Polynomial maps with applications to combinatorics and probability theory

by

Dan Port

B.S. Mathematics, University of California (1989)

Submitted to the Department of Applied Mathematics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 1994

Author ....
Department of Applied Mathematics
October 15, 1993

Certified by ....
Daniel Kleitman
Professor of Mathematics
Thesis Supervisor

Certified by ....
Richard Stanley
Chairman, Applied Mathematics

Accepted by ....
David Vogan
Chairman, Departmental Graduate Committee

Polynomial maps with applications to combinatorics and probability theory

by

Dan Port

Submitted to the Department of Applied Mathematics
on October 15, 1993, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract. This work considers broad classes of polynomial maps which generalize the exponential Bell polynomials. These maps correspond to various convolutions (such as Hadamard and Cauchy convolutions) and have been extensively studied in combinatorics, but relatively little in connection with probability theory.

It is shown that the exponential Bell Polynomials $Y_n(x_1, x_2, \dots)$ map the space of moments $\mathcal{M}$ to itself, a property which is called MP. The inverse image of $\mathcal{M}$ under (Y_n) is the space $\mathcal{K}$ of cumulants. By Hamburger's solution of the problem of moments, it follows that if the the Hankel determinants of the x_n are non-negative, then so are those of the Y_n. These latter determinants are independent of x_1, a property which is called HMI. This property is explored in some detail. Another application of exponential Bell polynomials is the determination of all random measures which arise from a compound Poisson process.

Next, the ordinary Bell polynomials $B_n^o(x_1, x_2, \cdots)$ are introduced and shown to have properties MP and HMI. The exponential and ordinary Bell polynomials are contained in a class $\mathcal{C}$ of polynomials introduced by Comtet in his book *Advanced Combinatorics*. The polynomials Y_n and B_n^o are characterized within $\mathcal{C}$ by having property HMI. This characterization has application to the problem of why exponential and ordinary generating functions are so ubiquitous in combinatorics.

Multidimensional analogs of the class $\mathcal{C}$ are next investigated. Analogs of properties MP and HMI are introduced and the above results are extended.

A detailed study is made of the polynomials $Y_n(x_1, x_2, \dots)$ when $x_i = t$ for $i \equiv a \pmod{m}$ and $x_i = 0$ otherwise. This leads to a two-parameter generalization of Stirling and Touchard numbers. The combinatorial, probabilistic, and congruential theory of these numbers is investigated.

It is shown that renewal theory and binomial posets give rise to polynomials in $\mathcal{C}$. The class $\mathcal{G}$ of polynomial maps arising from generalized compound Poisson processes is studied. These maps have property MP, but have property HMI only for Y_n, which is also the intersection of $\mathcal{G}$ with $\mathcal{C}$.

Thesis Supervisor: Dan Kleitman
Title: Proffesor of Mathematics

Acknowledgments

It is difficult to give justice to all those involved in the culmination of this work with such limited space, I hope that it will suffice them to know that they have my unending gratitude. This said, there are a particular few which I must mention. I wish to thank Dan Kleitman for supervising this work, his unique approach, and his assistance with the system. G.C. Rota for his many inspiring suggestions and stimulating courses. Richard Stanley for his terribly useful suggestions, unending patience, and fondness for Binomial posets. Dan Strook and Alar Toomre for help and kicks every so often. Special mention goes to my father, Sidney Port for leaving so many sheets of yellow paper with strange scribbles lying about the house for which I yearned to comprehend some day. Also, for his encouragement to finish, particularly when dark clouds hung low to my brow. The essential staff of the mathematics department, Phyllis Ruby (for putting up with my many strange requests) and Nancy Toscano for keeping everything together. Thanks to cousin Paula Corman and her husband Joel Corman for keeping me sane (as best they could) on the east coast. Finally, my deepest and fondest gratitude to Basil Gordon, without whom this thesis would not exist. His many hours of selfless guidance, encouragement, friendship, corrections, and the occasional scorpion sting, will stay with me forever.

Contents

Introduction

In this work we consider a broad class of polynomial transformations which generalize the exponential Bell polynomials. These transformations correspond to a variety of convolutions (such as Hadamard and Cauchy) and have been extensively studied in combinatorics (for example see Rota [31], or Comtet [10]), but relatively little in connection with probability theory; and then usually for particular cases such as the Stirling numbers of the second kind. See [6], [19], [26],[25],[34],[25],[32], and [7] as examples.

This apparent gap is curious as convolutions play a fundamental role in probability theory, and hence so do their associated polynomials; examples include sums of independent identically distributed random variables, the renewal equation, cumulants, and Hermite polynomials. Although this work concentrates more on using probability to develop combinatorial concepts, we hope it will lead to further consideration of the rich connection between combinatorics and probability beyond simple enumeration.

In this regard we find that the Bell polynomials have a natural place in the study of the moments of a random variable. Indeed, in Chapter I of the thesis it is shown that the exponential Bell polynomials $Y_n(x_1, x_2, \cdots)$, $n \geq 1$, map moment sequences to moment sequences. This is demonstrated by giving two constructions involving sums of iid random variables, first as the moments of a compound Poisson process and secondly as a limit of a sum of iid random variables. The distribution with moments $Y_n(x_1, \ldots)$ might not be unique as the moments do not always determine a unique distribution (see [33, p. viii]).

Using Hamburger's solution of the problem of moments, we find that if the Hankel determinants of the x_n are non-negative, so are the Hankel determinants of the Y_n. Surprisingly, these latter determinants are independent of x_1 (a property referred to as Hankel mean-independence), which we explore further. At present, Hankel mean-independence lacks a full explanation. We show that it can not be explained by transformations on a random variable. Perhaps an interpretation can be found in terms of cumulants, but more intriguing (as suggested by Stanley) is to interpret Hankel mean-independence in terms of Schur functions.

Another application of exponential Bell polynomials is the determination of all random measures which arise from a generalized compound Poisson process. It is shown in Chapter I that in addition to the usual axioms for an independent, stationary point process, a further divisibility property is required.

In Chapter II the ordinary Bell polynomials $B_n^o(x_1, x_2, \cdots)$, $n \geq 1$ are introduced, and their properties are studied. In particular it is shown that they also map moment sequences to moment sequences and are Hankel mean-independent. This is accomplished by showing that their Hankel determinants are obtained from those of

$x_1, x_2, \cdots$ by deleting the first row and column. Several mathematicians, including Kaluza [20], Horn [18], and Ligget [23] previously considered related problems. Although is is not always true that (x_n) is a moment sequence if (B_n^o) is, Horn's results conclude that various shifts on the indicies of the sequences provide nescessary and sufficient conditions. For example (x_{n+1}) is a moment sequence if and only if (B_{n+1}^o) is.

It is natural to ask if there are other polynomial transformations which are Hankel mean-independent or moment sequence preserving. In Chapter III we investigate this problem for the class of polynomials introduced by Comtet in his book *Advanced Combinatorics* [10]. What Comtet called Bell polynomials with respect to a sequence (Ω_n) will be referred to henceforth as Comtet polynomials. These constitute a broad class of transformations which include both the exponential and ordinary Bell maps. We prove that the only Hankel mean-independent Comtet polynomials are the exponential and ordinary Bell polynomials. This characterization theorem has application to the problem of why exponential and ordinary generating functions are so ubiquitous in combinatorics.

In Chapter V, multidimensional analogs of the Comtet polynomials are investigated; these of course include analogs of the exponential Bell polynomials. These analogs are shown to be moment sequence preserving; in fact a moment sequence is transformed by these polynomials into the moment sequence of a multidimensional compound Poisson process.

Returning to the one dimensional case, we make a detailed study of the polynomials $\Upsilon_{n;m,a}(t)$ which result from the exponential Bell polynomials when x_j is put equal to t for $j \equiv a \pmod{m}$ and zero otherwise. (These polynomials are are not of binomial type for $m > 1$.) Let $X_0, X_1, \ldots, X_{m-1}$ be iid Poisson random variables with parameter t, and let $\zeta = e^{2\pi i/m}$ be a primitive m'th root of unity. It is shown that the moments of $Z = (1/m)\sum_{h=0}^{m-1} \zeta^{ih} X_h$ are the polynomials $\Upsilon_{n;m,0}(t)$. We find that $\Upsilon_{n;1,0}(t) = \phi_n(t)$ are the exponential polynomials whose k'th coefficient is the Stirling number $S(n,k)$ of the second kind. This suggests a generalization of the Stirling numbers by looking at the coefficients of $\Upsilon_{n,m,a}(t)$. These numbers apparently have not been considered previously. Call these coefficients $S_{m,a}(n,k)$. It is shown that $S_{m,a}(n,k)$ can be interpreted as the number of partitions of an n-set into k parts where each block has cardinality $\equiv a \pmod{m}$. Several new recurrences are obtained by finding partial differential equations satisfied by the generating functions. In the Touchard case where $m = 2$ and $a = 0$, it turns out that $S_{2,0}(n,k) = (2k-1)!!C(n,k)$, where the $C(n,k)$ are the central factorial numbers. It is shown that when $m > 2$ there is no homogeneous linear recurrence analogous to the Stirling number recurrence. Other topics include congruential properties of these new sequences, and a generalization of Stirling numbers of the first kind.

In Chapter III it is shown that a binomial poset P gives rise to Comtet polynomials by introducing a stochastic process on P and defining an associated renewal sequence. In particular the ordinary Bell polynomials arise from a chain, and the exponential Bell polynomials arise from the Boolean algebra.

Chapter IV considers the compound polynomials; they are the moments of a compound random variable Y (not necessarily Poisson) defined by a sequence (p_n) which gives the probability of adding n iid copies of a random variable X. These polynomials are moment sequence preserving and are a modification of the exponential Bell polynomials when expressed in terms of factorial moments, but they are not Comtet polynomials. It is shown that they are Hankel mean-independent only in the case of the exponential Bell polynomials (the compound Poisson case). This leads to a further characterization theorem for exponential and ordinary generating functions in terms of the Hankel mean-independence property.

CHAPTER I

Exponential Bell polynomials

The exponential Bell polynomials constitute the most general sequence of binomial type (or sequence of *convolution polynomials* in the terminology of Knuth [22]). They are described in [10, p. 133] and [31, p. 118]. We will denote the n'th exponential Bell polynomial by $Y_n(x_1, x_2, x_3, \dots)$; the polynomials are formally defined as the coefficients in the expansion

$$G(z) = \sum_{n\geq 0} Y_n(x_1, x_2, x_3, \dots)\frac{z^n}{n!} = \exp\left(\sum_{m\geq 1} x_m \frac{z^m}{m!}\right) = \exp F(z). \tag{1}$$

One can obtain all sequences of binomial type from them by specializing the variables x_i; this follows from Lemma 1 below. Clearly $F(z)$ and $G(z)$ satisfy the formal differential equation

$$\frac{G'(z)}{G(z)} = F'(z).$$

The exponential Bell polynomials occur naturally in a large number of applications; see [10, p. 133] and [31] for a wealth of examples. Here we will discuss how they arise in connection with cumulants, moments of sums of independent identically distributed random variables, recurrent events on a Boolean lattice, and compound Poisson processes.

1. Properties

For future use we list the first few exponential Bell polynomials:

$$\begin{aligned}
Y_0 &= 1\\
Y_1 &= x_1\\
Y_2 &= x_2 + x_1^2\\
Y_3 &= x_3 + 3x_2x_1 + x_1^3\\
Y_4 &= x_4 + 4x_3x_1 + 3x_2^2 + 6x_1^2x_2 + x_1^4\\
Y_5 &= x_5 + 5x_4x_1 + 10x_3x_2 + 10x_3x_1^2 + 15x_2^2x_1 + 10x_2x_1^3 + x_1^5\\
Y_6 &= x_6 + 6x_5x_1 + 10x_3^2 + 15x_4x_2 + 15x_4x_1^2 + 60x_3x_2x_1 + 15x_2^3 + 20x_3x_1^3\\
&\quad + 45x_2^2x_1^2 + 15x_2x_1^4 + x_1^6
\end{aligned}$$

We note that the map $(x_1, x_2, \dots) \mapsto (Y_1, Y_2, \dots)$ is a bijection of the sequence space $\mathbb{R}^\infty$. Each term of degree k in Y_n has factors x_i whose indices form a partition of n into exactly k parts, say $n = c_1 + 2c_2 + 3c_3 + \cdots + nc_n$, with coefficient $n!/(c_1!c_2!\dots c_n!1!^{c_1}2!^{c_2}\dots n!^{c_n})$. Thus:

PROPOSITION 1.

$$\sum_{n\geq 0} Y_n \frac{z^n}{n!} = \sum_{c_1,c_2,\dots c_n \geq 0} \frac{(x_1^{c_1}x_2^{c_2}\dots x_n^{c_n})z^{c_1+c_2+\dots c_n}}{c_1!c_2!\dots c_n!1!^{c_1}2!^{c_2}\dots n!^{c_n}}$$

See [10, p. 134] for proof.

A fundamental property of exponential Bell polynomials is the following convolutional relation:

PROPOSITION 2.

$$Y_n(x_1+y_1, x_2+y_2, \dots) = \sum_{i=0}^{n} \binom{n}{i} Y_i(x_1, x_2, \dots) Y_{n-i}(y_1, y_2, \dots).$$

PROOF.

$$\begin{aligned}
&\sum_{n\geq 0} Y_n(x_1+y_1, x_2+y_2, \dots, x_n+y_n)\frac{z^n}{n!} = \exp\left(\sum_{m\geq 1}(x_m+y_m)\frac{z^m}{m!}\right)\\
&= \exp\left(\sum_{m\geq 1} x_m\frac{z^m}{m!} + \sum_{m\geq 1} y_m\frac{z^m}{m!}\right) = \exp\left(\sum_{m\geq 1} x_m\frac{z^m}{m!}\right)\exp\left(\sum_{m\geq 1} y_m\frac{z^m}{m!}\right)\\
&= \left(\sum_{m\geq 0} Y_m(x_1, x_2, \dots)\frac{z^m}{m!}\right)\left(\sum_{m\geq 0} Y_m(y_1, y_2, \dots)\frac{z^m}{m!}\right)\\
&= \sum_{n\geq 0}\sum_{i=0}^{n}\binom{n}{i} Y_i(x_1, x_2, \dots) Y_{n-i}(y_1, y_2, \dots).
\end{aligned}$$

□

We will often want to consider the t-Bell polynomials $Y_n(t) = Y_n(tx_1, tx_2, \ldots, tx_n)$ with a fixed sequence $(x_1, x_2, \ldots)$ as a function of the single variable t. Of particular importance are the *non-negative* t-Bell polynomials defined as follows:

DEFINITION 1. *$Y_n(t)$ is non-negative if $Y_n(t) \geq 0$ for all $t \geq 0$.*

Non-negativity of $Y_n(t)$ for all n is easily characterized as follows:

PROPOSITION 3. *$Y_n(t)$ is non-negative for all n if and only if $x_i \geq 0$ for all i.*

PROOF. If $x_i \geq 0$ for all i, then $Y_n(tx_1, tx_2, \ldots, tx_n) \geq 0$ when $t \geq 0$ because $Y_n(t)$ has non-negative coefficients. To prove the converse, note that $Y_n(tx_1, tx_2, \ldots, tx_n) = tx_n + O(t^2)$ as $t \to 0$. Hence if $Y_n(t) \geq 0$ for all $t \geq 0$, it follows that $x_n \geq 0$. Note that x_n is the only new variable which appears in $Y_n(t)$. Proceeding by induction on n, we have $x_i \geq 0$ for all i. $\square$

The concept of a *sequence of binomial type* is introduced in [31] and is defined as follows:

DEFINITION 2. *A sequence of polynomials $p_n(t)$ is of binomial type if* $\deg p_n(t) = n$ *and*

$$p_n(u+v) = \sum_{k=0}^{n} \binom{n}{k} p_k(u) p_{n-k}(v).$$

It is proved in [31, p. 162] that the exponential generating function for a sequence $(p_1(t), p_2(t), \ldots)$ of binomial type has the following form:

LEMMA 1.

$$\sum_{n \geq 0} p_n(t) \frac{z^n}{n!} = \exp\left(t \sum_{m \geq 1} a_m z^m\right).$$

From this it is clear that $p_n(t) = Y_n(1!ta_1, 2!ta_2, \ldots, n!ta_n)$.

We now refine the exponential Bell polynomials as follows. Note that

$$\sum_{n \geq 0} Y_n \frac{z^n}{n!} = \exp\left(\sum_{m \geq 1} x_m \frac{z^m}{m!}\right) = \sum_{k \geq 0} \frac{1}{k!} \left(\sum_{m \geq 1} x_m \frac{z^m}{m!}\right)^k.$$

This suggests the following notion of *partial* exponential Bell polynomials:

DEFINITION 3. *The* ***partial exponential Bell polynomials*** *$Y_{n,k}$ are defined by*

$$\sum_{n \geq k} Y_{n,k} \frac{z^n}{n!} = \frac{1}{k!} \left(\sum_{m \geq 1} x_m \frac{z^m}{m!}\right)^k. \tag{2}$$

We recover what can now be called the *complete* exponential Bell polynomials by

$$Y_n = \sum_{k=1}^{n} Y_{n,k}, \qquad Y_0 = 1.$$

This follows the approach of Comtet [10, p. 133]. We can think of the partial exponential Bell polynomials as a generalization of Stirling polynomials. It is easy to see that the choices $x_n = 1$ and $x_n = (-1)^{n-1}(n-1)!$ respectively produce $S(n,k)$ and $s(n,k)$, the Stirling numbers of the second and first kind.

2. Moments and cumulants

Let X be a random variable on $R \subseteq \mathbb{R}$ with probability distribution function $P(z)$. The *n'th moment* of X is

$$\mu_n = E[X^n] = \int_R z^n \, dP(z).$$

We define the *moment sequence* of X to be $(\mu_n)_{n\geq 1}$ (thus ignoring the zero'th moment $\mu_0 = 1$), and denote the set of all moment sequences by $\mathcal{M}$. Consider the Laplace transform

$$\phi(s) = E[e^{-sX}] = \int_R e^{-sz} \, dP(z),$$

It is easily seen that

$$\phi(-z) = 1 + \sum_{n\geq 1} \mu_n z^n / n!$$

is the exponential generating function of the moment sequence. The *cumulant* sequence $(\kappa_n)_{n\geq 1}$ is defined by

$$\log \phi(-z) = \psi(z) = \sum_{n\geq 1} \kappa_n \frac{z^n}{n!}.$$

Thus $\phi(-z) = \exp(\psi(z))$ and

$$1 + \sum_{n\geq 1} \mu_n(\kappa_1, \kappa_2, \ldots) \frac{z^n}{n!} = \exp\left(\sum_{m\geq 1} \kappa_m \frac{z^m}{m!} \right). \tag{3}$$

Hence μ_n is the n'th Bell polynomial in the cumulants. We define $\mathcal{K}$ to be the set of all cumulant sequences $(\kappa_n)_{n\geq 1}$. (Note that in defining $\mathcal{K}$ we do not include the zero'th cumulant $\kappa_0 = 0$, just as in defining $\mathcal{M}$ we did not include the zero'th moment $\mu_0 = 1$.) It follows from equation (3) that $\mathcal{K}$ is the inverse image of $\mathcal{M}$ under the exponential Bell map.

Until further notice, we will only consider random variables defined on the entire real line $(-\infty, \infty)$. In this case we can find necessary and sufficient conditions for a given sequence (κ_n) to be a cumulant sequence using the classical solution of the moment problem. A proof of the following theorem can be found in [33].

HAMBURGER'S THEOREM. *A sequence* (μ_n) *is a moment sequence if and only if* $\Delta_n = \det[\mu_{i+j}]_{i,j=0}^n \geq 0$ *for all* $n \geq 0$.

COROLLARY 1. $\mathcal{M}$ *is closed in the weak topology of the sequence space* $\mathbb{R}^\infty$.

COROLLARY 2. *A sequence* (κ_n) *is a cumulant sequence if and only if*

$$\delta_n = \det[Y_{i+j}(\kappa_1, \kappa_2, \kappa_3, \dots)]_{i,j=0}^n \geq 0 \quad \textit{for all} \quad n \geq 0.$$

The problem of cumulants (i.e. when a given sequence is the cumulant sequence of some distribution) is the problem of characterizing $\mathcal{K}$.

OPEN PROBLEM 1. *Are there more simple necessary and sufficient conditions than those of Corollary 2 which characterize cumulant sequences? What about just necessity or sufficiency?*

One might be tempted to conjecture that cumulant sequences are characterized by some analog of the Hamburger conditions, such as $\det[\kappa_{i+j+1}]_{i,j=0}^n \geq 0$ or perhaps $\det[\kappa_{i+j}]_{i,j=1}^n \geq 0$. However, neither of these conditions is either necessary or sufficient. For $\det[\kappa_{i+j+1}]_{i,j=0}^n$, consider the normal distribution with mean m and standard deviation σ^2. The cumulants are $\kappa_1 = m$, $\kappa_2 = \sigma^2$, $\kappa_n = 0$, for $n \geq 3$. We have

$$\det[\kappa_{i+j+1}]_{i,j=0}^1 = \begin{vmatrix} m & \sigma^2 \\ \sigma^2 & 0 \end{vmatrix} < 0,$$

contradicting necessity. On the other hand if $\kappa_n = -1$ for all $n \geq 1$, then $\det[\kappa_{i+j+1}]_{i,j=0}^n = 0$ for all n. But

$$\det[Y_{i+j+1}]_{i,j=0}^1 = \begin{vmatrix} 1 & -1 \\ -1 & 0 \end{vmatrix} = -1,$$

which contradicts sufficiency.

The question of whether cumulant sequences are characterized by $D_n = \det[\kappa_{i+j}]_{i,j=1}^n \geq 0$ is not so easily dispensed with. Consider the sequence

$$(\kappa_n)_{n\geq 1} = (0, 1, 1, 1, 1, 0, 0, \dots).$$

We have

$$\begin{aligned} D_1 &= 1 \\ D_2 &= 0 \\ D_3 &= 0 \\ D_4 &= 1 \\ D_n &= 0 \quad \text{for } n \geq 5, \end{aligned}$$

where the last entry follows from the appearance of a row of all zeros in the determinant D_n. Now

$$(Y_n)_{n\geq 0} = (1,0,1,1,4,11,40,161,686,3304,16716,91630,531916,\dots),$$

so if $\delta_n = \det[Y_{i+j}(\kappa_1,\kappa_2,\dots)]_{i,j=0}^n$, then

$$\begin{aligned}
\delta_0 &= 1\\
\delta_1 &= 1\\
\delta_2 &= 2\\
\delta_3 &= 10\\
\delta_4 &= -523\\
\delta_5 &= -1113510.
\end{aligned}$$

In particular $(Y_n)_{n\geq 1}$ is a moment sequence, so $(\kappa_n)_{n\geq 1}$ is not a cumulant sequence, contradicting sufficiency. The condition $D_n \geq 0$ is also not necessary, although it holds for many standard distributions. The binomial distribution $\binom{n}{k}p^k q^{n-k}$, for which $D_2 = -2n^2p^3q^3$, is a counterexample, but surprisingly the only one among the familiar discrete distributions.

