Cyclotomic Generating Functions

Sara C. Billey[a] Joshua P. Swanson[b]

Submitted: Jan 3, 2024; Accepted: Aug 15, 2024; Published: Oct 4, 2024

Abstract

It is a remarkable fact that for many statistics on finite sets of combinatorial objects, the roots of the corresponding generating function are each either a complex root of unity or zero. These and related polynomials have been studied for many years by a variety of authors from the fields of combinatorics, representation theory, probability, number theory, and commutative algebra. We call such polynomials *cyclotomic generating functions* (CGFs). With Konvalinka, we have studied the support and asymptotic distribution of the coefficients of severalfamilies of CGFs arising from tableau and forest combinatorics. In this paper, we survey general CGFs from algebraic, analytic, and asymptotic perspectives. We review some of the many known examples of CGFs in combinatorial representation theory; describe their coefficients, moments, cumulants, and characteristic functions; and give a variety of necessary and sufficient conditions for their existence arising from probability, commutative algebra, and invariant theory. As a sample result, we show that CGFs are "generically" asymptotically normal, generalizing a result of Diaconis on q-binomial coefficients using work of Hwang–Zacharovas. We include several open problems concerning CGFs.

Mathematics Subject Classifications: 05A16 (Primary), 60C05, 60F05 (Secondary)

1 Introduction

Many formulas in enumerative combinatorics express the cardinality of a finite set X as a product or quotient of integers. In many cases of interest, such formulas may be generalized to q-analogues with a corresponding refined count of X subject to the value of some statistic on X. Such q-analogues have nonnegative integer coefficients. A classic example is the ***q*-binomial coefficients**, which q-count integer partitions that fit in the $k \times (n-k)$ rectangle according to size for a given pair of nonnegative integers $k \leqslant n$. Here a partition $\lambda = (\lambda_1 \geqslant \cdots \geqslant \lambda_j \geqslant 0)$ fits in the rectangle provided $j \leqslant k$ and each

[a]Department of Mathematics, University of Washington, Seattle, WA, U.S.A. (`billey@uw.edu`).
[b]Department of Mathematics, University of Southern California, Los Angeles, CA, U.S.A. (`swansonj@usc.edu`).

λ_i is a nonnegative integer less than or equal to $(n-k)$. The size of λ is the sum of its parts, $|\lambda| = \sum \lambda_i$. These polynomials can be expressed in two ways,

$$\binom{n}{k}_q := \frac{[n]_q!}{[k]_q![n-k]_q!} = \sum_{\lambda \subset k \times (n-k)} q^{|\lambda|}, \tag{1.1}$$

where $[n]_q! := [n]_q[n-1]_q \cdots [1]_q$ is a ***q*-factorial** and $[n]_q := 1 + q + \cdots + q^{n-1} = (1-q^n)/(1-q)$ is a ***q*-integer**. The q-binomial coefficient $\binom{n}{k}_q$ is well known in topology and geometry as the Poincaré polynomial of the Grassmannian variety $\mathrm{Gr}(k,n)$ [Ful97]. Here X may be taken to be the set of Schubert cells of $\mathrm{Gr}(k,n)$ and the q-statistic is the dimension of the cell. Enumerative properties of the coefficients of $\binom{n}{k}_q$ have attracted attention since the mid 19th century [Syl78]. They are a prime example of unimodality [O'H90, Zei89].

In this paper, we study the class of such nonzero polynomials with nonnegative integer coefficients arising as quotients of q-integers from algebraic, analytic, and asymptotic perspectives. Such polynomials are closely associated with the cyclotomic polynomials Φ_n from number theory. See Section 2 for a review of their key properties. Since our examples all come from q-counting formulas for well-known combinatorial objects and their associated generating functions, we have chosen to call these polynomials **cyclotomic generating functions** or **CGFs** for short. We begin by stating several equivalent characterizations of this class. See Section 3 for the proof of their equivalence.

Theorem/Definition 1. Suppose $f(q) = \sum_{k=0}^{n} c_k q^k$ is a nonzero polynomial with nonnegative integer coefficients. The following are equivalent definitions for the polynomial f to be a **cyclotomic generating function (CGF)**.

(i) (Complex form.) The complex roots of $f(q)$ are all either roots of unity or zero.

(ii) (Kronecker form.) The complex roots of $f(q)$ all have modulus at most 1 and the leading coefficient is the greatest common divisor of all coefficients of f.

(iii) (Cyclotomic form.) The polynomial f can be written as a positive integer times a product of cyclotomic polynomials and factors of q.

(iv) (Rational form.) There are multisets $\{a_1, \ldots, a_m\}$ and $\{b_1, \ldots, b_m\}$ of positive integers and integers $\alpha \in \mathbb{Z}_{>0}$, $\beta \in \mathbb{Z}_{\geqslant 0}$ such that

$$f(q) = \alpha q^\beta \cdot \prod_{j=1}^{m} \frac{[a_j]_q}{[b_j]_q} = \alpha q^\beta \cdot \prod_{j=1}^{m} \frac{1-q^{a_j}}{1-q^{b_j}}.$$

Moreover, this factorization is unique if the multisets are disjoint.

(v) (Probabilistic form.) There is a discrete random variable $\mathcal{X}$ with probability distribution given by $P(\mathcal{X} = k) = c_k/f(1)$ such that the following equality in distribution holds:

$$\mathcal{X} + \mathcal{U}_{b_1} + \cdots + \mathcal{U}_{b_m} = \beta + \mathcal{U}_{a_1} + \cdots + \mathcal{U}_{a_m},$$

where the summands are all independent and β is maximal such that $q^\beta \mid f(q)$. Here $\mathcal{U}_a$ is a uniform random variable supported on $\{0, 1, \ldots, a-1\}$ and α is the greatest common divisor of all coefficients of f.

(vi) (Characteristic form.) There is a discrete random variable $\mathcal{X}$ on a uniform sample space supported on $\mathbb{Z}_{\geqslant 0}$ with probability generating function $\mathbb{E}[q^{\mathcal{X}}] = f(q)/f(1)$ such that the scaled characteristic function $\phi_{\mathcal{X}}(z) := \mathbb{E}[e^{2\pi i z \mathcal{X}}]$ has complex zeros only at rational $z \in \mathbb{Q}$.

Example 2. Allow us to indulge in a thought experiment. Imagine yourself as Percy MacMahon at the turn of the last century investigating **plane partitions** in an $x \times y \times z$ box for the first time. These are $x \times y$ matrices of nonnegative integers whose entries are at most z and weakly decrease along rows and columns. For $(x, y, z) = (3, 2, 2)$, you find there are 50 plane partitions. We have $50 = 2 \cdot 5^2$, and a wide-eyed optimist may hope for a product formula akin to that for $\binom{n}{k}$, though cancellations are difficult to uncover. Plane partitions come with the **size** statistic, which is the sum of all entries. The size generating function here is

$$f_{(3,2,2)}(q) = 1 + q + 3q^2 + 4q^3 + 6q^4 + 6q^5 + 8q^6 + 6q^7 + 6q^8 + 4q^9 + 3q^{10} + q^{11} + q^{12}.$$

You immediately notice the generating function is monic, palindromic, and even unimodal. Ever-optimistic, you try dividing off cyclotomic polynomial factors, which results in $f_{(3,2,2)}(q) = \Phi_6(q)\,\Phi_5(q)^2\,\Phi_4(q)$—you've found a CGF! Unique factorization as a reduced quotient of q-integers results in

$$f_{(3,2,2)}(q) = \frac{[6]_q\,[5]_q^2\,[4]_q}{[3]_q\,[2]_q^2\,[1]_q}.$$

We now see that most of the cancellations are "hidden" in the $q = 1$ specialization $2 \cdot 5^2$. It is not hard to imagine that further experimentation from here quickly leads you to the following version of MacMahon's famous formula,

$$f_{(x,y,z)}(q) = \prod_{i=1}^{x} \prod_{j=1}^{y} \frac{[i+j+z-1]_q}{[i+j-1]_q}, \tag{1.2}$$

which for $(x, y, z) = (3, 2, 2)$ yields

$$f_{(3,2,2)}(q) = \frac{[6]_q[5]_q[4]_q[5]_q[4]_q[3]_q}{[4]_q[3]_q[2]_q[3]_q[2]_q[1]_q}.$$

Definition 3. Among all CGFs, we focus on the family of **basic CGFs** with no α or q^β factors,

$$f(q) = \prod_{j=1}^{m} \frac{[a_j]_q}{[b_j]_q} = \prod_{j=1}^{m} \frac{1-q^{a_j}}{1-q^{b_j}}. \tag{1.3}$$

The **basic CGF monoid** Φ^+ is the monoid consisting of all basic CGFs under multiplication.

By the rational form, basic CGFs are always palindromic and monic. They need not be unimodal in general, e.g.

$$f(q) = q^6 + q^4 + q^3 + q^2 + 1 = \frac{[5]_q[6]_q}{[2]_q[3]_q} \in \Phi^+. \tag{1.4}$$

In Section 6.1, we study the submonoids Φ^{uni} and Φ^{lcc} of Φ^+ given by cyclotomic generating functions which are unimodal or log-concave with no internal zeros, respectively.

Testing if a polynomial $f \in \mathbb{Z}_{\geqslant 0}[q]$ is in Φ^+ may be done by repeatedly dividing off cyclotomic polynomial factors $\Phi_n(q)$. Note that we have the elementary bounds $\sqrt{n/2} \leqslant \deg \Phi_n(q) \leqslant n$, so there are only a finite number of basic CGFs of each given degree.

The main reason to study CGFs as a family of polynomials is that they are already prevalent in the literature. In addition to q-binomial coefficients, the standard q-analogues of $n!$ using inversions or the major index statistic on permutations and their generalizations to arbitrary words corresponding to q-multinomial generating functions are CGFs using MacMahon's classic formulas. Here are a (Catalan) number of examples from the literature. We emphasize that this list is not exhaustive.

1. Length generating functions of Weyl groups and their parabolic quotients [BB05].

2. The Hilbert series of all finite-dimensional quotient rings of the form

$$R \coloneqq B/(\theta_1, \ldots, \theta_m)$$

   where $\theta_1, \ldots, \theta_m$ is a homogeneous system of parameters in the polynomial ring $B = \Bbbk[x_1, \ldots, x_m]$, where $\deg(\theta_i) = a_i$ and $\deg(x_i) = b_i$. The ring of invariants of a finite reflection group has such a Hilbert series. See Section 6.3 for references and further discussion.

3. A q-analogue of Cayley's formula coming from the $t = q$ case of the diagonal harmonics, namely $q^{\binom{n}{2}} \operatorname{Hilb}(DH_n; q, q^{-1}) = [n+1]_q^{n-1}$ [Hai03, Thm. 4.2.4].

4. A q-analogue of the Catalan numbers [CWW08].

5. A q-analogue of the Fuß–Catalan numbers for irreducible well-generated complex reflection groups [KM13, (3.2), Thm. 25].

6. A q-analogue of Narayana numbers [RS18, Thm. 1.10].

7. A q-analogue of the Hook Length Formula for standard Young tableaux [Sta79, Prop. 4.11],

8. A q-analogue of the Hook Content Formula for semi-standard Young tableaux [Sta99, Thm. 7.21.2].

9. A q-hook length formula for linear extensions of forests [BW89].

10. A q-analogue of the Weyl dimension formula for highest weight modules of semisimple Lie algebras [Ste94, Lem. 2.5].

11. The q-analogues of the formulas enumerating alternating sign matrices [RSW14, p.171], cyclically symmetric plane partitions [MRR82], or totally symmetric plane partitions [KKZ11].

12. Rank generating functions of Bruhat intervals $[\mathrm{id}, w]$ for permutations w indexing smooth Schubert varieties in the complete flag manifolds [Gas98, GR02]. Similar results hold in other types as well. See [Slo15] for an extensive overview and recent results.

13. The statistic $\mathrm{baj} - \mathrm{inv}$ appeared in the context of extended affine Weyl groups and Hecke algebras in the work of Iwahori and Matsumoto in 1965 [IM65]. It is the Coxeter length generating function restricted to coset representatives of the extended affine Weyl group of type A_{n-1} mod translations by coroots. Stembridge and Waugh [SW98, Remarks 1.5 and 2.3] give a careful overview of this topic and further results. In particular, they prove the corresponding q-analogue of $n!$ is a cyclotomic generating function. See Zabrocki [Zab03] for the nomenclature and [BKS20a, §4] for an asymptotic description.

14. Rank generating functions of Gaussian posets and d-complete posets, with connections to Lie theory and order polynomials of posets, [Pro84, PS19, Ste94]. In recent work, Hopkins has explored general properties of posets with order polynomial product formulas in the context of cyclic sieving and other good dynamical behavior such as promotion and rowmotion [Hop24]. Furthermore, sometimes there is an associated cyclic sieving phenomenon associated to these order polynomials. These well behaved order polynomials are often cyclotomic generating functions. See also Stanton's online notes [Sta98] for conjectures related to CGFs generalizing rank generating functions of Gaussian posets, which he calls "fake Gaussian sequences".

Classes of polynomials similar to the class of cyclotomic generating functions have been studied for roughly a century in other contexts. A polynomial $f \in \mathbb{C}[q]$ of degree n is **self-inversive** if $f(q) = \alpha q^n f(q^{-1})$ where $|\alpha| = 1$. Equivalently, the zeros of f are symmetric about the unit circle. Self-inversive polynomials were studied by A. Cohn [Coh22] a century ago and by Bonsall–Marden [BM52] in the 1950s; they related them to the complex roots of f'.

Kedlaya [Ked08] and Hwang–Zacharovas [HZ15] refer to polynomials with roots all on the unit circle as **root-unitary**. Note that basic CGFs are in particular root-unitary with real coefficients. We will review key details from [HZ15] in Section 2 since they pertain to our focus on basic CGFs. In particular, they gave an elegant formula for the cumulants of the real-valued discrete random variables whose probability generating functions are basic CGFs and a test for asymptotic normality using the fourth cumulants. More recently, there has been interest in certain polynomials associated to numerical semigroups that have cyclotomic factorizations [CGSM16, CGSHPM22] and identifying criteria for central limit theorems for polynomials based on the geometry of their roots in the complex plane [HST24, MS19b, MS19a].

Another particularly well-known class of polynomials consists of elements of $\mathbb{R}_{\geqslant 0}[q]$ with all real roots [Brä15, Bre94, Pit97, Sta89]. Stirling numbers of the second kind are the motivating example [Har67]. A famous asymptotic characterization of their coefficients is due to Bender.

Theorem 4 ([Ben73, Thm. 2]; see also [Har67])**.** *If $\mathcal{X}_n$ is a sequence of real-valued discrete random variables whose probability generating functions are polynomials $f_n(q)$ with all real roots and standard deviation $\sigma_n \to \infty$, then $\mathcal{X}_n$ is asymptotically normal.*

In contrast to Theorem 4, CGF polynomials and their associated random variables have much more complex limiting behavior. In [BS22], we initiated the study of the metric space of all standardized CGF distributions in the Lévy metric and their asymptotic limits. For some families of CGFs, there are statistics that completely characterize their limiting behavior. Asymptotic normality is one common occurrence. While finding a complete description of the closure of all CGF distributions is still an open problem [BS22, Open Problem 1.19], we did show that all uncountably many **DUSTPAN distributions** (**d**istributions associated to a **u**niform **s**um for **t** **p**lus **a** **n**ormal distribution) can occur on the boundary. Furthermore, many more multimodal limiting distributions are possible, due to the following construction. Given a CGF $f(q)$ and positive integer N, note that $(1 + q^N)f(q)$ remains a CGF. If N is larger than the degree of $f(q)$, the result is a CGF with two disjoint copies of the distribution of $f(q)$. In this way, valid limits may be "duplicated" in a fractal pattern. The unimodal and log-concave submonoids of Φ^+ are not subject to this construction. We expect the unimodal basic CGFs will have the simplest limiting behavior.

Towards addressing the problem of characterizing when a limiting sequence of standardized CGF distributions in the Lévy metric approaches the standard normal distribution $\mathcal{N}(0,1)$, we present a new criterion in terms of the multisets in the rational presentation of the CGF polynomials. It is these multisets which have combinatorial significance in applications. We give a more complete characterization in the special case of polynomials which are products of q-integers, see Theorem 37. This approach builds on the work of Hwang–Zacharovas [HZ15], but emphasizes the role of the multisets of q-integers in the numerators. It also generalizes the work of Diaconis [Dia88, pp.128-129] who showed that the coefficients of a sequence of q-binomials $\binom{n}{k}_q$ are asymptotically normal provided both $k, n-k \to \infty$. See Section 4 on asymptotic considerations for the proof.

Theorem 5. *Let* $f_N(q) = \prod_{a \in a^{(N)}} [a]_q / \prod_{b \in b^{(N)}} [b]_q \in \Phi^+$ *for* $N = 1, 2, \ldots$ *be a sequence of basic cyclotomic generating functions expressed in terms of their associated multisets* $a^{(N)}, b^{(N)}$*. For each* N*, let* $\mathcal{X}_N$ *be the corresponding CGF random variable with* $\mathbb{E}[q^{\mathcal{X}_N}] = f_N(q)/f_N(1)$*. If*

$$\limsup_{N \to \infty} \frac{\sum_{b \in b^{(N)}} (b^2 - 1)}{\sum_{a \in a^{(N)}} (a^2 - 1)} < 1 \tag{1.5}$$

and

$$\lim_{N \to \infty} \sum_{a \in a^{(N)}} \left(\frac{a}{\max a^{(N)}} \right)^4 = \infty, \tag{1.6}$$

then $\mathcal{X}_1, \mathcal{X}_2, \ldots$ *is asymptotically normal.*

The ratios in (1.5) are related to the variances of the random variables obtained from sums of uniform random variables associated to the numerators and denominators respectively. The sum in (1.6) is derived using the fourth cumulant test due to Hwang–Zacharovas. Hence the appearance of the quadratic and quartic terms. When the limiting ratio in (1.5) is 1, massive cancellation occurs and obscures the asymptotic behavior. More careful analysis may still reveal limit laws, such as in the **Irwin–Hall** case of [BKS20a, Thm. 1.7].

