

involve

a journal of mathematics

Generalized exponential sums
and the power of computers

Francis N. Castro, Oscar E. González and Luis A. Medina



Generalized exponential sums and the power of computers

Francis N. Castro, Oscar E. González and Luis A. Medina

(Communicated by Kenneth S. Berenhaut)

Today's era can be characterized by the rise of computer technology. Computers have been, to some extent, responsible for the explosion of the scientific knowledge that we have today. In mathematics, for instance, we have the four color theorem, which is regarded as the first celebrated result to be proved with the assistance of computers. In this article we generalize some fascinating binomial sums that arise in the study of Boolean functions. We study these generalizations from the point of view of integer sequences and bring them to the current computer age of mathematics. The asymptotic behavior of these generalizations is calculated. In particular, we show that a previously known constant that appears in the study of exponential sums of symmetric Boolean functions is universal in the sense that it also emerges in the asymptotic behavior of all of the sequences considered in this work. Finally, in the last section, we use the power of computers and some remarkable algorithms to show that these generalizations are holonomic; i.e., they satisfy homogeneous linear recurrences with polynomial coefficients.

1. Introduction

Number theory and combinatorics often offer tantalizing objects that captivate the imaginations of mathematicians. Almost all of us have played with prime numbers, explored open problems like Goldbach's conjecture or drawn a lattice on a paper just to see how Catalan numbers work. Nowadays, computer technology allows us to extend the limits of our knowledge and explore these objects in a way that was almost unimaginable 40 years ago. In this work, we pay close attention to some binomial sums that come from the theory of Boolean functions. These binomial sums emerge when the problem of balancedness of these functions is considered. As it is a common practice in mathematics, the idea in this work is to study these binomial sums in a more general framework. Once the proper framework is established, we use the power of computers to expand our knowledge. We start this work with a

MSC2010: 11B37, 11T23, 06E30.

Keywords: Boolean functions, binomial sums, holonomic sequences.

short survey of Boolean functions and exponential sums in an effort to make the manuscript self-contained. The expert reader may skip the majority of it.

A Boolean function is a function from the vector space \mathbb{F}_2^n to \mathbb{F}_2 , where $\mathbb{F}_2 = \{0, 1\}$ is the binary field and n is some positive integer. These functions are beautiful combinatorial objects with applications to many areas of mathematics as well as outside the discipline. Some examples include combinatorics, electrical engineering, game theory, the theory of error-correcting codes, and cryptography. In the current era, efficient implementations of Boolean functions with many variables is a challenging problem due to memory restrictions of current technology. Because of this, symmetric Boolean functions are good candidates for efficient implementations.

It is known that every Boolean function can be identified with a multivariable polynomial. Let $F(\mathbf{X}) = F(X_1, \dots, X_n)$ be a polynomial in n variables over \mathbb{F}_2 . Assume that $F(\mathbf{X})$ is not a polynomial in some subset of the variables X_1, \dots, X_n . The exponential sum associated to F over \mathbb{F}_2 is

$$S(F) = \sum_{\mathbf{x} \in \mathbb{F}_2^n} (-1)^{F(\mathbf{x})}. \quad (1-1)$$

A Boolean function $F(\mathbf{X})$ is called *balanced* if $S(F) = 0$, i.e., the number of zeros and the number of ones are equal in the truth table of F . In many applications, especially ones related to cryptography, it is important for Boolean functions to be balanced. Balancedness of Boolean functions is an active area of research with open problems even for the relatively simple symmetric case [Adolphson and Sperber 1987; Cai et al. 1996; Canteaut and Videau 2005; Castro et al. 2015; Castro and Medina 2011; 2014; Cusick and Li 2005; Cusick et al. 2008; 2009; Gao et al. 2011; 2016; Su et al. 2013].

Our interest in this work lies in symmetric Boolean functions and therefore, an important step is to try to see what exponential sums of symmetric Boolean functions look like. Let $\sigma_{n,k}$ denote the elementary symmetric polynomial in n variables of degree k . This polynomial is formed by adding together all distinct products of k distinct variables. For example,

$$\sigma_{4,3} = X_1 X_2 X_3 + X_1 X_4 X_3 + X_2 X_4 X_3 + X_1 X_2 X_4. \quad (1-2)$$

Elementary symmetric polynomials are the building blocks of symmetric Boolean functions, as every such function can be identified with an expression of the form

$$\sigma_{n,k_1} + \sigma_{n,k_2} + \dots + \sigma_{n,k_s}, \quad (1-3)$$

where $0 \leq k_1 < k_2 < \dots < k_s$ are integers. For the sake of simplicity, we use the notation $\sigma_{n,[k_1, \dots, k_s]}$ to denote (1-3). For example,

$$\sigma_{3,[2,1]} = \sigma_{3,2} + \sigma_{3,1} = X_1 X_2 + X_3 X_2 + X_1 X_3 + X_1 + X_2 + X_3. \quad (1-4)$$

It turns out that exponential sums of symmetric polynomials have nice representations as binomial sums. Define A_j to be the set of all $(x_1, \dots, x_n) \in \mathbb{F}_2^n$ with exactly j entries equal to 1. Clearly, $|A_j| = \binom{n}{j}$ and by symmetry $\sigma_{n,k}(\mathbf{x}) = \binom{j}{k}$ for $\mathbf{x} \in A_j$. Therefore,

$$S(\sigma_{n,k}) = \sum_{j=0}^n \sum_{\mathbf{x} \in A_j} (-1)^{\sigma_{n,k}(\mathbf{x})} = \sum_{j=0}^n (-1)^{\binom{j}{k}} \binom{n}{j}. \tag{1-5}$$

In general, if $0 \leq k_1 < k_2 < \dots < k_s$ are fixed integers, then

$$S(\sigma_{n,[k_1, \dots, k_s]}) = \sum_{j=0}^n (-1)^{\binom{j}{k_1} + \binom{j}{k_2} + \dots + \binom{j}{k_s}} \binom{n}{j}. \tag{1-6}$$

Equation (1-6) is a clear computational improvement over (1-1). It also connects the problem of balancedness of symmetric Boolean functions to the intriguing problem of bisecting binomial coefficients; see [Mitchell 1990]. A solution $(\delta_0, \delta_1, \dots, \delta_n)$ to the equation

$$\sum_{j=0}^n \delta_j \binom{n}{j} = 0, \quad \delta_j \in \{-1, 1\}, \tag{1-7}$$

is said to give a *bisection of the binomial coefficients* $\binom{n}{j}$, $0 \leq j \leq n$. Observe that a solution to (1-7) provides us with two disjoint sets A, B such that $A \cup B = \{0, 1, 2, \dots, n\}$ and

$$\sum_{j \in A} \binom{n}{j} = \sum_{j \in B} \binom{n}{j} = 2^{n-1}. \tag{1-8}$$

The problem of bisecting binomial coefficients is an interesting problem in its own right; however, it is out of the scope of this work.