To state one of the main probabilistic features of the exponential Bell polynomials, we introduce the following definition:

DEFINITION 4. *Let $(p_n(x_1,x_2,\dots))$ be a sequence of polynomials. If $\delta_n = \det[p_{i+j}]_{i,j=0}^n$ is independent of x_1, we say that (p_n) is* ***Hankel mean–independent****.*

THEOREM 1. *The sequence $(Y_n(x_1,x_2,\dots))$ is Hankel mean–independent.*

PROOF. Using Proposition 2 we have

$$\begin{aligned}
Y_n(x_1,x_2,\dots) &= \sum_{i=0}^{n}\binom{n}{i}Y_i(x_1,0,\dots)Y_{n-i}(0,x_2,\dots)\\
&= \sum_{i=0}^{n}\binom{n}{i}(x_1)^i\,Y_{n-i}(0,x_2,\dots).
\end{aligned}$$

Inverting the last equation gives

$$Y_n(0,x_2,\dots) = \sum_{i=0}^{n}\binom{n}{i}(-x_1)^i\,Y_{n-i}(x_1,x_2,\dots). \tag{4}$$

Consider the $n\times n$ matrix $A=[a_{i,j}]$, where $a_{i,j}$ is the j'th term in the expansion of $(1-x_1)^i$, that is $a_{i,j} = \binom{i}{j}(-x_1)^{i-j}$ for $i\geq j$ and $a_{i,j}=0$ for $i<j$. Note that A is lower triangular with $a_{i,i}=1$ for all i. Thus $\det A = 1$, so multiplication by A or A^T does not change the value of $\det[Y_{i+j}]$. Premultiplication of $[Y_{i+j}]$ by A replaces row R_i by $\sum_{j=0}^{i}\binom{i}{j}(-x_1)^j R_j$. Post-multiplication by A^T operates in the same way on the

columns. Note that the i,j entry of A^{-1} is $|a_{i,j}|$ (the j'th term in the expansion of $(1+x_1)^i$). It follows easily that $A[Y_{i+j}]A^T = [c_{i+j}]$, where

$$c_{i,j} = \sum_{h=0}^{i+j} \binom{i+j}{h} (-x_1)^h Y_{i+j-h}.$$

By equation (4) we have $c_{i,j} = Y_{i+j}(0, x_2, \dots)$, which is independent of x_1. □

COROLLARY 3. *If $(\kappa_n)_{n\geq 1}$ is a cumulant sequence, then for any constant c, $(c, \kappa_2, \kappa_3, \dots)$ is a cumulant sequence.*

From the proof of Theorem 1 we have

COROLLARY 4. *The i,j minor $M_{i,j}$ of $[Y_{i+j}(x_1, x_2, \dots)]$ for $0 \leq i,j \leq n$ is a polynomial of degree $2n-i-j$ in the variable x_1.*

Corollary 3 and (3) show that every moment sequence gives rise to a one–parameter family of cumulant sequences, which are in general not moment sequences. We will see later that if (x_n) is a moment sequence, then (Y_n) is also one. Thus $\mathcal{K}$ properly contains $\mathcal{M}$.

The cumulants are called *semi-invariants* due to the following property, which gives a probabilistic explanation of Corollary 3.

PROPOSITION 4. *If $(\kappa_1, \kappa_2, \kappa_3, \dots)$ is the cumulant sequence of a random variable X, then $(\kappa_1 + c, \kappa_2, \kappa_3, \dots)$ is the cumulant sequence of the translated random variable $X + c$, for any constant c.*

PROOF. Consider $E[e^{zX}] = \sum_{n\geq 0} \mu_n z^n/n!$, the exponential moment generating function of X. Making use of (3) we get

$$E[e^{z(X+c)}] = \exp\left(zc + \sum_{n\geq 1} \kappa_n \frac{z^n}{n!}\right) = \exp\left((c+\kappa_1)z + \sum_{n\geq 2} \kappa_n \frac{z^n}{n!}\right).$$

Hence the cumulants κ_n for X and $X+c$ are the same for $n \neq 1$. □

As already noted, the cumulant space $\mathcal{K}$ is the inverse image of the moment space $\mathcal{M}$ under the exponential Bell map. Here is a list of the first few cumulants:

$$\begin{aligned}
\kappa_1 &= \mu_1 \\
\kappa_2 &= \mu_2 - \mu_1^2 \\
\kappa_3 &= \mu_3 - 3\mu_1\mu_2 + 2\mu_1^3 \\
\kappa_4 &= \mu_4 - 4\mu_1\mu_3 - 3\mu_2^2 + 12\mu_1^2\mu_2 - 6\mu_1^4 \\
\kappa_5 &= \mu_5 - 5\mu_1\mu_4 - 10\mu_2\mu_3 + 20\mu_1^2\mu_3 + 30\mu_1\mu_2^2 - 60\mu_1^3\mu_2 + 24\mu_1^5.
\end{aligned}$$

One might be tempted to suppose that Hankel mean–independence of the exponential Bell map should follow from Proposition 4. This, however, is not the case. For one thing, Theorem 1 is valid for sequences (x_n), not just moment sequences. Moreover, the diagonal but non-principal minors of δ_n are *not* independent of x_1.

3. Sums of independent random variables

In this section we will show that if (x_n) is a moment sequence, then (Y_n) is also a moment sequence. This leads to the concept of a *moment sequence preserving map,* defined formally as follows:

DEFINITION 5. *A map $(x_n) \mapsto (y_n)$ is called* ***moment sequence preserving*** *if $(x_n) \in \mathcal{M}$ implies that $(y_n) \in \mathcal{M}$.*

For simplicity we will also refer to moment sequence preserving maps as $\mathcal{M}$–*preserving.* Such maps are plentiful. For example, given any random variable X and a real function $f(t)$, we obtain a new random variable $Y = f(X)$, and thus f induces an $\mathcal{M}$–preserving map which sends $E[X^n]$ to $E[Y^n]$ (note that we are doing the reverse of the moment problem here). We will show in two ways that if (x_n) is the moment sequence of a random variable X, then $(Y_n(x_1, x_2, \dots))$ is the moment sequence of a random variable Y. On the one hand, Y can be obtained as the limit of a sum of independent identically distributed random variables, and secondly, Y can be obtained as a compound Poisson random variable. Either way we conclude that the map $(x_n) \mapsto (Y_n)$ is $\mathcal{M}$–preserving. The following lemma is not only of technical use, but also leads to many interesting questions regarding $\mathcal{M}$–preserving maps.

LEMMA 2. *If $(a_n)_{n\geq 1} \in \mathcal{M}$, then $(pa_n)_{n\geq 1} \in \mathcal{M}$ for all $p \in [0,1]$.*

PROOF. Let X be a random variable with moment sequence (a_n). If E_0 is the random variable that is always zero, then $Z = pX + (1-p)E_0$ is a random variable with moments $\mu_0 = 1$ and $\mu_n = pa_n$ for $n \geq 1$. This new variable Z is called a *convex mixture* of X and E_0. $\square$

THEOREM 2. *If X is a random variable with $E[X^n] = x_n$, there exists a random variable $Z = Z(p)$ such that $E[(Z_1 + Z_2 + \cdots + Z_m)^n] \to Y_n(tx_1, tx_2, \dots) = Y_n(t)$ as $m \to \infty, p \to 0, mp \to t$, where $Z_1, Z_2, \dots$ are independent identically distributed copies of Z.*

PROOF. By Lemma 2, for any $p \in [0,1]$ there is a random variable $Z = Z(p)$ such that $E[Z^n] = px_n$ for all $n \geq 1$. Let $Z_1, Z_2, \dots$ be independent identically distributed iid copies of Z. Now

$$E[(Z_1 + Z_2 + \cdots + Z_m)^n] = \sum_{i_1+i_2+\cdots=m} \binom{n}{i_1, i_2, \dots} E[Z_1^{i_1} Z_2^{i_2} \dots].$$

By independence of the Z_i, this gives

$$E[(Z_1 + Z_2 + \cdots + Z_m)^n] = \sum_{c_1+2c_2+\cdots+nc_n=n} n!\binom{m}{k}\binom{k}{c_1, c_2, \ldots}\left(\frac{px_1}{1!}\right)^{c_1}\left(\frac{px_2}{2!}\right)^{c_2}\cdots,$$

where $k = c_1 + c_2 + \cdots + c_n$. Now let $m \to \infty$ and $p \to 0$ in such a way that $mp \to t$. Since $\binom{m}{k}k! \sim m^k$ as $m \to \infty$ with k fixed, we have

$$E[(Z_1 + Z_2 + \cdots + Z_m)^n] \sim \sum_{c_1+2c_2+\cdots+nc_n=m} \frac{n!(mp)^k x_1^{c_1} x_2^{c_2}\cdots}{c_1!c_2!\ldots(1!)^{c_1}(2!)^{c_2}\ldots} \to Y_n(tx_1, tx_2, \ldots).$$

□

By weak closure of $\mathcal{M}$, it follows that $(Y_n(t)) \in \mathcal{M}$.

Theorem 2 is analogous to the derivation of the Poisson distribution as a limit of the binomial distribution, except that instead of operating with probabilities, we transform the moment sequence of X into another moment sequence via a similar limit process.

The second probabilistic interpretation of these polynomials involves *compounding* the random variable X, as the next theorem demonstrates.

DEFINITION 6. *Given copies $X_1, X_2, \ldots$ of a random variable X, the compound Poisson random variable $S_N(X,t)$ is defined by $S_N(X,t) = \sum_{k=1}^N X_k$, where N is Poisson distributed with parameter t.*

THEOREM 3. *If X is a random variable with $E[X^n] = x_n$, then $Y_n(tx_1, tx_2, \ldots)$ are the moments of the compound Poisson variable $S_N(X,t)$.*

PROOF. Let $F(z) = 1 + \sum_{n\geq 1} x_n z^n/n! = E[e^{zX}]$ be the exponential moment generating function of X. Put $S_m = X_1 + \cdots + X_m$, where the X_i are iid copies of X. Now $E[e^{zS_N(X,t)}|N = m] = E[e^{zS_m(X,t)}]$, which by independence is $(E[e^{zX}])^n = F(z)^n$. The exponential moment generating function for $S_N(X,t)$ is

$$\begin{aligned}
E[e^{zS_N(X,t)}] &= \sum_{m\geq 0} P(N = m)E[e^{zS_N}|N = m] = \sum_{m\geq 0} e^{-t}t^m\frac{F(z)^m}{m!} = e^{t(-1+F(z))} \\
&= \exp\left(t\left(\sum_{m\geq 1} x_m z^m/m!\right)\right) = \sum_{n\geq 0} Y_n(t)\frac{z^n}{n!}.
\end{aligned}$$

□

Theorems 2 and 3 both show that the map $(x_n) \mapsto (Y_n)$ is $\mathcal{M}$-preserving. In other words:

COROLLARY 5. *If $\Delta_n \geq 0$, then $\delta_n \geq 0$.*

However, the converse is false because of the Hankel mean–independence of (Y_n). Indeed, suppose we are given a moment sequence (Y_n). By Hankel mean–independence, (Y_n) is a moment sequence regardless of the value of x_1. But Δ_n is highly dependent on x_1 for $n \geq 1$, so Δ_n can be negative for $n \geq 1$. For example, consider $x_1 = c,\ x_n = 0$ for $n > 1$. We have linear dependence of the rows of $(Y_{i+j})_{i,j=0}^{n}$, for $n \geq 1$, so $\delta_n = 0$ for all such n, but $\Delta_1 = -c^2$. Thus (x_n) is not a moment sequence for $c \neq 0$.

4. Circular processes and partitions of an n-set

In this section we consider a class of random variables we call *circular processes.* We introduce these processes with two classical examples, the Poisson process and the excess of one Poisson process over another. As an application we will derive Dobinski's formula for the Bell numbers B_n and Touchard's formula for the Touchard numbers T_n. The methods presented easily extend and give rise to a generalization of the Stirling numbers of the second kind. These numbers have remarkable combinatorial and probabilistic properties which we explore in detail. This section expands on work of Touchard [38],[39], Bell [4], Hanlon et.al. [1], Rota [28], Riordan [27], and others.

The following lemma will be useful:

LEMMA 3. *For a random variable X on R, define the umbral variable β by $\beta^n \mapsto E[X^n]$. Then if $f(y) = a_0 + a_1 y + \cdots + a_m y^m$ is any real polynomial, we have $f(y+\beta) \mapsto E[f(y+X)]$.*

PROOF.

$$\begin{aligned} f(y+\beta) &= \sum_{k=0}^{m}\sum_{i=0}^{k}\binom{k}{i} a_k y^{k-i}\beta^k \mapsto \sum_{k=0}^{m}\sum_{i=0}^{k}\binom{k}{i} a_k y^{k-i}\int_R x^k \, d\mathrm{Pr}(x) \\ &= \int_R \sum_{k=0}^{m} a_k \sum_{i=0}^{k}\binom{k}{i} y^{k-i} x^k \, d\mathrm{Pr}(x) = \int_R \sum_{k=0}^{m} a_k (y+x)^k \, d\mathrm{Pr}(x) \\ &= \int_R f(y+x)\, d\mathrm{Pr}(x) = E[f(y+X)] \end{aligned}$$

□

Consider a random variable X, Poisson distributed with parameter t, and a polynomial $f(y)$. By Lemma 3,

$$f(y+\beta) \mapsto E[f(y+X)] = \sum_{i\geq 0} f(y+i)\mathrm{Pr}(X=i) = \frac{1}{e^t}\sum_{i\geq 0}\frac{f(y+i)t^i}{i!}.$$

For $y = 0$ and $f(x) = x^n$ we get

$$\beta^n \mapsto \frac{1}{e^t}\sum_{i\geq 0}\frac{i^n t^i}{i!}. \tag{5}$$

Dividing by $n!$ and summing over $n \geq 0$, we obtain the exponential generating function of the polynomials $\phi_n(t) = Y_n(t, t, \cdots)$ (the exponential polynomials):

$$\sum_{n\geq 0} \frac{\phi_n(t) z^n}{n!} = e^{t(e^z-1)}.$$

Since the Bell numbers satisfy $B_n = \phi_n(1)$, we obtain Dobinski's formula (see [28] and [39]) by putting $t = 1$ in (5). In terms of exponential Bell polynomials, $\beta^n \mapsto \phi_n(t)$. Hence the total number of partitions of an n-set is the n'th moment of a Poisson random variable (see also [39, p.316]. Here are the first few values of B_n and $\phi_n(t)$:

$$\begin{aligned}
(B_n) &= (1, 1, 2, 5, 15, 52, 203, \dots) \\
\phi_0(t) &= 1 \\
\phi_1(t) &= t \\
\phi_2(t) &= t + t^2 \\
\phi_3(t) &= t + 3t^2 + t^3 \\
\phi_4(t) &= t + 7t^2 + 6t^3 + t^4 \\
\phi_5(t) &= t + 15t^2 + 25t^3 + 10t^4 + t^5 \\
\phi_6(t) &= t + 31t^2 + 90t^3 + 65t^4 + 15t^5 + t^6
\end{aligned}$$

Next, let X and Y be iid Poisson variables with parameter t. Consider the excess $Z = Y - X$ of Y over X. We will show that $E[Z^n]$ is a polynomial $T_n(t)$ with $T_n(1) = T_n$, the n'th Touchard number. This extends the formula which expresses Touchard numbers in terms of Bell numbers, namely $(\beta_1 - \beta_2)^n \mapsto T_n$ where β_1, β_2 are distinct umbral variables with $\beta_1^n, \beta_2^n \mapsto B_n$ (see [39, p.309]). Define a new umbral variable γ by $\gamma^n \mapsto E[Z^n]$. Now

$$\Pr(Z = i) = \Pr(Y - X = i) = \sum_{\substack{k-j=i \\ k,j\geq 0}} \Pr(Y = k)\Pr(X = j) = \frac{1}{e^{2t}} \sum_{\substack{k-j=i \\ k,j\geq 0}} \frac{(t)^{k+j}}{k!j!},$$

so by Lemma 3 we have

$$f(x+\gamma) \mapsto E[f(y+Z)] = \sum_{i=-\infty}^{\infty} f(y+i)\Pr(Z = i) = \frac{1}{e^{2t}} \sum_{i=-\infty}^{\infty} \sum_{\substack{k-j=i \\ k,j\geq 0}} \frac{(t)^{k+j} f(y + k - j)}{k!j!}.$$

For $f(x) = x^n$ and $y = 0$ we obtain Touchard's formula ([39, p.317]):

$$\gamma^n = \frac{1}{e^{2t}} \sum_{i=-\infty}^{\infty} \sum_{\substack{k-j=i \\ k,j\geq 0}} \frac{(k - j)^n t^{k+j}}{k!j!}.$$

The Touchard polynomials can also be defined by $T_n(t) = Y_n(0, 2t, 0, 2t, \dots)$ (or by orthogonality to $\phi_n(t)$ as described in [39, p.313]). From this we can obtain the exponential generating function

$$\sum_{n\geq 0} \frac{T_n(t)z^n}{n!} = e^{t(e^z+e^{-z}-2)}.$$

Here are the first few values of T_n and $T_n(t)$:

$$\begin{aligned}
T_n &= (1,0,2,0,4,0,14,0,182,0,3614,0,\dots) \\
T_0(t) &= 1 \\
T_1(t) &= 0 \\
T_2(t) &= 2t \\
T_3(t) &= 0 \\
T_4(t) &= 2t + 12t^2 \\
T_5(t) &= 0 \\
T_6(t) &= 2t + 60t^2 + 120t^3 \\
T_7(t) &= 0 \\
T_8(t) &= 2t + 252t^2 + 1680t^3 + 1680t^4 \\
T_9(t) &= 0 \\
T_{10}(t) &= 2x + 1020t^2 + 17640t^3 + 50400t^4 + 30240t^5
\end{aligned}$$

Note that unlike $(\phi(t))$, the sequence $(T_n(t))$ is not a sequence of binomial type, since it violates the requirement that the n'th polynomial must have degree n. The Touchard polynomials will be said to have support 0 (mod 2).

Combinatorially it is better to consider the *semi-reduced* Touchard numbers $T_n^* = T_n^*(1)$, where $T_n^*(t) = Y_n(0,t,0,t,\dots)$ is obtained by replacing t by $t/2$ in the above formulas. It is clear from the construction of the exponential Bell polynomials that T_n^* is the total number of ways to partition an n-set into blocks of even cardinality. The random variable whose moments are $T_n^*(t)$ is $Z = (Y - X)/2$; we call this a *circular process of order 2.* The generating function of $(T_n^*(t))$ is

$$\sum_{n\geq 0} \frac{T_n^*(t)z^n}{n!} = e^{t(\cosh z - 1)}.$$

The first few values of T_n^* and $T_n^*(t)$ are as follows:

$$\begin{aligned}
T_n^* &= (1,0,1,0,4,0,31,0,379,0,6556,0,\dots) \\
T_0^*(t) &= 1 \\
T_1^*(t) &= 0 \\
T_2^*(t) &= t \\
T_3^*(t) &= 0 \\
T_4^*(t) &= t+3t^2 \\
T_5^*(t) &= 0 \\
T_6^*(t) &= t+15t^2+15t^3 \\
T_7^*(t) &= 0 \\
T_8^*(t) &= t+63t^2+210t^3+105t^4 \\
T_9^*(t) &= 0 \\
T_{10}^*(t) &= t+255t^2+2205t^3+3150t^4+945t^5
\end{aligned}$$

By generalizing the above techniques we can construct polynomial sequences which have support a mod m by considering $\Upsilon_{n;m,a}(t) = Y_n(x_1 t, x_2 t, x_3 t, \dots)$ where $x_i = 1$ for $i \equiv a \pmod{m}$ and $x_i = 0$ otherwise. If $a \equiv 0 \pmod{m}$, we call the $\Upsilon_{n;m,0}(t)$ the *circular polynomials of order* m. To each (m,a) there corresponds a Dobinski-type formula, and when $a \equiv 0 \pmod{m}$ we get the processes defined as follows. Let $X_0, X_1, X_2, \dots, X_{m-1}$ be iid Poisson variables with parameter t and $\zeta = e^{2\pi i/m}$ a primitive m'th root of unity. The circular process of order m is the random variable $Z = (1/m)\sum_{j=0}^{m-1} \zeta^j X_j$. We have $E[Z^n] = \Upsilon_{n;m,0}(t)$, with exponential generating function

$$\sum_{j\geq 0} \frac{\Upsilon_{n;m,0}(t) z^j}{j!} = \exp\left(\frac{t}{m}\left(\sum_{j=0}^{m-1} e^{\zeta^j z} - 1\right)\right).$$

Note that

$$\left(\frac{1}{m}\sum_{j=0}^{m-1} \zeta^j \beta_j\right)^n \mapsto \Upsilon_{n;m,0}(1),$$

where $\beta_0, \dots, \beta_{m-1}$ are distinct umbral variables with $\beta_j^n \mapsto B_n$.