Example 6. Continuing the plane partition example, suppose we have an infinite sequence of triples (x_N, y_N, z_N) for $N = 1, 2, \ldots$. For ease of notation, we will just write (x, y, z) for a general element in the sequence. In MacMahon's formula (1.2), we see intuitively that the numerator dominates the denominator, which is essentially condition (1.5). More rigorously, routine calculations show that the ratio (1.5) is the rational function

$$\frac{g(x,y)}{g(x,y) + z \cdot (x + y + z)}, \qquad \text{where } g(x,y) = (2x^2 + 3xy + 2y^2 - 7)/6.$$

When $x \leqslant y \leqslant z$, which may be arranged without loss of generality, and $z \to \infty$, we use the rational function above to observe (1.5) holds. Similarly we find that (1.6) holds if $y \to \infty$ as well. Hence with nothing more than basic computer algebra systems and routine calculations, we find that size on plane partitions in a box is asymptotically normal if $\mathrm{median}(x, y, z) \to \infty$, recovering a result from [BS22]. Indeed, this result is sharp in the sense that, when $\mathrm{median}(x, y, z)$ is bounded and $z \to \infty$, the possible limits are Irwin–Hall distributions; see [BS22]. The approach in this example to proving central limit theorems is a kindred spirit to Zeilberger's "automatic central limit theorem generator" [Zei09].

Remark 7. While Theorem 5 provides a precise asymptotic conclusion, [MS19a, Thm. 1.4] states the following quantitative bound after a lengthy technical argument. If $f(q)$ is a polynomial which is zero-free in the sector $\{z \in \mathbb{C} : |\arg(q)| < \delta\}$ and if $\mathcal{X}$ is a random variable with $\mathbb{E}[q^{\mathcal{X}}] = f(q)/f(1)$, then

$$\sup_{t \in \mathbb{R}} |\mathbb{P}(\mathcal{X}^* \leqslant t) - \mathbb{P}(\mathcal{Z} \leqslant t)| \leqslant \frac{C}{\delta\sigma}, \tag{1.7}$$

where C is an absolute constant and $\mathcal{Z}$ is the standard normal random variable. In the situation of Theorem 5, $\delta^{-1} = \max a^{(N)}/2\pi$ and $12\sigma^2 = \sum_{a\in a^{(N)}} a^2 - \sum_{b\in b^{(N)}} b^2$, so (1.7) becomes

$$\sup_{t\in\mathbb{R}} |\mathbb{P}(\mathcal{X}^* \leqslant t) - \mathbb{P}(\mathbb{Z} \leqslant t)| \leqslant \frac{C_0 \max a^{(N)}}{\sqrt{\sum_{a\in a^{(N)}} a^2 - \sum_{b\in b^{(N)}} b^2}}. \tag{1.8}$$

Using MacMahon's plane partition formula (1.2), $\max a^{(N)} = x+y+z-1$ and $12\sigma^2 = xyz(x+y+z)$. The right-hand side in (1.8) in this case may then be replaced with $C_0\sqrt{\frac{x+y+z}{xyz}}$. Note that this expression tends to 0 if and only if $\operatorname{median}(x,y,z) \to \infty$.

Towards addressing the problem of characterizing all limiting distributions for CGF random variables, we note the following special property. Roughly speaking, it states that the converse of the Frechét–Shohat Theorem holds for standardized CGF distributions. A proof is given in Section 5. The same statement holds with moments replaced by cumulants, and in fact, this paper focuses on the cumulants as the key quantities characterizing a CGF distribution.

Theorem 8. *Let $\mathcal{X}_1, \mathcal{X}_2, \ldots$ be a sequence of random variables corresponding to cyclotomic generating functions. Then the sequence of standardized random variables $\mathcal{X}_1^*, \mathcal{X}_2^*, \ldots$ converges in distribution if and only if for all $d \in \mathbb{Z}_{\geqslant 1}$, the limit of the d^{th} moments of the random variables*

$$\mu_d := \lim_{N\to\infty} \mu_d^{\mathcal{X}_N^*}$$

exists and is finite. In this case, $\mathcal{X}_n^* \Rightarrow \mathcal{X}$ where $\mu_d = \mu_d^{\mathcal{X}}$ is the d^{th} moment of $\mathcal{X}$, $\mathcal{X}$ is determined by its moments, and the moment-generating function of $\mathcal{X}$ is entire.

The paper is organized as follows. In Section 2, we state our notation and give background details from prior work. In Sections 3 to 5, we describe the algebraic, asymptotic and analytic properties of basic CGFs, their associated random variables, and characteristic functions. In Section 6, we examine the monoid of basic cyclotomic generating functions and several of its submonoids including unimodal and log-concave CGFs, CGFs whose rational form satisfies the Gale order, and CGFs that come from Hilbert series of polynomial rings modded out by homogeneous systems of parameters. We include several open problems about cyclotomic generating functions for future work based on experimentation.

2 Background

2.1 Cyclotomic polynomials

The **cyclotomic polynomials** are defined for all positive integers $n \geqslant 1$ by

$$\Phi_n(q) = \prod_{\substack{1\leqslant k\leqslant n \\ \gcd(k,n)=1}} (q - e^{2\pi i k/n}). \tag{2.1}$$

Each n^{th} root of unity $e^{2\pi ik/n}$ is in the cyclic group generated by $e^{2\pi i/d}$ for $d \in \{1, 2, \ldots, n\}$ where $d = n/\gcd(k, n)$. Thus, for $n \geqslant 1$, we have

$$q^n - 1 = \prod_{k=1}^{n}(q - e^{2\pi ik/n}) = \prod_{d|n} \Phi_d(q). \tag{2.2}$$

Since $q^n - 1 = (q-1)(1 + q + q^2 + \cdots + q^{n-1})$, we have

$$[n]_q = 1 + q + q^2 + \cdots + q^{n-1} = \prod_{1<d|n} \Phi_d(q). \tag{2.3}$$

Let $\mu(n)$ be the classical Möbius function. The Möbius function satisfies the recurrence $\sum_{d|n} \mu(n/d) = 0$ for all $n > 1$. Therefore, by Möbius inversion, we also have the identities for cyclotomic polynomials indexed by $n > 1$,

$$\Phi_n(q) = \prod_{d|n}(q^d - 1)^{\mu(n/d)} = \prod_{d|n}[d]_q^{\mu(n/d)}. \tag{2.4}$$

It follows by induction that $\Phi_n(q)$ is monic with integer coefficients.

The cyclotomic polynomials in the product (2.2) are the irreducible factors of $(q^n - 1)$ over the integers. For example, $\Phi_1(q) = q - 1$, $\Phi_2(q) = q + 1$, $\Phi_3(q) = 1 + q + q^2$, $\Phi_4(q) = q^2 + 1$, and $\Phi_{27}(q) = q^{18} + q^9 + 1$. See [Wik22] or [Coh03, Sect. 7.7] for many beautiful properties of cyclotomic polynomials including the formula

$$\Phi_{p^k}(q) = \sum_{j=0}^{p-1} q^{j\,p^{k-1}}.$$

Also, $\Phi_p(q) = 1 + q + q^2 + \cdots + q^{p-1} = [p]_q$ if and only if p is prime. Note, $\Phi_{p^k}(1) = p$, otherwise if $n > 1$ is not a prime power, then $\Phi_n(1) = 1$. The number of complex roots and therefore the degree of $\Phi_n(q)$ is given by Euler's totient function, $\varphi(n)$, so for $n \geqslant 3$ the degree of $\Phi_n(q)$ is even and bounded by $\sqrt{n/2} \leqslant \varphi(n) \leqslant n - 1$. For $n \geqslant 2$, the constant term is $\Phi_n(0) = 1$, and the coefficients are **palindromic** in the sense that $\Phi_n(q) = q^{\varphi(n)}\Phi_n(q^{-1})$.

2.2 Probabilistic generating functions and cumulants

We now briefly review the background from probability related to CGFs. See [Bil95] for more details or [BKS20a, §2] for a review aimed at a combinatorial audience.

The **probability generating function** of a discrete random variable $\mathcal{X}$ supported on $\mathbb{Z}_{\geqslant 0}$ is

$$G_{\mathcal{X}}(q) := \mathbb{E}[q^{\mathcal{X}}] = \sum_{k=0}^{\infty} P(\mathcal{X} = k)q^k.$$

If $\mathcal{X}$ is a nonnegative-integer-valued statistic on a set S which is sampled uniformly, then the probability generating function of $\mathcal{X}$ as a random variable is the same as the ordinary generating function of $\mathcal{X}$, up to a scale factor:

$$G_{\mathcal{X}}(q) = \frac{1}{\#S} \sum_{s \in S} q^{\mathcal{X}(s)}.$$

The **moment generating function** of $\mathcal{X}$ is obtained by substituting $q = e^t$ above so

$$M_{\mathcal{X}}(t) \coloneqq \mathbb{E}[e^{t\mathcal{X}}] = G_{\mathcal{X}}(e^t) = \sum_{d=0}^{\infty} \mu_d \frac{t^d}{d!},$$

where $\mu_d \coloneqq \mathbb{E}[\mathcal{X}^d]$ is the dth **moment**. The **central moments** of $\mathcal{X}$ are $\alpha_d \coloneqq \mathbb{E}[(\mathcal{X}-\mu)^d]$, where $\mu = \mu_1$ is the mean and $\alpha_2 = \sigma^2$ is the variance. The **characteristic function** of $\mathcal{X}$ is

$$\phi_{\mathcal{X}}(t) \coloneqq \mathbb{E}[e^{it\mathcal{X}}] = M_{\mathcal{X}}(it) = G_{\mathcal{X}}(e^{it}).$$

The characteristic function, for $t \in \mathbb{R}$, exists for all random variables and determines the distribution of $\mathcal{X}$. While moment generating functions in general are less well-behaved, all of the moment generating functions we will encounter converge in a complex neighborhood of 0 and the distributions will be determined by their moments.

The **cumulant generating function**, also known as the **second characteristic function**, of $\mathcal{X}$ is $\log \phi_{\mathcal{X}}(t) = \log M_{\mathcal{X}}(it)$. The **cumulants** $\kappa_1, \kappa_2, \ldots$ of $\mathcal{X}$ are defined to be the coefficients of the related exponential generating function

$$K_{\mathcal{X}}(t) \coloneqq \log M_{\mathcal{X}}(t) = \log \mathbb{E}[e^{t\mathcal{X}}] = \sum_{d=1}^{\infty} \kappa_d \frac{t^d}{d!}.$$

Cumulants and moments are polynomials in each other and are interchangeable for many purposes, though cumulants generally have more convenient formal properties. For example, for independent random variables $\mathcal{X}$ and $\mathcal{Y}$, $\kappa_d^{\mathcal{X}+\mathcal{Y}} = \kappa_d^{\mathcal{X}} + \kappa_d^{\mathcal{Y}}$ for all positive integers d.

The normal distribution with mean μ and standard deviation σ, denoted $\mathcal{N}(\mu, \sigma)$ is central in this work. It is the unique distribution with cumumlants $\kappa_1 = \mu, \kappa_2 = \sigma^2$, and $\kappa_d = 0$ for all positive integers $d \geqslant 3$.

The cumulants of random variables associated to cyclotomic generating functions have the following simple, explicit form due to Hwang–Zacharovas. This builds on work of Chen–Wang–Wang [CWW08, Thm. 3.1] and Sachkov [Sac97, §1.3.1]. From their formula it is easy to derive the formula for the moments and central moments of CGF random variables as explained in [BKS20a, §2.3].

Theorem 9. *[HZ15, §4.1] Suppose $\{a_1, \ldots, a_m\}$ and $\{b_1, \ldots, b_m\}$ are multisets of positive integers such that*

$$f(q) = \frac{\prod_{k=1}^{m} [a_k]_q}{\prod_{k=1}^{m} [b_k]_q} = \sum_{k=0}^{n} c_k q^k \in \Phi^+,$$

so in particular each $c_k \in \mathbb{Z}_{\geqslant 0}$ and $n = \sum a_k - b_k \geqslant 0$. Let $\mathcal{X}$ be a discrete random variable with $\mathbb{P}[\mathcal{X} = k] = c_k/f(1)$. Then the dth cumulant of $\mathcal{X}$ is

$$\kappa_d^{\mathcal{X}} = \frac{B_d}{d} \sum_{k=1}^{m} (a_k^d - b_k^d), \tag{2.5}$$

where B_d is the dth Bernoulli number (with $B_1 = \frac{1}{2}$). Moreover, the dth central moment of $\mathcal{X}$ is

$$\alpha_d = \sum_{\substack{|\lambda|=d \\ \text{has all parts even}}} \frac{d!}{z_\lambda} \prod_{i=1}^{\ell(\lambda)} \frac{B_{\lambda_i}}{\lambda_i!} \left[\sum_{k=1}^{m} \left(a_k^d - b_k^d \right) \right], \tag{2.6}$$

and the dth moment of $\mathcal{X}$ is

$$\mu_d = \sum_{\substack{|\lambda|=d \\ \text{has all parts either} \\ \text{even or size 1}}} \frac{d!}{z_\lambda} \prod_{i=1}^{\ell(\lambda)} \frac{B_{\lambda_i}}{\lambda_i!} \left[\sum_{k=1}^{m} \left(a_k^d - b_k^d \right) \right]. \tag{2.7}$$

Definition 10. Let $\mathcal{X}_1, \mathcal{X}_2, \ldots$ and $\mathcal{X}$ be real-valued random variables with cumulative distribution functions $F_1, F_2, \ldots$ and F, respectively. We say $\mathcal{X}_1, \mathcal{X}_2, \ldots$ **converges in distribution** to $\mathcal{X}$, written $\mathcal{X}_n \Rightarrow \mathcal{X}$, if for all $t \in \mathbb{R}$ at which F is continuous we have

$$\lim_{n\to\infty} F_n(t) = F(t).$$

For any real-valued random variable $\mathcal{X}$ with mean μ and variance $\sigma^2 > 0$, the corresponding **standardized random variable** is

$$\mathcal{X}^* := \frac{\mathcal{X} - \mu}{\sigma}.$$

Observe that $\mathcal{X}^*$ has mean $\mu^* = 0$ and variance $\sigma^{*2} = 1$. The moments and central moments of $\mathcal{X}^*$ agree for $d \geqslant 2$ and are given by

$$\mu_d^* = \alpha_d^* = \alpha_d/\sigma^d.$$

Similarly, the cumulants of $\mathcal{X}^*$ are given by $\kappa_1^* = 0$, $\kappa_2^* = 1$, and $\kappa_d^* = \kappa_d/\sigma^d$ for $d \geqslant 2$.

Definition 11. Let $\mathcal{X}_1, \mathcal{X}_2, \ldots$ be a sequence of real-valued random variables. We say the sequence is **asymptotically normal** if $\mathcal{X}_n^* \Rightarrow \mathcal{N}(0, 1)$.

We next describe two standard criteria for establishing asymptotic normality or more generally convergence in distribution of a sequence of random variables.

Theorem 12 (Lévy's Continuity Theorem, [Bil95, Theorem 26.3]). *A sequence $\mathcal{X}_1, \mathcal{X}_2, \ldots$ of real-valued random variables converges in distribution to a real-valued random variable $\mathcal{X}$ if and only if, for all $t \in \mathbb{R}$,*

$$\lim_{n\to\infty} \mathbb{E}[e^{it\mathcal{X}_n}] = \mathbb{E}[e^{it\mathcal{X}}].$$

Theorem 13 (Frechét–Shohat Theorem, [Bil95, Theorem 30.2])**.** *Let $\mathcal{X}_1, \mathcal{X}_2, \ldots$ be a sequence of real-valued random variables, and let $\mathcal{X}$ be a real-valued random variable. Suppose the moments of $\mathcal{X}_n$ and $\mathcal{X}$ all exist and the moment generating functions all have positive radius of convergence. If*

$$\lim_{n\to\infty} \mu_d^{\mathcal{X}_n} = \mu_d^{\mathcal{X}} \quad \forall d \in \mathbb{Z}_{\geqslant 1}, \tag{2.8}$$

then $\mathcal{X}_1, \mathcal{X}_2, \ldots$ converges in distribution to $\mathcal{X}$.

By Theorem 12, we may test for asymptotic normality by checking if the standardized characteristic functions tend pointwise to the characteristic function of the standard normal. Likewise by Theorem 13 we may instead perform the check on the level of individual standardized moments, which is often referred to as the **method of moments**. By the polynomial relationship between moments and cumulants, we may further replace the moment condition (2.8) with the cumulant condition

$$\lim_{n\to\infty} \kappa_d^{\mathcal{X}_n} = \kappa_d^{\mathcal{X}}. \tag{2.9}$$

For instance, we have the following explicit criterion.

Corollary 14. *A sequence $\mathcal{X}_1, \mathcal{X}_2, \ldots$ of real-valued random variables on finite sets is asymptotically normal if for all $d \geqslant 3$ we have*

$$\lim_{n\to\infty} \frac{\kappa_d^{\mathcal{X}_n}}{(\sigma^{\mathcal{X}_n})^d} = 0. \tag{2.10}$$

In fact, the converse of the Frechét–Shohat theorem holds for cyclotomic generating functions. See Theorem 40 below, which builds on [HZ15, Lem. 2.8]. Furthermore, we have the following simplified test for asymptotic normality due to Hwang and Zacharovas.

Theorem 15. *[HZ15, Thm. 1.1] Let $f_1(q), f_2(q), \ldots$ be a sequence of cyclotomic generating functions. Let $\mathcal{X}_1, \mathcal{X}_2, \ldots$ be a corresponding sequence of random variables with $\mathbb{E}[q^{\mathcal{X}_N}] = f_N(q)/f_N(1)$. Then, $\mathcal{X}_1, \mathcal{X}_2, \ldots$ is asymptotically normal if and only if the standardized fourth cumulants approach 0,*

$$\lim_{n\to\infty} \frac{\kappa_4^{\mathcal{X}_n}}{(\sigma^{\mathcal{X}_n})^4} = 0. \tag{2.11}$$

2.3 Formal cumulants

We may extend the notions of cumulants and moments to power series even when they do not necessarily have associated discrete random variables. Suppose that $f(q) \in R[[q]]$ is a formal power series with coefficients in a (commutative, unital) ring R of characteristic 0. If $f(1) = 1$, one may define the **formal cumulants** of f by the coefficients in the expansion of the generating function

$$\log f(e^t) = \sum_{d=1}^{\infty} \kappa_d(f) \frac{t^d}{d!} \tag{2.12}$$

where $e^t := \sum_{k=0}^{\infty} \frac{t^n}{n!} \in R[[t]]$. See [PW99] or [BKS20a, §2] for more details. If f is clear from context, we will often just write κ_d for $\kappa_d(f)$. Similarly, we write $\mu = \kappa_1(f)$ and $\sigma^2 = \kappa_2(f)$. If $f(1) \neq 0$ is invertible, we use $\kappa_d(f) := \kappa_d(f/f(1))$. If $R = \mathbb{C}$, we may also define the **formal characteristic function** of f by $\phi_f(t) := f(e^{it})/f(1)$.