The identity (1-6) was used by Castro and Medina [2011] to study exponential sums of symmetric Boolean functions from the point of view of integer sequences. As part of their study, they showed that the sequence $\{S(\sigma_{n,[k_1, \dots, k_s]})\}_{n \in \mathbb{N}}$ satisfies the homogeneous linear recurrence

$$a(n) = \sum_{j=1}^{2^r-1} (-1)^{j-1} \binom{2^r}{j} a(n-j), \tag{1-9}$$

where $r = \lfloor \log_2(k_s) \rfloor + 1$; this result was first proved by Cai, Green and Thierauf [Cai et al. 1996, Theorem 3.1, p. 248]. The characteristic polynomial of (1-9) is given by

$$(t-2)\Phi_4(t-1)\Phi_8(t-1)\dots\Phi_{2^r}(t-1), \tag{1-10}$$

where $\Phi_n(t)$ represents the n -th cyclotomic polynomial. This is very important, as it implies that (1-9) has an embedded nature. Before giving the formal definition

of what we mean by “embedded nature”, let us explore recurrence (1-9) in order to have a better understanding of where we want to go with this term. Observe that the exponential sum of every symmetric Boolean function of degree less than 4 satisfies

$$a(n) = \sum_{j=1}^3 (-1)^{j-1} \binom{4}{j} a(n-j), \quad (1-11)$$

the exponential sum of every symmetric Boolean function of degree less than 8 satisfies

$$a(n) = \sum_{j=1}^7 (-1)^{j-1} \binom{8}{j} a(n-j), \quad (1-12)$$

the exponential sum of every symmetric Boolean function of degree less than 16 satisfies

$$a(n) = \sum_{j=1}^{15} (-1)^{j-1} \binom{16}{j} a(n-j), \quad (1-13)$$

and so on. This means, for example, that $\{S(\sigma_{n,[7,2]})\}_{n \in \mathbb{N}}$, for which the first few values are given by

2, 4, 6, 8, 12, 24, 58, 144, 344, 784, 1716, 3632, 7464, 14928, 29128, 55680, \dots ,

must satisfy (1-12) and (1-13), but not (1-11). Next is the formal definition of *embedded recurrences*.

Definition 1.1. Let $\{a_{f(x)}(n)\}$ be a family of integer sequences indexed by some polynomial family $\{f(x)\}$. Suppose that every sequence $\{a_{f(x)}(n)\}$ satisfies a linear recurrence. We say that these recurrences are *embedded* if there is a sequence of integers $n_1 < n_2 < n_3 < \dots$ such that every sequence $\{a_{f(x)}(n)\}$ with the property $\deg(f) < n_l$ satisfies a global recurrence. For example, the sequences of exponential sums of symmetric Boolean functions satisfy recurrences that are embedded. In this case, $n_l = 2^l$ and the global recurrence is (1-9).

Castro and Medina [2011] also computed the asymptotic behavior of $S(\sigma_{n,[k_1, \dots, k_s]})$ as $n \rightarrow \infty$. To be specific, they showed that

$$\lim_{n \rightarrow \infty} \frac{1}{2^n} S(\sigma_{n,[k_1, \dots, k_s]}) = c_0(k_1, \dots, k_s), \quad (1-14)$$

where

$$c_0(k_1, \dots, k_s) = \frac{1}{2^r} \sum_{j=0}^{2^r-1} (-1)^{\binom{j}{k_1} + \dots + \binom{j}{k_s}}. \quad (1-15)$$

They used this limit to show that a conjecture by Cusick, Li and Stănică [Cusick et al. 2008] is true asymptotically. Some of these results, especially recurrence

(1-9) and limit (1-14), were extended to some perturbations of symmetric Boolean functions [Castro and Medina 2014].

In this manuscript, we generalize the concept of exponential sums of symmetric Boolean functions by virtue of the binomial sum in (1-6) and study some of its properties. Let d be a nonnegative integer. We define the d -generalized exponential sum of $\sigma_{n,[k_1,\dots,k_s]}$ as the power sum of binomial coefficients given by

$$S_d(\sigma_{n,[k_1,\dots,k_s]}) = \sum_{j=0}^n (-1)^{\binom{j}{k_1} + \dots + \binom{j}{k_s}} \binom{n}{j}^d. \tag{1-16}$$

In a similar manner, if $Q(x) = a_0 + a_1x + \dots + a_t x^t$ is a polynomial, then the $Q(x)$ -generalized exponential sum of $\sigma_{n,[k_1,\dots,k_s]}$ is defined as

$$S_{Q(x)}(\sigma_{n,[k_1,\dots,k_s]}) = \sum_{j=0}^n (-1)^{\binom{j}{k_1} + \dots + \binom{j}{k_s}} Q\left(\binom{n}{j}\right). \tag{1-17}$$

By linearity, the study of (1-17) is reduced to the study of (1-16). Thus, emphasis is made on d -generalized exponential sums.

It is clear that if $d = 1$, then the d -generalized exponential sum is just the regular exponential sum. However, we point out that d -generalized exponential sums generalize other combinatorial objects. For instance, when degree 0 is considered, we have $S_d(\sigma_{n,0}) = -f_{n,d}$, where

$$f_{n,d} = \sum_{j=0}^n \binom{n}{j}^d \tag{1-18}$$

is the d -th order Franel number. When $k = 1$,

$$S_d(\sigma_{n,1}) = \sum_{j=0}^n (-1)^j \binom{n}{j}^d \tag{1-19}$$

is the d -th order alternate Franel number, for which, when $d = 3$, we have the beautiful identity of Dixon

$$S_3(\sigma_{2n,1}) = \sum_{j=0}^{2n} (-1)^j \binom{2n}{j}^3 = (-1)^n \binom{2n}{n} \binom{3n}{n}. \tag{1-20}$$

Finally, the sequence $\{x_n\}$ defined by $x_0 = 0$, $x_1 = 3$ and $x_n = S_0(\sigma_{n-1,3})$ can be identified with sequence A018837 [Sloane and LeBrun 2008], which represents the minimum number of steps for a knight which starts at position $(0, 0)$ to reach $(n, 0)$ on an infinite chessboard.