As an example we consider $\Upsilon_{n,3,0}$. In this case we have

$$\zeta = e^{2\pi i/3} = \frac{-1+i\sqrt{3}}{2},$$

and the Dobinski-type formula is

$$\Upsilon_{n;3,0}(1) = \frac{1}{e}\sum_{i_0\geq 0}\sum_{i_1\geq 0}\sum_{i_2\geq 0} \frac{(i_0+\zeta i_1+\zeta^2 i_2)^n}{i_0! i_1! i_2! 3^{i_0+i_1+i_2}}$$

Here are the first few values of $\Upsilon_{n;3,0}(1)$ and $\Upsilon_{n;3,0}(t)$:

$$\begin{aligned}
\Upsilon_{n;3,0}(1) &= (1,0,0,1,0,0,11,0,0,365,0,0,25323,0,0,3068521,0,0,583027547,\dots) \\
\Upsilon_{0;3,0}(t) &= 1 \\
\Upsilon_{1;3,0}(t) &= 0 \\
\Upsilon_{2;3,0}(t) &= 0 \\
\Upsilon_{3;3,0}(t) &= t \\
\Upsilon_{4;3,0}(t) &= 0 \\
\Upsilon_{5;3,0}(t) &= 0 \\
\Upsilon_{6;3,0}(t) &= t+10t^2 \\
\Upsilon_{7;3,0}(t) &= 0 \\
\Upsilon_{8;3,0}(t) &= 0 \\
\Upsilon_{9;3,0}(t) &= t+84t^2+280t^3 \\
\Upsilon_{10;3,0}(t) &= 0 \\
\Upsilon_{11;3,0}(t) &= 0 \\
\Upsilon_{12;3,0}(t) &= t+682t^2+9240t^3+15400t^4
\end{aligned}$$

The circular polynomials for $m \geq 3$ appear not to have been considered previously. They have many combinatorial and algebraic properties analogous to those of the exponential and Touchard polynomials. Just as we can define the Stirling number $S(n,k)$ of the second kind as the coefficient of t^k in the polynomial $\phi_n(t)$, we can similarly define generalized Stirling numbers $S_{m,a}(n,k)$ of the second kind. We can express them using the partial exponential Bell polynomials $Y_{n,k}$ defined earlier. Recall that $S(n,k) = Y_{n,k}(1,1,1,\dots)$. In general, $S_{m,a}(n,k) = Y_{n,k}(x_1,x_2,x_3,\dots)$ where $x_i = 1$ for $i \equiv a \pmod{m}$ and $x_i = 0$ otherwise. Clearly $S_{m,a}(n,k)$ is equal to the number of partitions of an n-set into k non-empty subsets with cardinalities $\equiv a \pmod{m}$. Note that $S_{m,a}(n,k) = S_{m,b}(n,k)$ if $a \equiv b \pmod{m}$. Let $\zeta = e^{2\pi i/m}$ and define

$$g_{m,a}(x) = \frac{e^x + e^{\zeta x} + \cdots + e^{\zeta^{m-1}x}}{m} - 1$$

when $a \equiv 0 \pmod{m}$, and

$$g_{m,a}(x) = \frac{e^x + \zeta^{-a}e^{\zeta x} + \cdots + \zeta^{-a(m-1)}e^{\zeta^{m-1}x}}{m}$$

when $a \not\equiv 0 \pmod{m}$. We have

$$S_{m,a}(n,k) = \frac{1}{k!}\frac{d^n}{dx^n} g_{m,a}(x)\Big|_{x=0}.$$

Let $B_{m,a}(n) = \Upsilon_{n;m,a}(1)$ the total number of partitions of an n-set into subsets each of size congruent to $a \pmod m$. Thus

$$\sum_{k\geq 0} S_{m,a}(n,k) = B_{m,a}(n).$$

In particular the Bell numbers and semi-reduced Touchard numbers are given by $B_n = B_{1,0}(n)$ and $T^*_n = B_{2,0}(n)$. Now

$$g_{m,0}(x)^k = \sum_{i_0+\cdots+i_{m-1}+j=k} (-1)^j \binom{k}{i_0,\ldots i_{m-1},j} \frac{\exp(x(i_0+\zeta i_1+\cdots+\zeta^{m-1}i_{m-1}))}{m^{k-j}},$$

and similarly for $g_{m,a}(x)^k$. Hence for $a \equiv 0 \pmod m$ the general Dobinski-type formula is

$$B_{m,0}(n) = \sum_{i_0\geq 0}\sum_{i_1\geq 0}\cdots\sum_{i_{m-1}\geq 0}\sum_{j\geq 0} \frac{(-1)^j(i_0+\zeta i_1+\cdots+\zeta^{m-1}i_{m-1})^n}{i_0!i_1!\cdots i_{m-1}j!m^{i_0+i_2+\cdots+i_{m-1}}}.$$

Carrying out the summation over j, we get

$$B_{m,0}(n) = \frac{1}{e}\sum_{i_0\geq 0}\cdots\sum_{i_{m-1}\geq 0}\sum_{j\geq 0} \frac{(i_0+\zeta i_2+\cdots+\zeta^{m-1}i_{m-1})^n}{i_0!i_1!\cdots i_{m-1}j!m^{i_0+i_2+\cdots+i_{m-1}}}.$$

For the classical Stirling numbers $S(n,k) = S_{1,0}(n,k)$, the generating function

$$F(x,y) = \sum_{n,k\geq 0} S(n,k)\frac{x^n y^k}{n!} = e^{yg_{1,0}(x)} = e^{y(e^x-1)} \tag{6}$$

satisfies the partial differential equation

$$F_x = y(F_y + F). \tag{7}$$

Substituting the power series (6) into (7) and equating coefficients gives the well-known recurrence formula

$$S(n,k) = kS(n-1,k) + S(n-1,k-1). \tag{8}$$

Similarly, the generating function for the semi-reduced Touchard numbers $T^*(n,k) = S_{2,0}(n,k)$ is

$$F(x,y) = \sum_{n,k\geq 0} T^*(n,k)\frac{x^n y^k}{n!} = e^{yg_{2,0}(x)} = e^{y(\cosh x-1)}. \tag{9}$$

This satisfies the partial differential equation

$$F_{xx} = y(F_y+F) + y^2(F_{yy}+2F_y). \tag{10}$$

Substituting the power series (9) into (10) gives the recurrence

$$T^*(n,k) = k^2T^*(n-2,k) + (2k-1)T^*(n-2,k-1) \text{ for } n\geq 4. \tag{11}$$

This formula yields some alternative combinatorial interpretations for $T^*(n,k)$. For example, the number of partitions of $\{1,2,\ldots,n\}$ (n even) into k blocks where $2i$

and $2i-1$ are not both minimum elements of blocks satisfies the recurrence (11). For instance the partitions when $n=4$ and $k=2$ are $(12)(34), (123)(4), (124)(3)$. We pose the following two related problems:

OPEN PROBLEM 2. *For n even, find an explicit bijection between the partitions of $\{1,2,\ldots,n\}$ into k parts whose blocks have even cardinality, and partitions into k parts where no two blocks have $2i$ and $2i-1$ as minimum elements.*

OPEN PROBLEM 3. *Find a combinatorial proof that $T^*(n,k)$ satisfies the recurrence* (11).

For reasons to appear later, the semi-reduced Touchard number $T^*(n,k)$ is divisible by $1\cdot 3\cdot 5\cdots(2k-1)$. We introduce the *reduced Touchard numbers*

$$S^*_{2,0}(n,k) = T^{**}(n,k) = \frac{T^*(n,k)}{1\cdot 3\cdot 5\cdots(2k-1)}.$$

The recurrence (11) easily yields

$$T^{**}(n,k) = k^2T^{**}(n-2,k) + T^{**}(n-2,k-1) \text{ for } n \geq 4. \tag{12}$$

From this recurrence and the initial values of $T^{**}(n,k)$, we find that the reduced Touchard numbers are equal to the *central factorial numbers* discussed in [37, p.96] (except we must remove the zero's from our sequence; that is, consider $T^{**}(2n,k)$).

Associated with $T^*(n,k)$ are what we will call the sinh numbers. As $T^*(n,k)$ enumerates partitions of an n-set into k parts with even cardinality, $S_{2,1}(n,k)$ enumerates partitions of an n-set into k parts with odd cardinality. The generating function

$$F(x,y) = \sum_{n,k\geq 0} S_{2,1}(n,k)\frac{x^ny^k}{n!} = e^{yg_{2,1}(x)} = e^{y\sinh x}$$

satisfies the partial differential equation

$$F_{xx} = yF_y + y^2(F_{yy}+F_y). \tag{13}$$

This leads to in the recurrence

$$S_{2,1}(n,k) = k^2S_{2,1}(n-2,k) + S_{2,1}(n-2,k-2) \text{ for } n \geq 4.$$

It seems reasonable that there might exist linear recurrences like (8) and (11), for all values of m. After all, $S_{m,a}(n,k)$ has a combinatorial interpretation analogous to those of $S(n,k)$ and $T^*(n,k)$. However, we show in the next theorem that there are no recurrences of the form (8) and (11) for $m \geq 3$. This will later be strengthened by Theorem 6.

THEOREM 4. *If $m \geq 3$, then $F(x,y) = e^{yg_{m,0}(x)}$ does not satisfy a partial differential equation of the form*

(14)

$$\frac{\partial^m F(x,y)}{\partial x^m} = p_0(y)F + p_1(y)\frac{\partial F(x,y)}{\partial y} + p_2(y)\frac{\partial^2 F(x,y)}{\partial y^2} + \cdots + p_m(y)\frac{\partial^m F(x,y)}{\partial y^m},$$

where the $p_i(y)$ are polynomials of degree $\leq m$.

PROOF. Suppose on the contrary that an equation of the form (14) holds. Note that

$$\frac{\partial^n F(x,y)}{\partial x^n} = (g_{m,0}(x))^n F(x,y) \quad \text{and} \quad \frac{\partial^n F(x,y)}{\partial y^n} = Y_n(yf', yf'', yf''', \dots)F(x,y).$$

Equating coefficients of y^m in (14) yields

$$(g'_{m,0}(x))^m = a_0 + a_1 g_{m,0}(x) + a_2 g_{m,0}(x)^2 + \cdots + a_m g_{m,0}(x)^m \tag{15}$$

for constants $a_0, a_1, a_2, \dots, a_m$. When $m \geq 2$, $g'_{m,0}(0) = 0$ and thus $a_0 = 0$. Now

$$g'_{m,0}(x) = \frac{e^x + \zeta e^{\zeta x} + \zeta^2 e^{\zeta^2 x} \cdots + \zeta^{m-1} e^{\zeta^{m-1} x}}{m}.$$

Equating the coefficients of e^{mx} on both sides of (15) gives

$$\frac{1}{m^m} = \frac{a_m}{m^m},$$

so $a_m = 1$. Now we equate the coefficients of the term $e^{(m-1+\zeta)x}$. For $m \geq 3$ we have $\zeta \neq \pm 1$, so this term occurs only once on each side of equation (15), and we get

$$\frac{m\zeta}{m^m} = \frac{m a_m}{m^m}.$$

This yields $\zeta = 1$, a contradiction. □

In what follows, the term "recurrence" means "homogeneous linear recurrence". We now show that there is a recurrence in n for $S_{m,a}(n,k)$ when m, a and k are fixed. We will also obtain a practical method for computing such recurrences and gain further insight as to why a recurrence in n of fixed order does not exist for $m \geq 3$.

THEOREM 5. *For fixed k there are constants $b_1, b_2, \dots, b_d$ and such that*

$$S_{m,a}(n,k) + \sum_{i=1}^{d} b_i S_{m,a}(n-i,k) \equiv 0.$$

PROOF. It is clear that

$$\frac{d^m g_{m,a}(x)}{dx^m} = g_{m,a}(x) + \delta_{a,0}. \tag{16}$$

For $a \not\equiv 0 \pmod{m}$, let $C_{m,a}(k)$ be the set of all sums of k terms from the set $\{1, \zeta, \dots, \zeta^{m-1}\}$, and let $C_{m,0}(k)$ be the set of all sums of k terms from the set $\{0, 1, \zeta, \dots, \zeta^{m-1}\}$. For example,

$$C_{3,0}(2) = \{0,1,2,\zeta,2\zeta,\zeta^2,2\zeta^2,1{+}\zeta,1{+}\zeta^2,\zeta{+}\zeta^2\} = \{0,1,2,\zeta,2\zeta,\zeta^2,2\zeta^2,-\zeta^2,-\zeta,-1\}.$$

For k fixed we have

$$\frac{g_{m,a}(x)^k}{k!} = \begin{cases} \left(\left(e^x + e^{\zeta x} + \cdots + e^{\zeta^{m-1}x}/m\right) - 1\right)^k/k! & \text{if } a \equiv 0 \pmod{m}, \\ \left(e^x + \zeta^{-a}e^{\zeta x} + \cdots + \zeta^{-a(m-1)}e^{\zeta^{m-1}x}/m\right)^k/k! & \text{if } a \not\equiv 0 \pmod{m}, \end{cases}$$

so $g_{m,a}(x)^k$ is a linear combination of exponentials e^{cx} where $c \in C_{m,a}(k)$. Since

$$\left.\frac{d^n e^{cx}}{dx^n}\right|_{x=0} = c^n,$$

it follows that $S_{m,a}(n,k)$ is a linear combination of terms c^n, where $c \in C_{m,a}(k)$. Hence $S_{m,a}(n,k)$, considered as a function of n (with k fixed) satisfies the linear homogeneous recurrence whose characteristic polynomial is

$$P_{k;m,a}(x) = \prod_{c \in C_{m,a}(k)} (x - c).$$

In other words, if b_i is as the coefficient of x^{d-i} in the polynomial $P_{k;m,a}(x)$ of degree $d = |C_{m,a}(k)|$, we have

$$\sum_{i=0}^{d} b_i S_{m,a}(n-i,k) \equiv 0$$

as the desired recurrence. □

PROPOSITION 5. *The absolute values of the non-zero coefficients b_i of the characteristic polynomial $P_{k;m,a}(x)$ are unimodal for $m \le 2$ and non–unimodal for $m = 3$.*

PROOF. It is easily seen that if $0 \notin C_{m,a}(k)$, then $b_i = 0$ unless $m|i$. Moreover,

$$P_{k;m,a}(x) = \prod_{c \in C_{m,a}(k)} (x - c) = \sum_{j=0}^{\ell} (-1)^j e_{\ell-j}(r_1, r_2, \dots, r_d) x^{mj},$$

where $\ell = d/m$, where $r_1, \dots, r_\ell$ are the distinct m'th powers of the elements of $C_{m,a}(k)$, and $e_i(r_1, r_2, \dots, r_\ell)$ is the i'th elementary symmetric function.

On the other hand if $0 \in C_{m,a}(k)$, then $b_i = 0$ unless $i \equiv 1 \pmod{m}$, and

$$P_{k;m,a}(x) = \sum_{j=0}^{\ell} (-1)^j e_{\ell-j}(r_1, r_2, \dots, r_\ell) x^{mj+1},$$

where $\ell = (d-1)/m$. If $m \le 2$, then $r_1, r_2, \dots$ are non-negative and hence the following well known inequality holds (see [30]):

$$e_{i-1}/e_{i-2} \ge e_i/e_{i-1}.$$

This implies that the sequence $e_0, e_1, \ldots, e_d$ is logarithmically concave and therefore unimodal. When $m = 3$, we note that

$$P_{2;3,0}(x) = 8 - x^4 - 8x^7 + x^{10},$$

so the absolute values of the non-zero coefficients $P_{2;3,0}(x)$ are not unimodal. $\square$

CONJECTURE 1. *The absolute values of the non-zero coefficients b_i of $P_{k;m,a}(x)$ are monotone increasing for $m \geq 4$*

It is interesting to note that for $m = 3, 4, 6$, the set $\bigcup_{k=0}^{\infty} C_{m,a}(k)$ is a lattice in the complex plane. The cases $m = 1, 2$ are degenerate, and for $m = 5$ and $m > 6$, $\bigcup_{k=0}^{\infty} C_{m,a}(k)$ is everywhere dense. In general the convex hull of $C_{m,0}(k)$ is a regular m-gon. Because $C_{m,0}(k-1) \subset C_{m,0}(k)$, it is clear that $P_{k-1;m,0}(x)$ divides $P_{k;m,0}(x)$. If $|C_{m,0}(k) - C_{m,0}(k-1)|$ is bounded, it is conceivable that there might be a recurrence in (n, k) for $S_{m,0}(n, k)$ of fixed order in n, with coefficients depending on k as in equations (8), (11). To be specific, consider $m = 2$ (the Touchard case). Here we have $C_{2,0}(k) = \{-k, -k+1, \ldots, k-1, k\}$, which gives the characteristic polynomials

$$x(x^2-1),\ x(x^2-1)(x^2-4),\ x(x^2-1)(x^2-4)(x^2-9), \ldots .$$

These yield the recurrences

$$\begin{aligned}
S_{2,0}(n,1) &= S_{2,0}(n-1,1),\ n \geq 2, \\
S_{2,0}(n,2) &= 5S_{2,0}(n-1,2) - 4S_{2,0}(n-2,2),\ n \geq 3, \\
S_{2,0}(n,3) &= 14S_{2,0}(n-1,3) - 49S_{2,0}(n-2,3) + 36S_{2,0}(n-3,3),\ n \geq 4, \\
&\vdots
\end{aligned}$$

The orders of these recurrences increase with k, in contrast with (8) and (11).

The sinh numbers satisfy the same recurrence as the semi-reduced Touchard numbers. However these are in general not minimal. Indeed for $k = 1, 2, 3, \ldots$ the polynomials $P_{k;2,1}(x)$ are

$$x(x^2-1),\ x(x^2-4),\ x(x^2-1)(x^2-9), \ldots ,$$

which yield the recurrences

$$\begin{aligned}
S_{2,1}(n,1) &= S_{2,0}(n-2,1),\ n \geq 2, \\
S_{2,1}(n,2) &= 4S_{2,0}(n-2,2),\ n \geq 3, \\
S_{2,1}(n,3) &= 10S_{2,0}(n-2,3) - 9S_{2,0}(n-4,3),\ n \geq 4, \\
&\vdots
\end{aligned}$$

For $k \geq 2$, these are of smaller order than the recurrences for $S_{2,0}(n, k)$. Note that the k'th polynomial does not divide the $(k+1)$'th. Although the orders of these recurrences increase with k, there is still the recurrence (13) of fixed order in n.

THEOREM 6. *For $m \geq 3$ there is no non-trivial recurrence in (n,k) for $S_{m,a}(n,k)$ of fixed order in n with coefficients depending on k.*

PROOF. If such a recurrence did exist, then

$$\sum_{i=0}^{A}\sum_{j=0}^{k} g_{ij}(k) S_{m,a}(n-i,k-j) = 0$$

where A is a constant and $g_{ij}(k)$ is a function of k. Separating out the terms in $S_{m,a}(n-i,k)$ gives

$$\sum_{i=0}^{A} g_{i0}(k) S_{m,a}(n-i,k) = -\sum_{i=0}^{A}\sum_{j=1}^{k} g_{ij}(k) S_{m,a}(n-i,k-j). \tag{17}$$

Note that the homogeneous linear difference equation associated with the left hand side of (17) is

$$\sum_{i=0}^{A} g_{i0}(k) S_{m,a}(n-i,k) = 0. \tag{18}$$

For fixed k this has constant coefficients and finite order A in n; its solution in the complementary function of (17). According to the classical theory of linear difference equations, (17) has a particular solution which is a linear combination of exponentials e^{cn} with $c \in \bigcup_{j=0}^{k-1} C_{m,a}(j)$ and coefficients which are polynomials in n. The general solution is the sum of this solution and the complementary function. This contradicts the fact that for $m \geq 3$, we have $|C_{m,a}(k) \setminus \bigcup_{j=0}^{k-1} C_{m,a}(j)| \geq mk > A$ for sufficiently large k. $\square$

As an example of recurrences for $S_{m,a}(n,k)$ which are not of fixed order in n, consider the case $m=3, a=0$. For $k = 1,2,3,\ldots$ we have the polynomials $P_{k;3,0}$:

$$x(x^3-1),\ x(x^3-1)(x^3+1)(x^3-8),\ x(x^3-1)(x^3+1)(x^3-8)(x^3-27)(x^6+27),\ldots$$

These yield the recurrences

$$\begin{aligned}
S_{3,0}(n,1) &= S_{2,0}(n-3,1),\\
S_{3,0}(n,2) &= 8S_{2,0}(n-3,2) + S_{2,0}(n-6,2) - 8S_{2,0}(n-9,2),\\
S_{3,0}(n,3) &= 35S_{2,0}(n-3,3) - 242S_{2,0}(n-6,3) + 910S_{2,0}(n-9,3)\\
&\quad -5589S_{2,0}(n-12,3) - 945S_{2,0}(n-15,3) + 5832S_{2,0}(n-18,3),\\
&\ \ \vdots
\end{aligned}$$

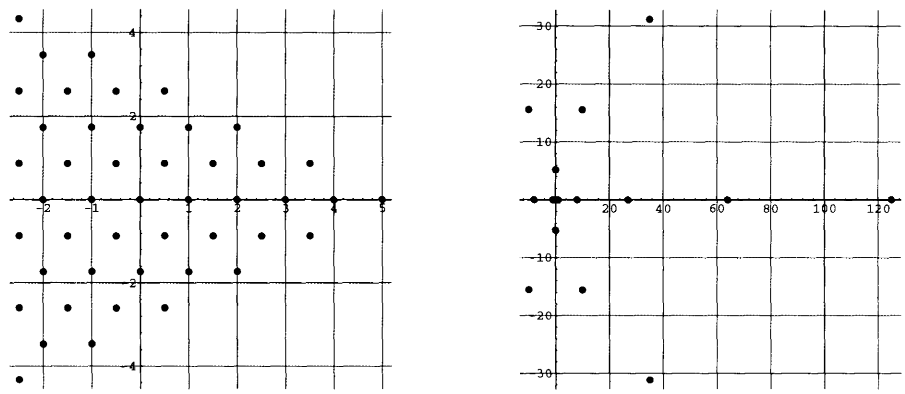
Note that the orders of these recurrences are not fixed, but grow with k.

We can produce further recurrences involving $S_{m,a}(n-i,k-j)$ by exploiting the recurrences for smaller values of k. For example, the new factors that enter into the polynomial $P_{k;3,0}(x)$ for $k=2$ are $(x^3-8)(x^3+1) = x^6 - 7x^3 - 8$. This leads to

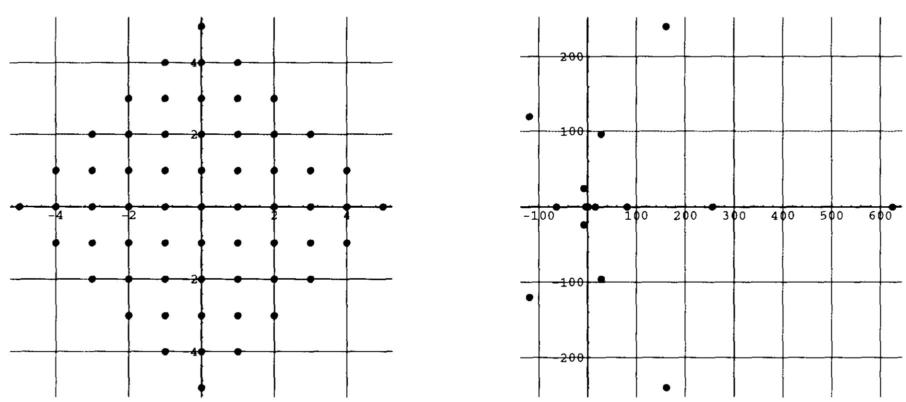
$$S_{3,0}(n,2) = 7S_{2,0}(n-3,2) + 8S_{2,0}(n-6,2) + 14S_{2,0}(n-6,1) + 10S_{2,0}(n-3,0).$$

Here are some plots of the sets $C_{m,0}(5)$ and $C_{m,0}(5)^m = \{c^m | c \in C_{m,0}(5)\}$:

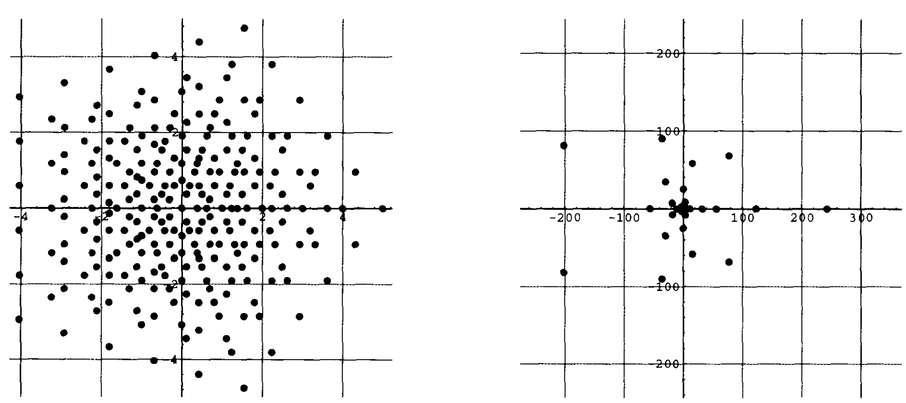
$C_{m,0}(5)$ $\qquad\qquad$ $C_{m,0}(5)^m$



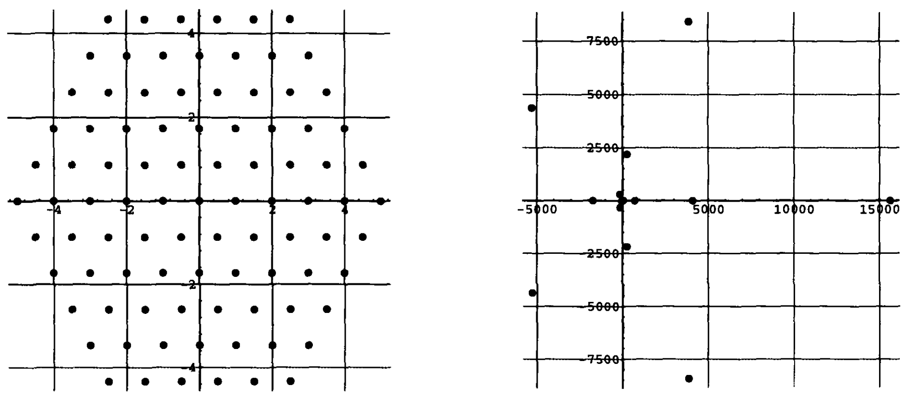
$m=3$



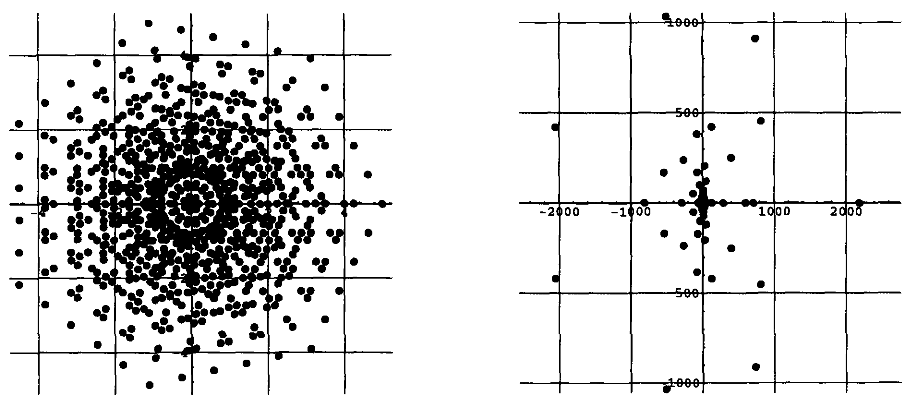
$m=4$



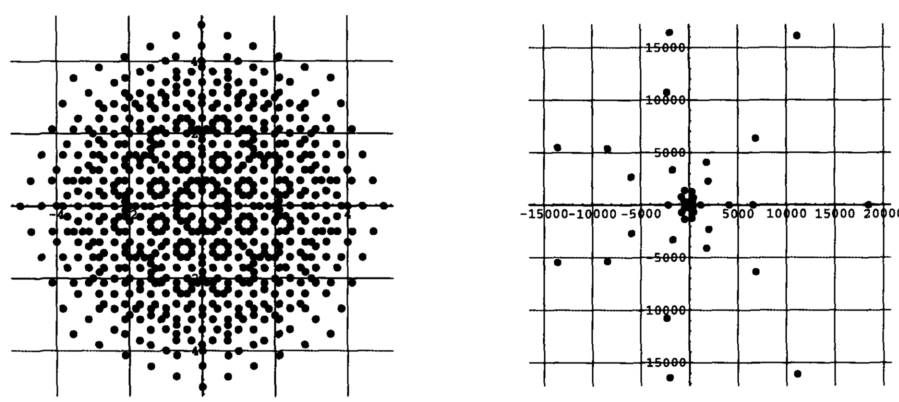
$m=5$



$m = 6$



$m = 7$



$m = 8$

We now consider congruential properties of the generalized Stirling numbers $S_{m,a}(n,k)$. We prove:

THEOREM 7. *Let p be a prime with $p \equiv 1 \pmod m$. Then*

(19)

$$S_{m,a}(n+p,k) \equiv \begin{cases} S_{m,a}(n+1,k) \pmod p & \text{if } a \not\equiv 1 \pmod m, \\ S_{m,1}(n+1,k) + S_{m,1}(n,k-p) \pmod p & \text{if } a \equiv 1 \pmod m, \end{cases}$$

PROOF. $S_{m,a}(n,k)$ enumerates partitions of an n-set into k parts of sizes $\equiv a \pmod{m}$. Call the set of such partitions $\mathcal{P}$. Consider the action of the group G generated by the permutation $(n+1, n+2, \ldots, n+p)$ on $\mathcal{P}$. This decomposes $\mathcal{P}$ into orbits of sizes p or 1 (since G has prime order p). Hence

$$S_{m,a}(n+p,k) \equiv \text{number of elements of } \mathcal{P} \text{ fixed by } G \pmod{p}.$$

(See [16, p. 40] for details on group actions.) The fixed points are of two types:

(1) $n+1, n+2, \ldots, n+p$ all in the same block,

(2) $n+1, n+2, \ldots, n+p$ all singletons.