For example, if $f(q) = \frac{\prod_{k=1}^m [a_k]_q}{\prod_{k=1}^m [b_k]_q} \in \mathbb{Z}[[q]]$, the cumulant formula (2.5) remains valid, so

$$\kappa_d(f) = \frac{B_d}{d} \left(\sum_{k=1}^m a_k^d - \sum_{k=1}^m b_k^d \right). \tag{2.13}$$

We also have two factored forms for the formal characteristic function

$$\phi_f(t) := f(e^{it})/f(1) = \frac{\prod_{k=1}^m [a_k]_{e^{it}}/a_k}{\prod_{k=1}^m [b_k]_{e^{it}}/b_k} = e^{-it\mu} \prod_{j=1}^m \frac{\operatorname{sinc}(a_j t/2)}{\operatorname{sinc}(b_j t/2)}, \tag{2.14}$$

where $[a]_{e^{it}} = (1 + e^{it} + e^{2it} + \cdots + e^{(a-1)it})$, $\operatorname{sinc}(x) := \frac{\sin x}{x}$ and $\mu = \kappa_1(f)$. This coincides with the actual characteristic function $\phi_{\mathcal{X}}(t)$ when $f(q)$ is a cyclotomic generating function with corresponding random variable $\mathcal{X}$.

In particular, the formal mean of the cyclotomic polynomial $\Phi_n(q)$ is

$$\mu = \kappa_1(\Phi_n(q)) = B_1 \sum_{k|n} \mu(n/k)k = B_1 \; n \prod_{\substack{p \text{ prime} \\ p|n}} \left(1 - \frac{1}{p}\right) = \varphi(n)/2$$

for any $n > 1$ by known formulas of the Möbius function and Euler's totient function $\varphi(n)$. Thus, μ is half the degree of $\Phi_n(q)$ as expected by (2.1). More generally, the formal cumulants of cyclotomic polynomials are closely related to Jordan's generalization of the Euler totient formula, given by

$$J_d(n) = n^k \prod_{\substack{p \text{ prime} \\ p|n}} \left(1 - \frac{1}{p^k}\right).$$

Lemma 16. *The formal cumulants of the cyclotomic polynomials for $n > 1$ satisfy*

$$\kappa_d(\Phi_n(q)) = \frac{B_d}{d} \sum_{k|n} \mu(n/k)k^d = \frac{B_d}{d} J_d(n). \tag{2.15}$$

Proof. By (2.4), the cyclotomic polynomials for $n > 1$ can be expressed in rational form as a ratio of q-integers so one can use (2.13) to compute $\kappa_d(\Phi_n(q)) = \frac{B_d}{d} \sum_{k|n} \mu(n/k)k^d$. This sum factors as

$$\sum_{k|n} \mu(n/k)k^d = n^d \prod_{\substack{p \text{ prime} \\ p|n}} \left(1 - \frac{1}{p^d}\right). \tag{2.16}$$

Indeed, (2.16) can be easily verified when n is a prime power, and it is straightforward to check that both sides are multiplicative functions. Thus, the second equality in (2.15) holds. □

We will use the next corollary which was first noted by Hwang–Zacharovas for root unitary polynomials. It applies to all polynomials which can be expressed as rational products of q-integers. The proof relies on (2.13) and the fact that the Bernoulli numbers B_n vanish for odd $n > 1$.

Corollary 17. *[HZ15, Cor. 3.1] Let $f \in \mathbb{R}_{\geqslant 0}[q]$ be any nonzero polynomial such that all of its complex roots have modulus 1. The corresponding odd cumulants $\kappa_{2d-1}(f)$ vanish after the first. The corresponding even cumulants $\kappa_{2d}(f)$ alternate in sign to make $(-1)^{d-1}\kappa_{2d}(f) \geqslant 0$.*

2.4 Generalized uniform sum distributions

Here we recall some of the notation of p-norms, the decreasing sequence space with finite p-norm, and the generalized uniform sum random variables. See [BS22] for more details.

Definition 18. Let $\mathbf{t} = (t_1, t_2, \ldots)$ be a sequence of nonnegative real numbers. For $p \in \mathbb{R}_{\geqslant 1}$, the ***p*-norm** of $\mathbf{t}$ is $|\mathbf{t}|_p \coloneqq (\sum_{k=1}^{\infty} t_k^p)^{1/p}$. We also set $|\mathbf{t}|_\infty \coloneqq \sup_k t_k$.

The p-norm has many nice properties. It is well-known (e.g. [MV97, Ex. 7.3, p.58]) that if $1 \leqslant p \leqslant q \leqslant \infty$, then $|\mathbf{t}|_p \geqslant |\mathbf{t}|_q$, and that if $|\mathbf{t}|_p < \infty$, then $\lim_{q\to\infty} |\mathbf{t}|_q = |\mathbf{t}|_\infty$. Thus, if $\mathbf{t}$ is weakly decreasing, $|\mathbf{t}|_\infty = \sup_k t_k = t_1$.

The **sequence space with finite *p*-norm** $\ell_p \coloneqq \{\mathbf{t} = (t_1, t_2, \ldots) \in \mathbb{R}_{\geqslant 0}^{\mathbb{N}} : |\mathbf{t}|_p < \infty\}$ is commonly used in functional analysis and statistics. Here we define a related concept for analyzing sums of central continuous uniform random variables.

Definition 19. The **decreasing sequence space with finite *p*-norm** is

$$\widetilde{\ell}_p \coloneqq \{\mathbf{t} = (t_1, t_2, \ldots) : t_1 \geqslant t_2 \geqslant \cdots \geqslant 0, |\mathbf{t}|_p < \infty\}.$$

The elements of $\widetilde{\ell}_p$ may equivalently be thought of as the set of **countable multisets of nonnegative real numbers with finite *p*-norm**. Any finite multiset of nonnegative real numbers can be considered as an element of $\widetilde{\ell}_p$ with finite support by sorting the multiset and appending 0's. We write $\widetilde{\ell}_{\leqslant m}$ for the resulting collection of decreasing sequences with at most m nonzero entries. The multisets in $\widetilde{\ell}_p$ are uniquely determined by their p-norms. In fact, any sequence of p-norm values injectively determines the multiset provided the sequence goes to infinity.

Definition 20. [BS22, §3] A **generalized uniform sum distribution** is any distribution associated to a random variable with finite mean and variance given as a countable sum of independent continuous uniform random variables. Such random variables are given by a constant overall shift plus a **uniform sum random variable**

$$\mathcal{S}_{\mathbf{t}} \coloneqq \mathcal{U}\left[-\frac{t_1}{2}, \frac{t_1}{2}\right] + \mathcal{U}\left[-\frac{t_2}{2}, \frac{t_2}{2}\right] + \cdots$$

for some $\mathbf{t} = (t_1, t_2, \ldots) \in \widetilde{\ell}_2$.

The uniform sum random variables have nice cumulant formulas which are similar to the CGF distributions. It was shown in [BS22, Lem. 3.11] that for $d \geqslant 2$ and $\mathbf{t} = (t_1, t_2, \ldots) \in \widetilde{\ell}_2$, we have

$$\kappa_d^{\mathcal{S}_\mathbf{t}} = \frac{B_d}{d} \sum_{k=1}^{\infty} (t_k)^d = \frac{B_d}{d} |\mathbf{t}|_d^d. \tag{2.17}$$

Theorem 21. *[BS22, Thm 3.13] Generalized uniform sum distributions are bijectively parameterized by $\mathbb{R} \times \widetilde{\ell}_2$. In particular, if $\mathbf{t}, \mathbf{u} \in \widetilde{\ell}_2$ with $\mathbf{t} \neq \mathbf{u}$, then $\mathcal{S}_\mathbf{t} \neq \mathcal{S}_\mathbf{u}$. Furthermore, $\mathcal{S}_\mathbf{t}^* = \mathcal{S}_\mathbf{u}^*$ if and only if $\mathbf{t}, \mathbf{u}$ differ by a scalar multiple.*

It is not known what all the possible limiting distributions of families of CGF polynomials are. By [BS22, Cor. 3.17], we know that all standardized uniform sum distributions $\mathcal{S}_\mathbf{t}^*$ do occur as limiting distributions coming from the hook length formulas for linear extensions of forests due to Björner and Wachs [BW89]. In fact, all standardized **DUSTPAN distributions** can occur as limits of CGF distributions. These are distributions of the form $\mathcal{S}_\mathbf{t} + \mathcal{N}(0, \sigma^2)$, assuming the two random variables are independent, $\mathbf{t} \in \widetilde{\ell}_2$, and $\sigma^2 := 1 - |\mathbf{t}|_2^2/12 \in \mathbb{R}_{\geqslant 0}$.

3 Algebraic considerations

3.1 Equivalent characterizations of CGFs

One algebraic justification for studying cyclotomic generating functions as a special class of polynomials is the following classical result of Kronecker from the 1850s. We include a proof similar to Kronecker's for completeness. See [Mat10] for further references.

Theorem 22. *[Kro57] Suppose $f(q) \in \mathbb{Z}[q]$ is monic and all of its complex roots have modulus at most 1. Then the roots of $f(q)$ are each either a root of unity or 0.*

Proof. If $f(q) = \prod_{j=1}^n (q - z_j)$ for $z_1, \ldots, z_n \in \mathbb{C}$, define $f_k(q) = \prod_{j=1}^n (q - z_j^k)$ for each positive integer k. The coefficients of $f_k(q)$ are elementary symmetric polynomials in the z_j^k and each $|z_j^k| \leqslant 1$ by hypothesis, so the coefficients of $f_k(q)$ are bounded in modulus by binomial coefficients. These coefficients are also symmetric functions of roots of f, so they belong to the fixed field of the Galois group of f, namely the base field $\mathbb{Q}$. Since f is monic, its roots $z_1, \ldots, z_n$ are algebraic integers as are all sums of products of the z_i's. Therefore, the coefficients of each $f_k(q)$ are also algebraic integers, and since the only algebraic integers in $\mathbb{Q}$ are the integers themselves, we observe $f_k(q) \in \mathbb{Z}[q]$. The list $f_1, f_2, \ldots$ must thus eventually repeat. We may as well suppose $f = f_1$ is repeated, so that taking kth powers for some $k > 1$ permutes the z_i's. Hence, taking kth powers $n!$ times implies $z_i = z_i^{kn!}$ for all i, and the result follows. □

We now prove the six equivalent characterizations of a CGF from the introduction. These results follow closely from properties of cyclotomic polynomials reviewed in Section 2.

Proof of Theorem/Definition 1. By the definition of cyclotomic polynomials, (i) $\Leftrightarrow$ (iii) and (iv) $\Rightarrow$ (i). For (iii) $\Rightarrow$ (iv), recall the identity for cyclotomic polynomials $\Phi_n(q) = \prod_{d|n}(q^d-1)^{\mu(n/d)}$ from (2.4). Furthermore, if $\Phi_n(q)$ is a divisor of $f \in \Phi^+$, then we can assume $n > 1$ since $f(1)$ is positive. By the well-known recurrence $\sum_{d|n} \mu(n/d) = 0$ for all $n > 1$, the number of factors in the numerator and denominator of $\Phi_n(q) = \prod_{d|n}(q^d-1)^{\mu(n/d)}$ are equal so

$$\Phi_n(q) = \prod_{d|n}(q^d-1)^{\mu(n/d)} = \prod_{d|n}[d]_q^{\mu(n/d)} = \prod_{j=1}^{m} \frac{1-q^{a_j}}{1-q^{b_j}} = \prod_{j=1}^{m} \frac{[a_j]_q}{[b_j]_q} \tag{3.1}$$

for some multisets $\{a_1, \ldots, a_m\}$ and $\{b_1, \ldots, b_m\}$ of positive integers. Hence (iv) is equivalent to (i) and (iii). The uniqueness claim follows from the uniqueness of polynomial factorizations.

The equivalence of (i) and (iii) implies (ii) because the cyclotomic polynomials are all monic. In the other direction, after dividing through by the leading coefficient which is also the greatest common divisor of all the coefficients by hypothesis, we can assume $f(q)$ is a monic polynomial with nonnegative integer coefficients. If $f(q)$ also has all of its complex roots with modulus at most 1, then the roots of f are each either a root of unity or 0 by Theorem 22. Hence, (ii) implies (i).

The equivalence of (iv) and (v) follows from the polynomial identity

$$f(q)\prod_{j=1}^{m}[b_j]_q = \alpha q^\beta \cdot \prod_{j=1}^{m}[a_j]_q.$$

Up to a choice of α and β, the equivalence of (i) and (vi) follows since the roots of unity $e^{2\pi ik/n}$ which are zeros of $f(q)/f(1) = \mathbb{E}[q^{\mathcal{X}}]$ are all determined by rational numbers k/n which give rise to all zeros of $\mathbb{E}[e^{2\pi it\mathcal{X}}]$. $\square$

3.2 Rational products of q,qintegers

Consider the general class of rational products of q-integers of the form

$$f(q) = \frac{\prod_{k=1}^{m}[a_k]_q}{\prod_{k=1}^{m}[b_k]_q} = \sum_{k=0}^{\infty} c_k q^k \tag{3.2}$$

as formal power series in $\mathbb{Z}[[q]]$. Such rational products include the set of basic CGFs. We examine several properties of such products.

In the next lemma, we state an explicit formula for the coefficients of the expansion of (3.2) generalizing work of Knuth for the number of permutations with $k \leqslant n$ inversions in S_n in [Knu73, p.16]. See also [Sta12, Ex. 1.124] and [OEI23, A008302], and the application to standard Young tableaux in [BKS20b]. Here we use

$$\binom{x}{k} := \frac{x(x-1)\cdots(x-k+1)}{k!}$$

for all $k \in \mathbb{Z}_{\geqslant 0}$ and $x \in \mathbb{Z}$, including $x < 0$. For all $x \in \mathbb{Z}$, define the empty product $\binom{x}{0} = 1$.

Lemma 23. *Assume* $f(q) = \frac{\prod_{k=1}^{m}[a_k]_q}{\prod_{k=1}^{m}[b_k]_q} = \sum_k c_k q^k \in \mathbb{Z}[[q]]$ *for multisets of positive integers* $\{a_1, \ldots, a_m\}$ *and* $\{b_1, \ldots, b_m\}$. *Set*

$$M_i \coloneqq \#\{k : b_k = i\} - \#\{k : a_k = i\}.$$

Then, for every k, *the coefficient* c_k *is a polynomial in* $M_1, \ldots, M_k$ *given by*

$$c_k = \sum_{|\mu|=k} \prod_{\substack{i \geqslant 1 \\ m_i(\mu)>0}} \binom{M_i + m_i(\mu) - 1}{m_i(\mu)} \tag{3.3}$$

where $m_i(\mu)$ *is the number of parts of* μ *of size* i. *Moreover, we may restrict the sum in* (3.3) *to only those* $|\mu| = k$ *where for all* $i \geqslant 1$, *either* $M_i > 0$ *or* $m_i(\mu) \leqslant |M_i|$.

Proof. By definition, we have after cancellation

$$f(q) = \prod_{M_i \neq 0} (1 - q^i)^{-M_i}.$$

Equation (3.3) follows using the expansion $(1 - q^i)^{-j} = \sum_{n=0}^{\infty} \binom{j+n-1}{n} q^{in}$, multiplication of ordinary generating functions, and the fact that $\binom{x}{0} = 1$. For the restriction, consider a term indexed by μ with $M_i \leqslant 0$ and $m_i(\mu) > |M_i| = -M_i$. It follows that $M_i \leqslant 0 \leqslant M_i + m_i(\mu) - 1$, so that

$$m_i(\mu)! \binom{M_i + m_i(\mu) - 1}{m_i(\mu)} = (M_i + m_i(\mu) - 1) \cdots (M_i + 2)(M_i + 1)M_i = 0. \qquad \square$$

Remark 24. The binomial coefficients $\binom{M_i + m_i(\mu) - 1}{m_i(\mu)}$ appearing in (3.3) are positive when $M_i > 0$. When $M_i \leqslant 0$ and $m_i(\mu) \leqslant |M_i|$, the binomial coefficient is nonzero and has sign $(-1)^{m_i(\mu)}$. Thus the negative summands in (3.3) are precisely those for which μ has an odd number of row lengths i such that $M_i \leqslant 0$.

Remark 25. In the particular case when $f(q)$ is a basic cyclotomic generating function, the fact that the expression in (3.3) is nonnegative, symmetric, and eventually 0 is quite remarkable. Is there a simplification of the expression in (3.3) that would directly imply nonnegativity of the coefficients for any $f(q) \in \Phi^+$? Even the case of $[n]_q!$ or any of the examples CGFs mentioned in the Introduction would be of interest.

3.3 Necessary conditions for $q$integer ratios to be polynomial

The most obvious necessary and sufficient condition for rational products of q-integers of the form

$$f(q) = \frac{\prod_{k=1}^{m}[a_k]_q}{\prod_{k=1}^{m}[b_k]_q} \tag{3.4}$$

to yield a polynomial (not necessarily with positive coefficients) is for the multiplicity of each primitive dth root of unity to be nonnegative. However, both numerator and denominator can be factored uniquely into cyclotomic polynomials, so the **polynomiality criterion** is equivalent to

$$\#\{k : \ell \mid a_k\} \geqslant \#\{k : \ell \mid b_k\} \qquad \forall\, \ell \in \mathbb{Z}_{\geqslant 2}. \tag{3.5}$$

Example 26. Recall from the introduction that Stanley's q-analogue of the hook length formula

$$q^{b(\lambda)} \frac{[n]_q!}{\prod_{c\in\lambda}[h_c]_q} = \sum_{T\in \mathrm{SYT}(\lambda)} q^{\mathrm{maj}(T)},$$

gives an important family of cyclotomic generating functions. In this case, the inequality (3.5) reduces to the fact that the ℓ-quotient of any partition λ with $|\lambda| = n$ has no more than $\lfloor n/\ell \rfloor$ cells. In particular, this gives a quick proof that $[n]_q!/\prod_{c\in\lambda}[h_c]_q \in \mathbb{Z}[q]$, though the stronger result that the coefficients are nonnegative is not clear from this approach. See [Sta79, Prop. 4.11] and [Sta99, Cor. 21.5] for more details and proofs for this equality. [BKS20a] for

3.4 Necessary conditions for a *q*integer ratio to be a CGF

In addition to satisfying the polynomiality conditions in (3.5) and (3.7), what further restrictions on $\{a_1, \ldots, a_m\}$ and $\{b_1, \ldots, b_m\}$ are required for a rational product of q-integers as in (3.4) to yield a cyclotomic generating function? Say $f(q) = \prod_{k=1}^m [a_k]_q / \prod_{k=1}^m [b_k]_q = \sum c_k q^k$. Using the expression for c_k in (3.3), $c_k \geqslant 0$ and $\lim_{k\to\infty} c_k = 0$ are by definition necessary and sufficient to prove $f(q) \in \Phi^+$, though these conditions are difficult to use for families of such rational products in practice. We will outline some more direct necessary conditions here. We begin with complete classifications for $m = 1, 2$.