In this article we extend some of the results that appear in [Castro and Medina 2011; 2014] to d -generalized exponential sums. In particular, we show that these

sequences satisfy recurrences and, as is the case for $d = 1$, there is an embedded component behind it. We also calculate the asymptotic behavior of these sequences and show that the constant $c_0(k_1, \dots, k_s)$ is universal in the sense that it appears in the asymptotic behavior of $S_{Q(x)}(\sigma_{n,[k_1, \dots, k_s]})$ for every polynomial $Q(x)$. The case $d = 0$ turns out to be relatively easy when compared to the case $d \neq 0$, and, as a result, we decided to discuss it now. First, it is clear that $S_0(\sigma_{n,[k_1, \dots, k_s]}) = O(n)$. Second, if $r = \lfloor \log_2(k_s) \rfloor + 1$, then it satisfies the linear recurrence

$$a(n) = a(n-1) + a(n-2^r) - a(n-2^r-1). \quad (1-21)$$

The characteristic polynomial of (1-21) is given by

$$(t-1)^2 \Phi_2(t) \Phi_4(t) \cdots \Phi_{2^r}(t), \quad (1-22)$$

and therefore, as in the case $d = 1$, these recurrences are embedded. Finally, if i_1, \dots, i_p are all the integers between 1 and $2^r - 1$ such that $\binom{i}{k_1} + \cdots + \binom{i}{k_s} \equiv 1 \pmod{2}$, then it is not hard to see that

$$S_0(\sigma_{n,[k_1, \dots, k_s]}) = n + 1 - 2 \left\lfloor \frac{n+1-i_1}{2^r} \right\rfloor - \cdots - 2 \left\lfloor \frac{n+1-i_p}{2^r} \right\rfloor. \quad (1-23)$$

The asymptotic behavior of d -generalized exponential sums is discussed in Section 2. Then, in Section 3, we use computer power to find recurrences for these sums. The reader is invited to use her favorite computer algebra system while reading this manuscript. This is not necessary, as we believe the manuscript is self-contained; however we encourage experimentation because it helps to build intuition and to cement and develop appreciation for mathematical knowledge.

2. Asymptotic behavior of the generalized exponential sum

The asymptotic behavior of $S(\sigma_{n,k})$ as $n \rightarrow \infty$ was used in [Castro and Medina 2011] to show a conjecture by Cusick, Li and Stănică [Cusick et al. 2008] is true for large n . This shows the importance of the behavior of $S(\sigma_{n,[k_1, \dots, k_s]})$ as n increases. In this section we discuss the asymptotic behavior of $\{S_d(\sigma_{n,[k_1, \dots, k_s]})\}_{n \in \mathbb{N}}$ and show that the behavior of $\{S_d(\sigma_{n,[k_1, \dots, k_s]})\}_{n \in \mathbb{N}}$, as n increases, is closely related to that of $\{S(\sigma_{n,[k_1, \dots, k_s]})\}_{n \in \mathbb{N}}$.

We start our discussion with the case $d = 2$ and $k = 3$; that is, we consider the sequence $\{S_2(\sigma_{n,3})\}_{n \in \mathbb{N}}$. The idea for doing this is to gain insight as to what is behind the asymptotic behavior of these sequences. A proof for the general case will be provided later in this section once our intuition is solidified.

The first few values of the sequence $\{S_2(\sigma_{n,3})\}_{n \in \mathbb{N}}$ are given by

$$2, 6, 18, 38, 52, 124, 980, 6470, 31916, 127156, \dots$$

It is not surprising, knowing already the behavior of $S(\sigma_{n,3})$, that the value of the n -th term of the sequence $\{S_2(\sigma_{n,3})\}_{n \in \mathbb{N}}$ increases quite rapidly as $n \rightarrow \infty$. Now, by previous knowledge we have that

$$\lim_{n \rightarrow \infty} \frac{1}{2^n} S(\sigma_{n,3}) = \frac{1}{2},$$

where 2^n is the number of n -tuples with 0, 1 entries; thus, it is natural to consider the behavior of $S_2(n, 3)/2^n$. The reader can check via computer experimentation that $S_2(n, 3)/2^n$ seems to diverge to ∞ , which, if true, it would imply that our sequence increases a rate that is faster than 2^n . Taking into consideration that in this case $d = 2$, it is not a wild idea to check the behavior of $S_2(\sigma_{n,3})/2^{2n}$. In this case, the reader can convince herself that $S_2(\sigma_{n,3})/4^n \rightarrow 0$ as $n \rightarrow \infty$. Moreover, experiments on a computer suggest that

$$\lim_{n \rightarrow \infty} \frac{1}{4^n} S_2(\sigma_{n,k}) = 0 \tag{2-1}$$

for any positive integer k . For example, the values of $S_2(\sigma_{n,7})/4^n$ for $n = 10, 100$, and 1000 are given by

$$0.148731, \quad 0.0426647, \quad \text{and} \quad 0.0133793,$$

respectively. Thus, it appears that $S_2(\sigma_{n,k})$ increases faster than 2^n , but slower than 4^n . So, what is the appropriate behavior?

To answer the question, we start by analyzing the reason behind the behavior of the regular exponential sum $S(\sigma_{n,3})$. Using the definition of $S(\sigma_{n,k})$ in terms of binomial coefficients, we see that

$$\begin{aligned} S(\sigma_{n,3}) &= \sum_{j=0}^n (-1)^{\binom{j}{3}} \binom{n}{j} = \sum_{j=0}^n \binom{n}{j} - 2 \sum_{j=0}^n \binom{n}{4j+3} \\ &= 2^n - 2 \sum_{j=0}^n \binom{n}{4j+3}. \end{aligned} \tag{2-2}$$

Observe that when we divide $S(\sigma_{n,3})$ by 2^n , we control the contribution of the negative terms. We now do the analogous thing for $S_2(\sigma_{n,3})$. Observe that

$$\begin{aligned} S_2(\sigma_{n,3}) &= \sum_{j=0}^n (-1)^{\binom{j}{3}} \binom{n}{j}^2 = \sum_{j=0}^n \binom{n}{j}^2 - 2 \sum_{j=0}^n \binom{n}{4j+3}^2 \\ &= \binom{2n}{n} - 2 \sum_{j=0}^n \binom{n}{4j+3}^2. \end{aligned} \tag{2-3}$$

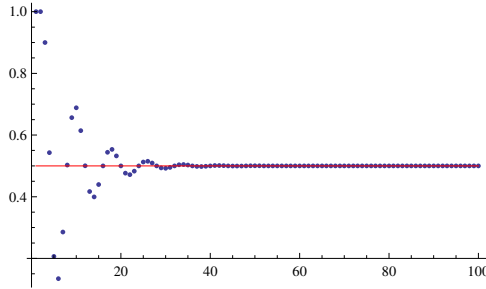


Figure 1. Graphical representation of $S_2(\sigma_{n,3})/\binom{2n}{n}$.