To count the fixed points of type (1), we observe that for them,

$$A = \{n+1, \cdots, n+p\}$$

can be treated as a single element of an $(n+1)$–set $\{1, 2, \ldots, n, A\}$. (Here we have used the fact that $|A| = p \equiv 1 \pmod{m}$.) Therefore the number of fixed points of type (1) is $S_{m,a}(n+1,k)$. Fixed points of type (2) can occur only if $a \equiv 1 \pmod{m}$, and in that case the number of them is $S_{m,a}(n, k-p)$. The theorem follows from this. $\square$

By summing (19) on k, we get:

COROLLARY 6. *Let p be a prime with $p \equiv 1 \pmod{m}$. Then*

$$B_{m,a}(n+p) \equiv \begin{cases} B_{m,a}(n+1) \pmod{p} & \text{if } a \not\equiv 1 \pmod{m}, \\ B_{m,1}(n+1) + B_{m,1}(n) \pmod{p} & \text{if } a \equiv 1 \pmod{m}, \end{cases}$$

The second congruence with $m = 1$ is the classical Bell number congruence. The others are believed to be new. The above readily shows that the period of $B_{m,1}(n) \pmod{p}$ for $p \equiv 1 \pmod{m}$ divides

$$\frac{p^p - 1}{p - 1},$$

and that of $B_{m,a}(n)$ with $a \not\equiv 1 \pmod{m}$ divides $p - 1$.

In conclusion, we note that when $m = 1$, one way to define the Stirling numbers $s(n,k)$ of the first kind is by the condition that

$$\sum_{\ell} S(n,\ell) s(\ell,k) = \begin{cases} 1 & n = k \\ 0 & n \neq k \end{cases}$$

We will investigate this in more detail later. Along these lines we also note that it is no coincidence that $s(k+1,\ell)$ is the coefficient of x^ℓ in

$$P_{k;1,0}(x) = (x)_{k+1} = x(x-1)(x-2)\cdots(x-k).$$

Define $S^*_{1,0}(n,k) = S_{1,0}(n,k) = S(n,k)$ and $s^*_{1,0}(n,k) = s(n,k)$. We present a method which can be used to define general reduced Stirling numbers $s^*_{m,0}(n,k)$ of the first kind. For clarity, we discuss the case $m = 2$, but it will be clear how this example can be extended to the general case. Set $s^*_{2,0}(n,k) = 0$ if k is odd. For k even, take two copies of the set $\{1, 2, \ldots, n\}$, say $\mathcal{N} = \{1, 2, \ldots, n\}$, and $\mathcal{N}' = \{1', 2', \ldots, n'\}$. Decompose $\mathcal{N}$ into $k/2$ non-empty blocks and $\mathcal{N}'$ into the corresponding blocks with the primes. Construct two permutations with the blocks as cycles, independently for the primes and unprimes. If the part sizes of a such pair of partitions π are $n_1, n_2, \ldots, n_{k/2}$, then the number of pairs of permutations is

$$w(\pi) = \left[\prod_{i=1}^{k/2} (n_i - 1)!\right]^2$$

We define $d^*_{2,0}(n,k)$ to be the total number of such pairs of permutations:

$$d^*_{2,0}(n,k) = \sum_{\pi} w(\pi),$$

and we define

$$s^*_{2,0}(n,k) = (-1)^{n+k} d^*_{2,0}(n,k).$$

For $m = 1$ or 2, there is a close connection between $S^*_{m,0}(n,k)$ and $s^*_{m,0}(n,k)$, namely:

PROPOSITION 6. *If $m = 1$ or 2, then $s^*_{m,0}(k,\ell)$ is the coefficient of $x^{m\ell+1}$ in $P_{k;m,0}(x)$. Moreover*

$$\sum_{\ell} S^*_{m,0}(n,\ell) s^*_{m,0}(\ell,k) = \begin{cases} 1 & n = k \equiv \pmod{m} \\ 0 & \text{otherwise} \end{cases}.$$

PROOF. As noted earlier, $P_{k;1,0}(x) = (x)_{k+1}$; hence for $m = 1$ the result follows from the well-known identities

$$(x)_n = \sum_{k=0}^{n} s(n,k) x^k,$$

and

$$x^n = \sum_{k=0}^{n} S(n,k)(x)_k,$$

where $s(n,k) = s_{1,0}(n,k)$, and $S(n,k) = S_{1,0}(n,k)$ are the Stirling numbers of the first and second kind respectively.

Now suppose $m = 2$ and consider $d^*_{2,0}(k,\ell) = |s^*_{2,0}(k,\ell)|$. By definition, $d^*_{2,0}(k,\ell)$ counts the total number of pairs of permutations of $\{1, 2, \ldots, n\}$ and $\{1', 2', \ldots, n'\}$. Consider the elements n, and n' and count the total number of ways they can be placed in such a pair. Each of n and n' is either in a cycle containing other numbers or in a cycle by itself. In the former case, n can be placed after any of the elements

$\{1, \ldots, n-1\}$ in the cycle decomposition of a pair enumerated by $d^*_{2,0}(n-1,k)$, and similarly for n'; a total of $(n-1)^2 d^*_{2,0}(n-1,k)$ places. In the latter case, n and n' are fixed points and so there are a total of $d^*_{2,0}(n-1,k-2)$ such pairs. This yields the relation

$$d^*_{2,0}(k,\ell) = (k-1)^2 d^*_{2,0}(k-1,\ell) + d^*_{2,0}(k-1,\ell-2).$$

Note that $d^*_{2,0}(0,0) = 1$ and $d^*_{2,0}(k,\ell) = 0$ if $k \leq 0$ or $\ell \leq 0$. Set

$$p_k(x) = x(x^2+1)\cdots(x^2+(k-1)^2) = \sum_{\ell=0}^{k} c(k,\ell)x^{2\ell}. \tag{20}$$

We have $c(0,0) = 1$ and $c(k,\ell) = 0$ if $k \leq 0$ or $\ell \leq 0$. Now

$$p_k(x) = (x^2+(k-1)^2)p_{k-1}(x) = \sum_{\ell=1}^{k} c(k-1,\ell-2)x^{2\ell} + (k-1)^2 \sum_{\ell=0}^{k-1} c(k-1,\ell)x^{2\ell},$$

which yields the recurrence

$$c(k,\ell) = c(k-1,\ell-2) + (k-1)^2 c(k-1,\ell).$$

This is the same as the the recurrence for $d^*_{2,0}(k,\ell)$ with the same initial conditions; hence $c(k,\ell) = d^*_{2,0}(k,\ell)$. Substituting $-x^2$ for x^2 and multiplying by $(-1)^{k-1}$ in (20) gives

$$\sum_{\ell=0}^{k} s^*_{2,0}(k,\ell)x^{m\ell+1} = x(x^2-1)(x^2-4)\cdots(x^2-(k-1)^2).$$

It is clear from the definition of $C_{2,0}(k)$ that

$$P_{k;2,0}(x) = x(x^2-1)(x^2-4)\cdots(x^2-k^2),$$

and thus

$$P_{k;2,0}(x) = \sum_{\ell=0}^{k} s^*_{2,0}(k,\ell)x^{m\ell+1}.$$

Recall that $S_{2,0}(n,k) = T^*(n,k)$, the semi–reduced Touchard number, and hence $S^*_{2,0}(n,k) = T^{**}(n,k)$. We will prove that

$$x^{k+1} = \sum_{\ell=0}^{k} T^{**}(k,\ell)P_{\ell;2,0}(x) \quad \text{for } k \text{ even,} \tag{21}$$

from which the result follows. With the obvious convention that $P_{0;2,0}(x) = x$, (21) holds for $k=0$. Now assume that $k \geq 2$ and that (21) holds for $k-2$. Then

$$\begin{aligned} x^{k+1} &= x^2 \cdot x^{k-1} \\ &= x^2 \sum_{\ell=0}^{k-1} T^{**}(k-2,\ell)P_{\ell;2,0}(x) = \sum_{\ell=0}^{k-1} T^{**}(k-2,\ell)P_{\ell;2,0}(x)(x^2-\ell^2+\ell^2) \\ &= \sum_{\ell=0}^{k-1} T^{**}(k-2,\ell)P_{\ell;2,0}(x)(x^2-\ell^2) + \sum_{\ell=0}^{k-1} \ell^2 T^{**}(k-2,\ell)P_{\ell;2,0}(x). \end{aligned}$$

Using $T^{**}(k-2,k-1) = T^{**}(k-2,-1) = 0$ and shifting indices, we get

$$\begin{aligned} x^k &= \sum_{\ell=0}^{k} T^{**}(k-2,\ell-1)P_{\ell-1;2,0}(x)(x-\ell^2) + \sum_{\ell=0}^{k-1} \ell^2 T^{**}(k-2,\ell)P_{\ell;2,0}(x) \\ &= \sum_{\ell=0}^{k} T^{**}(k-2,\ell-1)P_{\ell;2,0}(x) + \sum_{\ell=0}^{k} \ell^2 T^{**}(k-2,\ell)P_{\ell;2,0}(x). \end{aligned}$$

Here we have used the fact that $P_{\ell-1;2,0}(x)(x-\ell^2) = P_{\ell;2,0}(x)$. From (12) we have

$$T^{**}(k,\ell) = \ell^2 T^{**}(k-2,\ell) + T^{**}(k-2,\ell-1),$$

and the result follows. $\square$

5. Further Properties of δ_n

In previous sections we have seen the significance of the quantities Δ_n and δ_n and mentioned a few their properties. We now explore them in greater detail. For this purpose we consider $\Delta_n(t) = \det[tx_{i+j}]_{i,j=0}^n$ (except for $i = j = 0$, in which case the matrix element is 1) and $\delta_n(t) = \det[Y_{i+j}(tx_1, tx_2, tx_3, \dots)]_{i,j=0}^n$, where t is a real variable. Here are the first few values of $\Delta_n(t)$ and $\delta_n(t)$:

$$\begin{aligned} \Delta_0(t) &= 1 \\ \Delta_1(t) &= tx_2 - t^2x_1^2 \\ \Delta_2(t) &= t^2(x_2x_4 - x_3^2) + t^3(2x_1x_2x_3 - x_1^2x_4 - x_2^3) \\ \Delta_3(t) &= t^3(x_2x_4x_6 - x_3^2x_6 - x_2x_5^2 + 2x_3x_4x_5 - x_4^3) + t^4(x_3^4 - 3x_2x_3^2x_4 + x_2^2x_4^2 + x_1^2x_5^2 \\ &\quad +2x_1x_3x_4^2 + 2x_2^2x_3x_5 - 2x_1x_3^2x_5 - 2x_1x_2x_4x_5 - x_2^3x_6 + 2x_1x_2x_3x_6 - x_1^2x_4x_6) \\ \delta_0(t) &= 1 \\ \delta_1(t) &= tx_2 \\ \delta_2(t) &= t^2(x_2x_4 - x_3^2) + t^3(2x_2^3) \\ \delta_3(t) &= t^3(x_2x_4x_6 - x_3^2x_6 - x_2x_5^2 + 2x_3x_4x_5 - x_4^3) + t^4(12x_2x_3^2x_4 - 9x_3^4 + 7x_2^2x_4^2 \\ &\quad -12x_2^2x_3x_5 + 2x_2^3x_6) + t^5 24x_2^3(x_2x_4 - x_3^2) + t^6(12x_2^6) \end{aligned}$$

PROPOSITION 7. *The polynomial $\delta_n(t)$ has lower degree n.*

PROOF. We recall that $Y_n(t) = tx_n + O(t^2)$ as $t \to 0$. Clearly the conclusion holds for $\delta_0(t)$. Now assume it holds for δ_{n-1}. By Laplace expansion on the last row, $\delta_n(t) = (tx_n + O(t^2))\delta_{n-1}(t)$ as $t \to 0$ and the result follows. $\square$

PROPOSITION 8. *The polynomial $\Delta_n(t)$ has lower degree n, and the coefficients of t^n in $\Delta_n(t)$ and $\delta_n(t)$ are equal.*

PROOF. The first part follows from proposition 7. Since $Y_n(t) \sim tx_n$ as $t \to 0$, it follows that $\delta_n(t) \sim \Delta_n(t)$ as $t \to 0$. This implies that the lowest order non-zero terms of $\delta_n(t)$ and $\Delta_n(t)$ are equal. $\square$

COROLLARY 7. *If $\delta_n(t) \geq 0$ for sufficiently $t \geq 0$, then $\Delta_n(t) \geq 0$ for all sufficiently small $t \geq 0$.*

Note that for fixed t the corollary is false. For example if $x_1 = x_2 = 1$, $x_3 = x_4 = 2$, then $\delta_2(1) = 0$ but $\Delta_2(1) = -1$. The corollary states that if $(Y_n(tx_1, tx_2, \dots))_{n\geq 1}$ is a moment sequence for all sufficiently small $t \geq 0$, then $(tx_1, tx_2, \dots)$ is a moment sequence for all sufficiently small $t \geq 0$. To discuss this in more detail, we first prove the following proposition.

PROPOSITION 9. *$\Delta_n = 0$ for all $n \geq 1$ if and only if $x_n = \lambda^n$ for some constant λ.*

PROOF. Clearly if $x_n = \lambda^n$ then $\Delta_n = 0$ for all $n \geq 1$, since its second row is λ times its first row. Suppose conversely that $\Delta_n = 0$ for all $n \geq 1$. Then

$$\Delta_1 = \begin{vmatrix} 1 & x_1 \\ x_1 & x_2 \end{vmatrix} = 0,$$

so putting $x_1 = \lambda$ we get $x_2 = \lambda^2$. Now suppose $n \geq 2$ and that $x_k = \lambda^k$ for $k \leq n$ has been proved. Then

$$\Delta_n = \begin{vmatrix} 1 & \lambda & \lambda^2 & \dots & \lambda^{n-1} & \lambda^n \\ \lambda & \lambda^2 & \lambda^3 & \dots & \lambda^n & x_{n+1} \\ \lambda^2 & \lambda^3 & \lambda^4 & \dots & x_{n+1} & x_{n+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \lambda^{n-1} & \lambda^n & x_{n+1} & \dots & x_{2n-2} & x_{2n-1} \\ \lambda^n & x_{n+1} & x_{n+2} & \dots & x_{2n-1} & x_{2n} \end{vmatrix} = 0. \tag{22}$$

Multiply the first row of the determinant by λ^{i-1} and subtract the result from the i'th row ($i = 2, 3, \dots, n+1$). The result is the equation

$$\begin{vmatrix} 1 & \lambda & \lambda^2 & \dots & \lambda^{n-1} & \lambda^n \\ 0 & 0 & 0 & \dots & 0 & x_{n+1} - \lambda^{n+1} \\ 0 & 0 & 0 & \dots & x_{n+1} - \lambda^{n+1} & x_{n+2} - \lambda^{n+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & x_{n+1} - \lambda^{n+1} & x_{n+2} - \lambda^{n+2} & \dots & x_{2n-1} - \lambda^{2n+1} & x_{2n} - \lambda^{2n} \end{vmatrix} = 0. \tag{23}$$

For example when $n = 3$, this equation is

$$\begin{vmatrix} 1 & \lambda & \lambda^2 & \lambda^3 \\ 0 & 0 & 0 & x_4 - \lambda^4 \\ 0 & 0 & x_4 - \lambda^4 & x_5 - \lambda^5 \\ 0 & x_4 - \lambda^4 & x_5 - \lambda^5 & x_6 - \lambda^6 \end{vmatrix}. \tag{24}$$

There is only one non–zero term in the complete expansion of the determinant (23), and it is equal to $(\lambda^{n+1} - x_{n+1})^n$. Hence $x_{n+1} = \lambda^{n+1}$, completing the induction. $\square$

PROPOSITION 10. *The polynomial $\delta_n(t)$ is of degree $\binom{n+1}{2}$ in t, and the coefficient of $t^{\binom{n+1}{2}}$ is*

$$x_2^{\binom{n+1}{2}} \prod_{i=1}^{n} i!.$$

PROOF. By Hankel mean–independence we can assume $x_1 = 0$. Then every term in the complete expansion of $\delta_n(t)$ has degree at most $\binom{n+1}{2}$ in t, and the only terms which attain this have the form $cx_2^{\binom{n+1}{2}}$. We can therefore suppose that $x_i = 0$ for all $i \neq 2$. Setting x_2 and t equal to 1, we have

$$Y_{2n}(0,1,0,\dots) = \frac{(2n)!}{2^n n!} = 1 \cdot 3 \cdot 5 \cdots (2n-1), \quad Y_{2n+1}(0,1,0,\dots) = 0.$$

We now substitute these values into the determinant $\delta_n(1) = \det[Y_{i+j}]_{i,j=0}^{n}$. For clarity, we illustrate with the case $n = 5$:

$$\delta_5(1) = \det[Y_{i+j}]_{i,j=0}^{5} = \begin{vmatrix} 1 & 0 & 1 & 0 & 3 & 0 \\ 0 & 1 & 0 & 3 & 0 & 3\cdot5 \\ 1 & 0 & 3 & 0 & 3\cdot5 & 0 \\ 0 & 3 & 0 & 3\cdot5 & 0 & 3\cdot5\cdot7 \\ 3 & 0 & 3\cdot5 & 0 & 3\cdot5\cdot7 & 0 \\ 0 & 3\cdot5 & 0 & 3\cdot5\cdot7 & 0 & 3\cdot5\cdot7\cdot9 \end{vmatrix}.$$

We first pull out the odd factors $1 \cdot 3 \cdot 5 \cdots (2\lceil \frac{i}{2} \rceil - 1)$ from the i'th row $(0 \le i \le n)$. The total factor thus pulled out is

$$\prod_{i=0}^{n} \prod_{\substack{j \le i \\ j \text{ odd}}} j.$$

In the present example we get

$$\delta_5(1) = 3^3 5 \begin{vmatrix} 1 & 0 & 1 & 0 & 3 & 0 \\ 0 & 1 & 0 & 3 & 0 & 3\cdot5 \\ 1 & 0 & 3 & 0 & 3\cdot5 & 0 \\ 0 & 1 & 0 & 5 & 0 & 5\cdot7 \\ 1 & 0 & 5 & 0 & 5\cdot7 & 0 \\ 0 & 1 & 0 & 7 & 0 & 7\cdot9 \end{vmatrix}.$$

Subtracting each row from the one two below it gives

$$3^3 5 \begin{vmatrix} 1 & 0 & 1 & 0 & 3 & 0 \\ 0 & 1 & 0 & 3 & 0 & 15 \\ 0 & 0 & 2 & 0 & 12 & 0 \\ 0 & 0 & 0 & 2 & 0 & 20 \\ 0 & 0 & 2 & 0 & 20 & 0 \\ 0 & 0 & 0 & 2 & 0 & 28 \end{vmatrix}.$$

Starting with the 3'rd row, we subtract each from the one two below it, getting

$$3^3 5 \begin{vmatrix} 1 & 0 & 1 & 0 & 3 & 0 \\ 0 & 1 & 0 & 3 & 0 & 15 \\ 0 & 0 & 2 & 0 & 12 & 0 \\ 0 & 0 & 0 & 2 & 0 & 20 \\ 0 & 0 & 0 & 0 & 8 & 0 \\ 0 & 0 & 0 & 0 & 0 & 8 \end{vmatrix} = 34560 = \prod_{k=1}^{5} k!.$$

For arbitrary n note that after the odd factors are removed from the $2i$'th and $(2i+1)$'th rows are equal to $2i+1$ rows the column has entries

$$p_j(i) = \begin{cases} \prod_{k=0}^{i-1} 2(m+i)+1 & \text{if } i = j \pmod 2, \\ 0 & \text{if } i \neq j \pmod 2, \end{cases}$$

and similarly for the $2i$'th columns. Subtracting a row from one two below it amounts to forming

$$\Delta p_j(i) = p_j(i) - p_j(i-2).$$

We have

$$\frac{p_j(i)}{p_{j-1}(i)} = 2i + 2j - 1 \quad \text{and} \quad \frac{p_j(i-1)}{p_{j-1}(i)} = 2i - 1,$$

so

$$\Delta p_j(i) = 2j p_{j-1}(i).$$

Iterating this n times is equivalent to multiply subtracting successive rows which gives

$$\Delta^j p_j(i) = 2 \cdot 4 \cdots 2j.$$

The final matrix will be upper triangular and hence the determinant is the product of the diagonal terms. This along with the factored out odd terms gives the desired result. □

We could continue our study of the coefficients of $\delta_n(t)$ in this manner. For example, we make the following conjecture:

CONJECTURE 2. *For $n \geq 3$, the coefficient of $t^{\binom{n+1}{2}-1}$ in $\delta_n(t)$ is*

$$\frac{\prod_{i=1}^{n+1} i!}{12(n-2)}(x_2x_4 - x_3^2)x_2^{\binom{n+1}{2}-3}.$$

The above suggests among other things that simpler conditions than the non-negativity of δ_n for a given sequence to be a cumulant sequence may not exist. For fixed $x_1, x_2, \dots$ with $x_2 > 0$, Proposition 10 shows that $\delta_n(t) > 0$ for all sufficiently large t.