Lemma 27. *A rational product of q-integers $[a]_q/[b]_q$ is a CGF if and only if $b \mid a$.*

Proof. If the ratio is a CGF, it must evaluate to a positive integer at $q = 1$, so $b \mid a$ is necessary. It is also sufficient since when $b \mid a$, we have $[a]_q/[b]_q = [a/b]_{q^b}$, which is a CGF. ☐

Lemma 28. *Consider a rational product of q-integers $f(q) = [a_1]_q[a_2]_q/[b_1]_q[b_2]_q$.*

1. *Then, $f(q)$ is a polynomial if and only if*
   - *(i) $b_1 \mid a_1$ and $b_2 \mid a_2$; or*
   - *(ii) $b_1 \mid a_2$ and $b_2 \mid a_1$; or*
   - *(iii) $b_1, b_2 \mid a_1$ and $\gcd(b_1, b_2) \mid a_2$; or*
   - *(iv) $b_1, b_2 \mid a_2$ and $\gcd(b_1, b_2) \mid a_1$.*
2. *If $f(q)$ is a power series with nonnegative coefficients, then $a_1, a_2 \in \mathrm{Span}_{\mathbb{Z}_{\geqslant 0}}\{b_1, b_2\}$.*

3. Moreover, $f(q) \in \Phi^+$ if and only if both the divisibility and span conditions hold.

Proof. First suppose $f(q)$ is a polynomial. If either $b_i = 1$, then $[1]_q = 1$ and the conditions follow easily by (2.3), so assume each $b_i > 1$. By (2.3), $\Phi_{b_i}(q)$ divides $[b_i]$. Since $\Phi_{b_i}(q)$ appears in the denominator of $f(q)$, it must be canceled in the numerator, forcing $b_i \mid a_1$ or a_2 for $i = 1, 2$. If we do not have cases (i) or (ii), we have the first clause of (iii) or (iv). Without loss of generality suppose $b_1, b_2 \mid a_1$ and $\gcd(b_1, b_2) > 1$. Then, $\Phi_d(q)$ for $d = \gcd(b_1, b_2)$ appears twice in the denominator, and so must divide both $[a_1]_q$ and $[a_2]_q$ in order for $f(q)$ to be polynomial, giving $d \mid a_2$ by (2.3) again.

Conversely, suppose one of (i)-(iv) holds. Then, $f(q)$ is a polynomial if conditions (i) or (ii) hold by Lemma 27, and (iv) is equivalent to (iii), so it suffices to assume (iii) holds. Use the cyclotomic expansion of $[n]_q$ again from (2.3). Since (iii) holds, every cyclotomic divisor of either $[b_1]_q$ or $[b_2]_q$ but not both is canceled by a cyclotomic divisor of $[a_1]_q$, and every divisor of both $[b_1]_q$ and $[b_2]_q$ is canceled by divisors of $[a_1]_q$ and $[a_2]_q$ together. Thus, (1) holds.

The span condition in statement (2) is proved for all basic CGFs in Lemma 30 below. To prove (3), observe that $f(q) \in \Phi^+$ implies both the divisibility and span conditions hold by (1) and (2). Conversely, if any one of the divisibility conditions hold, then $f(q)$ is a polynomial with integer coefficients. If divisibility conditions (i) or (ii) hold, then $f(q) \in \Phi^+$ by Lemma 27, so suppose divisibility condition (iii) and the span condition in (2) holds. Again the case (iv) is similar.

By canceling out common cyclotomic factors, we may suppose without loss of generality that $\gcd(b_1, b_2) = 1$ and that $a_1 = b_1 b_2$. Since $a_2 \in \mathrm{Span}_{\mathbb{Z}_{\geqslant 0}}\{b_1, b_2\}$, there exist positive integers u, v such that $a_2 = ub_1 + vb_2$. So, we must show that the polynomial $f(q) = [b_1 b_2]_q [ub_1 + vb_2]_q / [b_1]_q [b_2]_q$ has nonnegative coefficients.

Given the assumptions above, $f(q)$ may be expressed as

$$f(q) = (1 - q^{b_1 b_2})(1 - q^{ub_1 + vb_2})(1 + q^{b_1} + q^{2b_1} + \cdots)(1 + q^{b_2} + q^{2b_2} + \cdots).$$

Let $N(n) = \#\{x, y \in \mathbb{Z}_{\geqslant 0} : xb_1 + yb_2 = n\}$. The coefficient of q^n in $f(q)$ is hence

$$N(n) - N(n - b_1 b_2) - N(n - ub_1 - vb_2) + N(n - b_1 b_2 - ub_1 - vb_2).$$

Let $N(n; \alpha, \beta) = \#\{\alpha \leqslant x, \beta \leqslant y : xb_1 + yb_2 = n\}$. The previous expression becomes

$$N(n; 0, 0) - N(n; b_2, 0) - N(n; u, v) + N(n; u + b_2, v).$$

Furthermore, let $N(n; \alpha, \beta; \delta) = \#\{\alpha \leqslant x < \alpha + \delta, \beta \leqslant y : xb_1 + yb_2 = n\}$. Thus, the expression for the coefficient of q^n in $f(q)$ becomes

$$N(n; 0, 0; b_2) - N(n; u, v; b_2).$$

Assume there exist $(x_0, y_0) \in \mathbb{Z}^2$ such that $x_0 b_1 + y_0 b_2 = n$, since otherwise the coefficient of q^n in $f(q)$ is 0. Then, all integer solutions of $xb_1 + yb_2 = n$ are of the form $x = x_0 - tb_2, y = y_0 + tb_1$ for some $t \in \mathbb{Z}$ by the theory of linear Diophantine equations.

In particular, there exist unique solutions $(x_1, y_1), (x_2, y_2) \in \mathbb{Z}^2$ with $0 \leqslant x_1 < b_2$ and $u \leqslant x_2 < u + b_2$. Hence, $N(n; 0, 0; b_2) = \delta_{y_1 \geqslant 0}$ and $N(n; u, v; b_2) = \delta_{y_2 \geqslant v}$. Since $x_1 \leqslant x_2$, we have $y_1 \geqslant y_2$ as solutions to the linear equation. Thus, $\delta_{y_2 \geqslant v} = 1 \Rightarrow \delta_{y_1 \geqslant 0} = 1$. Therefore, $N(n; 0, 0; b_2) - N(n; u, v; b_2)$ is nonnegative, which completes the proof. □

Example 29. For $m = 2$, the polynomiality condition alone does not imply nonnegative integer coefficients. For example, $[1]_q[6]_q/[2]_q[3]_q = \Phi_6(q) = q^2 - q + 1$. Here $6 \in \mathrm{Span}_{\mathbb{Z}_{\geqslant 0}}\{2, 3\}$, but 1 is not. For $m = 3$ and higher, the polynomiality and span conditions do not imply nonnegative integer coefficients. For example,

$$\frac{[4]_q[4]_q[15]_q}{[2]_q[3]_q[5]_q} = q^{13} + q^{11} + q^{10} - q^9 + 2q^8 + 2q^5 - q^4 + q^3 + q^2 + 1. \tag{3.6}$$

However, we do have the following necessary conditions for basic CGFs related to spans and sums.

In the next lemma, we will use the big Theta notation for asymptoticly bounding a function. If $f(n)$ is a function of n, we write $f = \Theta(g)$ provided there exist constants c, d, a function $g(n)$, and a positive integer n_0 such that for all $n \geqslant n_0$ we have $cg(n) \leqslant f(n) \leqslant dg(n)$. Thus, the function g is a reasonable approximation to f for large n. We will extend the big Theta notation to other families of functions not necessarily indexed by positive integers if such constants exist.

Lemma 30. *If $f(q) = \prod_{k=1}^{m} [a_k]_q/[b_k]_q \in \Phi^+$, then the following hold.*

(i) For each $1 \leqslant k \leqslant m$, $a_k \in \mathrm{Span}_{\mathbb{Z}_{\geqslant 0}}\{b_1, \ldots, b_m\}$.

(ii) If $a_1 \leqslant \cdots \leqslant a_m$ and $b_1 \leqslant \cdots \leqslant b_m$, then $a_1 \geqslant b_1 \geqslant 1$ and $a_m \geqslant b_m$.

(iii) Let $\mathcal{X}$ be the random variable corresponding to $f(q)/f(1)$ with mean $\mu = \frac{1}{2}\sum_{k=1}^{m}(a_k - b_k)$ and variance $\sigma^2 = \frac{1}{6}\sum_{k=1}^{m}(a_k^2 - b_k^2)$. Then, $\log \sigma = \Theta(\log \mu)$. More precisely,

$$\mu/2 \leqslant \sigma^2 \leqslant \mu^2.$$

(iv) We have

$$\frac{\sum_{i=1}^{m}(a_i^4 - b_i^4)}{\left(\sum_{i=1}^{m}(a_i^2 - b_i^2)\right)^2} \leqslant \frac{5}{3}.$$

(v) For all positive even integers d and $d = 1$, we have

$$\sum_{k=1}^{m} a_k^d \geqslant \sum_{k=1}^{m} b_k^d, \tag{3.7}$$

where the inequality is strict if the polynomial is non-constant. For example, when $d = 1$ this inequality is equivalent to noting the degree of $f(q)$ is $\sum_{k=1}^{m} a_k - b_k \geqslant 0$.

Proof. (i) From the definition of $f(q)$, we have

$$\prod_{k=1}^{m} 1/(1-q^{b_k}) = f(q) \prod_{k=1}^{m} 1/(1-q^{a_k}).$$

The support of the power series $\prod_{k=1}^{m} 1/(1-q^{b_k})$ is $\mathrm{Span}_{\mathbb{Z}_{\geqslant 0}}\{b_1, \ldots, b_m\}$. Since $f(q)$ has nonnegative coefficients and constant term 1, comparing supports gives

$$\mathrm{Span}_{\mathbb{Z}_{\geqslant 0}}\{a_1, \ldots, a_m\} \subset \mathrm{Span}_{\mathbb{Z}_{\geqslant 0}}\{b_1, \ldots, b_m\},$$

which yields (i).

(ii) From (i), it follows that $a_1 \geqslant \min\{b_1, \ldots, b_m\} = b_1$. From (3.5) at $\ell = b_m$, there is some a_k such that $b_m \mid a_k \leqslant a_m$, so $b_m \leqslant a_m$.

(iii) The argument in [HZ15, Lem. 2.5] applies here.

(iv) By definition, the *kurtosis* of $\mathcal{X}$ is $\mathbb{E}[(\frac{X-\mu}{\sigma})^4]$. Jensen's Inequality for convex functions gives the inequality of central moments

$$\alpha_2^2 = \mathbb{E}[(\mathcal{X}-\mu)^2]^2 \leqslant \mathbb{E}[(\mathcal{X}-\mu)^4] = \alpha_4.$$

In terms of cumulants, this says $\kappa_2^2 \leqslant \kappa_4 + 3\kappa_2^2$, which simplifies to the stated expression by (2.5).

(v) The inequality follows directly from the formula for cumulants of a rational product of q-integers (2.5) and the sign constraints in Corollary 17. $\square$

Remark 31. Warnaar–Zudilin [WZ11, Conj. 1] conjecture that a family of inequalities similar to (3.5) for ratios of q-factorials implies positivity. They prove the conjecture for q-super Catalan numbers.

Remark 32. One might wonder if $f(q) = \prod_{k=1}^{m} [a_k]_q/[b_k]_q \in \Phi^+$ implies $a_k \geqslant b_k$ for all k when sorted as in Lemma 30(ii). However, this is **false**, though empirically counterexamples are rare. For instance,

$$\frac{[2]_q[3]_q[3]_q[8]_q[12]_q}{[1]_q[1]_q[4]_q[4]_q[6]_q} = q^{12} + 2q^{11} + 2q^{10} + 2q^7 + 4q^6 + 2q^5 + 2q^2 + 2q + 1$$

is one of only three counterexamples with 5 q-integers in the numerator and denominator and maximum entry $\leqslant 14$, out of 956719 cyclotomic generating functions satisfying these conditions. On the other hand, we propose the following.

Conjecture 33. Suppose $f(q) = \prod_{k=1}^{m} [a_k]_q/[b_k]_q \in \Phi^+$, $a_1 \leqslant \cdots \leqslant a_m$, $b_1 \leqslant \cdots \leqslant b_m$. Then $\sum_{k=1}^{\ell} a_k \geqslant \sum_{k=1}^{\ell} b_k$ and $\sum_{k=\ell}^{m} a_k \geqslant \sum_{k=\ell}^{m} b_k$ for all ℓ. That is, $\{a_1, \ldots, a_m\}$ (weakly) **majorizes** $\{b_1, \ldots, b_m\}$ from both sides.

The majorization inequalities hold for $\ell = 1$ by Lemma 30(ii) and $\ell = m$ by (3.7). These majorization inequalities have also been checked for multisets with 7 elements and largest entry at most 15. Out of the $\binom{15+7-1}{7}^2 = 13521038400$ pairs of multisets, 70653669 or roughly 0.5% yield cyclotomic generating functions. Of these, 2713 do not satisfy $a_k \geqslant b_k$ for all k, or roughly 0.004% of the cyclotomic generating functions. Each of these nonetheless satisfy the majorization inequalities in Conjecture 33. Moreover, Conjecture 33 holds for all f of degree $\leqslant 42$, of which there are 10439036. Finally, the majorization condition is preserved under multiplication.

Remark 34. The affirmative answer to Conjecture 33 together with Karamata's inequality would give $\sum_{k=1}^{\ell} \psi(a_k) \geqslant \sum_{k=1}^{\ell} \psi(b_k)$ and $\sum_{k=\ell}^{m} \psi(a_k) \geqslant \sum_{k=\ell}^{m} \psi(b_k)$ for every nondecreasing convex function $\psi\colon [1,\infty) \to \mathbb{R}$, in particular strengthening (3.7) and leading to more necessary conditions for CGFs.

3.5 CGFs with a fixed denominator

In a different direction, one may fix the denominator $\prod_{k=1}^{m}[b_k]_q$ and consider which numerators $\prod_{k=1}^{m}[a_k]_q$ yield cyclotomic generating functions. Fixing the multiset $B = \{b_1, \ldots, b_m\}$, let G_B be the graph whose vertices are multisets $\{a_1, \ldots, a_m\}$ for which $\prod_{k=1}^{m}[a_k]_q/[b_k]_q \in \Phi^+$ and two vertices are connected by an edge if their multisets differ by a single element.

Lemma 35. *Fix a multiset $B = \{b_1, \ldots, b_m\}$ of positive integers. Suppose $\{a_1, \ldots, a_m\}$ and $\{a'_1, \ldots, a'_m\}$ are multisets of positive integers such that both $\prod_{k=1}^{m}[a_k]_q/[b_k]_q$ and $\prod_{k=1}^{m}[a'_k]_q/[b_k]_q \in \Phi^+$ are cyclotomic generating functions. Then there exists a multiset $\{h_1, \ldots, h_m\}$ of positive integers such that for all $1 \leqslant i \leqslant m$,*

$$\frac{\prod_{k=1}^{i}[h_k]_q \prod_{k=i+1}^{m}[a_k]_q}{\prod_{k=1}^{m}[b_k]_q} \in \Phi^+ \qquad \textit{and} \qquad \frac{\prod_{k=1}^{i}[h_k]_q \prod_{k=i+1}^{m}[a'_k]_q}{\prod_{k=1}^{m}[b_k]_q} \in \Phi^+.$$

Proof. By long division, one may observe $[xy]_q/[y]_q = [x]_{q^y} \in \Phi^+$. So, we may replace any $[a_k]_q$ in the numerator of a cyclotomic generating function with $[\ell a_k]_q$ for any positive integer ℓ and get another cyclotomic generating function. Hence, defining $h_k = a_k a'_k$ for $1 \leqslant k \leqslant m$ has the desired property. $\square$

Corollary 36. *The graph G_B is nonempty, connected and has diameter at most $2m$.*

4 Asymptotic considerations

Given a sequence of cyclotomic generating functions and their corresponding random variables, we can put their distributions in one common frame of reference by standardizing them to have mean 0 and standard deviation 1. Under what conditions does such a sequence converge? Can we classify all possible standardized limiting distributions of random variables coming from CGFs? Both problems have been completely solved for certain families of CGFs coming from q-hook length formulas, but the complete classification is not known, see [BS22, Open Problem 1.19]. In several "generic" regimes, they are asymptotically normal [BKS20a, Thm. 1.7], [BS22, Thm. 1.13]. In mildly degenerate regimes, they are related to independent sums of uniform distributions [BS22, Thm. 1.8]. Our aim in this section is to give another "generic" asymptotic normality criterion for sequences of CGFs, which is sufficient to quickly identify the limit in many cases of interest.

Throughout the rest of the section, we will use the following notation. Let $a^{(N)}$ and $b^{(N)}$ for $N = 1, 2, \ldots$ denote two sequences of multisets of positive integers of the same size. By Lemma 30(ii), we can always assume the values in $a^{(N)}$ are strictly greater than 1. The multisets $b^{(N)}$ may contain 1's in order to have the same size as $a^{(N)}$. Let $f_N(q) = \prod_{a \in a^{(N)}}[a]_q / \prod_{b \in b^{(N)}}[b]_q$, and let $\mathcal{X}_N$ be the random variable associated to $\mathbb{E}[q^{\mathcal{X}_N}] = f_N(q)/f_N(1)$. We will denote the dth moment and cumulant of $\mathcal{X}_N$ by $\mu_d^{(N)}$ and $\kappa_d^{(N)}$ for $N = 1, 2, \ldots$ respectively. Similarly, the standard deviation and mean of $\mathcal{X}_N$ are denoted by $\sigma^{(N)}$ and $\mu^{(N)}$. Recall the notation for uniform sum distributions $\mathcal{S}_{\mathbf{t}}$ and p-norms from Section 2.4.