Therefore, it is now natural to see that dividing $S_2(\sigma_{n,3})$ by the central binomial coefficient controls the contribution of the negative terms. Figure 1 is a graphical representation of this fact. The dots correspond to $S(\sigma_{n,3})/\binom{2n}{n}$. The line corresponds to $y = \frac{1}{2}$.

It is clear now that $S_2(\sigma_{n,3})$ increases faster than 2^n , but a bit slower than 4^n . Its behavior is somewhat similar to that of the central binomial coefficient and by Stirling’s formula we know that

$$\binom{2n}{n} \sim \frac{4^n}{\sqrt{\pi n}}. \tag{2-4}$$

Moreover, observe that

$$\lim_{n \rightarrow \infty} \binom{2n}{n}^{-1} S_2(\sigma_{n,3}) = \frac{1}{2} = c_0(3). \tag{2-5}$$

Equation (2-5) is not a coincidence, as we will show that $c_0(k_1, \dots, k_s)$ appears in the behavior of $S_d(\sigma_{n,[k_1, \dots, k_s]})$. We are now ready to discuss the general case.

Let d be a nonnegative integer. Define $G(n, d)$ as the d -th order Franel number

$$G(n, d) = \sum_{j=0}^n \binom{n}{j}^d. \tag{2-6}$$

For $d = 0, 1, 2$, the value of $G(n, d)$ is given by

$$G(n, 0) = n + 1, \quad G(n, 1) = 2^n \quad \text{and} \quad G(n, 2) = \binom{2n}{n}. \tag{2-7}$$

Sadly, there is not a nice closed formula for $G(n, d)$ when $d > 2$. Instead, the value of $G(n, d)$ is given by the hypergeometric function

$$G(n, d) = {}_dF_{d-1}(-n, -n, \dots, -n; 1, 1, \dots, 1; (-1)^n). \tag{2-8}$$

The asymptotic behavior of $G(n, d)$ is already known [Pólya and Szegő 1976]:

$$G(n, d) \sim \frac{2^{dn}}{\sqrt{d}} \left(\frac{2}{\pi n} \right)^{(d-1)/2}. \tag{2-9}$$

A formal proof of (2-9) was given by Farmer and Leth [2005]. A treatment for $G(2n, d)$ using Euler’s summation formula and the tail-exchange trick appears in [Graham et al. 1994]. Also, a proper adjustment to the proof of Farmer and Leth leads to the following result.

Lemma 2.1. *Let m and d be fixed natural numbers and i an integer such that $0 \leq i \leq m$. Then, as n increases, we have*

$$\sum_{j=0}^n \binom{n}{mj+i}^d \sim \frac{2^{dn}}{m\sqrt{d}} \left(\frac{2}{\pi n}\right)^{(d-1)/2} \sim \frac{1}{m} G(n, d). \tag{2-10}$$

With Lemma 2.1 at hand, we are now ready to provide the asymptotic behavior of $S_d(\sigma_{n,[k_1, \dots, k_s]})$.

Theorem 2.2. *Let d and $k_1 < \dots < k_s$ be fixed positive integers. Then,*

$$\lim_{n \rightarrow \infty} \frac{S_d(\sigma_{n,[k_1, \dots, k_s]})}{G(n, d)} = c_0(k_1, \dots, k_s). \tag{2-11}$$

Proof. Let $r = \lfloor \log_2(k_s) \rfloor + 1$. Let i_1, \dots, i_p be all the integers between 1 and $2^r - 1$ such that $\binom{i}{k_1} + \dots + \binom{i}{k_s} \equiv 1 \pmod{2}$. It is known, see [Castro and Medina 2011], that the sequence $\left\{ \binom{n}{k_1} + \dots + \binom{n}{k_s} \pmod{2} \right\}_{n \in \mathbb{N}}$ is periodic and the period is a divisor of 2^r . Therefore, $\binom{i}{k_1} + \dots + \binom{i}{k_s} \equiv 1 \pmod{2}$ if and only if $i \equiv i_l \pmod{2^r}$ for some $i_l \in \{i_1, \dots, i_p\}$.

Using the definition of $S_d(\sigma_{n,[k_1, \dots, k_s]})$ we observe that

$$S_d(\sigma_{n,[k_1, \dots, k_s]}) = G(n, d) - 2 \sum_{j=0}^n \left[\binom{n}{2^r \cdot j + i_1}^d + \dots + \binom{n}{2^r \cdot j + i_p}^d \right]. \tag{2-12}$$

Therefore, as $n \rightarrow \infty$, we have

$$S_d(\sigma_{n,[k_1, \sigma, k_s]}) \sim G(n, d) - \frac{2p}{2^r} G(n, d) = (1 - p \cdot 2^{1-r}) G(n, d). \tag{2-13}$$

It is not hard to show that $c_0(k_1, \dots, k_s) = 1 - p \cdot 2^{1-r}$. □

Using the asymptotic behavior (2-9), we obtain the following corollary.