A matrix is called *totally non-negative* if all its minors are non-negative (see [21]). In closing we prove a simple non-negativity property of the Hankel matrices $[Y_{i+j}]_{i,j=0}^n$.

PROPOSITION 11. *Δ_n and $\det[x_{i+j+1}]_{i,j=0}^n$ are non-negative if and only if $[x_{i+j}]_{i,j=0}^n$ and $[Y_{i+j}]_{i,j=0}^n$ are totally non-negative.*

PROOF. By earlier work we know that the map $(x_n) \mapsto (Y_n)$ is $\mathcal{M}$–preserving; hence $\delta_n = \det[Y_{i+j}]_{i,j=0}^n \geq 0$. Since $\det[x_{i+j+1}]_{i,j=0}^n \geq 0$ we have the x_n are the Stieltjes moments for a non-negative random variable (see [33] or [21]) and the Y_n must also be moments of a non-negative random variable as they are derived by compounding the orginal non-negative random variable. There exist by definition of discrete moments values $0 < \alpha_1 < \alpha_2 < \cdots$ and $0 < \beta_1 < \beta_2 < \cdots < 1$ such that $Y_i = \sum_{j=1}^m \alpha_j^i \beta_j$. Now define $b_{ik} = \alpha_k^i \sqrt{\beta_k}$ and $B = [b_{ij}]$. Then $[x_{i+j}] = BB^T$ and we know that the Vandermond matrix $[\alpha_j^{i-1}]$ is totally non-negative (see [21]), since $\sqrt{\beta_k} \geq 0$ we have that B and B^T are totally non-negative, and hence $[x_{i+j}]$ is totally non-negative. This same argument applies to $[Y_{i+j}]$. The converse is clear as Δ_n and $\det[x_{i+j+1}]_{i,j=0}^n$ are particular minors of $[x_{i+j}]_{i,j=0}^{n+1}$. $\square$

The assumption that $\Delta_n \geq 0$ in the above proposition can perhaps be weakened, but this alone is not enough, even if we assume Y_n is non-negative. For example, take $x_{2n} = 1$ and all other $x_{2n+1} = 0$ (these are the moments of a Bernoulli random variable whose values are $-1, 1$). Then

$$[Y_{i+j}]_{0 \leq i,j \leq 2} \geq 0 = \begin{vmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 4 \end{vmatrix}$$

has the negative 2×2 minor

$$M_{1,3} = \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} = -1.$$

Also total non-negativity of $[Y_{i+j}]$ does not imply that (x_n) is a moment sequence. To see this, take $x_1 = x_2 = 1$ and all other $x_n = 2^{n-2}$. Then $[Y_{i+j}]$ totally non-negative (see [21]), but

$$\begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \\ 1 & 2 & 4 \end{vmatrix} = -1.$$

Furthermore, if (x_n) is not a moment sequence, $[Y_{i+j}]$ can still be totally non-negative.

6. Measures on Poisson lattices

In section 3 we introduced the general compound Poisson random variable and saw that its moments are the exponential Bell polynomials. In this section we introduce the concept of a *Poisson lattice pair,* which leads to a combinatorial interpretation of a compound Poisson random variable, and explore its relation to exponential Bell polynomials. We start with a distributive lattice $\mathcal{L}$ which we can assume is contained in the power set of some set S by virtue of Birkoff's representation theorem (see [5]). We assume that $\mathcal{L}$ contains the empty set $\emptyset$; on $\mathcal{L}$ we define the concept of a *measure* as follows:

DEFINITION 7. *A* ***measure*** *on a distributive lattice* $\mathcal{L}$ *is a function* $\mu : \mathcal{L} \to \mathbb{R}^+$ *such that (i)* $\mu(A) = 0$ *if and only if* $A = \emptyset$, *and (ii)* $\mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B)$ *for all* $A, B \in \mathcal{L}$.

We fix a measure μ on $\mathcal{L}$ and say that $A \in \mathcal{L}$ is *μ–finite* if $\mu(A) < \infty$. We proceed to define a *point process* on $\mathcal{L}$ by considering another lattice $\mathcal{R}$ contained in the power set of S with the property that for all $\rho \in \mathcal{R}$, the cardinality $N(A, \rho) = |\rho \cap A|$ is finite whenever A is μ–finite. We call ρ a *rare set* and its elements *blips* (see [2, p. 287]). The pair $(\mathcal{L}, \mathcal{R})$ is called *compatible.* Let $\mathcal{B}$ be the σ–ring generated by the sets

$$S_A(n) = \{\rho : |\rho \cap A| = n\} \subset \mathcal{R},$$

where A runs through $\mathcal{L}$. We can now define the concept of a Poisson lattice pair.

DEFINITION 8. *A* ***Poisson lattice pair*** *is a pair of compatible lattices* $(\mathcal{L}, \mathcal{R})$ *contained in the power set of* S *with a measure* μ *on* $\mathcal{L}$ *and a countably additive measure* Pr *on* $\mathcal{B}$ *which satisfy the following:*

i. *If* $0 < \mu(A) < \infty$, *then* $\sum_A Pr(S_n(A) = 1$.
ii. *If* $A \in \mathcal{L}$ *and* $0 < \mu(A) < \infty$, *then* $N(A, \rho)$ *is not identically* 0.
iii. *(independence) If* A *and* B *are disjoint and* μ*–finite, then the events* $S_n(A)$ *and* $S_n(B)$ *are independent. That is,*

$$Pr(S_n(A) \cap S_m(B)) = Pr(S_n(A))\, Pr(S_m(B)).$$

iv. (μ*-invariance) If* $\mu(A) = \mu(B)$, *then* $Pr(S_m(A)) = Pr(S_m(B))$.

v. *(divisibility) Given $A \in \mathcal{L}$ and $n > 0$, there exist n pairwise disjoint sets $A_i \in \mathcal{L}$ with $A_1 \cup A_2 \cup \cdots \cup A_n = A$, and $\mu(A_i) = \mu(A)/n$.*

From this we get

THEOREM 8. *Let $(\mathcal{L}, \mathcal{R})$ be a Poisson lattice pair. Suppose $A \in \mathcal{L}$ satisfies $\mu(A) = t < \infty$. Then there is a constant λ such that*

$$G(t,z) = E[z^{N(A,\rho)}] = \sum_{n \geq 0} Pr(S_A(n))\, z^n = e^{\lambda t(f(z)-1)} = e^{-\lambda t} \sum_{n \geq 0} Y_n(t) \frac{z^n}{n!},$$

where $f(z) = \sum_{k \geq 1} p_k z^k$ is the probability generating function for a discrete distribution with $p_0 = 0$, and $Y_n(t)$ are non-negative exponential Bell polynomials.

PROOF. First note that for A and B disjoint,

$$S_{A \cup B}(n) = \bigcup_{k=0}^{n} \left(S_A(k) \bigcap S_B(n-k)\right).$$

Hence if $\mu(B) = s$, then

$$\begin{aligned} G\,(t+s,z) &= G\,(\mu(A \cup B), z) = \sum_{n=0}^{\infty} \Pr\,(S_{A \cup B}(n))\, z^n \\ &= \sum_{n=0}^{\infty} \sum_{k=0}^{n} \Pr\,(S_A(k)) \Pr\,(S_B(n-k))\, z^n \\ &= \left(\sum_{n=0}^{\infty} \Pr\,(S_A(n))\, z^n\right) \left(\sum_{n=0}^{\infty} \Pr\,(S_B(n))\, z^n\right) \\ &= G\,(t,z)\, G\,(s,z). \end{aligned} \tag{25}$$

From (25) and (v), it follows that $\log G(t+s,z) = \log G(t,z) + \log G(s,z)$ for s and t in everywhere dense subsets of an interval $[0, L]$.

If $A \subset B$, then $N(B,\rho) \leq n$ implies $N(A,\rho) \leq n$; therefore

$$\Pr(N(A,\rho) \leq n) \geq \Pr(N(B,\rho) \leq n).$$

Hence if $A \subset B$, each coefficient of the series

$$\frac{1}{1-z} G(\mu(A), z) = \sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} \Pr(N(A,\rho) = k)\right) z^n = \sum_{n=0}^{\infty} \Pr(N(A,\rho) \leq n) z^n$$

dominates the corresponding coefficient of

$$\frac{1}{1-z} G(\mu(B), z).$$

Therefore

$$G(\mu(A), z) \geq G(\mu(B), z) \qquad \text{for } z \geq 0,$$

Hence $\log G(t,z)$ is a decreasing function of t for t in the range of μ. As indicated above, these values of t are everywhere dense in an interval $[0, L]$. Hence the monotone solution $\log G(t,z)$ of Cauchy's functional equation is of the form $tC(z)$. That is,

$$G(t,z) = e^{tC(z)} \quad \text{for all } t \text{ in the range of } \mu. \tag{26}$$

Here $C(z)$ is a function of z to be determined next.

Set $\lambda = C(0)$ and

$$f(z) = \frac{C(0) - C(z)}{C(0)}.$$

From (26) we have $G(t,z) = 1 + tC(z) + o(t)$ as $t \to 0$, so

$$C(z) = \lim_{t \to 0} \frac{G(t,z) - 1}{t}.$$

Hence

$$f(z) = \lim_{t \to 0} \frac{\frac{G(t,0)-1}{t} - \frac{G(t,z)-1}{t}}{\frac{G(t,0)-1}{t}} = \lim_{t \to 0} \frac{G(t,z) - G(t,0)}{1 - G(t,0)}.$$

By Abel's continuity theorem and the fact that $G(t,z)$ is uniformly continuous in t, the above expression is continuous at $z = 1$ and is the limit as $t \to 0$ of the probability generating functions for

$$p_k(t) = \Pr(N(A,\rho) = k | N(A,\rho) > 0).$$

Therefore $f(z)$ is a probability generating function by the continuity theorem for probability generating functions (see [14]). Let

$$f(z) = \sum_{k=0}^{\infty} p_k z^k, \tag{27}$$

where

$$p_k = \lim_{t \to 0} p_k(t) = \lim_{\mu(A) \to 0} \Pr(N(A,\rho) = k | N(A,\rho) > 0);$$

then $p_0 = f(0) = 0$. We have $C(z) = \lambda(1 - f(z))$, so (26) becomes

$$G(t,z) = e^{\lambda t (f(z) - 1)}.$$

The substitution $x_n = \lambda p_n n!$ gives

$$G(t,z) = e^{-\lambda t} \sum_{n \geq 0} Y_n(t) \frac{z^n}{n!}. \tag{28}$$

Since the x_n are non-negative, $Y_n(t)$ is non-negative by Proposition 3. $\square$

Note that Theorem 8 describes all measures $\Pr(\cdot)$ which can be defined on a Poisson lattice pair, that is measures satisfying (i)–(v) (this also may characterize all measures on Poisson lattice pairs; see open problem 4). In particular if $A \in \mathcal{L}$, the distribution of $N(A, \rho)$ is given by

$$\mathrm{Pr}(N(A, \rho) = n) = e^{-\lambda\mu(A)} \frac{Y_n(\mu(A))}{n!}. \tag{29}$$

We recognize $G(t, z)$ as the probability generating function for a compound Poisson random variable as defined in section 3. Consider $N(t)$, a Poisson random variable with parameter λ, and $Z = \sum_{k=1}^{N(t)} X_k$, where X_k are iid random variables independent of $N(t)$. We assume that $\mathrm{Pr}(Z = 0) = 0$; then

$$\begin{aligned} \mathrm{Pr}(Z = n) &= \mathrm{Pr}(\sum_{k=1}^{N(t)} X_k = n) \\ &= \sum_{m \geq 0} \mathrm{Pr}(\sum_{k=1}^{m} X_k = n | N(t) = m) \mathrm{Pr}(N(t) = m) \\ &= \sum_{m \geq 0} \mathrm{Pr}(\sum_{k=1}^{m} X_k = n) e^{-\lambda t} \frac{(\lambda t)^m}{m!}. \end{aligned}$$

Thus we can write the above in terms of equation (29) and interpret Poisson lattices as a general structure for compound Poisson processes.

It is natural to ask whether every non-negative integer-valued random variable with a compound Poisson distribution arises in this way. We pose this formally as an open problem:

OPEN PROBLEM 4. *For every sequence of non-negative exponential Bell polynomials $(Y_n(t))$, does there exist a corresponding Poisson lattice?*

The dual role that the exponential Bell polynomials play as moments and probabilities is noteworthy. In section 3 we saw how the the exponential Bell polynomials arise as *moments* of a suitable limit of a sum of random variables. On the other hand, the derivation of the exponential Bell polynomials as *probabilities,* using a Poisson lattice pair, is reminiscent of the way in which the Poisson distribution arises as a limit of the binomial distribution. In some sense the limit of the moment sequence of a sum of independent random variables comes from a distribution defined on a Poisson lattice. It would be interesting to find a sequence of what might be called binomial lattices whose limit is a given Poisson lattice. Later we will explore the poset structure of more general Bell polynomials and a general duality principle.

CHAPTER II

Ordinary Bell polynomials

Although the exponential Bell polynomials arise naturally, encompass many well-known polynomial sequences, and have noteworthy probabilistic properties, they are by no means the only class of polynomials with such distinctions. In this chapter we shall explore another class of polynomials, the *ordinary Bell polynomials,* which have similar properties, and are perhaps even more basic in nature.

Equation (2) suggests that the exponential Bell polynomials are useful when dealing with k-fold products of exponential generating functions. If we desire instead to use ordinary generating functions, it is natural to study $\left(\sum_{n\geq 1} x_n z^n\right)^k$. Hence we make the following definition:

DEFINITION 9. *The* ***partial ordinary Bell polynomials*** $B^o_{n,k}(x_1, x_2, \dots)$ *are defined by*

$$\left(\sum_{n\geq 1} x_n z^n\right)^k = \sum_{n\geq k} B^o_{n,k}\, z^n. \tag{30}$$

It is natural to define the complete ordinary Bell polynomials as follows:

DEFINITION 10. *The* ***complete ordinary Bell polynomials*** B^o_n *are given by*

$$B^o_n = \sum_{k=1}^{n} B^o_{n,k}, \qquad B^o_0 = 1.$$

As with the exponential Bell polynomials, we will often be interested in $B^o_n(t) = B^o_n(tx_1, tx_2, \dots)$, and will discuss the case where $B^o_n(t)$ is non-negative for $t \geq 0$. We will adopt the convention that the term "Bell polynomial" always refers to the complete Bell polynomial.

1. Properties

We now derive some basic properties of the ordinary Bell polynomials analogous to those of the exponential Bell polynomials. Our first result is:

PROPOSITION 12. *If $B^o_{n,k}(x_1, x_2, \dots)$ are the partial ordinary Bell polynomials, then*

$$\exp\left(\sum_{m\geq 1} x_m z^m\right) = \sum_{n\geq 0}\sum_{k=0}^{n} B^o_{n,k}(x_1, x_2, \dots)\frac{z^n}{k!}.$$

PROOF. Multiplying 30 by $1/k!$ and summing over $k \geq 0$, we obtain the desired result. See also [10, p. 136]. □

As a consequence we have the following simple relations between partial ordinary and complete exponential polynomials:

$$\sum_{n\geq 0}\sum_{k=0}^{n} B^o_{n,k}(x_1, x_2, \dots)\frac{z^n}{k!} = \sum_{n\geq 0} Y_n(1!x_1, 2!x_2, 3!x_3, \dots)\frac{z^n}{n!}.$$

For the partial exponential Bell polynomials, there is a refinement of the convolution property in Proposition (2) (see [10, p. 136]):

PROPOSITION 13.

$$Y_{n,k}(x_1 + y_1, x_2 + y_2, \dots) = \sum_{i\leq n,\, j\leq k} \binom{n}{i} Y_{i,j}(x_1, x_2, \dots) Y_{n-i,k-j}(y_1, y_2, \dots).$$

The corresponding property for the partial ordinary Bell polynomials is the following:

PROPOSITION 14. *For $F(t) = \sum_{n\geq 1} f_n z^n$ and $G(t) = \sum_{n\geq 1} g_n z^n$, define $H(t) = \sum_{n\geq 1} h_n z^n = F \circ G(t)$. Then*

$$\sum_{n=k}^{\infty} B^o_{n,k}(h_1, h_2, \dots) z^n = \sum_{n=0}^{\infty}\sum_{\ell=k}^{n} B^o_{n,\ell}(g_1, g_2, \dots) B^o_{\ell,k}(f_1, f_2, \dots) z^n.$$

PROOF. The above is a special case of a proposition which will be proved in chapter III, Proposition 23. □

Finally we note:

PROPOSITION 15. *The partial ordinary Bell polynomials $B^o_{n,k}$ are given explicitly by*

$$B^o_{n,k}(x_1, x_2, \dots) = \sum_{\substack{i_1+2i_2+\dots+ni_n=n \\ i_1+i_2+\dots+i_n=k}} \frac{k!}{i_1!i_2!\cdots i_n!} x_1^{i_1} x_1^{i_1} \cdots x_n^{i_n}.$$

Note that the quantity $k!/(i_1!i_2!\cdots i_n!)$ is the number of k-letter words with i_ν letters equal to x_ν ($1 \leq \nu \leq n$). This gives a rapid way to compute the ordinary Bell polynomials. For example, to compute B^o_4 we the write monomials in the variables x_i associated with the partitions of 4 :

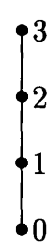
$$x_4,\ x_3x_1,\ x_2^2,\ x_2x_1^2,\ x_1^4.$$

There is only one way to write the monomial x_4; hence its coefficient is 1. The same holds for x_1^4. There are two ways to write x_3x_1, namely x_3x_1 and x_1x_3; hence we add in $2x_3x_1$. There is only one way to write x_2^2. Finally, there are three ways to write the $x_2x_1^2$ term, namely $x_1x_1x_2$, $x_1x_2x_1$, and $x_2x_1x_1$. Hence we add in $3x_2x_1^2$. The sum of these five terms is $B_4^o = x_4 + 2x_3x_1 + x_2^2 + 3x_1^2x_2 + x_1^4$. Here is a list of the first few polynomials $B_n^o(x_1, x_2, \dots)$

$$
\begin{aligned}
B_0^o &= 1 \\
B_1^o &= x_1 \\
B_2^o &= x_2 + x_1^2 \\
B_3^o &= x_3 + 2x_2x_1 + x_1^3 \\
B_4^o &= x_4 + 2x_3x_1 + x_2^2 + 3x_1^2x_2 + x_1^4 \\
B_5^o &= x_5 + 2x_4x_1 + 2x_3x_2 + 3x_3x_1^2 + 3x_2^2x_1 + 4x_2x_1^3 + x_1^5 \\
B_6^o &= x_6 + 2x_5x_1 + x_3^2 + 2x_4x_2 + 3x_4x_1^2 + 6x_3x_2x_1 + x_2^3 + 4x_3x_1^3 \\
&\quad + 6x_2^2x_1^2 + 5x_2x_1^4 + x_1^6
\end{aligned}
$$

2. Recurrent events and the renewal equation

As noted in the last section, the coefficients of the ordinary Bell polynomials B_n^o can be interpreted as the number of ways of writing the monomials associated to each partition of n. This can be thought of as a process on a chain of length n in which we want to count the total number of ways of jumping from the minimal element to the maximal element. Consider for example a chain of length 3 :



To get to 3 from 0 we can make a direct jump of 3, which gives the term x_3. Alternatively we could make a jump of 1, then a jump of 2, or a jump of 2, then a jump of 1; this gives the term $2x_2x_1$. Finally, we can jump one unit three times; this gives the term x_1^3. We interpret this as a probability model by considering the event E_n that we travel through an interval of length n one or more jumps. The events E_n are clearly recurrent, since each element is minimal with respect to the elements above it. In fact, the entire poset above that element looks like the original poset; hence once the event E_n has occurred, the process starts over again and so the event E_n is recurrent. We define $g_n = \Pr(E_n)$ and we know from prior discussion that E_n is the disjoint union of products of the events J_k, where J_k is the event of making a direct jump of k. Letting $f_n = \Pr(J_k)$, we have:

THEOREM 9 (RENEWAL EQUATION). *For a recurrent event E, let $f_k = Pr(E$ first occurs at time $n)$ and $g_n = Pr(E$ occurs at time $n)$. Then*

$$g_n = \sum_{k=0}^{n-1} g_n f_{n-k} + \delta_{n,0},$$

where $\delta_{n,0}$ is the Kronecker delta.

A proof of this result can be found in [14]. In a more compact form, we see that if $F(z) = \sum_{n\geq 0} f_n z^n$ and $G(z) = \sum_{n\geq 0} g_n z^n$, then the above relation between f_n and g_n can be written in the form

$$G(z) = F(z)G(z) + 1.$$

Now we are prepared for the following result:

PROPOSITION 16. *If $G(z) = F(z)G(z)+1$ with $F(z)$ as above, then $g_n = B_n^o(f_1, f_2, \dots)$.*

PROOF.

$$\begin{aligned}
\sum_{n\geq 0} g_n z^n &= G(z) = \frac{1}{1-F(z)} = \sum_{k\geq 0} F(z)^k = \sum_{k\geq 0} \left(\sum_{n\geq 1} f_n z^n \right)^k \\
&= \sum_{k\geq 0} \sum_{n\geq k} B_{n,k}^o(f_1, f_2, \dots) z^n = \sum_{n\geq 0} \sum_{k=0}^{n} B_{n,k}^o(f_1, f_2, \dots) z^n \\
&= \sum_{n\geq 0} B_n^o(f_1, f_2, \dots) z^n.
\end{aligned}$$

Note that $B_{n,k}^o = 0$ for $k > n$. □

3. Shift polynomials, Hankel mean–independence

We now prove a very useful property of the ordinary Bell polynomials, which we call the *2-shift theorem*.