Theorem 37. *Let $a^{(N)}$ be a sequence of multisets of positive integers where $a \geqslant 2$ for each $a \in a^{(N)}$. Let $f_N(q) = \prod_{a \in a^{(N)}}[a]_q$ and $\mathcal{X}_N$ be the associated sequence of CGF polynomials and random variables.*

(i) The sequence of random variables $\mathcal{X}_1, \mathcal{X}_2, \ldots$ is asymptotically normal if

$$\sum_{a \in a^{(N)}} \left(\frac{a}{\max a^{(N)}} \right)^4 \to \infty. \tag{4.1}$$

(ii) Suppose the cardinality of the multisets $a^{(N)}$ is bounded by some m and $\max a^{(N)} \to \infty$. Then $\mathcal{X}_1^*, \mathcal{X}_2^*, \ldots$ converges in distribution if and only if the rescaled multisets $a^{(N)}/|a^{(N)}|_2 \in \widetilde{\ell}_{\leqslant m}$ converge pointwise to some multiset $\mathbf{t} \in \widetilde{\ell}_{\leqslant m}$ in the sense of [BS22, §3.2]. In that case, the limiting distribution is the standardized uniform sum $\mathcal{S}_{\mathbf{t}}^*$.

(iii) Suppose the cardinality of $a^{(N)}$ is bounded and $\max a^{(N)}$ is also bounded. Then $\mathcal{X}_1^*, \mathcal{X}_2^*, \ldots$ converges in distribution if and only if $a^{(N)}$ is eventually a constant multiset.

Proof. The polynomial $f_N(q) = \prod_{a \in a^{(N)}}[a]_q$ can also be expressed in rational form as $= \prod_{a \in a^{(N)}}[a]_q/[1]_q$. Thus, we obtain the corresponding sequence of multisets $b^{(N)} = (1, 1, \ldots, 1)$ from the denominators where the size of $b^{(N)}$ equals the size of $a^{(N)}$ as multisets. By Theorem 9 and the scaling property of cumulants, the standardized cumulants of $\mathcal{X}_N$ are given by

$$(\kappa_d^{(N)})^* = \frac{\kappa_d^{(N)}}{(\sigma^{(N)})^{d/2}} = \frac{\frac{B_d}{d} \sum_{a \in a^{(N)}}(a^d - 1)}{\left(\frac{B_2}{2} \sum_{a \in a^{(N)}}(a^2 - 1)\right)^{d/2}}. \tag{4.2}$$

Since $a \geqslant 2$ for all $a \in a^{(N)}$, $\frac{1}{2}a^d \leqslant a^d - 1 \leqslant a^d$, so $a^d - 1 = \Theta(a^d)$ uniformly. Hence for $d \geqslant 2$ even,

$$(\kappa_d^{(N)})^* = \Theta\left(\frac{\sum_{a \in a^{(N)}} a^d}{(\sum_{a \in a^{(N)}} a^2)^{d/2}} \right) = \Theta\left(\frac{|a^{(N)}|_d^d}{|a^{(N)}|_2^d} \right) = \Theta\left(\frac{|a^{(N)}|_d}{|a^{(N)}|_2} \right)^d.$$

Therefore by Theorem 15, asymptotic normality is equivalent to $|a^{(N)}|_4/|a^{(N)}|_2 \to 0$, or equivalently

$$\frac{|a^{(N)}|_2}{|a^{(N)}|_4} \to \infty.$$

By [BS22, (3.7)],

$$|a^{(N)}|_4 \leqslant |a^{(N)}|_\infty^{1/2}\ |a^{(N)}|_2^{1/2}.$$

Hence, squaring both sides and rearranging the factors we have

$$\frac{|a^{(N)}|_4}{|a^{(N)}|_\infty} \leqslant \frac{|a^{(N)}|_2}{|a^{(N)}|_4}.$$

Thus, $|a^{(N)}|_4/|a^{(N)}|_\infty \to \infty$ implies asymptotic normality. By Definition 18

$$\left(|a^{(N)}|_4/|a^{(N)}|_\infty\right)^4 = \sum_{a \in a^{(N)}} \left(\frac{a}{\max a^{(N)}}\right)^4$$

so (i) holds.

For (ii), let $\hat{a}^{(N)} := \sqrt{12} \cdot a^{(N)}/|a^{(N)}|_2$ be the rescaled multiset such that $\mathcal{S}_{\hat{a}^{(N)}}$ is standardized for each N according to (2.17). Since $|a^{(N)}|_2 \geqslant |a^{(N)}|_\infty := \max a^{(N)} \to \infty$, we have by (4.2) that

$$\begin{aligned}
(\kappa_d^{(N)})^* &= \frac{\frac{B_d}{d}\sum_{a\in a^{(N)}}(a^d-1)}{\left(\frac{B_2}{2}\sum_{a\in a^{(N)}}(a^2-1)\right)^{d/2}} \\
&\sim \frac{\frac{B_d}{d}\sum_{a\in a^{(N)}}(a/|a^{(N)}|_2)^d}{\left(\frac{B_2}{2}\sum_{a\in a^{(N)}}(a/|a^{(N)}|_2)^2\right)^{d/2}} = \frac{\frac{B_d}{d}}{\left(\frac{B_2}{2}\right)^{d/2}}\frac{|\hat{a}^{(N)}|_d^d}{|\hat{a}^{(N)}|_2^{d/2}}.
\end{aligned}$$

Here and elsewhere we write $f(N) \sim g(N)$ to mean $\lim_{N\to\infty} f(N)/g(N) = 1$. This last expression is the dth cumulant of $\mathcal{S}_{\hat{a}^{(N)}}$ by (2.17). Therefore, by the cumulant version of Theorem 13 and Theorem 40, we know $\mathcal{X}_1^*, \mathcal{X}_2^*, \dots$ converges in distribution if and only if the sequence $\mathcal{S}_{\hat{a}^{(N)}}$ converges in distribution, and to the same limit. By [BS22, Thm. 1.15], $\mathcal{S}_{\hat{a}^{(N)}}$ converges in distribution if and only if the multisets $\hat{a}^{(N)}$ converge pointwise in the sense of [BS22, §3] to some $\mathbf{t} \in \tilde{\ell}_2$, with limiting distribution $\mathcal{S}_{\mathbf{t}} + \mathcal{N}(0, \sigma^2)$ where $\sigma = \sqrt{1 - |\mathbf{t}|_2^2/12}$. Since the cardinality of $a^{(N)}$ is bounded, $|\hat{a}^{(N)}|_2^2 = 12$ for each N, so $|\mathbf{t}|_2^2 = \lim_{N\to\infty} |a^{(N)}|_2^2 = 12$. This implies $\sigma = 0$ by definition, so the limiting distribution of the standardized random variables $\mathcal{X}_N^*$ is just the standardized generalized uniform sum $\mathcal{S}_{\mathbf{t}} = \mathcal{S}_{\mathbf{t}}^*$ as given in the statement.

For (iii), if both the cardinality of $a^{(N)}$ and $|a^{(N)}|_\infty$ are bounded, there are only finitely many possible choices for $a^{(N)}$. Hence, convergence in distribution happens if and only if the sequence $a^{(N)}$ is eventually constant. $\square$

Building on the case of finite products of q-integers above, we can now address the characterization of asymptotic normality for basic CGFs stated in the introduction. Intuitively, the idea is that if the numerator of the rational form of the CGF leads to asymptotic normality, then the rational form does as well since it is dominated by its numerator.

Proof of Theorem 5. Observe that for any positive integer n, the polynomial $[n]_q/n$ is the probability generating function of a discrete uniform random variable $\mathcal{U}_n$ supported on $\{0, 1, \ldots, n-1\}$. Let $\mathcal{A}_N = \sum_{a\in a^{(N)}} \mathcal{U}_a$ and $\mathcal{B}_N = \sum_{b\in b^{(N)}} \mathcal{U}_b$ denote the CGF distributions corresponding to the numerator and denominator of $f_N(q) = \prod_{a\in a^{(N)}}[a]_q / \prod_{b\in b^{(N)}}[b]_q$. By Theorem 9, $\sigma^2_{\mathcal{A}_N} = \frac{1}{12}\sum_{a\in a^{(N)}}(a^2-1)$ and $\sigma^2_{\mathcal{B}_N} = \frac{1}{12}\sum_{b\in b^{(N)}}(b^2-1)$.

Let $c_N := \sigma_{\mathcal{B}_N}/\sigma_{\mathcal{A}_N}$, so each $c_N \geqslant 0$. By the hypothesis in (1.5),

$$\limsup_{N\to\infty} c_N^2 = \limsup_{N\to\infty} \frac{\sum_{b\in b^{(N)}}(b^2-1)}{\sum_{a\in a^{(N)}}(a^2-1)} < 1.$$

Recall that convergence in distribution of real-valued random variables can be metrized using the Lévy metric. Therefore, it suffices to show that every subsequence of $\mathcal{X}_1, \mathcal{X}_2, \ldots$ itself has an asymptotically normal subsequence. Hence, without loss of generality, we may assume that the sequence c_N converges to some $0 \leqslant c < 1$.

By construction, we have the following equality in distribution:

$$\mathcal{X}_N + \mathcal{B}_N = \mathcal{A}_N, \tag{4.3}$$

where all summands are independent. It follows from independence that

$$\sigma^2_{\mathcal{X}_N} = \sigma^2_{\mathcal{A}_N} - \sigma^2_{\mathcal{B}_N}.$$

By standardizing random variables and simplifying (4.3), we have

$$\mathcal{X}_N^* \sqrt{1-c_N^2} + \mathcal{B}_N^*\, c_N = \mathcal{A}_N^*. \tag{4.4}$$

In terms of characteristic functions, (4.4) gives

$$\phi_{\mathcal{X}_N^*}(t) = \frac{\phi_{\mathcal{A}_N^*}\left(t\,/\sqrt{1-c_N^2}\right)}{\phi_{\mathcal{B}_N^*}\left(t\,c_N/\sqrt{1-c_N^2}\right)}. \tag{4.5}$$

Since the hypothesis in (1.6) implies (4.1) holds, Theorem 37 says the sequence $\mathcal{A}_N$ is asymptotically normal, so by Lévy's continuity Theorem 12,

$$\lim_{N\to\infty} \phi_{\mathcal{A}_N^*}(t) = \exp(-t^2/2)$$

for all $t \in \mathbb{R}$. This convergence is in fact uniform on bounded subsets of $\mathbb{R}$ (see e.g. [Bil95, Exercise 26.15(b)]). Consider the cases for $0 \leqslant c < 1$.

(i) Suppose $c > 0$. By Lemma 30(ii), we have $|b^{(N)}|_\infty \leqslant |a^{(N)}|_\infty$. This observation and the fact that $c > 0$ imply that (1.6) holds with $a^{(N)}$ replaced by $b^{(N)}$. Hence

$$\lim_{N\to\infty} \phi_{\mathcal{B}_N^*}(t) = \exp(-t^2/2).$$

Since $c < 1$ and convergence is uniform on bounded subsets, (4.5) gives

$$\lim_{N\to\infty} \phi_{\mathcal{X}_N^*}(t) = \frac{\exp(-t^2/2(1-c^2))}{\exp(-t^2c^2/2(1-c^2))} = \exp(-t^2/2).$$

Therefore, by Lévy's continuity theorem again, $\mathcal{X}_1, \mathcal{X}_2, \ldots$ is asymptotically normal.

(ii) Suppose $c = 0$. Note that the variance of $\mathcal{B}_N^* c_N$ is c_N^2 and the mean is 0. Hence, $c_N \to c = 0$ implies $\mathcal{B}_N^* c_N$ converges to the constant random variable 0 by Chebyshev's Inequality. Thus,

$$\lim_{N\to\infty} \phi_{\mathcal{B}_N^* c_N}(t) = 1,$$

and the result follows from (4.5) and the calculations from the previous case. $\square$

Example 38. Consider a sequence of q-binomial coefficients

$$f_N(q) = \binom{n}{k}_q = \prod_{j=1}^{k} \frac{[n-k+j]_q}{[j]_q}$$

where k, n represent sequences indexed by N. Assume both k and $n-k$ approach infinity as $N \to \infty$. In the notation of the proof above, we have $\mathcal{A}_N = \sum_{j=1}^{k} \mathcal{U}_{n-k+j}$, $\mathcal{B}_N = \sum_{j=1}^{k} \mathcal{U}_j$, and

$$c_N^2 = \frac{\sigma_{\mathcal{B}_N}^2}{\sigma_{\mathcal{A}_N}^2} = \frac{\sum_{j=1}^{k} j^2 - 1}{\sum_{j=n-k+1}^{n} j^2 - 1} \sim \frac{k^3}{n^3 - (n-k)^3} = \frac{(k/n)^3}{1 - (1-k/n)^3}.$$

Let $x = k/n$. Since both $k, n-k \to \infty$ as $N \to \infty$, we may suppose $k \leqslant n-k$ for all N since we can replace k with $n-k$ if necessary without changing $f_N(q)$, so $0 \leqslant x \leqslant 1/2$. It is easy to check that $x^3/(1-(1-x)^3) \leqslant 1/7$ for $0 \leqslant x \leqslant 1/2$, so that (1.5) holds. We also see

$$\begin{aligned}
\frac{\sum_{j=n-k+1}^{n} j^4}{n^4} &\sim \frac{n^5 - (n-k)^5}{n^4} \\
&= \frac{k(5n^4 - 10n^3k + 10n^2k^2 - 5nk^3 + k^4)}{n^4} \\
&= k(5 - 10x + 10x^2 - 5x^3 + x^4),
\end{aligned}$$

so (1.6) holds as $k \to \infty$ since $(5 - 10x + 10x^2 - 5x^3 + x^4)$ is positive for all real values in the range $0 \leqslant x \leqslant 1/2$. Thus, Theorem 5 applies to show the sequence of CGF random variables corresponding to $f_N(q)$ is asymptotically normal provided $k, n-k \to \infty$. The details of this calculation are left implicit in [Dia88].

5 Analytic considerations

In this section, we consider the problem of classifying all standardized CGF distribution via analytic considerations of their characteristic functions $\phi_{\mathcal{X}^*}(t)$ and their second characteristic functions, $\log \phi_{\mathcal{X}^*}(t)$. Given the many different characterizations of CGFs introduced in Theorem/Definition 1, we have a rich set of tools for studying these complex functions. In particular, we will show that the limiting standardized characteristic functions are entire, see Corollary 41. This allows us to complete the proof of Theorem 8.

We begin by spelling out the connections between the CGFs in rational and cyclotomic form with the first and second (standardized) characteristic functions.

Recall from Section 2.2 that for any CGF $f(q)$ with corresponding random variable $\mathcal{X}$, we have $\mathbb{E}[q^{\mathcal{X}}] = f(q)/f(1)$, so the characteristic function of $\mathcal{X}$ is

$$\phi_{\mathcal{X}}(t) \coloneqq \mathbb{E}[e^{it\mathcal{X}}] = f(e^{it})/f(1),$$

and the cumulant generating function of $\mathcal{X}$ is

$$\log \phi_{\mathcal{X}}(t) = \log \mathbb{E}[e^{it\mathcal{X}}] = \sum_{d=1}^{\infty} \kappa_d(f) \frac{(it)^d}{d!}.$$

Furthermore, $\kappa_d(f) = 0$ for $d > 1$ odd, $\mu = \kappa_1(f)$ is the mean of $\mathcal{X}$ and $\sigma^2 = \kappa_2(f)$ is its variance.

Note, both $f(q)$ and $\alpha q^{\beta} f(q)$ for positive integers α, β give rise to the same standardized random variable $\mathcal{X}^*$ so in order to study all standardized CGF distributions, it suffices to assume $f(q) \in \Phi^+$. Thus, if $f(q) = \Phi_{i_1}(q)\Phi_{i_2}(q)\cdots\Phi_{i_k}(q) = \prod_{j=1}^{k} \frac{[a_j]_q}{[b_j]_q}$, then we can extend the corresponding characteristic function to the complex plane by setting

$$\phi_{\mathcal{X}}(z) = f(e^{iz})/f(1) = \Phi_{i_1}(e^{iz})\Phi_{i_2}(e^{iz})\cdots\Phi_{i_k}(e^{iz})/f(1) = \prod_{j=1}^{k} \frac{[a_j]_{e^{iz}}}{[b_j]_{e^{iz}}} \tag{5.1}$$

and

$$\begin{aligned} \phi_{\mathcal{X}^*}(z) &= \mathbb{E}[e^{iz\mathcal{X}^*}] = \mathbb{E}[e^{iz(X-\mu)/\sigma}] = e^{-iz\mu/\sigma}\mathbb{E}[e^{izX/\sigma}] \\ &= e^{-iz\mu/\sigma} f(e^{iz/\sigma})/f(1) = e^{-iz\mu/\sigma}\phi_{\mathcal{X}}(z/\sigma). \end{aligned} \tag{5.2}$$

Furthermore, we have a convergent power series representation of the cumulant generating function which can be expressed in terms of the multisets in the rational form of $f(q)$ as

$$\begin{aligned} \log \phi_{\mathcal{X}^*}(z) &= \sum_{d=1}^{\infty} (-1)^d \kappa^*_{2d}(f) \frac{z^{2d}}{(2d)!} \\ &= \sum_{d=1}^{\infty} (-1)^d \left(\frac{B_{2d} \sum_{k=1}^{m} (a_k^{2d} - b_k^{2d})}{2d((B_2/2)\sum_{k=1}^{m}(a_k^2 - b_k^2))^d} \right) \frac{z^{2d}}{(2d)!} \end{aligned} \tag{5.3}$$

since $\kappa^*_{2d}(f) = \kappa_{2d}(f)/\kappa_2(f)^d$ and $\kappa_d(f)$ is given by (2.5) for all positive integers d. Furthermore, $(-1)^d \kappa^*_{2d}(f)$ is a non-positive real number by Corollary 17.

5.1 CGF characteristic functions

Consider the set of standardized characteristic functions of random variables associated to cyclotomic generating functions. As mentioned above, it suffices to consider only basic CGFs, so consider the set of all standardized CGF characteristic functions on the complex plane,

$$\mathcal{C}_{\mathrm{CGF}} \coloneqq \{\phi_{\mathcal{X}^*}(z) : \mathbb{E}[q^{\mathcal{X}}] = f(q)/f(1) \text{ for } f(q) \in \Phi^+\}.$$

We first show that $\mathcal{C}_{\mathrm{CGF}}$ is a **normal family** of continuous functions. This means that every infinite sequence $(\phi_{\mathcal{X}_N^*}(z) : N = 1, 2, \dots)$ in $\mathcal{C}_{\mathrm{CGF}}$ contains a subsequence which converges uniformly on compact subsets of $\mathbb{C}$.