Corollary 2.3. *Let d and $k_1 < \dots < k_s$ be positive integers. Then,*

$$\lim_{n \rightarrow \infty} \frac{(\sqrt{n})^{d-1} \cdot S_d(\sigma_{n,[k_1, \dots, k_s]})}{2^{dn}} = \frac{1}{\sqrt{d}} \left(\frac{2}{\pi}\right)^{(d-1)/2} c_0(k_1, \dots, k_s). \tag{2-14}$$

More generally, if $Q(x) = a_0 + a_1x + \dots + a_t x^t$ is a polynomial and

$$A_{Q(x)}(n) = a_0 \cdot (n+1) + a_1 \cdot 2^n + \frac{a_2}{\sqrt{2}} \left(\frac{2}{\pi \cdot n}\right)^{1/2} 2^{2n} + \dots + \frac{a_t}{\sqrt{t}} \left(\frac{2}{\pi \cdot n}\right)^{(t-1)/2} 2^{tn}, \tag{2-15}$$

then,

$$\lim_{n \rightarrow \infty} \frac{S_{Q(x)}(\sigma_{n,[k_1, \dots, k_s]})}{A_{Q(x)}(n)} = c_0(k_1, \dots, k_s). \tag{2-16}$$

Proof. This is a direct consequence of Theorem 2.2 and the asymptotic behavior of $G(n, d)$. □

Example 2.4. Consider the case $d = 4$ and $k = 7$. We know that $c_0(7) = \frac{3}{4}$. Thus,

$$\frac{1}{\sqrt{d}} \left(\frac{2}{\pi}\right)^{(d-1)/2} c_0(k) = \frac{3}{8} \left(\frac{2}{\pi}\right)^{3/2} \approx 0.1904809078 \dots \tag{2-17}$$

Note that

n	$(\sqrt{n})^3 S_4(\sigma_{n,7})/2^{4n}$
1	0.1250000000
10	0.2280899652
100	0.2021752897
1000	0.1903737868
10000	0.1904701935
100000	0.1904798364

Example 2.5. Let $k_1 = 2, k_2 = 3, k_3 = 4,$ and $k_4 = 5$. Consider the polynomial $Q(x) = x^3 + 5x + 2$. The reader can check that in this case we have

$$A_{Q(x)}(n) \cdot c_0(2, 3, 4, 5) = \frac{1}{2} \left(5 \cdot 2^n + 2(n + 1) + \frac{2^{3n+1}}{\sqrt{3\pi n}} \right). \tag{2-18}$$

Corollary 2.3 states that

$$\lim_{n \rightarrow \infty} \frac{S_{Q(x)}(\sigma_{n,[2,3,4,5]})}{A_{Q(x)}(n) \cdot c_0(2, 3, 4, 5)} = 1. \tag{2-19}$$

Figure 2 is a graphical representation of (2-19).

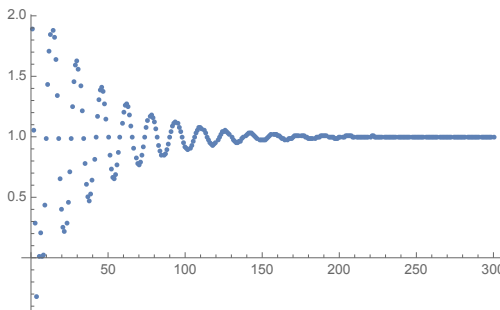


Figure 2. Graphical representation of $S_{Q(x)}(\sigma_{n,[2,3,4,5]}) / (A_{Q(x)}(n) \cdot c_0(2, 3, 4, 5))$ when $Q(x) = x^3 + 5x + 2$.

We conclude this section with the observation that Theorem 2.2 and Corollary 2.3 imply that the constant $c_0(k_1, \dots, k_s)$ is universal in the sense that it appears in the asymptotic behavior of d -generalized exponential sums. Moreover, Theorem 2.2 is the natural generalization of limit (1-14). In the next section, we explore a generalization to recurrence (1-9).

3. Recurrence relations: some experiments

In this section we discuss recurrence relations for the sequences $\{S_d(\sigma_{n,[k_1, \dots, k_s]})\}_{n \in \mathbb{N}}$. We already know that for $d = 1$, i.e., for $\{S(\sigma_{n,[k_1, \dots, k_s]})\}_{n \in \mathbb{N}}$, we have the homogeneous linear recurrence with constant coefficients

$$a(n) = \sum_{m=1}^{2^r-1} (-1)^{m-1} \binom{2^r}{m} a(n-m), \tag{3-1}$$

where $r = \lfloor \log_2(k_s) \rfloor + 1$. See [Cai et al. 1996; Castro and Medina 2011; 2014] for more details. Experiments show that something similar happens for $\{S_d(\sigma_{n,[k_1, \dots, k_s]})\}_{n \in \mathbb{N}}$ when $d > 1$; i.e., these sequences satisfy linear recurrences. However, as we will see, the coefficients of these recurrences are no longer constant; instead, they are polynomials in n . In other words, these sequences seems to be *holonomic* (this should not come as a surprise to the expert reader or to the reader aware of the work of Franel [1894; 1895] and Cusick [1989] on power sums of binomial coefficients). Therefore, for $d > 1$, the problem of finding the minimal recurrence is a hard one. Once again, the reader is encouraged to open her favorite computer algebra system while reading this section.

To show the difficulty of the problem at hand, let us consider (once again) the rather simple example $\{S_2(\sigma_{n,3})\}_{n \in \mathbb{N}}$. Note that

$$S_2(\sigma_{n,3}) = \sum_{j=0}^n (-1)^{\binom{j}{3}} \binom{n}{j}^2 = \binom{2n}{n} - 2 \sum_{j=0}^n \binom{n}{4j+3}^2. \tag{3-2}$$

We already know that the central binomial coefficient satisfies a linear recurrence with nonconstant coefficients; i.e., it satisfies the recurrence

$$(n+1)a(n+1) - (4n+2)a(n) = 0. \tag{3-3}$$

Thus, it is natural to expect that if this sequence satisfies a linear recurrence, then the coefficients of the recurrence are nonconstant.

In order to find such a recurrence, we emulate what we already know about the case $d = 1$. In that case, we have

$$S(\sigma_{n,3}) = \sum_{j=0}^n (-1)^{\binom{j}{3}} \binom{n}{j} = 2^n - 2 \sum_{j=0}^n \binom{n}{4j+3}. \tag{3-4}$$

The “negative” part of it, i.e., $\sum_{j=0}^n \binom{n}{4j+3}$, satisfies the homogeneous recurrence

$$a(n) = 4a(n-1) - 6a(n-2) + 4a(n-3). \quad (3-5)$$

It is not hard to see that 2 is a root of the characteristic polynomial of recurrence (3-5) and so 2^n also satisfies it. Thus, $\{S(\sigma_{n,3})\}_{n \in \mathbb{N}}$ satisfies (3-5).