THEOREM 10. *The ordinary Bell polynomials $B_n^o(x_1, x_2, x_3, \dots)$, satisfy*

$$\det[B_{i+j}^o]_{i,j=0}^n = \det[x_{i+j+2}]_{i,j=0}^{n-1}.$$

PROOF. Put $B = [B_{i+j}^o]_{i,j=0}^n$,

$$S = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & x_2 & x_3 & \dots & x_n \\ 0 & x_3 & x_4 & \dots & x_{n+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & x_n & x_{n+1} & \dots & x_{2n} \end{bmatrix},$$

and

$$a_{i,j} = \begin{cases} -x_{i-j} & i > j \\ 0 & i < j \\ 1 & i = j \end{cases} \qquad A = [a_{i,j}] = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ -x_1 & 1 & 0 & \dots & 0 \\ -x_2 & -x_1 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -x_n & -x_{n-1} & -x_{n-2} & \dots & 1 \end{bmatrix}.$$

Note that

$$A^{-1} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ B_1^o & 1 & 0 & \dots & 0 \\ B_2^o & B_1^o & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ B_n^o & B_{n-1}^o & B_{n-2}^o & \dots & 1 \end{bmatrix}.$$

Clearly $ABA^T = S$ if and only if $B = A^{-1}S\left(A^T\right)^{-1}$. If $(AB)_{i,j}$ is the i,j entry of the product AB, then

$$(AB)_{i,j} = \sum_{k=0}^{n} a_{i,k}B_{k+j}^o = \sum_{k=0}^{i-1} -x_{i-k}B_{k+j}^o + B_{i+j}^o = B_{i+j}^o - \sum_{h=j}^{i+j-1} B_h^o x_{i+j-h}.$$

Now

$$x_n = B_n^o - \sum_{k=1}^{n-1} B_k^o x_{n-k} = B_n^o - \left(\sum_{k=1}^{j-1} B_k^o x_{j-k} + \sum_{m=j}^{i+j-1} B_m^o x_{m-i-j}\right).$$

Hence,

$$x_{i+j} + \sum_{k=1}^{j-1} B_k^o x_{j+k} = (AB)_{i,j}.$$

Multiplying AB by A^T subtracts B_j^o times the j'th column from the i'th column. That is,

$$(ABA^T)_{i,j} = \begin{cases} (AB)_{i,j} - \sum_{k=1}^{j-1} B_k^o x_{j-k} = x_{i+j} & i \neq 0,\ j \neq 0 \\ (AB)_{i,j} - B_j^o = 0 & i = 0,\ j \neq 0 \\ 0 & i > 0,\ j = 0 \\ 1 & i = 0,\ j = 0, \end{cases}$$

which is the desired conclusion. □

By the above theorem, the Hankel determinants of the ordinary Bell polynomials are obtained from the Hankel determinants of $x_1, x_2, \dots$ by replacing x_i by x_{i+2}. Noting that x_1 is thereby shifted out, we have our first application of the 2-shift theorem:

COROLLARY 8. *The ordinary Bell polynomials $B_n^o(x_1, x_2, \dots)$ are Hankel mean-independent.*

Trivially we also have:

COROLLARY 9. *If* $f_n = \sum_{k=0}^{n-1} f_k g_{n-k} + \delta_{n,0}$, *then* $\det[f_{i+j}]_{i,j=0}^{n} = \det[g_{i+j+2}]_{i,j=0}^{n-1}$.

Thus if $\delta_n^o = \det[B_{i+j}^o]_{i,j=0}^n$, the first few values of δ_n^o are:

$$\begin{aligned}
\delta_0^o &= 1\\
\delta_1^o &= x_2\\
\delta_2^o &= x_2x_4 - x_3^2\\
\delta_3^o &= x_2x_4x_6 - x_3^2x_6 - x_2x_5^2 + 2x_3x_4x_5 - x_4^3.
\end{aligned}$$

We can use the 2-shift theorem to study cumulants by considering the *exponential shift polynomials* S_n defined by $S_n = Y_n - \sum_{k=1}^{n-1} Y_k S_{n-k}$. The 2-shift theorem states that $\det[Y_{i+j}]_{i,j=0}^n = \det[S_{i+j}]_{i,j=1}^n$. Since the $Y_n(x_1, x_2, \ldots)$ are Hankel mean–independent, we can set $x_1 = 0$ in Y_n. Here are a few values of the resulting polynomials S_n :

$$\begin{aligned}
S_0 &= 1\\
S_1 &= 0\\
S_2 &= x_2\\
S_3 &= x_3\\
S_4 &= x_4 + 2x_2^2\\
S_5 &= x_5 + 8x_2x_3\\
S_6 &= x_6 + 13x_2x_4 + 9x_3^2 + 10x_2^3
\end{aligned}$$

There are other results similar to the 2-shift theorem, however we will make good use of the following result which we call the *1-shift* theorem:

THEOREM 11. *The ordinary Bell polynomials* $B_n^o(x_1, x_2, x_3, \ldots)$, *satisfy*

$$\det[B_{i+j+1}^o]_{i,j=0}^n = \det[x_{i+j+1}]_{i,j=0}^n.$$

PROOF. Put $B = [B_{i+j+1}^o]_{i,j=0}^n$, and $S = [x_{i+j+1}]_{i,j=0}^n$. Take A as defined in the proof of Theorem 10, and by similar reasoning as there we have $S = ABA^T$. We omit the details. $\square$

4. Moment sequence preserving maps, ordinary cumulants

The ordinary Bell polynomials share many of the properties of their exponential counterparts. For instance,

PROPOSITION 17.

$$\delta_n^o(t) = \det\left([B_{i+j}^o]_{i,j=0}^n\right) \sim \Delta_n(t) \quad \textit{as } t \to 0.$$

More importantly, the shift theorem easily gives

PROPOSITION 18. *The ordinary Bell polynomials* $B_n^o(x_1, x_2, \ldots)$ *are* $\mathcal{M}$*–preserving.*

PROOF. What needs to be shown is that if $(x_n) \in \mathcal{M}$, then $(B_n^o) \in \mathcal{M}$. Via the Hamburger condition, this is true if $\Delta_n \geq 0$ implies $\delta_n^o \geq 0$ for all non-negative integers n. By hypothesis the quadratic form with matrix

$$\Delta_n = \begin{vmatrix} 1 & x_1 & \dots & x_n \\ x_1 & x_2 & \dots & x_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ x_n & x_{n+1} & \dots & x_{2n} \end{vmatrix}$$

is non-negative definite. Hence all its principal minors are non–negative. In particular $M_{1,1} \geq 0$; by the shift theorem we have $M_{1,1} = \delta_n^o \geq 0$. □

COROLLARY 10. *If $\Delta_n(t) \geq 0$, then $\delta_{n-1}^o(t) \geq 0$.*

A particular consequence of the above proposition implies a result obtained by Kaluza [20] in 1928.

PROPOSITION 19. *Let $\mathcal{M}^+$ be the set of all moment sequences whose terms are non-negative. The sequence (x_n) is in $\mathcal{M}^+$ if and only if (B_n^o) is in $\mathcal{M}^+$.*

Note that the moment sequence preserving property of the ordinary Bell polynomials gives a stronger result than the above in one direction, as there is no non-negativity restriction. As in the exponential case, $(B_n^o) \in \mathcal{M}$ does not always imply $(x_n) \in \mathcal{M}$. Consider for example $x_1 = c$ and $x_n = 0$ for $n \geq 2$. Then $B_n^o = c^n$ and so as before, (B_n^o) is a moment sequence whereas (x_n) is not. This disproves an earlier claim by Liggett [23] in which he asserts that $(B_n^o) \in \mathcal{M}$ if and only if $(x_n) \in \mathcal{M}$.

As further application of the shift theorems, we present new proofs of some related work by Horn in [18]). Our approach will be fully algebraic, whereas Horn relied on analysis. For this purpose we extend our consideration of probability moment sequences to general moment sequences in which the zero'th moment may take on values other than 1. This is reasonable as we can always normalize any general moment sequence to a probability moment sequence by dividing each moment by the zero'th moment. The following theorem is presented as Theorem 1 in [18] which we now re-prove.

THEOREM 12. *The sequence (x_{n+1}) is a Hamburger moment sequence if and only if (B_{n+1}^o) is a Hamburger moment sequence.*

PROOF. We have by the 1-shift theorem that $\det[x_{i+j+1}] = \det[B_{i+j+1}^o]$ and hence the result follows by Hamburgers theorem. □

Recall that a Stieltjes moment sequence is a moment sequence whose distribution has support on $\mathbb{R}^+$. Horn presents the following theorem:

THEOREM 13. *The sequence (x_{n+1}) is a Stieltjes moment sequence if and only if (B_n^o) is a Stieltjes moment sequence.*

PROOF. The Stieltjes condition as presented in [33, p.5] states that (μ_n) is a Stieltjes moment sequence if and only if $\det[\mu_{i+j}] > 0$ and $\det[\mu_{i+j+1}] > 0$. Suppose that (x_{n+1}) is a Stieltjes moment sequence, so $\det[x_{i+j+1}] > 0$ and $\det[x_{i+j+2}] > 0$. By the 1-shift theorem $\det[x_{i+j+1}] = \det[B_{i+j+1}^o]$, so $\det[B_{i+j+1}^o] > 0$. By the 2-shift theorem $\det[x_{i+j+2}] = \det[B_{i+j}^o]$, hence $\det[B_{i+j}^o] > 0$ and the result follows. $\square$

Several other of Horns results follow just as easily.

In contrast with the exponential Bell polynomials, the total non-negativity property is discussed here in the context of renewal theory (see Theorem 1 of [23]). However the proof is essentially the same as the one presented earlier in consideration of the results of Horn above.

PROPOSITION 20. *Δ_n and $[x_{i+j+1}]_{i,j=0}^n$ are non-negative if and only if $[x_{i+j}]_{i,j=0}^n$ and $[B_{i+j}^o]_{i,j=0}^n$ are totally non-negative.*

There is an as yet unresolved philosophical matter concerning how we arrived at this point. We found that the exponential Bell polynomials are $\mathcal{M}$–preserving by finding a random variable of which they are the moments. For the ordinary Bell polynomials we have only shown the existence of a random variable, without finding it explicitly. As a partial step to find such a random variable, set $F(z) = E[Z^X] = \sum_{n\geq 1} \Pr(X = n)z^n = \sum_{n\geq 1} x_n z^n$ where $\Pr(X = 0) = 0$, and let $X_1, X_2, \dots$ be iid copies of X. We have:

PROPOSITION 21. *If $S_n = X_1 + X_2 + \cdots + X_n$, $S_0 = 1$, then*

$$E[z^Y] = \sum_{n\geq 0} B_n^o(x_1, x_2, \dots)z^n = E[\sum_{k\geq 1} z^{S_k}].$$

PROOF. Putting $x_0 = 0$ and summing equation (15) over k, we obtain

$$\sum_{k\geq 1}\left(\sum_{n\geq 0} x_n z^n\right)^k = \sum_{k\geq 1}\sum_{n\geq k} B_{n,k}^o z^n = \sum_{n\geq 0}\sum_{k=1}^{n} B_{n,k}^o z^n = \sum_{n\geq 0} B_n^o z^n = E[z^Y]$$

for a new random variable Y. Now set $F(z) = E[z^X] = \sum_{n\geq 0} x_n z^n$, where $\Pr(X = n) = x_n$. Then $F(z)^k = E[z^{S_k}]$ by convolution of probabilities (see [14]), so

$$E[z^Y] = \sum_{k\geq 1} E[z^{S_k}] = E[\sum_{k\geq 1} z^{S_k}]$$

by linearity of expectations. $\square$

Earlier we saw that the cumulant space $\mathcal{K}$ is the inverse image of $\mathcal{M}$ under the exponential Bell map. Similarly, we can define the *ordinary cumulant space* as the inverse image of $\mathcal{M}$ under the ordinary Bell map. Here are the first few ordinary cumulants as functions of the moment sequence (μ_n) :

$$\begin{aligned}
\kappa_1^o &= \mu_1 \\
\kappa_2^o &= \mu_2 - \mu_1^2 \\
\kappa_3^o &= \mu_3 - 2\mu_1\mu_2 + \mu_1^3 \\
\kappa_4^o &= \mu_4 - 2\mu_1\mu_3 - \mu_2^2 + 3\mu_1^2\mu_2 - 4\mu_1^4 \\
\kappa_5^o &= \mu_5 - 2\mu_1\mu_4 - 2\mu_2\mu_3 + 3\mu_1^2\mu_3 + 3\mu_1\mu_2^2 - 4\mu_1^3\mu_2 + \mu_1^5
\end{aligned}$$

It is natural to ask if the ordinary cumulants have an invariance property analogous to that given by Proposition 4. Specifically, what (if any) transformation on a random variable X leaves all ordinary cumulants but the first invariant? It can be shown that any such transformation must be non-linear. We submit this as an open problem:

OPEN PROBLEM 5. *Do the ordinary cumulants have an invariance property?*

CHAPTER III

Comtet polynomials and binomial posets

Up to this point we have studied two polynomial maps, the exponential and ordinary Bell maps, which have analogous properties. We now embed these two maps into in a much broader class. Following Comtet [10, p. 137], we define the partial Bell polynomials with respect to a sequence (Ω_n) by

$$\Omega_k \left(\sum_{m \geq 1} x_m z^m \right)^k = \sum_{n \geq k} B_{n,k}^{\Omega} \Omega_n z^n. \tag{31}$$

We assume here that $\Omega_0 = \Omega_1 = 1$ and $\Omega_n \neq 0$ for all n. However, we continue to write Ω_1 instead of 1 when this more clearly indicates the pattern of some of our formulas. We call the polynomials $B_{n,k}^{\Omega}$ *partial Comtet polynomials* with respect to (Ω_n). Note that $\Omega_n = 1/n!$ and $\Omega_n = 1$ give the partial exponential and partial ordinary Bell polynomials respectively. We define the *complete Comtet polynomials* B_n^{Ω} by

$$Y_n^{\Omega} = \sum_{k=0}^{n} B_{n,k}^{\Omega}.$$

1. Properties

Here are some examples of partial Comtet polynomials:

$$\begin{aligned}
B_{0,0}^{\Omega} &= 1 \\
B_{n,1}^{\Omega} &= x_n \\
B_{n,n}^{\Omega} &= x_1^n \\
B_{n,n-1}^{\Omega} &= (n-1)\Omega_n^{-1}\Omega_2\Omega_{n-1}x_2x_1^{n-2}
\end{aligned}$$

$$B_{n,n-2}^{\Omega} = \Omega_n^{-1}\Omega_2\Omega_{n-1}x_1^{n-4}\left(\binom{n-2}{2}\Omega_2^2x_2^2 + (n-2)\Omega_3x_3x_1\right)$$

$$B_{n,n-3}^{\Omega} = \Omega_n^{-1}\Omega_2\Omega_{n-1}x_1^{n-6}\left(\binom{n-3}{3}\Omega_2^3x_2^3 + 2\binom{n-3}{2}\Omega_3x_3x_2x_1 + (n-3)\Omega_4x_4x_1^2\right)$$

In general, each $B^{\Omega}_{n,k}$ consists of terms x_i whose indices form a partition of n into exactly k parts, say $n = r_1 + 2r_2 + \cdots + nr_n$, and whose coefficient is $\binom{k}{r_1, r_2, \ldots, r_n}$ times $\Omega_k(\Omega_1 x_1)^{r_1}(\Omega_2 x_2)^{r_2}\cdots$. Thus:

PROPOSITION 22.

$$\Omega_n B^{\Omega}_{n,k} = \sum_{\substack{r_1+r_2+r_3+\cdots+r_n=k \\ r_1+2r_2+3r_3+\cdots+nr_n=n}} \binom{k}{r_1, r_2, \ldots, r_n} \Omega_k \Omega_1^{r_1}\Omega_2^{r_2}\cdots\Omega_n^{r_n} x_1^{r_1} x_2^{r_2}\cdots x_n^{r_n}. \tag{32}$$

A fundamental property of Comtet polynomials is the following convolutional formula:

PROPOSITION 23. *If* $h = f \circ g$ *where* $f = \sum_{n\geq 1} \Omega_n f_n z^n$ *and* $g = \sum_{n\geq 1} \Omega_n g_n z^n$, *then*

$$\sum_{n\geq k} B^{\Omega}_{n,k}(h_1, h_2, \ldots)\Omega_n z^n = \sum_{n\geq \ell\geq k} B^{\Omega}_{n,\ell}(f_1, f_2, \ldots) B^{\Omega}_{\ell,k}(g_1, g_2, \ldots)\Omega_n z^n.$$

PROOF. See [10, p. 146]. □

The first five complete Comtet polynomials are:

$$\begin{aligned}
B^{\Omega}_0 &= 1 \\
B^{\Omega}_1 &= \Omega_1 x_1 \\
B^{\Omega}_1 &= \Omega_1^2 x_2 + \Omega_1 x_2 \\
B^{\Omega}_3 &= \Omega_1^3 x_1^3 + \frac{2\Omega_1\Omega_2^2 x_1 x_2}{\Omega_3} + \Omega_1 x_3 \\
B^{\Omega}_4 &= \Omega_1^4 x_1^4 + \frac{3\Omega_1^2\Omega_2\Omega_3 x_1^2 x_2}{\Omega_4} + \frac{\Omega_2^3 x_2^2}{\Omega_4} + \frac{2\Omega_1\Omega_2\Omega_3 x_1 x_3}{\Omega_4} + \Omega_1 x_4.
\end{aligned}$$

2. Recurrent events and binomial posets

In previous sections we saw how the ordinary Bell polynomials arise from the renewal equation as probabilities for a recurrent event. Here we introduce a more general form of this concept and discuss its relation to the Comtet polynomials.

DEFINITION 11. *A poset P is called a **binomial poset** if it satisfies the following three conditions:*

- ***a:*** *P is locally finite with a $\hat{0}$.*
- ***b:*** *The intervals $[x, y]$ of P are graded. If the length of the interval $[x, y]$ is $\ell(x, y) = n$, we call $[x, y]$ an **n-interval.***
- ***c:*** *For all $n \in \mathbb{N}$, any two n-intervals have the same number $B(n)$ of maximal chains. We call $B(n)$ the **factorial function** of P.*

See [36, p. 140] for more details. The basic convolutional property of binomial posets is the following theorem found in [36, p. 144]:

THEOREM 14. *Let P be a binomial poset with factorial function $B(n)$ and incidence algebra $I(P)$ over $\mathbb{C}$. Define*

$$R(P) = \{f \in I(P) : f(x,y) = f(x',y') \text{ if } \ell(x,y) = \ell(x',y')\}.$$

If $f \in R(P)$, write $f(n)$ for $f(x,y)$ when $\ell(x,y) = n$. Then $R(P)$ is a subalgebra of $I(P)$, and there is an algebra isomorphism $\phi : R(P) \to \mathbb{C}[[z]]$ given by

$$\phi(f) = \sum_{n \geq 0} f(n) z^n / B(n).$$

As in [36, p. 143], we denote the number of elements z of rank i in an n-interval $[x,y]$ by

$$\begin{bmatrix} n \\ i \end{bmatrix} = \frac{B(n)}{B(i)B(n-i)}.$$

This is easily extended to

$$\begin{bmatrix} n \\ i_1, i_2, \ldots \end{bmatrix} = \frac{B(n)}{B(i_1)B(i_2)\cdots},$$

which counts the number of sequences of elements of rank $i_1, i_2, \ldots$ (where $i_1 + i_2 + \cdots = n$) in an n-interval.

We are now in a position to see how binomial posets give rise to Comtet polynomials. As a start, consider $c_k(n)$, the number of chains $x = w_0 < w_1 < \cdots < w_k = y$ of length k in any n-interval $[x,y]$ of a binomial poset P. We have $c_k(n) = (\zeta - 1)^k(x,y)$, where $(\zeta - 1)^k$ is the k–fold convolution of $(\zeta - 1)$ in $I(P)$; here $\zeta \in I(P)$ is the *zetafunction* on P defined by $\zeta(x,y) = 1$, for all $x \leq y$ in P, and $1(x,y) = 1$ or 0 according as $x = y$ or $x \neq y$. We have

$$\sum_{n \geq 0} c_k(n) z^n / B(n) = \left(\sum_{n \geq 1} z^n / B(n) \right)^k.$$

If we consider Comtet polynomials where $x_n = 1$ for all n and $\Omega_n = 1/B(n)$, the above relation becomes equation (31), where $c_k(n) = B^{\Omega}_{n,k}/\Omega_k$. This suggests that we replace $\zeta - 1$ by the function

$$\theta(x,y) = \begin{cases} x_n & x < y \text{ and } \ell(x,y) = n \\ 0 & \text{otherwise.} \end{cases}$$

Then consider $d_k(n) = \theta^k(x,y)B(k)$; this gives the partial Comtet polynomials as defined in equation (31). The difficulty with this approach lies in interpreting this function. To accomplish this, we consider the case for $x_n \in [0,1]$. Define a stochastic process on P, by putting x_n as the probability of jumping directly to a particular rank n element in a maximal chain. Then derive the reciprocal probability f_n that the process ever arrives at a particular element of rank n in a maximal chain, but perhaps through other elements first. Clearly

$$f_n = \sum_I \Pr(\text{choose the maximal chain for } c)\Pr(\text{jump through elements of } c).$$

Now if we assume each path is equally likely to be traversed, then

$$\Pr(\text{choose the maximal chain for } c) = 1/B(n),$$

z is of rank n and $B(n)$ is the number of maximal chains in the interval $[0, z]$. Under these circumstances, if c is a maximal chain that jumps through elements of ranks $i_1, i_1 + i_2, \ldots, i_1 + i_2 + \cdots i_k$ then

$$\Pr(\text{jump through the elements in path } c) = x_{i_1} x_{i_2} \cdots x_{i_k}.$$

Since this is invariant under permutations of $i_1, i_2, \ldots, i_k$, we obtain:

THEOREM 15. *If x_n is the probability of jumping directly to a particular rank n element in a maximal chain in a binomial poset P, and f_n is the probability of ever arriving at a particular rank n element in a maximal chain, then $f_n = Y_n^{\Omega}$, where $\Omega_n = 1/B(n)$.*

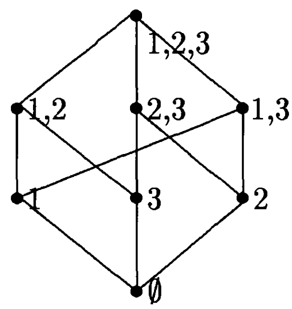
PROOF. Consider the interpretation of $B_{n,k}^{\Omega}$ given at the beginning of this chapter. $\square$

Example 1. $P = \mathbb{N}$ and $B(n) = 1$. We compute f_3, the probability of ever reaching 3. There are four ways in which we can arrive at 3. One, by jumping directly to 3 with probability x_3 Two, jumping to 1, then to 3, with probability $x_1 x_2$. Third, jumping to 2, then to 3, with probability $x_2 x_1$. Fourth, jumping to 1, then 2, then 3, with probability x_1^3. Hence

$$\begin{aligned} f_3 &= x_3/1 + x_2 x_1/1 + x_1 x_2/1 + x_1^3/1 \\ &= x_3 + 2x_2 x_1 + x_1^3 = Y_3^{\Omega}(x_1, x_2, x_3). \end{aligned}$$

In general this leads to $f_n = Y_n^{\Omega}(x_1, x_2, \ldots, x_n)$, where $\Omega_n = 1/B(n) = 1$, the ordinary Bell polynomials.

Example 2. $P = \mathbb{B}$, the Boolean lattice of subsets of $\{1, 2, \ldots, n\}$ under inclusion. Here $B(n) = n!$ and we find f_3, the probability of ever jumping from $\emptyset$ to $\{1, 2, 3\}$. There are four types of paths which arrive at $\{1, 2, 3\}$. One, by jumping directly to $\{1,2,3\}$; this has probability x_3. Two, jumping a rank of 1 with probability x_1, then a rank of 2 with probability x_2. There are three paths to a rank 1 element, then only one path from there to $\{1,2,3\}$, giving three possible routes. Third, jumping to a rank 2 element (of which there are three), then jumping to the top, giving three possible paths. Fourth, jumping by 1 three times; there are three ways for the first jump, two for the second, one for the third, so six paths in all. Hence



$$\begin{aligned} f_3 &= x_3/1! + 3x_1x_2/2! + 3x_2x_1/2! + 6x_1^3/3! \\ &= x_3 + 3x_2x_1 + x_1^3 = Y_3^{\Omega}(x_1, x_2, x_3), \end{aligned}$$

where $\Omega_n = 1/B(n) = 1/n!$ for all n. This gives the exponential Bell polynomials.

