wt}(w)$. Hence, the two sides of Equation (2) are equal, completing the proof. □

4. Proofs by the WZ Method

We note that the main theorems of this paper can also be verified by the WZ method. For background on the WZ method and related algorithmic proof techniques, see $A = B$ [6]. The relevant recurrences can be generated, for example, by Maple's built-in package `SumTools` or by Zeilberger's Maple package `EKHAD` [11]. Since the main focus of this paper is on combinatorial proofs, we illustrate the method only for Theorem 2, as it applies in exactly the same way to Theorem 1.

Let $L(n,j)$ and $R(n,j)$ denote the left-hand side and right-hand side, respectively, of Equation (2). A standard application of Zeilberger's algorithm shows that both satisfy the same recurrence equation

$$S(n+2,j)-4S(n+1,j)+3S(n,j)=B(n,j),$$

where

$$B(n,j)=\frac{2^n}{n+2}\left(16\cdot 2^n j+6\cdot 2^n n+12\cdot 2^n+4j-n-2\right)\binom{2j+n+1}{n+1}.$$

Moreover, $L(0,j)=R(0,j)=1$ and $L(1,j)=R(1,j)=6j+5$. It follows that $L(n,j)=R(n,j)$ for all $n,j\geq 0$, yielding a WZ proof of Theorem 2.

5. Conclusion

In this paper, we have established new combinatorial identities involving alternating binomial sums weighted by powers of two and rational binomial factors. Our proofs are based on sign-reversing involutions on words over finite alphabets. The main significance of this work lies in providing combinatorial proofs that serve as transparent alternatives to analytic or hypergeometric methods, while also yielding natural combinatorial interpretations of the identities. The results obtained here were inspired by earlier work in [10]. We also note that the WZ method can be used to verify the main results.

To the best of our knowledge, identities in exactly these forms have not appeared previously in the literature, including in [3]. Their structure suggests the existence of broader families of binomial and hypergeometric-type identities involving alternating sums. In addition, our identities give rise to integer sequences suitable for inclusion in the OEIS. Future work may include the development of q-analogues and multivariate extensions of the results presented here.

Acknowledgement. The authors thank the editors and the anonymous referee for their helpful suggestions and valuable comments.

References

[1] A. T. Benjamin and J. J. Quinn, *Proofs That Really Count: The Art of Combinatorial Proof*, Mathematical Association of America, Washington, DC, 2003.

[2] A. T. Benjamin and J. J. Quinn, An alternate approach to alternating sums: a method to DIE for, *College Math. J.* **39** (2008), no. 3, 191–201.

[3] H. W. Gould, *Combinatorial Identities: A Standardized Set of Tables Listing 500 Binomial Coefficient Summations*, revised ed., Morgantown, WV, 1972.

[4] R. L. Graham, D. E. Knuth, and O. Patashnik, *Concrete Mathematics: A Foundation for Computer Science*, 2nd ed., Addison–Wesley, Reading, MA, 1994.

[5] OEIS Foundation Inc., The On-Line Encyclopedia of Integer Sequences, `https://oeis.org`.

[6] M. Petkovšek, H. S. Wilf, and D. Zeilberger, $A = B$, A K Peters, Wellesley, MA, 1996.

[7] H. Prodinger, Knuth's old sum—a survey, *EATCS Bull.* **52** (1994), 232–245.

[8] H. Prodinger, Some information about the binomial transform, *Fibonacci Quart.* **32** (1994), no. 5, 412–415.

[9] A. K. Rathie, I. Kim, and R. B. Paris, A note on a generalization of two well-known combinatorial identities via hypergeometric series approach, *Integers* **22** (2022), #A28.

[10] A. Tefera and A. Zeleke, On proofs of generalized Knuth's old sum, *Integers* **23** (2023), #A99.

[11] D. Zeilberger, *EKHAD* (Maple package), 2018. Available at `https://sites.math.rutgers.edu/~zeilberg/tokhniot/EKHAD`.