In general, if $1 \leq k_1 < \dots < k_s$ are integers, $r = \lfloor \log_2(k_s) \rfloor + 1$, and i_1, \dots, i_p are all integers between 1 and $2^r - 1$ such that $\binom{i}{k_1} + \dots + \binom{i}{k_s} \equiv 1 \pmod{2}$, then

$$S(\sigma_{n,[k_1, \dots, k_s]}) = 2^n - 2 \sum_{j=0}^n \left(\binom{n}{2^r j + i_1} + \dots + \binom{n}{2^r j + i_p} \right), \quad (3-6)$$

and the negative part of (3-6) satisfies (3-1). Since 2 is a root of the characteristic polynomial of (3-1), we know 2^n , and therefore $\{S(\sigma_{n,[k_1, \dots, k_s]})\}_{n \in \mathbb{N}}$, satisfy (3-1).

Emulating what we did in the above paragraph, we start by looking for a recurrence for

$$\sum_{j=0}^n \binom{n}{4j+3}^2. \quad (3-7)$$

It is at this stage that we use the power of computers. This power, of course, is assisted by the ingenuity of a great mathematician, in this case, the great combinatorialist Doron Zeilberger [1990a]. Zeilberger’s algorithm is already a built-in function in Maple and a version for Mathematica can be found at <http://www.risc.jku.at/research/combinat/risc/software>. Using it we obtain (with an automated proof!) that (3-7) satisfies the homogeneous linear recurrence with nonconstant coefficients

$$\sum_{j=0}^7 p_j(n) a(n+j) = 0, \quad (3-8)$$

where the polynomials $p_j(n)$ can be found in the online supplement. Analogous to 2^n for $d = 1$, the central binomial coefficient satisfies (3-8). Thus, $\{S_2(\sigma_{n,3})\}_{n \in \mathbb{N}}$ satisfies (3-8).

Zeilberger’s algorithm also proves that the sequences

$$\sum_{j=0}^n \binom{n}{4j+i}^2 \quad (3-9)$$

for $i = 0, 1, 2, 3$, satisfy (3-8) too. Since we have that

$$S_2(\sigma_{n,2}) = \binom{2n}{n} - 2 \sum_{j=0}^n \left(\binom{n}{4j+2}^2 + \binom{n}{4j+3}^2 \right),$$

$$S_2(\sigma_{n,[2,1]}) = \binom{2n}{n} - 2 \sum_{j=0}^n \left(\binom{n}{4j+1}^2 + \binom{n}{4j+2}^2 \right),$$

$$\begin{aligned}
 S_2(\sigma_{n,[3,2]}) &= \binom{2n}{n} - 2 \sum_{j=0}^n \binom{n}{4j+2}^2, \\
 S_2(\sigma_{n,[3,1]}) &= \binom{2n}{n} - 2 \sum_{j=0}^n \binom{n}{4j+1}^2, \\
 S_2(\sigma_{n,[3,2,1]}) &= \binom{2n}{n} - 2 \sum_{i=1}^3 \sum_{j=0}^n \binom{n}{4j+i}^2,
 \end{aligned}$$

all of them satisfy (3-8). In fact, for $4 \leq k_s \leq 7$, the sequence $\{S_2(\sigma_{n,[k_1,\dots,k_s]})\}$ satisfies a recurrence of order 15 with polynomial coefficients. Moreover, every sequence $\{S_2(\sigma_{n,[k_1,\dots,k_s]})\}$ with $1 \leq k_s \leq 7$ satisfies this recurrence of order 15. This pattern seems to hold for higher k_s and any $d > 2$. If this holds true, then, as in the cases $d = 0$ and $d = 1$, these sequences satisfy recurrences that are embedded.

The expert reader may notice that it is not hard to show that d -generalized exponential sums, and therefore $Q(x)$ -generalized exponential sums, are indeed holonomic. This follows from the fact that binomial coefficients are holonomic in both variables and from some closure properties of these sequences; a great read on this subject is [Zeilberger 1990b]. A formal proof, however, will require a proper discussion on holonomic sequences and this is out of the scope of this work.

The natural question now is: can we show that the recurrences are embedded? The answer is yes! Suppose that a sequence $\{a(n)\}$ is holonomic; that is, suppose that there exist polynomials $p_0(n), p_1(n), \dots, p_l(n) \in \mathbb{C}[n]$ such that

$$p_l(n)a(n+l) + p_{l-1}(n)a(n+l-1) + \dots + p_0(n)a(n) = 0. \tag{3-10}$$

Let E be the *shift operator* that maps $a(n)$ to $a(n+1)$. Equation (3-10) can be written as $A(E)(a(n)) = 0$, where

$$A(E) = \sum_{j=0}^l p_j(n)E^j. \tag{3-11}$$

The operator $A(E)$ is called an *annihilating operator* of the sequence $\{a(n)\}$. The number l is called the *order* of the annihilating operator. It is not hard to see that the set of all annihilating operators of $\{a(n)\}$ forms an ideal of the ring $\mathbb{C}[n][E]$.

Consider the sequence

$$a_{d,r,i}(n) = \sum_{j=0}^n \binom{n}{2^r j+i}^d. \tag{3-12}$$

Let $A_{d,r,i}(E) \in \mathbb{C}[n][E]$ be an annihilating operator for $\{a_{d,r,i}(n)\}$. Define

$$A_{d,r}(E) = \prod_{i=0}^{2^r-1} A_{d,r,i}(E). \tag{3-13}$$

Since the set of all annihilating operators of a sequence $\{a(n)\}$ is an ideal, we know $A_{d,r}(E)(a_{d,r,i}(n)) = 0$ for every r, d and i . Also, since

$$G(n, d) = \sum_{i=0}^{2^r-1} a_{d,r,i}(n), \quad (3-14)$$

we have $A_{d,r}(E)(G(n, d)) = 0$. Finally, if $r = \lfloor \log_2(k_s) \rfloor + 1$, then $S_d(\sigma_{n,[k_1, \dots, k_s]})$ is a linear combination of $G(n, d)$ and some terms $a_{d,r,i}(n)$; therefore

$$A_{d,r}(E)(S_d(\sigma_{n,[k_1, \dots, k_s]})) = 0 \quad (3-15)$$

and so the recurrences are embedded. To be specific, for every symmetric Boolean function of degree less than 4, the d -generalized exponential sum satisfies

$$A_{d,2}(E)(a(n)) = 0, \quad (3-16)$$

for every symmetric Boolean function of degree less than 8, the d -generalized exponential sum satisfies

$$A_{d,3}(E)(a(n)) = 0, \quad (3-17)$$

and so on.