Example 3. $P = \{S \times T \subseteq \mathbb{N} \times \mathbb{N} | |S| = |T|\}$ (the cubical lattice). We have $B(n) = (n!)^2$, so:

$$\begin{aligned} f_3 &= x_3/1!^2 + 9x_1x_2/2!^2 + 9x_2x_1/2!^2 + 36x_1^3/3!^2 \\ &= x_3 + \frac{9}{2}x_2x_1 + x_1^3 = Y_3^{\Omega}(x_1, x_2, x_3). \end{aligned}$$

where $\Omega_n = 1/B(n) = 1/(n!)^2$ for all n. This gives the *double exponential* Bell polynomials.

3. Characterization of Hankel mean–independence

The Comtet polynomials form a broad class of transformations with useful compositional properties and application to the enumeration of natural objects. We have seen that the Hankel mean–independence property for the exponential and ordinary Bell polynomials is closely related to their moment sequence preserving properties. Superficially it might seem that there should be a great many Comtet polynomial sequences which are Hankel mean–independent. However in this section we show that there are only two. This fact will be called the *characterization theorem;* it has obvious implications for the question of what makes the exponential and ordinary generating functions so special. We require the following technical lemma:

LEMMA 4. *The recurrence* $c_1-(n-1)c_{n-1}+(n-2)c_n=0$, $(n\geq 2)$ *has the general solution*

$$c_n=(n-1)c_2+(n-2)c_1, \tag{33}$$

where c_2 *can be arbitrarily prescribed.*

PROOF. Clearly (33) holds if $n=2$. Assume it holds for some $n\geq 2$. From 4 we have $c_1-nc_n+(n-1)c_{n+1}=0$, so

$$c_{n+1}=\frac{nc_n+c_1}{n-1}.$$

Substituting (33) into this, we get

$$c_{n+1}=\frac{n}{n-1}[(n-1)c_2-(n-2)c_1]-\frac{c_1}{n-1}=nc_2-(n-1)c_1,$$

completing the induction. □

Before proving the characterization theorem, we first observe that if $\Omega=(\Omega_n)$ is a Comtet sequence, then the sequence $(a^{n-1}\Omega_n)$, where $a\neq 0$, gives rise to the same Comtet polynomials. In view of this, we sometimes find it convenient to *normalize* Ω so that Ω_2 is a given constant. The general case then results by scaling.

THEOREM 16. $\delta_n^\Omega=\det[Y_{i+j}^\Omega]_{i,j=0}^n$ *is independent of* x_1 *if and only if* $\Omega_n=1$ *or* $\Omega_n=1/n!$ *after the normalizations* $\Omega_2=1$ *and* $\Omega_2=1/2$ *respectively.*

PROOF. From previous sections we know that the Comtet polynomials with $\Omega_n=1$ or $\Omega_n=1/n!$ are Hankel mean–independent. For necessity we will show that the term in $x_1x_{n-1}x_n^{n-2}x_{2n}$ of δ_n^Ω has coefficient

$$\frac{\Omega_1}{\Omega_2}-(n-1)\frac{\Omega_{n-1}}{\Omega_n}+(n-2)\frac{\Omega_n}{\Omega_{n+1}}. \tag{34}$$

For δ_n^Ω to be independent of x_1, this coefficient must be zero. Setting $c_n=\Omega_n/\Omega_{n+1}$, we obtain the recurrence

$$c_1-(n-1)c_{n-1}+(n-2)c_n=0. \tag{35}$$

It is easily verified that the term $x_1^2x_2^2x_6$ in δ_3^Ω has coefficient

$$-3\Omega_1^5-4\frac{\Omega_1^3\Omega_2^4}{\Omega_3^2}+4\frac{\Omega_1^4\Omega_2^2}{\Omega_3}+3\frac{\Omega_1^4\Omega_2\Omega_3}{\Omega_4}.$$

The vanishing of this coefficient translates to the condition

$$-3c_1^2-4c_2^2+4c_2c_1+3c_1c_3=0.$$

Combining this with the equation $c_3=2c_2-c_1$, we obtain

$$(c_2-c_1)(2c_2-c_1)=0.$$

If $c_2=c_1$, then $c_n=c_1$ for all n, while if $2c_2-3c_1=0$, then $c_n=c_1(n+1)/2$.

To show that x_1-independence of δ_n^Ω implies the recurrence (35), we set $x_i = 0$ for $i \neq 1, n-1, n, 2n$ and determine the coefficient of $x_1 x_{n-1} x_n^{n-2} x_{2n}$ in δ_n^Ω. This term has the factor x_{2n} which occurs only in $Y_{2n}^{\Omega_n}$ (the bottom right element of δ_n^Ω). Hence by expanding δ_n^Ω on its last row, we see that we need only determine the coefficient of $x_1 x_{n-1} x_n^{n-2}$ in δ_{n-1}^Ω. The first occurrence of x_{n-1} in $\delta_{n-1}^{\Omega_n}$ is on the main anti-diagonal. This term must accumulate $n-1$ of the x_n factors; these can only be acquired from the anti-diagonal just below the main anti-diagonal. We can ignore any higher powers of x_1, x_{n-1} and powers of x_n greater than the $(n-2)$'nd as they clearly do not contribute. We then obtain the following reduction of δ_{n-1}^Ω :

$$A = \begin{vmatrix} 1 & d & 0 & \dots & 0 & 0 & 0 & 0 & a \\ d & 0 & 0 & \dots & 0 & 0 & 0 & a & b \\ 0 & 0 & 0 & \dots & 0 & 0 & a & b & c \\ 0 & 0 & 0 & \dots & 0 & a & b & c & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & a & b & \dots & 0 & 0 & 0 & 0 & 0 \\ a & b & c & \dots & 0 & 0 & 0 & 0 & 0 \end{vmatrix}, \tag{36}$$

where $a = \Omega_1 x_{n-1}$, $b = \Omega_1 x_n + 2\Omega_1\Omega_2\Omega_{n-1}/\Omega_n x_1 x_{n-1}$, $c = 2\Omega_1\Omega_2\Omega_n/\Omega_{n+1} x_1 x_n$ and $d = \Omega_1 x_1$. To compute a we note that Y_{n-1}^Ω can only contribute factors of x_{n-1}, which occurs in $B_{n-1,1}^\Omega$ For b we note that Y_n^Ω can contribute x_n and $x_1 x_{n-1}$, which occur in $B_{n,1}^\Omega$. and $B_{n,2}^\Omega$ respectively. For c we note that Y_{n+1}^Ω can contribute an $x_1 x_n$ factor. For d we note that $B_1 = \Omega_1 x_1$.

Since we must avoid picking up any higher powers of x_1 and x_{n-1}, we can ignore any term of A in a^2, c^2, d^2, cd and more generally, any term $a^{i_1} b^{i_2} c^{i_3} d^{i_3}$ with $i_1 > 1$, $i_2 > n-1$, $i_3 > 1$, $i_4 > 1$, $i_3 + i_4 > 1$, $i_1 + i_2 > n-1$, $i_1 + i_2 + i_3 > n-1$. The only three surviving terms in A are $2ab^{n-2}d$, $(n-2)ab^{n-3}c$, and b^{n-1}. For b^{n-1} we are only interested in the term which gives $x_1 x_{n-1} x_n^{n-2}$, namely

$$\binom{n-1}{n-2} (\Omega_1 x_n)^{n-2} (2\Omega_1\Omega_2\Omega_{n-1}\Omega_n^{-1} x_1 x_{n-1}) = \frac{2(n-1)\Omega_1^{n-1}\Omega_2\Omega_{n-1} x_1 x_{n-1} x_n^{n-2}}{\Omega_n}.$$

This term has the opposite sign from the other two surviving terms due to its position under the main anti-diagonal. Similarly, for b^{n-1} we must take only the x_n^{n-2} term, which has coefficient Ω_1^{n-2}. Hence $2ab^{n-2}d$ contributes a term $x_1 x_{n-1} x_n^{n-2}$ with coefficient $2\Omega_1^{n+1}$. Lastly, in b^{n-3} we need the x_n^{n-3} term, which is Ω_1^{n-3}. Thus the $x_1 x_{n-1} x_n^{n-2}$ term of $(n-2)ab^{n-3}c$ has coefficient

$$\frac{2(n-2)\Omega_1^{n-1}\Omega_2\Omega_n}{\Omega_{n+1}}.$$

In all, the coefficient of $x_1x_{n-1}x_n^{n-2}$ in A is

$$2\Omega_1^{n+1} - \frac{2(n-1)\Omega_1^{n-1}\Omega_2\Omega_{n-1}x_1x_{n-1}x_n^{n-2}}{\Omega_n} + \frac{2(n-2)\Omega_1^{n-1}\Omega_2\Omega_n}{\Omega_{n+1}}. \tag{37}$$

This must vanish if δ_n^{Ω} is to be independent of x_1, and its vanishing is equivalent to (35). □

4. Moment sequence preserving maps

As in previous sections, we would like to know which Comtet polynomials are $\mathcal{M}$-preserving. Since

$$Y_n^{\Omega}(t) = \Omega_1 x_n t + O(t^2)$$

and $\Omega_1 = 1$, it is clear that $\delta_n^{\Omega}(t) \sim \Delta_n(t)$ as $t \to 0$. From this we have the following simple proposition:

PROPOSITION 24. *If $\delta_n^{\Omega}(t)$ is non-negative for sufficiently small t, then $\Delta_n(t)$ is non-negative for sufficiently small t.*

We have seen that the above fails if t is fixed. In general Comtet polynomials are not moment sequence preserving, even for $\Omega_n = 1/B(n)$, where $B(n)$ is the factorial function for a binomial poset. For example, consider $x_n = p^n$ $(n \geq 0)$ for $0 \leq p \leq 1$; these are the moments of a degenerate random variable. Clearly $\Delta_n = 0$ for $n \geq 1$. Applying the Comtet transformation, we get

$$\delta_2^{\Omega} = p^6\left(\Omega_1^3 + 3\Omega_1^4 - \frac{4\Omega_1^2\Omega_2^4}{\Omega_3^2} - \frac{4\Omega_1^2\Omega_2^2}{\Omega_3} + \frac{4\Omega_1^3\Omega_2^2}{\Omega_3} + \frac{\Omega_1\Omega_2^3}{\Omega_4} + \frac{2\Omega_1^2\Omega_2\Omega_3}{\Omega_4} + \frac{3\Omega_1^3\Omega_2\Omega_3}{\Omega_4}\right).$$

If $B(n) = A(1)A(2)\cdots A(n) = 2^{n-2}$ (as is the case for the butterfly binomial poset), we have

$$c_n = A(n+1) = \begin{cases} 1 & n = 1 \\ 2 & \text{otherwise} \end{cases}.$$

This gives $\delta_2^{\Omega} = -4p^6$, and hence this Comtet map is not $\mathcal{M}$-preserving. Note that if $\Omega_n = 1$ (a chain), then $\delta_n^{\Omega} = 0$ for $n > 1$ by the shift theorem (since two rows are equal). If $\Omega_n = 1/n!$ (the Boolean algebra), then

$$\delta_n^{\Omega} = \det\left[\phi_{i+j}(p)\right]_{i,j=0}^{n} = \left(\prod_{k=1}^{n} k!\right) p^{\binom{n-1}{2}} \geq 0.$$

See [21] for details.

OPEN PROBLEM 6. *What are the conditions on Ω_n so that the resulting Comtet map is $\mathcal{M}$-preserving?*

Up to this point we have only encountered $\mathcal{M}$-preserving Comtet maps which are also Hankel mean–independent. It is natural to inquire whether Hankel mean–independence is necessary for preservation of moment sequences. A solution to the following conjecture would be some progress in this direction:

CONJECTURE 3. *The Comtet map with* $\Omega_n = 1/n!^2$ *is* $\mathcal{M}$*-preserving but not Hankel mean–independent.*

It is not difficult to verify that the Comtet map with $\Omega_n = 1/n!^2$ above is not Hankel mean-independent. It appears difficult to prove that it is $\mathcal{M}$-preserving.

Many questions that can be asked about binomial posets, which are beyond the scope of the present work. For example, we may ask for a characterization of the factorial functions $B(n)$ for binomial posets. In particular, is $B(n)$ always a moment sequence? Can every finite binomial poset be extended to an infinite binomial poset?

CHAPTER IV

Compound polynomials

1. Properties

We saw in the last chapter that some Comtet maps are $\mathcal{M}$-preserving and some are not. In this chapter we study a different class of polynomial maps which are always $\mathcal{M}$-preserving and are easily interpreted probabilistically. Let A be a non-negative, integer-valued random variable with $\Pr(A = i) = p_i$. Furthermore, let X be any random variable with moments $E[X^i] = x_i$ and put $Y = S_A = X_1 + X_2 + \cdots + X_A$. The first few moments of Y are:

$$\begin{aligned} E[1] &= 1 \\ E[Y] &= p_1x_1 + 2p_2x_2 + 3p_3x_3 + \cdots \\ E[Y^2] &= p_1x_2 + 2p_2x_2 + 2p_2x_1^2 + 3p_3x_2 + 6p_2x_1^2 + \cdots \\ E[Y^3] &= p_1x_3 + 2p_2x_3 + 3x_2x_1 + 3p_3x_3 + 6x_2x_1 + 2x_1^3 + \cdots . \end{aligned}$$

Let $E[Y^n] = C_n(x_1, x_2, \ldots, x_n)$. If a_i and f_i are the i'th moment and i'th factorial moment of A respectively, then $C_n(x_1, x_2, \ldots, x_n)$ can be expressed as a polynomial in either the a_i and x_i, or in the f_i and x_i. The first few values are as follows:

$$\begin{aligned} C_0 &= 1 \\ C_1 &= a_1x_1 = f_1x_1 \\ C_2 &= a_1x_2 + (a_2 - a_1)x_1^2 = f_1x_2 + f_2x_1^2 \\ C_3 &= a_1x_3 + (a_2 - a_1)x_2x_1 + (a_3 - 3a_2 - 2a_1)x_1^3 = f_1x_3 + 3f_2x_2x_1 + f_3x_1^3. \end{aligned}$$

By construction the map (C_n) is $\mathcal{M}$-preserving; we call the polynomials C_n *compound polynomials.*

They agree with the exponential Bell polynomials when expressed in terms of the factorial moments, except that each term in C_n has a further factor of f_d, where d is the total degree of the term. To prove this, we first require the following lemma:

LEMMA 5. *If A is a Poisson variable with parameter t, then A has i'th factorial moment $f_i = t^i$.*

PROOF. First note from previous work that $E[A^n] = \phi_n(t)$, where $\phi_n(t)$ is the n'th exponential polynomial. Thus

$$E[A^n] = \phi_n(t) = \sum_{k=0}^{n} S(n,k)t^k,$$

where $S(n,k)$ is the Stirling number of the second kind. The factorial moments of A are

$$f_n = E[(A)_n] = \sum_{k=0}^{n} s(n,k)E[A^k] = \sum_{k=0}^{n} s(n,k)\phi_k(t),$$

where $s(n,k)$ is the Stirling number of the first kind. Inverting the above gives

$$\phi_n(t) = \sum_{k=0}^{n} S(n,k)f_k = \sum_{k=0}^{n} S(n,k)t^k.$$

An easy induction shows from this that $f_i = t^i$ as desired. □

The following proposition relates the compound polynomials to the exponential Bell polynomials:

PROPOSITION 25. *If $f_i = t^i$ then $C_n(x_1, x_2, \ldots, x_n) = Y_n(tx_1, tx_2, \ldots, tx_n)$.*

PROOF. From previous work we know that if the x_i are the moments of a variable X, then $Y_n(tx_1, tx_2, \ldots, tx_n)$ are the moments of the compound Poisson variable S_N, where N is a Poisson random variable with parameter t. Thus $f_i = t^i$ by lemma 5, and the result follows from the definition of the compound polynomials. □

Since the compound polynomials are always $\mathcal{M}$-preserving and superficially resemble Comtet polynomials, it is natural to investigate the intersection of the two classes. The following theorem shows settles this problem:

THEOREM 17. *If $C_n(x_1, x_2, \ldots, x_n) = Y_n^{\Omega}(x_1, x_2, \ldots, x_n)$, then $\Omega_n = 1/n!$.*

PROOF. Matching coefficients of x_1^n in C_n and Y_n^{Ω} yields $\Omega_1^n = f_n$. By Proposition 25 this gives $C_n(x_1, x_2, \ldots, x_n) = Y_n(\Omega_1 x_1, \Omega_1 x_2, \ldots, \Omega_1 x_n)$. □

The question of when the compound polynomials are Hankel mean–independent is settled by the following:

THEOREM 18. *The matrix $[C_{i+j}]_{i,j=0}^n$ is Hankel mean-independent if and only if $f_i = t^i$ for t a constant.*

PROOF. We have $C_n(x_1, 0, 0, \ldots, 0) = f_n x_1^n$. Therefore $\det[f_{i+j}]_{i,j=0}^n$ as the coefficient of $x_1^{n(n+1)}$ in the expansion of $\det[C_{i+j}]_{i,j=0}^n$. Hence if $[C_{i+j}]_{i,j=0}^n$ is Hankel mean-independent, then $\det[f_{i+j}]_{i,j=0}^n \equiv 0$. By Proposition 9 this holds if and only if $f_i = t^i$ for some constant t. □

We can interpret the above theorem as saying that the only Hankel mean–independent compound map is the intersection of the compound maps with the Comtet polynomials, that is, the exponential Bell map.

CHAPTER V

Multidimensional extensions

We have seen that the Comtet polynomials have a number of interesting properties and applications. Thus far these were essentially one-dimensional. In this chapter we investigate the extent to which they carry over to higher-dimensional analogs of the Comtet polynomials. We consider a variety of such analogs and study their properties and applications. We find that many analogous properties hold, but some do not. For example, the determinantal conditions for the one-dimensional Hamburger moment problem are not sufficient in general. Earlier we applied this sufficiency condition to prove that the one-dimensional ordinary Bell map is $\mathcal{M}$-preserving. In higher dimensions this condition is no longer sufficient.

1. Properties

As with many generalizations, there are several choices for how to extend the definition of the Comtet polynomials to higher dimensions. Before selecting one of these, we require some definitions.

DEFINITION 12. *A* ***composition*** *of a non-negative integer n is a sequence of non-negative integers $\lambda = (\lambda_1, \lambda_2, \dots,)$, such that $n = \lambda_1 + \lambda_2 + \cdots$.*

DEFINITION 13. *A d–****composition*** *of a non-negative integer n is a composition $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_d, 0, 0, \dots)$ of n.*

Given a sequence of variables $X = (x_1, x_2, \dots)$ and compositions $\lambda = (\lambda_1, \lambda_2, \dots,)$ and $\mu = (\mu_1, \mu_2, \dots)$ we define

$$\lambda! = \lambda_1!\lambda_2!\cdots, \qquad |\lambda| = \lambda_1 + \lambda_2 + \cdots, \qquad (\lambda)_\mu = (\lambda_1)_{\mu_1}(\lambda_2)_{\mu_2}\cdots$$

$$X^\lambda = x_1^{\lambda_1}x_2^{\lambda_2}\cdots \qquad \binom{\lambda}{\mu} = \binom{\lambda_1}{\mu_1}\binom{\lambda_2}{\mu_2}\cdots \qquad \begin{array}{l} \mu \le \lambda \text{ iff } \mu_i \le \lambda_i \text{ for all } i \\ \mu < \lambda \text{ if } \mu \le \lambda \text{ and } \mu \ne \lambda \end{array}$$

Let $T = (t_\lambda)$ be a d-dimensional vector where λ runs through all d-compositions, and similarly for $X^{(1)}, \dots, X^{(e)}$. Fix two vectors (Ω_λ) and (ω_ν), of dimension d and e respectively.

DEFINITION 14. *The **multidimensional partial Comtet polynomials** $B_{\lambda,\mu}^{\Omega,\omega}(X^{(1)},\dots,X^{(e)})$ are defined by*

$$\omega_\nu \left(\sum_\lambda x_\lambda^{(1)}\Omega_\lambda T^\lambda\right)^{\nu_1} \cdots \left(\sum_\lambda x_\lambda^{(e)}\Omega_\lambda T^\lambda\right)^{\nu_e} = \sum_\lambda B_{\lambda,\nu}^{\Omega,\omega}(X^{(1)},\dots,X^{(e)})\Omega_\lambda T^\lambda.$$

DEFINITION 15. *The **multidimensional Comtet polynomials** $B_\lambda^{\Omega,\omega}(X^{(1)},\dots,X^{(e)})$ are defined by*

$$B_\lambda^{\Omega,\omega}(X^{(1)},\dots,X^{(e)}) = \sum_{\mu\leq\lambda} B_{\lambda,\mu}^{\Omega,\omega}(X^{(1)},\dots,X^{(e)}).$$

In this chapter we restrict attention to the case $e = 1$. We can then write ω_n instead of ω_ν. In general, it should be clear how to extend results for $d = 2$ to arbitrary d; hence for most of the discussion we will take $d = 2$.

A useful check is to note that the d-dimensional Comtet polynomials $B_\lambda^{\Omega,\omega}(X)$ subsume the $(d-1)$-dimensional polynomials by putting some x_λ equal to 0. For example, $B_{\lambda_1,0,\lambda_3}^{\Omega,\omega}(X) = B_{\lambda_1,\lambda_3}^{\Omega,\omega}(X)$.

Particular forms of the multidimensional Comtet polynomials have been considered previously. For example W. Chen [9] considers polynomials of *compositional type* $p_\lambda(X)$ which have generating function

$$\sum_{\lambda\geq 0} p_\lambda(X)\frac{T^\lambda}{\lambda!} = e^{x_1g_1(T)+x_2g_2(T)+\cdots},$$

where the $g_i(T)$ are power series with no constant term. These are generalizations of sequences of binomial type. The multidimensional analog $Y_\lambda(X)$ of the exponential Bell polynomials arises when $e = 1$, $\Omega_\lambda = 1/\lambda!$, $\omega_n = 1$. Thus

$$\sum_{|\lambda|\geq 0} Y_\lambda(X)\frac{T^\lambda}{\lambda!} = e^{\sum_{|\mu|\geq 1} x_\mu \frac{T^\mu}{\mu!}}.$$

In analogy with Proposition 1, an easy calculation shows:

PROPOSITION 26.

$$Y_\lambda(X) = \lambda!\sum\prod_\mu \frac{(x_\mu)^{c_\mu}}{c_\mu!(\mu!)^{c_\mu}}$$

where the sum is taken over $\lambda = \sum_{0<\mu<\lambda} c_\mu\mu$.