Theorem 39. *The set $\mathcal{C}_{\mathrm{CGF}}$ is a normal family of entire functions.*

Proof. Given $\phi_{\mathcal{X}^*}(z) \in \mathcal{C}_{\mathrm{CGF}}$, let $f(q) \in \Phi^+$ be the associated CGF, and let $\mathcal{X}$ be the corresponding random variable with $\mathbb{E}[q^{\mathcal{X}}] = f(q)/f(1)$, mean μ and standard deviation σ. Let $\tilde{f}(q) = q^{-\mu} f(q)/f(1) \in \mathbb{R}_{\geqslant 0}[q^{\pm 1/2}]$. Then, by (5.2)

$$\phi_{\mathcal{X}^*}(z) = e^{-iz\mu/\sigma} f(e^{iz/\sigma})/f(1) = \tilde{f}(e^{iz/\sigma}). \tag{5.4}$$

Thus, since $f(q) \in \Phi^+$ is a finite product of cyclotomic polynomials, $\phi_{\mathcal{X}^*}(z)$ is a finite sum of products of exponential functions by (5.1) and (5.2), hence it is entire.

From Montel's theorem in complex analysis, $\mathcal{C}_{\mathrm{CGF}}$ is a normal family of entire functions if and only if it is bounded on all complex disks $|z| \leqslant R$. Hence we will bound $|\phi_{\mathcal{X}^*}(z)|$ in terms of R. Note that $f(q) = 1$ if and only if $\sigma = 0$ by Lemma 30(iii), in which case $\phi_{\mathcal{X}^*}(z) = 1$ is bounded by any function greater than 1. So, we will assume $f(q)$ is not constant and $\sigma > 0$. By [HZ15, Lem. 2.8], for all real $t \geqslant 0$,[1]

$$\mathbb{E}[e^{t\mathcal{X}^*}] \leqslant \exp\left(\frac{3}{2} t^2 e^{2t/\sigma}\right). \tag{5.5}$$

Since $\mathbb{E}[e^{z\mathcal{X}^*}] = \phi_{\mathcal{X}^*}(-iz)$, we can use this inequality to bound $|\phi_{\mathcal{X}^*}(z)|$ for all complex $|z| \leqslant R$ as follows.

For all $|z| \leqslant R$, we claim that $|\tilde{f}(e^{iz/\sigma})| \leqslant 2\tilde{f}(e^{R/\sigma})$. Indeed, since $f(q) \in \Phi^+$, we have $\tilde{f}(q) = \sum_{k=-N}^{N} a_k q^{k/2}$ for $a_k \geqslant 0$ satisfying $a_k = a_{-k}$. Therefore,

$$\begin{aligned}
\left|\tilde{f}(e^{iz/\sigma})\right| &= \left|\sum_{k=-N}^{N} a_k e^{izk/2\sigma}\right| \leqslant \sum_k a_k \left|e^{izk/2\sigma}\right| = \sum_k a_k e^{\mathrm{Re}(izk/2\sigma)} \\
&\leqslant \sum_k a_k e^{R|k|/2\sigma} \leqslant \sum_k a_k (e^{Rk/2\sigma} + e^{-Rk/2\sigma}) = \sum_k a_k e^{Rk/2\sigma} + \sum_k a_{-k} e^{-Rk/2\sigma} \\
&= 2\tilde{f}(e^{R/\sigma}).
\end{aligned}$$

By Lemma 30(iii), $\sigma^2 \geqslant 1/4$ is bounded away from 0 since we assumed $f(q)$ is not constant. Therefore, by the claim, (5.4), and (5.5), we have the required uniform bound

$$|\phi_{\mathcal{X}^*}(z)| = |\tilde{f}(e^{iz/\sigma})| \leqslant 2\tilde{f}(e^{R/\sigma}) = 2\phi_{\mathcal{X}^*}(-iR) = 2\mathbb{E}[e^{R\mathcal{X}^*}] \leqslant 2\exp\left(\frac{3}{2} R^2 e^{8R}\right). \qquad \square$$

Theorem 40 (Converse of Frechét–Shohat for CGFs)**.** *Suppose $\mathcal{X}_1, \mathcal{X}_2, \dots$ is a sequence of random variables corresponding to cyclotomic generating functions such that $\mathcal{X}_n^* \Rightarrow \mathcal{X}$ for some random variable $\mathcal{X}$. Then $\mathcal{X}$ is determined by its moments and, for all $d \in \mathbb{Z}_{\geqslant 1}$,*

$$\lim_{n\to\infty} \mu_d^{\mathcal{X}_n^*} = \mu_d^{\mathcal{X}}.$$

[1] [HZ15, Lem. 2.8] is incorrectly stated for all $t \in \mathbb{R}$, though one of the last steps in the argument requires $t \geqslant 0$.

Proof. By Theorem 39, we may pass to a subsequence and assume $\phi_{\mathcal{X}_n^*}(z)$ converges uniformly on compact subsets of $\mathbb{C}$. Hence, they converge to an entire function $\phi(z)$. On the other hand, by Lévy continuity, $\phi_{\mathcal{X}_n^*}(t) \to \phi_{\mathcal{X}}(t)$ pointwise for all $t \in \mathbb{R}$. Thus, $\phi_{\mathcal{X}}(t) = \phi(t)$ for $t \in \mathbb{R}$.

A priori, a characteristic function $\phi_{\mathcal{X}}(t) := \mathbb{E}[e^{it\mathcal{X}}]$ exists only for $t \in \mathbb{R}$. However, in this case we have an entire function $\phi(z)$ which coincides with $\phi_{\mathcal{X}}(z)$ for all $z \in \mathbb{R}$. By [Luk70, Thm. 7.1.1, pp.191-193], agreement on the real line suffices to show $\phi_{\mathcal{X}}(z)$ exists and agrees with $\phi(z)$ for all $z \in \mathbb{C}$. Since $\phi(z)$ is entire, it can be represented by a power series that converges everywhere in the complex plane. Hence, the moment-generating function $\mathbb{E}[e^{t\mathcal{X}}] = \phi_{\mathcal{X}}(-it) = \phi(-it)$ exists for all $t \in \mathbb{R}$ and it can be expressed as a convergent power series with finite coefficients. Hence $\mathcal{X}$ has moments of all orders and is determined by its moments by [Bil95, Thm. 30.1]. □

Corollary 41. *If $\mathcal{X}_1, \mathcal{X}_2, \ldots$ is a sequence of CGF random variables and $\mathcal{X}_n^* \Rightarrow \mathcal{X}$, then $\phi_{\mathcal{X}}(z)$ is entire.*

We may now prove Theorem 8 from the introduction.

Proof of Theorem 8. Suppose $\mathcal{X}_N^* \Rightarrow \mathcal{X}$ converges in distribution. Then the result follows by Theorem 40. Conversely, suppose $\lim_{N\to\infty} \mu_d^{\mathcal{X}_N^*}$ exists and is finite for all $d \in \mathbb{Z}_{\geqslant 1}$. By Theorem 39, we may pass to a subsequence on which $\phi_{\mathcal{X}_N^*}(z)$ converges uniformly on compact subsets of $\mathbb{C}$. The limiting entire function $\phi(z)$ then has power series coefficients determined by the $\mu_d := \lim_{N\to\infty} \mu_d^{\mathcal{X}_N^*}$, so the limit is independent of the subsequence we passed to and $\lim_{N\to\infty} \phi_{\mathcal{X}_N^*}(t) = \phi(t)$ for all $t \in \mathbb{R}$. Moreover, $\phi(t)$ is continuous at 0, so there exists $\mathcal{X}$ such that $\mathcal{X}_N^* \Rightarrow \mathcal{X}$ by [Bil95, Cor. 26.1]. □

5.2 Formal cumulants and cumulant generating functions

Recall formal cumulants from Section 2.3. The formal cumulants of cyclotomic polynomials have the following explicit growth rate.

Lemma 42. *For each fixed $d \geqslant 1$, the formal cumulants of the cyclotomic polynomials for $n > 1$ satisfy*

$$\frac{(2d)!}{d}\left(\frac{n}{2\pi}\right)^{2d} \leqslant |\kappa_{2d}| \leqslant \frac{|B_{2d}|n^{2d}}{2d}. \tag{5.6}$$

Proof. The upper bound in (5.6) is clear from the factored form of the Jordan totient formula in (2.15). For the lower bound, the Euler product for the Riemann zeta function is

$$\prod_{p \text{ prime}} (1 - p^{-s}) = \frac{1}{\zeta(s)}.$$

See [Wik23] for a sketch of the proof. Applying the expression for the zeta function at positive even integers gives

$$\frac{|\kappa_{2d}|}{n^{2d}} = \frac{|B_{2d}|}{2d} \prod_{\substack{p \text{ prime} \\ p|n}} (1 - p^{-2d})$$

$$\begin{aligned}
&\geqslant \frac{|B_{2d}|}{2d}\frac{1}{\zeta(2d)}\\
&= \frac{|B_{2d}|}{2d}\frac{2(2d)!}{(2\pi)^{2d}|B_{2d}|}\\
&= \frac{1}{d}\frac{(2d)!}{(2\pi)^{2d}}. \qquad \square
\end{aligned}$$

While cyclotomic generating functions are typically given in rational form, Lemma 16 allows their cumulants to be described in terms of the cyclotomic form as follows. One may for instance use Example 26 to describe the q-hook formula in cyclotomic form combinatorially.

Corollary 43. *If $f(q) = \prod_{j=1}^{k} \Phi_{n_j}(q)$, then for each $d \geqslant 1$ we have*

$$\kappa_{2d}(f(q)) = \sum_{j=1}^{k} \kappa_{2d}(\Phi_{n_j}(q)) = \sum_{j=1}^{k} \frac{B_{2d}}{2d} n_j^{2d} \prod_{\substack{p \text{ prime} \\ p \mid n_j}} \left(1 - \frac{1}{p^{2d}}\right)$$

and uniformly

$$|\kappa_{2d}(f(q))| = \Theta\left(\sum_{j=1}^{k} n_j^{2d}\right).$$

The asymptotic behavior of sequences of CGF random variables can be determined by their standardized cumulant generating function. For instance, $\mathcal{X}_1, \mathcal{X}_2, \dots$ is asymptotically normal if and only if the corresponding sequence of standardized cumulants $(\kappa_d^{(N)})^* \to 0$ for all $d \geqslant 3$ as $N \to \infty$, in which case by (5.3), $\log \phi_{\mathcal{X}_N^*}(t) \to -t^2/2$ for all $t \in \mathbb{R}$ as $N \to \infty$. Standardized second characteristic functions of CGFs are particularly well-behaved in the following sense.

Proposition 44. *Let $\mathcal{X}$ be a random variable corresponding to a cyclotomic generating function. Then the standardized cumulant generating function $\log \phi_{\mathcal{X}^*}(z)$ is analytic in the complex disk $|z| < \sqrt{2}$. Furthermore, for real values $-\sqrt{2} < t < \sqrt{2}$ we have*

$$\log \phi_{\mathcal{X}^*}(t) \leqslant -\frac{t^2}{2} = \log \phi_{\mathcal{N}(0,1)}(t). \tag{5.7}$$

Proof. By Theorem 39, $\phi_{\mathcal{X}}(z)$ is entire, so the singularities of $\log \phi_{\mathcal{X}}(z)$ come about from the zeros of $\phi_{\mathcal{X}}(z)$. To determine the zeros, it suffices to assume $\mathcal{X}$ is determined by $f(q) = \frac{\prod_{k=1}^{m}[a_k]_q}{\prod_{k=1}^{m}[b_k]_q} \in \Phi^+$ since the results pertain to the standardized random variable corresponding to $\mathcal{X}$. Furthermore, since $\phi_{\mathcal{X}}(z) = 1$ if $f(q) = 1$, which satisfies both claims, we may assume that $m = \max\{a_k\} > 1$ and m does not appear in the denominator multiset. Hence, $\Phi_m(q)$ is a factor of $f(q)$. Furthermore, from the product formula (2.14), we observe that the closest zero of $\phi_{\mathcal{X}}(z) = f(e^{iz})$ to the origin occurs at $z = 2\pi/m$. Hence $\phi_{\mathcal{X}}(z)$ is analytic on the simply connected domain $|z| < 2\pi/m$, so $\log \phi_{\mathcal{X}}(z)$ is analytic on the disk $|z| < 2\pi/m$.

By (5.2), $\phi_{\mathcal{X}^*}(z) = e^{-i\mu z/\sigma}\phi_{\mathcal{X}}(z/\sigma)$, where $\mu = \kappa_1(f)$ and $\sigma^2 = \kappa_2(f)$ are the mean and variance of $\mathcal{X}$. Therefore, to prove $\log \phi_{\mathcal{X}^*}(z)$ is analytic in the complex disk $|z| < \sqrt{2}$, it suffices to show that $|z| < \sqrt{2}$ implies $\left|\frac{z}{\sigma}\right| < 2\pi/m$.

Since $\Phi_m(q)$ is a factor of $f(q)$, we have $0 < \kappa_2(\Phi_m) \leqslant \sigma^2$ by Lemma 42 and Corollary 43. By the lower bound in (5.6), $\sqrt{\kappa_2(\Phi_m)} \geqslant \frac{m}{\pi\sqrt{2}}$. Hence if $|z| < \sqrt{2}$,

$$\left|\frac{z}{\sigma}\right| < \frac{\sqrt{2}}{m/\pi\sqrt{2}} = \frac{2\pi}{m}$$

as required.

To prove the "furthermore" statement, we use the expansion

$$\log \phi_{\mathcal{X}^*}(z) = -\sum_{d\geqslant 1} \frac{(-1)^{d-1}\kappa_{2d}}{(2d)!}\left(\frac{z}{\sigma}\right)^{2d} \tag{5.8}$$

to compute $\log \phi_{\mathcal{X}^*}(t)$ for real $-\sqrt{2} < t < \sqrt{2}$. The inequality (5.7) comes from truncating this power series expansion after the first term and applying Corollary 17 to see that all terms on the right side are nonpositive. □

6 CGF monoids and related open problems

As mentioned in Definition 3, the set of basic cyclotomic generating functions forms a monoid under multiplication, Φ^+. We also consider the larger monoid generated by all cyclotomic polynomials. Since many families of polynomials of interest are also either unimodal or log-concave with no internal zeros, we consider these submonoids as well, together with another variant using the Gale order on multisets. We conclude with a monoid associated to Hilbert series of polynomial rings quotiented by regular sequences and several open problems.

6.1 Basic, unimodal, and log-concave CGF monoids

Definition 45. The **cyclotomic monoid** is the monoid $\Phi^\pm$ generated by the cyclotomic polynomials under multiplication, graded by polynomial degree. The polynomials in $\Phi^\pm$ can have both positive and negative integer coefficients. Recall that the basic CGF monoid is the submonoid Φ^+ of $\Phi^\pm$ consisting of all basic cyclotomic generating functions, which is clearly closed under multiplication.

Similarly, let Φ^{uni} and Φ^{lcc} denote the submonoids of Φ^+ given by cyclotomic generating functions which are unimodal or log-concave with no internal zeros, respectively. These properties are preserved under multiplication [Sta89, Prop. 1-2].

Lemma 46. *We have $\Phi^{\mathrm{lcc}} \subset \Phi^{\mathrm{uni}} \subset \Phi^+ \subset \Phi^\pm$, and for each monoid $\mathcal{M}$ in $\{\Phi^{\mathrm{lcc}}, \Phi^{\mathrm{uni}}, \Phi^+, \Phi^\pm\}$ the following facts hold.*

(i) Each $\mathcal{M}$ has a unique minimal set of generators under inclusion, namely the set of irreducible elements $x \in \mathcal{M}$ where $x = yz$ for $y, z \in \mathcal{M}$ implies $y = 1$ or $z = 1$.

(ii) There are only a finite number of polynomials in $\mathcal{M}$ of any given degree n.

(iii) Each $\mathcal{M}$ cannot be generated by a finite subset.

(iv) Any element $1 \neq f \in \Phi^+$ has a cyclotomic factor of the form $\Phi_{p^k}(q)$ for some prime p.

(v) A cyclotomic polynomial $\Phi_n(q)$ is in Φ^+ if and only if n is a prime power.

(vi) Any $f \in \Phi^+$ with odd degree has $\Phi_2(q)$ as a factor.

Proof. The list of inclusions follows directly from the definitions and the fact that log-concave polynomials with no internal zeros are always unimodal. Property (i) follows easily from classical factorization in $\mathbb{C}[q]$ and the fact that every polynomial in $\Phi^\pm$ is monic. For (ii), the $\mathcal{M} = \Phi^\pm$ case follows from the fact that Euler's totient function has finite fibers, and the rest are submonoids. For (iii), note that $[n]_q \in \mathcal{M}$ for each such $\mathcal{M}$, and $[n]_q$ has primitive nth roots of unity as roots, so as $n \to \infty$ a finite set of generators cannot yield all $[n]_q$.

For (iv), first note that $f(1) \geqslant 1 \neq 0$ for $f \in \Phi^+$, so $\Phi_1(q) = q - 1$ is not a factor of f. Further recall that $\Phi_n(0) = \Phi_n(1) = 1$ for $n \geqslant 2$ not a prime power. Hence, if f had no factors of the form $\Phi_{p^k}(q)$, we would have $f(0) = f(1) = 1$, forcing $f(q) = 1$ since the coefficients of f are nonnegative. The forwards implication in (v) is given by (iv), and the backwards implication follows from the fact that

$$\Phi_{p^k}(q) = 1 + q^{p^k} + q^{2p^k} + \cdots + q^{(p-1)p^k} \in \Phi^+.$$

Finally, for (vi), $\varphi(n)$ is even for all $n \geqslant 3$, so any product of cyclotomic polynomials of odd degree must contain Φ_2. □

Example 47. We have

$$\begin{aligned}\Phi_5(q)\Phi_6(q) &= (q^4 + q^3 + q^2 + q + 1)(q^2 - q + 1)\\ &= q^6 + 0q^5 + q^4 + q^3 + q^2 + 0q^1 + 1 \in \Phi^+ - \Phi^{\mathrm{uni}}.\end{aligned}$$

Consequently, $\Phi_5(q)\Phi_6(q)$ is irreducible in Φ^+ and hence belongs to its minimal set of generators. The smallest degree polynomial in $\Phi^{\mathrm{uni}} - \Phi^{\mathrm{lcc}}$ is

$$\Phi_3(q)\Phi_4(q) = q^4 + q^3 + 2q^2 + q^1 + 1.$$

Intuitively, multiplying by $\Phi_p(q) = 1 + q + \cdots + q^{p-1}$ for p prime tends to smooth out chaotic coefficients. Computationally, it appears to be necessary to include such a factor in unimodal or log-concave CGFs. More precisely, we conjecture the following analogue of Lemma 46(iv), which has been checked up to degree 50.