We finish this section by noticing that the recurrences included in this work are not necessarily the minimal ones. For instance, we know that $\{S_2(\sigma_{n,2})\}_{n \in \mathbb{N}}$ satisfies (3-8). However, using the Mathematica implementation `GuessMinRE`, which is part of the package `Guess.m` written by Manuel Kauers, available at <http://www.risc.jku.at/research/combinat/risc/software>, we guess that $\{S_2(\sigma_{n,2})\}_{n \in \mathbb{N}}$ satisfies the recurrence

$$\sum_{j=0}^4 q_j(n) a(n+j) = 0, \quad (3-18)$$

where

$$\begin{aligned} q_0(n) &= 424 + 924n + 692n^2 + 216n^3 + 24n^4, \\ q_1(n) &= 1280 + 2352n + 1576n^2 + 456n^3 + 48n^4, \\ q_2(n) &= 1600 + 2780n + 1756n^2 + 480n^3 + 48n^4, \\ q_3(n) &= -960 - 1604n - 968n^2 - 252n^3 - 24n^4, \\ q_4(n) &= 276 + 449n + 263n^2 + 66n^3 + 6n^4. \end{aligned}$$

This has been checked for values of n up to 20000.

Acknowledgments

González was partially supported as a student by NSF-DUE 1356474 and the Mellon Mays Undergraduate Fellowship. Medina acknowledges the partial support of UPR-FIPI 1890015.00.

References

- [Adolphson and Sperber 1987] A. Adolphson and S. Sperber, “ p -adic estimates for exponential sums and the theorem of Chevalley–Warning”, *Ann. Sci. École Norm. Sup.* (4) **20**:4 (1987), 545–556. MR Zbl
- [Cai et al. 1996] J.-Y. Cai, F. Green, and T. Thierauf, “On the correlation of symmetric functions”, *Math. Systems Theory* **29**:3 (1996), 245–258. MR Zbl
- [Canteaut and Videau 2005] A. Canteaut and M. Videau, “Symmetric Boolean functions”, *IEEE Trans. Inform. Theory* **51**:8 (2005), 2791–2811. MR Zbl
- [Castro and Medina 2011] F. N. Castro and L. A. Medina, “Linear recurrences and asymptotic behavior of exponential sums of symmetric Boolean functions”, *Electron. J. Combin.* **18**:2 (2011), art. id. 8, 21 pp. MR Zbl
- [Castro and Medina 2014] F. N. Castro and L. A. Medina, “Asymptotic behavior of perturbations of symmetric functions”, *Ann. Comb.* **18**:3 (2014), 397–417. MR Zbl
- [Castro et al. 2015] F. N. Castro, O. E. González, and L. A. Medina, “A divisibility approach to the open boundary cases of Cusick–Li–Stănică’s conjecture”, *Cryptogr. Commun.* **7**:4 (2015), 379–402. MR Zbl
- [Cusick 1989] T. W. Cusick, “Recurrences for sums of powers of binomial coefficients”, *J. Combin. Theory Ser. A* **52**:1 (1989), 77–83. MR Zbl
- [Cusick and Li 2005] T. W. Cusick and Y. Li, “ k -th order symmetric SAC Boolean functions and bisecting binomial coefficients”, *Discrete Appl. Math.* **149**:1-3 (2005), 73–86. MR Zbl
- [Cusick et al. 2008] T. W. Cusick, Y. Li, and P. Stănică, “Balanced symmetric functions over $\text{GF}(p)$ ”, *IEEE Trans. Inform. Theory* **54**:3 (2008), 1304–1307. MR Zbl
- [Cusick et al. 2009] T. W. Cusick, Y. Li, and P. Stănică, “On a conjecture for balanced symmetric Boolean functions”, *J. Math. Cryptol.* **3**:4 (2009), 273–290. MR Zbl
- [Farmer and Leth 2005] J. D. Farmer and S. C. Leth, “An asymptotic formula for powers of binomial coefficients”, *Math. Gazette* **89**:516 (2005), 385–391.
- [Franel 1894] J. Franel, in reply to question 42 by Laisant, *L’Int. Math.* **1** (1894), 45–47. In French.
- [Franel 1895] J. Franel, in reply to question 170 by Laisant, *L’Int. Math.* **2** (1895), 33–35. In French.
- [Gao et al. 2011] G.-P. Gao, W.-F. Liu, and X.-Y. Zhang, “The degree of balanced elementary symmetric Boolean functions of $4k + 3$ variables”, *IEEE Trans. Inform. Theory* **57**:7 (2011), 4822–4825. MR
- [Gao et al. 2016] G. Gao, Y. Guo, and Y. Zhao, “Recent results on balanced symmetric Boolean functions”, *IEEE Trans. Inform. Theory* **62**:9 (2016), 5199–5203. MR
- [Graham et al. 1994] R. L. Graham, D. E. Knuth, and O. Patashnik, *Concrete mathematics: a foundation for computer science*, 2nd ed., Addison-Wesley, Reading, MA, 1994. MR Zbl
- [Mitchell 1990] C. Mitchell, “Enumerating Boolean functions of cryptographic significance”, *J. Cryptology* **2**:3 (1990), 155–170. MR Zbl
- [Pólya and Szegő 1976] G. Pólya and G. Szegő, *Problems and theorems in analysis, II*, Die Grundlehren der Math. Wissenschaften **216**, Springer, New York, 1976. MR Zbl
- [Sloane and LeBrun 2008] N. Sloane and M. LeBrun, “Number of steps for knight to reach $(n, 0)$ on infinite chessboard”, pp. A018837 in *The online encyclopedia of integer sequences*, 2008.
- [Su et al. 2013] W. Su, X. Tang, and A. Pott, “A note on a conjecture for balanced elementary symmetric Boolean functions”, *IEEE Trans. Inform. Theory* **59**:1 (2013), 665–671. MR

[Zeilberger 1990a] D. Zeilberger, “A fast algorithm for proving terminating hypergeometric identities”, *Discrete Math.* **80**:2 (1990), 207–211. MR Zbl

[Zeilberger 1990b] D. Zeilberger, “A holonomic systems approach to special functions identities”, *J. Comput. Appl. Math.* **32**:3 (1990), 321–368. MR Zbl

Received: 2016-08-26 Revised: 2017-01-12 Accepted: 2017-02-04

franciscastr@gmail.com *Department of Mathematics, University of Puerto Rico,
San Juan, Puerto Rico*

oscar.gonzalez3@upr.edu *Department of Mathematics, University of Puerto Rico,
San Juan, Puerto Rico*

luis.medina17@upr.edu *Department of Mathematics, University of Puerto Rico,
San Juan, Puerto Rico*

involve

msp.org/involve

INVOLVE YOUR STUDENTS IN RESEARCH

Involve showcases and encourages high-quality mathematical research involving students from all academic levels. The editorial board consists of mathematical scientists committed to nurturing student participation in research. Bridging the gap between the extremes of purely undergraduate research journals and mainstream research journals, *Involve* provides a venue to mathematicians wishing to encourage the creative involvement of students.