A short list for $d = 2$ is given below. We list $Y_{a_1,a_2}(x_{1,0},x_{0,1},x_{1,1},\dots)$ only for $a_1 \geq a_2$; clearly $Y_{a_2,a_1}(x_{1,0},x_{0,1},x_{1,1},\dots)$ results from $Y_{a_2,a_1}(x_{1,0},x_{0,1},x_{1,1},\dots)$ by replacing

$x_{i,j}$ by $x_{j,i}$.

$$
(38)\quad
\begin{aligned}
Y_{0,0} &= 1\\
Y_{1,0} &= x_{1,0}\\
Y_{1,1} &= x_{0,1}x_{1,0} + x_{1,1}\\
Y_{2,0} &= x_{1,0}^2 + x_{2,0}\\
Y_{2,1} &= x_{0,1}x_{1,0}^2 + 2x_{1,0}x_{1,1} + x_{0,1}x_{2,0} + x_{2,1}\\
Y_{2,2} &= x_{0,1}^2x_{1,0}^2 + x_{0,2}x_{1,0}^2 + 4x_{0,1}x_{1,0}x_{1,1} + 2x_{1,1}^2 + 2x_{1,0}x_{1,2} + x_{0,1}^2x_{2,0}\\
&\quad + x_{0,2}x_{2,0} + 2x_{0,1}x_{2,1} + x_{2,2}
\end{aligned}
$$

Similarly, we can define the multi-dimensional ordinary Bell polynomials by setting $e = 1$, $\Omega_\lambda = 1$ and $\omega_n = 1$. Here is a short list for $d = 2$ of the ordinary Bell polynomials:

$$
(39)\quad
\begin{aligned}
B^o_{0,0} &= 1\\
B^o_{1,0} &= x_{1,0}\\
B^o_{1,1} &= 2x_{0,1}x_{1,0} + x_{1,1}\\
B^o_{2,0} &= x_{1,0}^2 + x_{2,0}\\
B^o_{2,1} &= 3x_{0,1}x_{1,0}^2 + 2x_{1,0}x_{1,1} + 2x_{0,1}x_{2,0} + x_{2,1}\\
B^o_{2,2} &= x_{0,1}^2x_{1,0}^2 + 3x_{0,2}x_{1,0}^2 + 6x_{0,1}x_{1,0}x_{1,1} + x_{1,1}^2 + 2x_{1,0}x_{1,2} + 3x_{0,1}^2x_{2,0}\\
&\quad + 2x_{0,2}x_{2,0} + 2x_{0,1}x_{2,1} + x_{2,2}
\end{aligned}
$$

The following convolutional property of the multidimensional exponential Bell polynomials holds:

PROPOSITION 27.

$$Y_\lambda(U+V) = \sum_\mu \binom{\lambda}{\mu} Y_\mu(U)Y_{\lambda-\mu}(V).$$

PROOF. The proof is a straightforward generalization of the one given earlier for Proposition 2. □

2. Moments and Hankel mean-independence

Let X be a random variable taking values in $\mathcal{R} \subseteq \mathbb{R}^d$ with a d-dimensional probability distribution function $\Pr(Z)$. If λ is a d-composition, the *λ'th moment* of X is

$$m_\lambda = E[X^\lambda] = \int_{\mathcal{R}} Z^\lambda \, d\Pr(Z).$$

For example, if $X = (U, V)$ is a 2-dimensional random vector, then $m_{i,j} = E[U^iV^j]$.

We define the *moment sequence* of X to be $(m_\lambda)_{\lambda>0}$ (thus ignoring the zero'th moment $m_{0,0,\ldots} = 1$), and denote the set of all moment sequences in $\mathbb{R}^d$ by $\mathcal{M}^d$. With the above definitions it should be clear what is meant by the d-dimensional moment problem on $\mathcal{R}$. The concept of moment generating function extends in the expected way:

DEFINITION 16. *The moment generating function $G(T)$ for a d-dimensional random vector X is*

$$G(T) = \sum_{\lambda \geq 0} \frac{1}{\lambda!} E[X^\lambda] T^\lambda.$$

PROPOSITION 28. *If $X = (U, V)$ is a 2-dimensional random vector with moment generating function $G(T)$ with $T = (t_1, t_2)$, then*

$$G(T) = E[e^{T \cdot X}] = E[e^{t_1 U + t_2 V}].$$

PROOF.

$$\begin{aligned} E[e^{T \cdot X}] &= E[e^{t_1 U + t_2 V}] = E\left[\sum_{n=0}^{\infty} \frac{1}{n!}(t_1 U + t_2 V)^n\right] \\ &= E\left[\sum_{n=0}^{\infty} \frac{1}{n!} \sum_{i,j=n} \binom{n}{i} t_1^i U^i t_2^j V^j\right] = \sum_{i,j=0}^{\infty} \frac{1}{i!j!} E[U^i V^j] t_1^i t_2^j = G(T). \end{aligned}$$

□

We now extend the idea of Hankel mean-independence to higher dimensions. It is not immediately clear how to define a Hankel matrix for d-dimensional moments; however by looking in more detail at the one-dimensional case we are led to a meaningful definition.

In previous chapters we exploited the fact the non-negativity of the Hankel determinants is necessary and sufficient for solvability of the moment problem on $\mathbb{R}$. This criterion is an application of the following theorem found in [33]:

THEOREM 19. *A necessary and sufficient condition for the d-dimensional moment problem on $\mathcal{R} \subseteq \mathbb{R}^d$ defined by a sequence (m_λ) to have a solution is that the functional*

$$L(p) = \sum_{\lambda} c_\lambda m_\lambda$$

be non-negative whenever

$$p(U) = \sum_{\lambda} c_\lambda U^\lambda$$

is any a polynomial with $P(U) \geq 0$ for all $U \in \mathcal{R}$.

COROLLARY 11. *The moment space $\mathcal{M}^d$ is closed under the weak topology.*

In one dimension, the Hamburger Theorem follows by from the fact that every polynomial $p(u)$ non-negative on $\mathbb{R}$ is a sum of squares. If $p(u) = (c_0 + c_1 u + \cdots + c_n u^n)^2$, the resulting functional L is the Hankel quadratic form

$$\sum_{i,j=0}^{n} m_{i+j} c_i c_j,$$

It is well known from the theory of quadratic forms that this is non-negative for all $c_i \in \mathbb{R}$ if and only if the principal minors of the matrix $[m_{i+j}]_{i,j=0}^{n}$ are non-negative.

This suggests that we try to find an analog of the Hamburger Theorem by considering the resulting functional L applied to $p(U) = q(U)^2$. For simplicity, from here on we consider only the case $d = 2$, with obvious extensions to arbitrary d. Consider the non-negative polynomial

$$p(U) = \left[\sum_{|\lambda| \leq n} c_\lambda U^\lambda \right]^2 .$$

For definiteness, we order the coefficients c_λ in the same manner as with a two-variable Taylor series expansion, first in increasing order of $|\lambda|$, then in increasing order of λ_2 for fixed $|\lambda|$. The resulting matrix is symmetric. By the theory of quadratic forms, the functional L is non-negative if and only if the principal minors of the matrix

$$H_n = [m_{\lambda+\mu}]_{|\lambda|,|\mu|=0}^{n}$$

are all non-negative. Technically, we should not refer to H_n as a Hankel matrix, since $m_{\lambda+\mu}$ does not depend solely on $|\lambda + \mu|$. Nevertheless, we will make a slight abuse of terminology by referring to

$$\Delta_n = \det H_n$$

as the n'th Hankel determinant. For example, consider the square of a linear polynomial:

$$p(u_1, u_2) = (a + bu_1 + cu_2)^2 = a^2 + 2abu_1 + b^2u_1^2 + 2acu_2 + 2bcu_1u_2 + c^2u_2^2.$$

We have

$$L(p) = a^2 m_{0,0} + 2ab m_{1,0} + b^2 m_{2,0} + 2ac m_{0,1} + 2bc m_{1,1} + c^2 m_{0,2},$$

and thus labeling the rows and columns a, b, c, we get

$$H_2 = \begin{array}{c c} & \begin{array}{ccc} a & b & c \end{array} \\ \begin{array}{c} a \\ b \\ c \end{array} & \begin{bmatrix} 1 & m_{1,0} & m_{0,1} \\ m_{1,0} & m_{2,0} & m_{1,1} \\ m_{0,1} & m_{1,1} & m_{0,2} \end{bmatrix} \end{array}$$

For the square of a quadratic polynomial the determinant is

$$H_3 = \begin{bmatrix} 1 & m_{1,0} & m_{0,1} & m_{2,0} & m_{1,1} & m_{0,2} \\ m_{1,0} & m_{2,0} & m_{1,1} & m_{3,0} & m_{2,1} & m_{1,2} \\ m_{0,1} & m_{1,1} & m_{0,2} & m_{2,1} & m_{1,2} & m_{0,3} \\ m_{2,0} & m_{3,0} & m_{2,1} & m_{4,0} & m_{3,1} & m_{2,2} \\ m_{1,1} & m_{2,1} & m_{1,2} & m_{3,1} & m_{2,2} & m_{1,3} \\ m_{0,2} & m_{1,2} & m_{0,3} & m_{2,2} & m_{1,3} & m_{0,4} \end{bmatrix} . \tag{40}$$

We note that the non-negativity of all the principal minors of H_n is not sufficient for the solvability of the d-dimensional moment problem when $d \geq 2$. This is because for $d \geq 2$ there exist non-negative polynomials which cannot be expressed as a sum of squares, as noted by Hilbert [17]. In particular, the Motzkin polynomial

$$(u_1^2 + u_2^2 - 3)u_1^2u_2^2 + 1$$

is such a polynomial on $\mathbb{R}^2$ (see [24]). Recall that we used the sufficiency of the Hamburger Theorem to prove the remarkable fact that the one-dimensional ordinary Bell polynomials are $\mathcal{M}$-preserving. Such methods will not apply in the higher-dimensional case, and we must resort to other means.

If $V = (p_\nu(X))$ is any 2-dimensional vector with polynomial coordinates, we can define the Hankel matrix $H_n(V)$ and $\Delta_n(V)$ by replacing x_λ by v_λ in H_n and Δ_n.

We can now present an analog of Hankel mean-independence.

DEFINITION 17. *Let $V = (p_\nu(X))$ be a 2-dimensional vector with polynomial coordinates. If the Hankel determinants for $\Delta(p_\nu(U))$ are independent of x_μ, for $|\mu| = 1$, we say that $(p_\lambda(U))$ is* ***Hankel mean–independent***.

THEOREM 20. *The sequence $(Y_\nu(X))$ is Hankel mean–independent.*

PROOF. Let $X = (x_{i_1,i_2})$. Define
$X' = (x_{1,0}, x_{0,1}, 0, \dots)$, and $X'' = (0, 0, x_{2,0}, x_{1,1}, x_{0,2}, \dots)$. It is easily seen that $Y_\lambda(X') = x_{1,0}^{\lambda_1}x_{0,1}^{\lambda_2}$. By Proposition 27 we have

$$Y_\lambda(X) = \sum_\mu \binom{\lambda}{\mu} Y_\mu(X')Y_{\lambda-\mu}(X'') = \sum_\mu \binom{\lambda}{\mu}(X')^\mu Y_{\lambda-\mu}(X'').$$

Inverting the above gives

$$Y_\lambda(X'') = \sum_\mu \binom{\lambda}{\mu}(-X')^\mu Y_{\lambda-\mu}(X) = \sum_\mu \binom{\lambda}{\mu}(-x_{1,0})^{\mu_1}(-x_{0,1})^{\mu_2} Y_{\lambda-\mu}(X). \tag{41}$$

We now replace the λ'th row R_λ of H_n by

$$\sum_\mu \binom{\lambda}{\mu}(-X')^\mu R_{\lambda-\mu},$$

where R_μ is the μ'th row in the Hankel matrix. Similarly, we replace the λ'th column by

$$\sum_\mu \binom{\lambda}{\mu}(-X')^\mu C_{\lambda-\mu}$$

where C_μ represents the μ'th column. Such elementary row and column operations do not change the value of the determinant Δ_n, and hence the result follows from equation (41). □

Note that the above proof shows that the other principal minors of H_n are also independent of x_μ for all $|\mu| = 1$.

By applying the functional L to the Motzkin polynomial we get

$$m_{4,2} + m_{2,4} - 3m_{2,2} + m_{0,0}. \tag{42}$$

If $m_{i,j} = Y_{i,j}(X)$, equation (42) is not independent of $x_{1,0}$ and $x_{0,1}$, as $Y_{4,2}(X)$ contains the term $x_{1,0}^2 x_{0,1}^4$ which can only be eliminated by a term in $Y_{2,2}(X)$ which clearly does not exist. This implies that Hankel mean-independence is not merely a consequence of shifting the random variable so that its first moments $x_{1,0}$ and $x_{0,1}$ are zero.

The following proposition is of note:

PROPOSITION 29. *The $B_\lambda(X)$ are not Hankel mean-independent.*

PROOF. An easy calculation using 39 shows that

$$\Delta_2(B^o) = \begin{vmatrix} 1 & B^o_{1,0} & B^o_{0,1} \\ B^o_{1,0} & B^o_{2,0} & B^o_{1,1} \\ B^o_{0,1} & B^o_{1,1} & B^o_{0,2} \end{vmatrix} = x_{0,2}x_{2,0} - x_{1,1}^2 - 2x_{0,1}x_{1,0}x_{1,1} - x_{0,1}^2 x_{1,0}^2. \tag{43}$$

□

3. Moment sequence preserving maps

We now look at some of the $\mathcal{M}^d$-preserving properties of the multi-dimensional Comtet polynomials. To start off, it is remarkable that the ordinary d-dimensional Bell polynomials are no longer moment sequence preserving for $d \geq 2$:

PROPOSITION 30. *The $(B^o_\lambda(X))$ are not $\mathcal{M}^d$-preserving for $d \geq 2$.*

PROOF. Since each d-dimensional polynomial vector $B^o_\lambda(X)$ contains the previous $(d-1)$-dimensional vector, derived by putting some x_{λ_i} equal to 0, we need only prove the proposition for $d = 2$. Consider a 2-dimensional random vector where each component is an independent Bernoulli process with probability of success p. It is easily seen that the moments are

$$m_{i,j} = \begin{cases} p^2 & \text{if } i \neq 0 \text{ and } j \neq 0, \\ p & \text{if } i = 0 \text{ or } j = 0, \end{cases}.$$

Substituting the above values into equation (43) gives $p^2 - 4p^4$. This is negative for $p > 1/2$, and thus the map is not moment preserving. □

We now prove an extension of Theorem 2:

THEOREM 21. *Let $X = (S,T)$ be a 2-dimensional random vector with moments*

$$E[X^\lambda] = E[S^{\lambda_1}T^{\lambda_2}] = x_\lambda.$$

Then there exists a 2-dimensional random vector $Z = (U(p), V(p))$ such that if $Z_1, Z_2, \ldots$ are iid copies of Z, then

$$E[(Z_1 + \cdots + Z_m)^\lambda] = E[(U_1 + \cdots + U_m)^{\lambda_1}(V_1 + \cdots + V_m)^{\lambda_2}]$$

tends to $Y_\lambda(tX)$ as $m \to \infty$, $p \to 0$ and $mp \to t$, where t is a scalar.

PROOF. By an obvious extension of Lemma 2, if $p \in [0,1]$ there is a 2-dimensional random vector $Z(p)$ with $E[Z^\lambda] = pE[X^\lambda]$ for all $\lambda > 0$. Let $Z_1, \ldots, Z_m$ be iid copies of Z. Then

$$\begin{aligned} E[(Z_1 + \cdots + Z_m)^\lambda] &= = E[(U_1 + \cdots + U_m)^{\lambda_1}(V_1 + \cdots + V_m)^{\lambda_2}] \\ &= \sum_{\substack{i_1+i_2+\cdots=m \\ j_1+j_2+\cdots=m}} \binom{\lambda_1}{i_1 i_2 \cdots}\binom{\lambda_2}{j_1 j_2 \cdots} E[U_1^{i_1}V_1^{j_1}U_2^{i_2}V_2^{j_2}\cdots] \end{aligned}$$

By independence of the Z_i, this gives

$$\begin{aligned} &E[(U_1 + \cdots + U_m)^{\lambda_1}(V_1 + \cdots + V_m)^{\lambda_2}] \\ = &\sum_{\lambda=\sum_{0<\mu<\lambda} c_\mu \mu} \frac{m!}{(m - \sum_\mu c_\mu)!} \frac{\lambda! p^{\sum_\mu c_\mu}}{\prod_\mu c_\mu! (\mu!)^{c_\mu}} \prod_\mu E[S^{\mu_1}T^{\mu_2}]. \end{aligned}$$

We have

$$\frac{m!}{(m - \sum_\mu c_\mu)!} p^{\sum_\mu c_\mu} \sim (mp)^{\sum_\mu c_\mu} \sim t^{\sum_\mu c_\mu}.$$

Hence

$$E[(U_1 + \cdots + U_m)^{\lambda_1}(V_1 + \cdots + V_m)^{\lambda_2}] \sim \lambda! \sum_{\lambda=\sum_{0<\mu<\lambda} c_\mu \mu} \prod_\mu \frac{(tx_\mu)^{c_\mu}}{c_\mu!(\mu!)^{c_\mu}} = Y_\lambda(t),$$

by Proposition 26. $\square$

We now have that $Y_\lambda(X)$ are moments by the closure of $\mathcal{M}^d$.

As in Theorem 3, we can also interpret the multi-dimensional exponential Bell polynomials as moments of multidimensional compound Poisson process.

DEFINITION 18. *Given a multidimensional random variable U and independent identically distributed copies $U_1, U_2, \ldots$ of it, the compound Poisson random variable $S_N(U,t)$ is defined by $S_N(U,t) = \sum_{k=1}^N U_k$, where N is Poisson distributed with parameter t.*

THEOREM 22. *If U is a random variable with moments $X = (x_\lambda)$, then $Y_\lambda(tX)$ are the moments of the compound Poisson random variable $S_N(U,t)$.*

PROOF. Let $F(z) = 1 + \sum_\lambda x_\lambda Z^\lambda/\lambda!$ be the exponential moment generating function of U. Put $S_n = U_1 + \cdots + U_n$, where the U_i are iid copies of U. It is easy to see that as in the one-dimensional case, the moment generating function for the sum of iid random vectors is the product of their individual moment generating functions. Hence by independence we have

$$F(Z)^n = \sum_\nu E[S_N|N=n]\frac{Z^\mu}{\nu!}.$$

The exponential moment generating function for $S_N(U,t)$ is

$$\begin{aligned} E[e^{Z\cdot S_N}] &= \sum_{n\geq 0}\Pr(N=n)\sum_\nu E[S_N|N=n]\frac{Z^\mu}{\nu!} = \sum_{n\geq 0} e^{-t}\frac{F(Z)^n}{n!} = e^{-t+F(Z)} \\ &= e^{t\sum_{\lambda>0} X_\lambda Z_\lambda/\lambda!} = \sum_\mu Y_\mu(X)\frac{Z^\mu}{\mu!}. \end{aligned}$$

□

COROLLARY 12. *The $Y_\lambda(X)$ are $\mathcal{M}$-preserving.*

4. Concluding remarks

We have only touched on the vast number of possible extensions to multiple variables in this last chapter. For example we could easily have developed a theory for the multi-dimensional Comtet polynomials using multi-dimensional binomial posets along the lines suggested by algebras of Dirichlet type found in [12], extended the results on circular processes and generalizations of the Stirling numbers, examined relations to polynomials of compositional and plethystic type, derived a multi-dimensional characterization theorem, and much more.

Further sacrifices had to be made even in the one-dimensional work. Topics such as extension of finite binomial posets, orthogonality (as described by Touchard in [39]), higher invariance of cumulants, species, Schur functions, exponential structures [35], and such will have to be dealt with in later works.

It is hoped that this work will inspire new approaches to the investigation into the interplay of combinatorics and probability. In particular, we hope to have brought some new light into the area of research of polynomials derived from compositions, as expounded by Rota [31], Comtet [10], Touchard [38], Bell [4], Stanley [37] and more.

Bibliography

[1] P. Hanlon A.R. Calderbank and R.W. Robinson. Partitions into even and odd block size and some unusual characters of the symmetric groups. *Proc. Lond. Math. Soc. Series 3*, 70, 1986.

[2] K. Baclawski, G. C. Rota, and S. Billey. *An Introduction to the Theory of Probability.* Unpublished, 1990.

[3] E. T. Bell. Exponential numbers. *Amer. Math. Monthly*, 41:411–419, 1934.

[4] E. T. Bell. Exponential polynomials. *Amer. Math. Monthly, 41*, 41:258–277, 1934.

[5] G. Birkoff. *Lattice Theory, 3rd ed.* Amer. Math. Soc., 1967.

[6] Ch. A. Charalambides and J. Singh. A review of the Stirling numbers, their generalizations and statistical applications. *Commun. Statist.-Theory Meth.*, 17:2533–2595, 1988.

[7] Ch. A. Charlambides. A new kind of numbers appearing in the n-fold convolution of truncated binomial and negative binomial distributions. *Siam J. Appl. Math.*, 33:279–288, 1977.

[8] W. Chen. Compositional calculus. *Preprint*, 1991.

[9] W. Chen. The theory of compositionals. *Preprint*, 1991.

[10] L. Comtet. *Advanced Combinatorics.* Reidel, Dordrecht, 1974.

[11] D.J. Daley and D. Vere-Jones. *An Introduction to the Theory of Point Processes.* Springer-Verlag, 1988.

[12] P. Doubilet, G.C. Rota, and R. Stanley. On the foundations of combinatorial theory (vi): the idea of generating function. *Sixth Berkeley Symp. on Math Stat. and Prob., vol 2:Probability Theory*, 1972.

[13] R. Ehrenborg and M. Mendez. A bijective proof of inifinite variated Good's inversion. *Preprint*, 1992.

[14] W. Feller. *An Introduction To Probability Theory and Its Applications, 3rd ed.* John Wiley and Sons, Inc., 1968.

[15] G. Polya and G. Szegö. Über den zentralen Grenzwertsatz der Wahrscheinlichkeitrechnung. *Mathematische Zeitschrift 8*, 8, 1920.

[16] C.F. Gardiner. *Algebraic Structures.* Ellis Horwood Limited, 1986.

[17] D. Hilbert. Über Die Darstelluug definiter formen als Summe von Formenquadrate. *Math. Ann.*, 32, 1933.

[18] R.A Horn. On moment sequences and renewal sequences. *Journal of Mathematical Analysis and Applications*, 31:131–135, 1970.

[19] F. T. Howard. A theorem relating potential and Bell polynomials. *Discrete Mathematics*, 39:129–143, 1982.

[20] T. Kaluza. Über die Koefficienten reziproker Potenzreihen. *Math. Z.*, 28:161–170, 1928.

[21] S. Karlin. *Total Positivity, Volume I.* Stanford University Press, 1968.

[22] D. Knuth. Convolution polynomials. *Preprint*, 1991.

[23] T.M. Liggett. Total positivity and renewal theory. *Probability, Statistics, and Mathematics. Papers in Honor of Samuel Karlin*, 1989.

[24] T. S. Motzkin. *The arithmetic–geometric inequality.* Wright–Patterson AFB, 1967.

[25] S. B. Nandi and S. K. Dutta. Some discrete distributions involving Stirling numbers. *Sankya: The Journal of Statistics*, 48:301–314, 1986.

[26] H. Niederhausen. Sheffer polynomials for computing exact Kolmogorov-Smirnov and Renyi type distributions. *The Annals of Statistics*, 9 no.5:923–944, 1981.

[27] J. Riordan and H. W. Becker. The arithmatic of Bell and Stirling numbers. *Amer. J. Math.*, 70, 1948.

[28] G.C. Rota. The number of partitions of a set. *Amer. Math. Monthly Vol.*, 71, no. 5, 1964.

[29] G.C. Rota. Geometric probability. *Unpublished*, 1991.

[30] G.C. Rota. Personal communication. *Course 18.315, Fall*, 1991.

[31] G.C. Rota and S. Roman. The umbral calculus. *Advances in Mathematics*, 1967.

[32] A.J. Roth. A probabilistically attained set of polynomials that generate Stirling numbers of the second kind. *Discrete Mathematics*, 32:77–84, 1980.

[33] J. A. Shohat and J. D. Tamarkin. *The Problem of Moments.* Amer. Math. Soc., 1943.

[34] M. Sibuya. Log-Concavity of Stirling numbers and unimodality of Stirling distributions. *Ann. Inst. Statist. Math.*, 40:693–714, 1988.

[35] R. Stanley. Exponential structures. *Studies in Applied Mathematics*, 59:73–82, 1978.

[36] R. Stanley. *Enumerative Combinatorics, Volume I.* Wadsworth and Brooks/Cole, 1986.

[37] R. Stanley. *Enumerative Combinatorics, Volume II.* Unpublished notes, 1992.

[38] J. Touchard. Sur les cycles des substitutions. *Acta Math. 70*, 70:243–297, 1939.

[39] J. Touchard. Nombres exponentials et nombres de Bernoulli. *Canad. J. Math.*, 8:305–320, 1956.