Conjecture 48. Any element $1 \neq f \in \Phi^{\mathrm{uni}}$ has a cyclotomic factor of the form $\Phi_p(q)$ for some prime p.

In 1988, Boyd–Montgomery [BM90] described some of the properties of the cyclotomic monoid $\Phi^{\pm}_n$. In particular, they showed that $\log|\Phi^{\pm}_n| \sim \sqrt{105\zeta(3)n}/\pi$. This result is attributed to Vaclav Kotesovec in [OEI23, A120963], but no citation is given. One could ask for similar results for the other cyclotomic monoids.

Problem 49. For a monoid of polynomials $\mathcal{M}$, let $\mathcal{M}_n$ be the set of degree n polynomials in $\mathcal{M}$. Identify the growth rate of $|\mathcal{M}_n|$ as $n \to \infty$ for the other cyclotomic monoids $\mathcal{M} \in \{\Phi^{\mathrm{lcc}}, \Phi^{\mathrm{uni}}\}$.

Problem 50. Classify the minimal generating set $\mathcal{M} \in \{\Phi^{\mathrm{lcc}}, \Phi^{\mathrm{uni}}, \Phi^{+}\}$. Give an efficient algorithm for identifying the generators up to any desired degree. Find the asymptotic growth rate of the generating set as a function of degree.

We use the cyclotomic form for basic CGFs and the fact that we know a lower bound for the degree of the cyclotomic polynomial $\Phi_n(q)$ to find all basic CGFs of a given degree. Some initial data and OEIS identifiers [OEI23] can be found in Section 7. The sequence corresponding to $\Phi^{\pm}_n$ [OEI23, A120963] was created in 2006. We added the corresponding pages for the other cyclotomic monoids, so we believe the other cyclotomic monoids have not been studied in the literature.

6.2 Gale order and the associated CGF monoid

Given two multisets of the same size, say $A = \{a_1 \leqslant a_2 \leqslant \cdots \leqslant a_m\}$ and $B = \{b_1 \leqslant b_2 \leqslant \cdots \leqslant b_m\}$ both sorted into increasing order, we say $A \preceq B$ in **Gale order** provided $a_k \leqslant b_k$ for all $1 \leqslant k \leqslant m$. This partial order is known by many other names; we are following [ARW16] for consistency. Gale studied this partial order in the context of matroids on m-subsets of $\{1, 2, \ldots, n\}$ in the 1960's [Gal68].

Definition 51. Let Φ^{Gale} denote the **Gale** submonoid of Φ^{+} given by cyclotomic generating functions $f(q) = \frac{\prod_{k=1}^{m}[a_k]_q}{\prod_{k=1}^{m}[b_k]_q} \in \Phi^{+}$ such that $\{b_1, b_2, \ldots, b_m\} \preceq \{a_1, a_2, \ldots, a_m\}$ in Gale order. Note that Gale order holds independent of the representation of the rational expression chosen since $\{b_1, b_2, \ldots, b_m\} \preceq \{a_1, a_2, \ldots, a_m\}$ if and only if $\{i, b_1, b_2, \ldots, b_m\} \preceq \{i, a_1, a_2, \ldots, a_m\}$ for any positive integer i. Furthermore, the Gale property is again preserved under multiplication, hence Φ^{Gale} is closed under multiplication.

The properties in Lemma 46 also hold for the Gale monoid. In particular, it has a finite number of elements of each degree and a unique minimal set of generators which can be explored computationally. Data is given in Section 7 for the number of elements in Φ^{Gale} of degree n up to $n = 18$ along with the pointer to the corresponding OEIS entry. The number of generators of each degree are also noted in Section 7.

As noted in Remark 32, not all basic CGFs are in the Gale monoid. They agree up to degree 10, there are two basic CGFs which don't satisfy the Gale property in degree 11, namely

$$q^{11}+4q^{10} + 8q^{9} + 9q^{8} + 5q^{7} + 5q^{4} + 9q^{3} + 8q^{2} + 4q^{1} + 1 = \frac{[12]_q[3]_q^3[2]_q^2}{[6]_q[4]_q[1]_q^6} = \Phi_{12}\Phi_3^3\Phi_2$$

$$q^{11}+4q^{10}+5q^9+q^8+5q^6+5q^5+q^3+5q^2+4q^1+1=\frac{[12]_q[2]_q^5}{[4]_q[3]_q[1]_q^4}=\Phi_{12}\Phi_6\Phi_2^5.$$

There are 4 non-Gale basic CGFs in degree 12, and so on.

The Gale monoid and unimodal CGF monoids are not comparable. The Gale monoid includes $[4]_q/[2]_q=\Phi_4=1+q^2$, which is not unimodal. The smallest degree unimodal CGFs which don't satisfy the Gale property have degree 20,

$$f(q)=\frac{[12]_q[8]_q[3]_q[3]_q[3]_q[3]_q[3]_q[3]_q[2]_q}{[6]_q[4]_q[4]_q[1]_q[1]_q[1]_q[1]_q[1]_q[1]_q}=\Phi_{12}\Phi_8\Phi_3^6$$

$$g(q)=\frac{[12]_q[8]_q[3]_q[3]_q[3]_q[3]_q[3]_q[2]_q[2]_q[2]_q}{[6]_q[4]_q[4]_q[1]_q[1]_q[1]_q[1]_q[1]_q[1]_q[1]_q}=\Phi_{12}\Phi_8\Phi_3^5\Phi_2^2.$$

The Gale monoid also does not contain the log-concave monoid. The unique smallest degree log-concave CGF which does not satisfy the Gale property has degree 25,

$$\frac{[12]_q[2]_q^{19}}{[4]_q[3]_q}=\Phi_{12}\Phi_6\Phi_2^{19}.$$

6.3 CGF monoid from regular sequences

In this subsection, we discuss one more submonoid of the basic CGF monoid coming from certain Hilbert series of quotients of polynomial rings that naturally arise in commutative algebra. As described below, this monoid is also a submonoid of the Gale monoid.

Fix a pair of multisets of positive integers of the same size, say $\{a_1,\dots,a_m\}$ and $\{b_1,\dots,b_m\}$. Let $\Bbbk[x_1,\dots,x_m]$ be a (free) polynomial ring over a field $\Bbbk$ with grading determined by $\deg(x_j)=b_j$ for $1\leqslant j\leqslant m$. Let $\theta_1,\dots,\theta_m$ be a sequence of non-constant homogeneous polynomials in $\Bbbk[x_1,\dots,x_m]$ with $\deg(\theta_j)=a_j$. Then, $\theta_1,\dots,\theta_m$ is a **regular sequence** if θ_i is not a zero-divisor in $\Bbbk[x_1,\dots,x_m]/(\theta_1,\dots,\theta_{i-1})$ for all $1\leqslant i\leqslant m$. Furthermore, $\theta_1,\dots,\theta_m\in\Bbbk[x_1,\dots,x_n]$ form a **homogeneous system of parameters** (HSOP) if $\Bbbk[x_1,\dots,x_m]$ is a finitely generated $\Bbbk[\theta_1,\dots,\theta_m]$-module. Consider the corresponding quotient rings,

$$R:=\Bbbk[x_1,\dots,x_m]/(\theta_1,\dots,\theta_m).$$

Such rings play an important role in commutative algebra and the study of affine and projective varieties [SKKT00]. The following equivalences are well-known and are stated explicitly for completeness.

Lemma 52. *Let $\Bbbk[x_1,\dots,x_m]$ be a polynomial ring over a field $\Bbbk$ with $\deg(x_j)=b_j\in\mathbb{Z}_{\geqslant 1}$. Suppose $\theta_1,\dots,\theta_m$ are homogeneous elements. Then the following are equivalent:*

(i) $\theta_1,\dots,\theta_m$ is a regular sequence;

(ii) $\theta_1,\dots,\theta_m$ is a homogeneous system of parameters;

(iii) R *is finite-dimensional over* $\Bbbk$*;*

(iv) *for all* $1 \leqslant i \leqslant m$*, there is some* N *such that* $x_i^N \in (\theta_1, \ldots, \theta_m)$*;*

(v) R *has Krull dimension* 0*; and*

(vi) *the only common zero of* $\theta_1, \ldots, \theta_m$ *over the algebraic closure* $\overline{\Bbbk}$ *is the origin.*

Proof. For (i)-(v), see for instance [Sta78, §3] and [Sta79, §3], and [CLO15] for the equivalence of (iii)-(vi). In brief, (ii) $\Leftrightarrow$ (iii) is elementary, and (i) $\Leftrightarrow$ (ii) since $\Bbbk[x_1, \ldots, x_m]$ is *Cohen–Macaulay*, so regular sequences and homogeneous systems of parameters coincide. The equivalence (iii) $\Leftrightarrow$ (iv) is also elementary. For (iv) $\Leftrightarrow$ (v), the Krull dimension is the order of the pole of the Hilbert series of R at 1, which is 0 if and only if R is finite-dimensional over $\Bbbk$. For (vi), the Krull (and vector space) dimension of R is preserved by extension of scalars, and in the algebraically closed case we may apply the Nullstellensatz. $\square$

Let R_k be the linear span of the homogeneous degree k elements in R. Then the **Hilbert series** of R is

$$\mathrm{Hilb}(R; q) \coloneqq \sum_{k \geqslant 0} (\dim R_k) q^k$$

If R is finite-dimensional over $\Bbbk$, then $\mathrm{Hilb}(R; q)$ is a polynomial.

Given a homogeneous system of parameters $\theta_1, \ldots, \theta_m$ with corresponding ring R, is such a Hilbert series given by a CGF? The affirmative answer given in the following theorem is based on the work of Macaulay [Mac94] from 100 years ago. The proof uses the natural short exact sequence $0 \to (\theta) \to R \to R/(\theta) \to 0$ and induction. Stanley built on this theory in his work on Hilbert functions of graded algebras [Sta78] with connections to invariant theory.

Theorem 53 (c.f. [Sta78, Cor. 3.2-3.3])**.** *If* $\theta_1, \ldots, \theta_m$ *is a regular sequence with* $\deg(\theta_j) = a_j$ *in the polynomial ring* $\Bbbk[x_1, \ldots, x_m]$ *with* $\deg(x_j) = b_j$*, then*

$$R = \Bbbk[x_1, \ldots, x_m]/(\theta_1, \ldots, \theta_m)$$

is a finite-dimensional $\Bbbk$*-vector space, and its Hilbert series is the basic cyclotomic generating function*

$$\mathrm{Hilb}(R; q) \coloneqq \sum_{k \geqslant 0} (\dim R_k) q^k = \prod_{k=1}^{m} \frac{[a_k]_q}{[b_k]_q} \in \Phi^+. \tag{6.1}$$

Moreover, the converse holds: if (6.1) *holds, then* $\theta_1, \ldots, \theta_m$ *is a regular sequence.*

Example 54. Consider the ring

$$\frac{\mathbb{C}[x_1, \ldots, x_k]^{S_k}}{\langle h_{\ell+1}, \ldots, h_{\ell+k} \rangle} \tag{6.2}$$

where h_i denotes the ith homogeneous symmetric polynomial in k variables $x_1, \dots, x_k$. This ring is well-known to be a presentation of the cohomology ring of the complex Grassmannian of k-planes in $\ell + k$-space. The polynomials $h_{\ell+1}, \dots, h_{\ell+k}$ form a homogeneous system of parameters, so we have the corresponding cyclotomic generating function

$$\frac{[\ell+1]_q \cdots [\ell+k]_q}{[1]_q \cdots [k]_q} = \binom{\ell+k}{k}_q,$$

recovering the q-binomial coefficients. See [CKW09, Prop. 2.9] for more regular sequences of a similar flavor.

Example 55. Bessis introduced an HSOP associated to a well-generated complex reflection groups W [Bes15, Thm. 5.3]. The corresponding Hilbert series is a q-analogue of the W-noncrossing partition number,

$$\prod_{i=1}^{n} \frac{[ih]_q}{[d_i]_q} \in \Phi^{\mathrm{HSOP}}.$$

CGFs are often intimately related to the cyclic sieving phenomenon [RSW04], and Douvropolous proved a CSP in this context [Dou18].

Suppose two basic CGFs f, g arise as the Hilbert series of quotient rings R_1 and R_2 coming from homogeneous systems of parameters. Then fg arises in the same way by taking the tensor product of R_1 and R_2. Hence we may consider the **HSOP monoid** Φ^{HSOP} as a submonoid of Φ^+ consisting of all Hilbert series of quotients $\mathbb{C}[x_1, \dots, x_m]/(\theta_1, \dots, \theta_m)$ for homogeneous systems of parameters.

Lemma 56. *The HSOP monoid Φ^{HSOP} is a submonoid of the Gale monoid Φ^{Gale}.*

Proof. Suppose $\theta_1, \dots, \theta_m$ is a regular sequence for $\Bbbk[x_1, \dots, x_m]$ where $\deg(\theta_i) = a_i$, $\deg(x_i) = b_i$, and we have sorted the elements so that $a_1 \leqslant \cdots \leqslant a_m$ and $b_1 \leqslant \cdots \leqslant b_m$. If $a_k < b_k$, then $\theta_1, \dots, \theta_k$ have degree at most a_k, and $\deg(x_k) = b_k > a_k$, so $\theta_1, \dots, \theta_k \in \Bbbk[x_1, \dots, x_{k-1}]$. By Theorem 53, $\Bbbk[x_1, \dots, x_{k-1}]/(\theta_1, \dots, \theta_{k-1})$ is finite-dimensional, so it contains a power of θ_k, contradicting the zero-divisor condition. Hence, $a_k \geqslant b_k$ for all $k = 1, \dots, m$ and the result follows from the definition of Gale order. ☐

Which basic CGFs can be realized as the Hilbert series of a quotient ring by a homogeneous sequence of parameters? In the special case $b_1 = \cdots = b_m = 1$, i.e. $\deg(x_j) = 1$ for all j, a homogeneous system of parameters corresponds to a sequence of generators for a classical **complete intersection** X. The corresponding cyclotomic generating function is the Hilbert series of the projective coordinate ring of X,

$$\mathrm{Hilb}(X; q) = \prod_{k=1}^{m} [a_k]_q.$$

Here the multiset of degrees $\{a_1, \dots, a_m\}$ can clearly be chosen arbitrarily.

For general denominator multisets $\{b_1, \ldots, b_m\}$, such homogeneous systems of parameters yield complete intersections inside **weighted projective space** $\mathbb{P}^{\{b_1,\ldots,b_m\}}$. However, it is not at all clear which multisets $\{a_1, \ldots, a_m\}$ are realizable degrees for some homogeneous system of parameters. The requirement that $\prod_{k=1}^m [a_k]_q/[b_k]_q \in \mathbb{Z}_{\geqslant 0}[q]$ is a significant hurdle, as the next example shows.

Example 57. One may check that

$$\frac{[3]_q[5]_q[14]_q}{[2]_q[3]_q[7]_q} = 1 + q^2 + q^4 - q^5 + q^6 + q^8 + q^{10}$$

is not a cyclotomic generating function since it has a negative coefficient. Therefore, $\Bbbk[x_1, x_2, x_3]$ with $\deg(x_1) = 2, \deg(x_2) = 3, \deg(x_3) = 7$ has no homogeneous system of parameters with degrees $3, 5, 14$. Indeed, in this case, the only monic homogeneous elements of degree 3 and 5 are x_2 and x_1x_2, and one may see in a variety of ways (regular sequences, dimension counting, nontrivial vanishing) that these two elements cannot belong to a homogeneous system of parameters. This example also satisfies all of the necessary conditions in Lemma 30.

Example 58. Let $\Bbbk[x_1, \ldots, x_m]$ be a polynomial ring with $\deg(x_k) = b_k$ for each k. Let $c_1, \ldots, c_m$ be any positive integers. Then taking $\theta_k = x_k^{c_k}$ gives a homogeneous system of parameters whose elements have degrees $a_k = c_k b_k$. Hence

$$\prod_{i=1}^m \frac{[c_k b_k]_q}{[b_k]_q} = \prod_{i=1}^m [c_k]_{q^{b_k}} \in \Phi^{\mathrm{HSOP}}$$

for all $c_1, \ldots, c_m, b_1, \ldots, b_m \in \mathbb{Z}_{\geqslant 1}$. Therefore, every sequence of positive integers $b_1, \ldots, b_k$ gives rise to a finite dimensional quotient ring with Hilbert series given by a CGF with denominator given by the product of $[b_i]_q$'s.

By Lemma 30, we know that for any cyclotomic generating function

$$f(q) = \prod_{j=1}^m [a_k]_q/[b_k]_q \in \Phi^+,$$

including $f \in \Phi^{\mathrm{HSOP}}$, we have $a_k \in \mathrm{Span}_{\mathbb{Z}_{\geqslant 0}}\{b_1, \ldots, b_m\}$. It is not immediately clear how to characterize which sequences $a_1, \ldots, a_k$ can actually be realized as degree sequences of HSOP's.

Example 59. Take $\deg(x_1) = 2, \deg(x_2) = 3$ in $\Bbbk[x_1, x_2]$. Set $\theta_1 = x_1x_2, \theta_2 = x_1^3 + x_2^2$, so that $\deg(\theta_1) = 5, \deg(\theta_2) = 6$. It is easy to see that $x_1^4, x_2^4 \in (\theta_1, \theta_2)$, so they form an HSOP. The corresponding Hilbert series is

$$\frac{[6]_q[5]_q}{[3]_q[2]_q} = q^6 + q^4 + q^3 + q^2 + 1 \in \Phi^{\mathrm{HSOP}}.$$

Problem 60. Besides identifying a specific homogeneous system of parameters in a graded ring, how can one test membership in the HSOP monoid?

Problem 61. How can one efficiently characterize the minimal set of generators of Φ^{HSOP}?

Problem 62. Identify the growth rate of $\log|\Phi_n^{\mathrm{HSOP}}|$ as $n \to \infty$.