MANAGING EDITOR

Kenneth S. Berenhaut Wake Forest University, USA

BOARD OF EDITORS

Colin Adams	Williams College, USA	Suzanne Lenhart	University of Tennessee, USA
John V. Baxley	Wake Forest University, NC, USA	Chi-Kwong Li	College of William and Mary, USA
Arthur T. Benjamin	Harvey Mudd College, USA	Robert B. Lund	Clemson University, USA
Martin Bohner	Missouri U of Science and Technology, USA	Gaven J. Martin	Massey University, New Zealand
Nigel Boston	University of Wisconsin, USA	Mary Meyer	Colorado State University, USA
Amarjit S. Budhiraja	U of North Carolina, Chapel Hill, USA	Emil Minchev	Ruse, Bulgaria
Pietro Cerone	La Trobe University, Australia	Frank Morgan	Williams College, USA
Scott Chapman	Sam Houston State University, USA	Mohammad Sal Moslehian	Ferdowsi University of Mashhad, Iran
Joshua N. Cooper	University of South Carolina, USA	Zuhair Nashed	University of Central Florida, USA
Jem N. Corcoran	University of Colorado, USA	Ken Ono	Emory University, USA
Toka Diagana	Howard University, USA	Timothy E. O'Brien	Loyola University Chicago, USA
Michael Dorff	Brigham Young University, USA	Joseph O'Rourke	Smith College, USA
Sever S. Dragomir	Victoria University, Australia	Yuval Peres	Microsoft Research, USA
Behrouz Emamizadeh	The Petroleum Institute, UAE	Y.-F. S. Pétermann	Université de Genève, Switzerland
Joel Foisy	SUNY Potsdam, USA	Robert J. Plemmons	Wake Forest University, USA
Errin W. Fulp	Wake Forest University, USA	Carl B. Pomerance	Dartmouth College, USA
Joseph Gallian	University of Minnesota Duluth, USA	Vadim Ponomarenko	San Diego State University, USA
Stephan R. Garcia	Pomona College, USA	Bjorn Poonen	UC Berkeley, USA
Anant Godbole	East Tennessee State University, USA	James Propp	U Mass Lowell, USA
Ron Gould	Emory University, USA	József H. Przytycki	George Washington University, USA
Andrew Granville	Université Montréal, Canada	Richard Rebarber	University of Nebraska, USA
Jerold Griggs	University of South Carolina, USA	Robert W. Robinson	University of Georgia, USA
Sat Gupta	U of North Carolina, Greensboro, USA	Filip Saidak	U of North Carolina, Greensboro, USA
Jim Haglund	University of Pennsylvania, USA	James A. Sellers	Penn State University, USA
Johnny Henderson	Baylor University, USA	Andrew J. Sterge	Honorary Editor
Jim Hoste	Pitzer College, USA	Ann Trenk	Wellesley College, USA
Natalia Hritonenko	Prairie View A&M University, USA	Ravi Vakil	Stanford University, USA
Glenn H. Hurlbert	Arizona State University, USA	Antonia Vecchio	Consiglio Nazionale delle Ricerche, Italy
Charles R. Johnson	College of William and Mary, USA	Ram U. Verma	University of Toledo, USA
K. B. Kulasekera	Clemson University, USA	John C. Wierman	Johns Hopkins University, USA
Gerry Ladas	University of Rhode Island, USA	Michael E. Zieve	University of Michigan, USA

PRODUCTION

Silvio Levy, Scientific Editor


Cover: Alex Scorpan

See inside back cover or msp.org/involve for submission instructions. The subscription price for 2018 is US \$190/year for the electronic version, and \$250/year (+\$35, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscriber address should be sent to MSP.

Involve (ISSN 1944-4184 electronic, 1944-4176 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840, is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

Involve peer review and production are managed by EditFLOW® from Mathematical Sciences Publishers.

PUBLISHED BY

 **mathematical sciences publishers**
nonprofit scientific publishing

<http://msp.org/>

© 2018 Mathematical Sciences Publishers

involve

2018 vol. 11 no. 1

On halving-edges graphs	1
TANYA KHOVANOVA AND DAI YANG	
Knot mosaic tabulation	13
HWA JEONG LEE, LEWIS D. LUDWIG, JOSEPH PAAT AND AMANDA PEIFFER	
Extending hypothesis testing with persistent homology to three or more groups	27
CHRISTOPHER CERICOLA, INGA JOHNSON, JOSHUA KIERS, MITCHELL KROCK, JORDAN PURDY AND JOHANNA TORRENCE	
Merging peg solitaire on graphs	53
JOHN ENGBERS AND RYAN WEBER	
Labeling crossed prisms with a condition at distance two	67
MATTHEW BEAUDOUIN-LAFON, SERENA CHEN, NATHANIEL KARST, JESSICA OEHRLEIN AND DENISE SAKAI TROXELL	
Normal forms of endomorphism-valued power series	81
CHRISTOPHER KEANE AND SZILÁRD SZABÓ	
Continuous dependence and differentiating solutions of a second order boundary value problem with average value condition	95
JEFFREY W. LYONS, SAMANTHA A. MAJOR AND KAITLYN B. SEABROOK	
On uniform large-scale volume growth for the Carnot–Carathéodory metric on unbounded model hypersurfaces in \mathbb{C}^2	103
ETHAN DLUGIE AND AARON PETERSON	
Variations of the Greenberg unrelated question binary model	119
DAVID P. SUAREZ AND SAT GUPTA	
Generalized exponential sums and the power of computers	127
FRANCIS N. CASTRO, OSCAR E. GONZÁLEZ AND LUIS A. MEDINA	
Coincidences among skew stable and dual stable Grothendieck polynomials	143
ETHAN ALWAISE, SHULI CHEN, ALEXANDER CLIFTON, REBECCA PATRIAS, ROHIL PRASAD, MADELINE SHINNERS AND ALBERT ZHENG	
A probabilistic heuristic for counting components of functional graphs of polynomials over finite fields	169
ELISA BELLAH, DEREK GARTON, ERIN TANNENBAUM AND NOAH WALTON	



1944-4176(2018)11:1;1-8