One way to explore the HSOP monoid with a fixed denominator multiset comes from an analogue of Lemma 35. In particular, the analogous HSOP subgraph of the graph described in Section 3.5 remains connected with diameter at most $2m$.

Lemma 63. *[Hoc07, p.7 "Discussion"] Suppose $\Bbbk[x_1, \ldots, x_m]$ is a polynomial ring with homogeneous regular sequences $\theta_1, \ldots, \theta_m$ and $\theta'_1, \ldots, \theta'_m$. Then there exist homogeneous elements $\gamma_1, \ldots, \gamma_m$ such that*

$$\gamma_1, \ldots, \gamma_i, \theta_{i+1}, \ldots, \theta_m \qquad \text{and} \qquad \gamma_1, \ldots, \gamma_i, \theta'_{i+1}, \ldots, \theta'_m$$

are regular sequences for all $1 \leqslant i \leqslant m$.

Remark 64. The argument in Theorem 53 works more generally when $\Bbbk[x_1, \ldots, x_m]$ is a **Cohen–Macaulay** $\mathbb{N}$-graded $\Bbbk$-algebra of Krull dimension m in the sense of [Sta79, §3], except that the stated expression for $\mathrm{Hilb}(R; q)$ will be multiplied by some polynomial $P(q) \in \mathbb{Z}[q]$. Intuitively, $P(q)$ arises from syzygies of $x_1, \ldots, x_m$ and may have negative coefficients.

The study of Cohen–Macaulay rings has been a rich subject. It would be interesting to consider properties of the set of all Hilbert series of Cohen–Macaulay rings as we have done here for CGFs.

Problem 65. What interesting properties do Hilbert series of Cohen–Macaulay rings have from the algebraic, probabilistic and analytic perspectives?

7 Appendix

The data below gives the size of $\mathcal{M}_n$ for each monoid discussed in Section 6.1 and Section 6.3. The sequence $|\Phi_n^{\pm}|$ has a nice generating function and Vaclav gives an asymptotic approximation in [OEI23, A120963]. We do not know of any generating function formulas or asymptotics for the other sequences, which we added to the OEIS in conjunction with this paper. Code to produce this data is available at

https://sites.math.washington.edu/~billey/papers/CGFs/.

$\mathcal{M}$	$\lvert\mathcal{M}_n\rvert$ for $n = 1, \ldots, 18$	OEIS
Φ^{lcc}	$1, 2, 3, 5, 7, 12, 16, 26, 35, 53, 70, 109, 142, 217, 285, 418, 548, 799$	*A360622*
Φ^{uni}	$1, 2, 3, 6, 8, 14, 20, 34, 48, 72, 100, 162, 214, 309, 437, 641, 860, 1205$	*A360621*
Φ^{Gale}	$1, 3, 4, 10, 12, 27, 33, 68, 82, 154, 187, 346, 410, 714, 857, 1460, 1722, 2860$	*A362553*
Φ^{+}	$1, 3, 4, 10, 12, 27, 33, 68, 82, 154, 189, 350, 417, 728, 874, 1492, 1767, 2937$	*A360620*
$\Phi^{\pm}$	$2, 6, 10, 24, 38, 78, 118, 224, 330, 584, 838, 1420, 2002, 3258, 4514, 7134, 9754, 15010$	*A120963*

The data below gives the number of generators for each monoid discussed in Section 6.1 and Section 6.3. The sequence of the number of minimal generators of $\Phi^{\pm}$ by degree in [OEI23, A014197] has a nice Dirichlet generating function and other formulas. We do not know of any generating function formulas or asymptotics for the other sequences below.

$\mathcal{M}$	Number of generators of $\mathcal{M}$ in degree n for $n = 1, \dots, 20$	
Φ^{lcc}	$1, 1, 1, 1, 1, 2, 2, 4, 4, 7, 8, 18, 19, 37, 42, 66, 87, 132, 157, 252$	*A*361439
Φ^{uni}	$1, 1, 1, 2, 2, 3, 4, 7, 10, 9, 15, 28, 30, 34, 66, 82, 125, 126, 222, 294$	*A*361440
Φ^{Gale}	$1, 2, 1, 3, 1, 4, 1, 6, 1, 5, 1, 14, 2, 9, 4, 28, 1, 33, 14, 61$	*A*362554
Φ^{+}	$1, 2, 1, 3, 1, 4, 1, 6, 1, 5, 3, 16, 5, 14, 6, 37, 9, 46, 33, 87$	*A*361441
$\Phi^{\pm}$	$2, 3, 0, 4, 0, 4, 0, 5, 0, 2, 0, 6, 0, 0, 0, 6, 0, 4, 0, 5$	*A*014197

Acknowledgments

We would like to thank Theo Douvropoulos, Darij Grinberg, Sam Hopkins, Matjaž Konvalinka, Svante Janson, Neil Sloane, Karen Smith, Dennis Stanton, Richard Stanley, Garcia Sun, Nathan Williams, and Doron Zeilberger for insightful comments and discussions. We also thank the anonymous referees that provided helpful suggestions.

Billey was partially supported by the Washington Research Foundation and NSF DMS-1764012. Swanson was partially supported by NSF DMS-2348843.

References

[ARW16] F. Ardila, F. Rincón, and L. Williams. Positroids and non-crossing partitions. *Trans. Amer. Math. Soc.*, 368(1):337–363, 2016.

[BB05] A. Björner and F. Brenti. *Combinatorics of Coxeter Groups*, volume 231 of *Graduate Texts in Mathematics*. Springer, New York, 2005.

[Ben73] E. A. Bender. Central and local limit theorems applied to asymptotic enumeration. *J. Combinatorial Theory Ser. A*, 15:91–111, 1973.

[Bes15] D. Bessis. Finite complex reflection arrangements are $K(\pi, 1)$. *Ann. of Math. (2)*, 181(3):809–904, 2015.

[Bil95] P. Billingsley. *Probability and Measure*. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, third edition, 1995. A Wiley-Interscience Publication.

[BKS20a] S. C. Billey, M. Konvalinka, and J. P. Swanson. Asymptotic normality of the major index on standard tableaux. *Adv. in Appl. Math.*, 113:101972, 36, 2020.

[BKS20b] S. C. Billey, M. Konvalinka, and J. P. Swanson. Tableau posets and the fake degrees of coinvariant algebras. *Adv. Math.*, 371:107252, 46, 2020.

[BM52] F. F. Bonsall and M. Marden. Zeros of self-inversive polynomials. *Proc. Amer. Math. Soc.*, 3:471–475, 1952.

[BM90] D. W. Boyd and H. L. Montgomery. Cyclotomic partitions. In *Number Theory (Banff, AB, 1988)*, pages 7–25. de Gruyter, Berlin, 1990.

[Brä15] P. Brändén. Unimodality, log-concavity, real-rootedness and beyond. In *Handbook of Enumerative Combinatorics*, Discrete Math. Appl. (Boca Raton), pages 437–483. CRC Press, Boca Raton, FL, 2015.

[Bre94] F. Brenti. Log-concave and unimodal sequences in algebra, combinatorics, and geometry: an update. In *Jerusalem combinatorics '93*, volume 178 of *Contemp. Math.*, pages 71–89. Amer. Math. Soc., Providence, RI, 1994.

[BS22] S. C. Billey and J. P. Swanson. The metric space of limit laws for q-hook formulas. *Comb. Theory*, 2(2):Paper No. 5, 58, 2022.

[BW89] Anders Björner and Michelle L. Wachs. q-Hook length formulas for forests. *J. Combin. Theory Ser. A*, 52(2):165–187, 1989.

[CGSHPM22] A. Ciolan, P. A. García-Sánchez, A. Herrera-Poyatos, and P. Moree. Cyclotomic exponent sequences of numerical semigroups. *Discrete Math.*, 345(6):Paper No. 112820, 22, 2022.

[CGSM16] A. Ciolan, P. A. García-Sánchez, and P. Moree. Cyclotomic numerical semigroups. *SIAM J. Discrete Math.*, 30(2):650–668, 2016.

[CKW09] A. Conca, C. Krattenthaler, and J. Watanabe. Regular sequences of symmetric polynomials. *Rend. Semin. Mat. Univ. Padova*, 121:179–199, 2009.

[CLO15] D. A. Cox, J. Little, and D. O'Shea. *Ideals, Varieties, and Algorithms*. Undergraduate Texts in Mathematics. Springer, Cham, fourth edition, 2015. An introduction to computational algebraic geometry and commutative algebra.

[Coh22] A. Cohn. Über die Anzahl der Wurzeln einer algebraischen Gleichung in einem Kreise. *Math. Z.*, 14(1):110–148, 1922.

[Coh03] P. M. Cohn. *Basic Algebra*. Springer-Verlag London, Ltd., London, 2003. Groups, rings and fields.

[CWW08] W. Y. C. Chen, C. J. Wang, and L. X. W. Wang. The limiting distribution of the coefficients of the q-Catalan numbers. *Proc. Amer. Math. Soc.*, 136(11):3759–3767, 2008.

[Dia88] P. Diaconis. *Group Representations in Probability and Statistics*, volume 11 of *Institute of Mathematical Statistics Lecture Notes—Monograph Series*. Institute of Mathematical Statistics, Hayward, CA, 1988.

[Dou18] T. Douvropoulos. Cyclic sieving for reduced reflection factorizations of the Coxeter element. *Sém. Lothar. Combin.*, 80B:Art. 86, 12, 2018.

[Ful97] W. Fulton. *Young Tableaux; With Applications To Representation Theory And Geometry*, volume 35 of *London Mathematical Society Student Texts*. Cambridge University Press, New York, 1997.

[Gal68] D. Gale. Optimal assignments in an ordered set: An application of matroid theory. *J. Combinatorial Theory*, 4:176–180, 1968.

[Gas98] V. Gasharov. Factoring the Poincaré polynomials for the Bruhat order on S_n. *Combinatorial Theory, Series A*, 83:159–164, 1998.

[GR02] V. Gasharov and V. Reiner. Cohomology of smooth Schubert varieties in partial flag manifolds. *Journal of the London Mathematical Society (2)*, 66(3):550–562, 2002.

[Hai03] M. Haiman. Combinatorics, symmetric functions, and Hilbert schemes. In *Current Developments in Mathematics, 2002*, pages 39–111. Int. Press, Somerville, MA, 2003.

[Har67] L. H. Harper. Stirling behavior is asymptotically normal. *Ann. Math. Statist.*, 38:410–414, 1967.

[Hoc07] Hochster. Math 711: Lecture of September 5, 2007. [Online; accessed 23-April-2023]. URL: https://dept.math.lsa.umich.edu/~hochster/711F07/L09.05.pdf.

[Hop24] S. Hopkins. Order polynomial product formulas and poset dynamics. *Proceedings of Symposia in Pure Mathematics*, 110, 2024.

[HST24] N. Heerten, H. Sambale, and C. Thäle. Probabilistic limit theorems induced by the zeros of polynomials. *Math. Nachr.*, 297(5):1772–1792, 2024.

[HZ15] H.-K. Hwang and V. Zacharovas. Limit distribution of the coefficients of polynomials with only unit roots. *Random Structures Algorithms*, 46(4):707–738, 2015.

[IM65] N. Iwahori and H. Matsumoto. On some Bruhat decomposition and the structure of the Hecke rings of $\mathfrak{p}$-adic Chevalley groups. *Inst. Hautes Études Sci. Publ. Math.*, 25:5–48, 1965.

[Ked08] K. S. Kedlaya. Search techniques for root-unitary polynomials. In *Computational arithmetic geometry*, volume 463 of *Contemp. Math.*, pages 71–81. Amer. Math. Soc., Providence, RI, 2008.

[KKZ11] C. Koutschan, M. Kauers, and D. Zeilberger. Proof of George Andrews's and David Robbins's q-TSPP conjecture. *Proc. Natl. Acad. Sci. USA*, 108(6):2196–2199, 2011.

[KM13] C. Krattenthaler and T. W. Müller. Cyclic sieving for generalised non-crossing partitions associated with complex reflection groups of exceptional type. In *Advances in combinatorics*, pages 209–247. Springer, Heidelberg, 2013.

[Knu73] D. E. Knuth. *The Art of Computer Programming*, volume 3. Addison–Wesley, Reading, MA, 1973.

[Kro57] L. Kronecker. Zwei Sätze über Gleichungen mit ganzzahligen Coefficienten. *J. Reine Angew. Math.*, 53:173–175, 1857.

[Luk70] E. Lukacs. *Characteristic Functions.* Hafner Publishing Co., New York, 1970. Second edition.

[Mac94] F. S. Macaulay. *The Algebraic Theory of Modular Systems.* Cambridge Mathematical Library. Cambridge University Press, Cambridge, 1994. Revised reprint of the 1916 original.

[Mat10] MathOverflow. English reference for a result of Kronecker?, 2010. [Online; accessed 23-April-2023]. URL: https://mathoverflow.net/questions/10911/english-reference-for-a-result-of-kronecker.

[MRR82] W. H. Mills, D. P. Robbins, and H. Rumsey, Jr. Proof of the Macdonald conjecture. *Invent. Math.*, 66(1):73–87, 1982.

[MS19a] M. Michelen and J. Sahasrabudhe. Central limit theorems and the geometry of polynomials, 2019. arXiv:1908.09020.

[MS19b] M. Michelen and J. Sahasrabudhe. Central limit theorems from the roots of probability generating functions. *Adv. Math.*, 358:106840, 27, 2019.

[MV97] R. Meise and D. Vogt. *Introduction to Functional Analysis*, volume 2 of *Oxford Graduate Texts in Mathematics*. The Clarendon Press, Oxford University Press, New York, 1997.

[OEI23] OEIS Foundation Inc. The On-Line Encyclopedia of Integer Sequences, 2023. Online. http://oeis.org.

[O'H90] K. M. O'Hara. Unimodality of Gaussian coefficients: a constructive proof. *J. Combin. Theory Ser. A*, 53(1):29–52, 1990.

[Pit97] J. Pitman. Probabilistic bounds on the coefficients of polynomials with only real zeros. *J. Combin. Theory Ser. A*, 77(2):279–303, 1997.

[Pro84] R. A. Proctor. Bruhat lattices, plane partition generating functions, and minuscule representations. *European J. Combin.*, 5(4):331–350, 1984.

[PS19] R. A. Proctor and L. M. Scoppetta. d-Complete posets: local structural axioms, properties, and equivalent definitions. *Order*, 36(3):399–422, 2019.

[PW99] G. Pistone and H. P. Wynn. Finitely generated cumulants. *Statist. Sinica*, 9(4):1029–1052, 1999.

[RS18] V. Reiner and E. Sommers. Weyl group q-Kreweras numbers and cyclic sieving. *Ann. Comb.*, 22(4):819–874, 2018.

[RSW04] V. Reiner, D. Stanton, and D. White. The cyclic sieving phenomenon. *J. Combin. Theory Ser. A*, 108(1):17–50, 2004.

[RSW14] V. Reiner, D. Stanton, and D. White. What is ... cyclic sieving? *Notices Amer. Math. Soc.*, 61(2):169–171, 2014.

[Sac97] V. N. Sachkov. *Probabilistic methods in combinatorial analysis*, volume 56 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1997.

[SKKT00] K. E. Smith, L. Kahanpää, P. Kekäläinen, and W. Traves. *An Invitation to Algebraic Geometry*. Universitext. Springer-Verlag, New York, 2000.

[Slo15] W. Slofstra. Rationally smooth Schubert varieties and inversion hyperplane arrangements. *Adv. Math.*, 285:709–736, 2015.

[Sta78] R. P. Stanley. Hilbert functions of graded algebras. *Advances in Math.*, 28(1):57–83, 1978.

[Sta79] R. P. Stanley. Invariants of finite groups and their applications to combinatorics. *Bull. Amer. Math. Soc. (N.S.)*, 1(3):475–511, 1979.

[Sta89] R. P. Stanley. Log-concave and unimodal sequences in algebra, combinatorics, and geometry. In *Graph theory and its applications: East and West (Jinan, 1986)*, volume 576 of *Ann. New York Acad. Sci.*, pages 500–535. New York Acad. Sci., New York, 1989.

[Sta98] D. Stanton. Fake Gaussian sequences, Dec. 4, 1998. [Online; accessed 10-May-2023]. URL: https://www-users.cse.umn.edu/~stant001/PAPERS/macmahon.pdf.

[Sta99] R. P. Stanley. *Enumerative Combinatorics. Vol. 2*, volume 62 of *Cambridge Studies in Advanced Mathematics.* Cambridge University Press, Cambridge, 1999.

[Sta12] R. P. Stanley. *Enumerative Combinatorics. Vol. 1*, volume 49 of *Cambridge Studies in Advanced Mathematics.* Cambridge University Press, Cambridge, second edition, 2012.

[Ste94] J. R. Stembridge. On minuscule representations, plane partitions and involutions in complex Lie groups. *Duke Math. J.*, 73(2):469–490, 1994.

[SW98] J. R. Stembridge and D. J. Waugh. A Weyl group generating function that ought to be better known. *Indag. Math. (N.S.)*, 9(3):451–457, 1998.

[Syl78] J. J. Sylvester. Proof of the hitherto undemonstrated fundamental theorem of invariants. *Phil. Mag.*, 5(30):178–188, 1878.

[Wik22] Wikipedia contributors. Cyclotomic polynomial — Wikipedia, the free encyclopedia, 2022. [Online; accessed 23-April-2023]. URL: https://en.wikipedia.org/w/index.php?title=Cyclotomic_polynomial&oldid=928883454.

[Wik23] Wikipedia contributors. Proof of the Euler product formula for the Riemann zeta function — Wikipedia, the free encyclopedia. https://en.wikipedia.org/w/index.php?title=Proof_of_the_Euler_product_formula_for_the_Riemann_zeta_function&oldid=1185992088, 2023. [Online; accessed 3-September-2024].

[WZ11] S. Ole Warnaar and W. Zudilin. A q-rious positivity. *Aequationes Math.*, 81(1-2):177–183, 2011.

[Zab03] M. Zabrocki. A bijective proof of an unusual symmetric group generating function, 2003. arXiv:math/0310301.

[Zei89] D. Zeilberger. Kathy O'Hara's constructive proof of the unimodality of the Gaussian polynomials. *Amer. Math. Monthly*, 96(7):590–602, 1989.

[Zei09] D. Zeilberger. The automatic central limit theorems generator (and much more!). In *Advances in combinatorial mathematics*, pages 165–174. Springer, Berlin, 2009.