7-7-2014

Recursive Methods in Number Theory, Combinatorial Graph Theory, and Probability

Jonathan Burns
University of South Florida, jtburns@mail.usf.edu

Recursive Methods in Number Theory, Combinatorial Graph Theory, and Probability

by

Jonathan Burns

A thesis submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
in Mathematics
with a concentration in Pure and Applied Mathematics
Department of Mathematics & Statistics
College of Arts and Sciences
University of South Florida

Co-Major Professor: Arcadii Grinshpan, Ph.D.
Co-Major Professor: Nataša Jonoska, Ph.D.
Scott Campbell, Ph.D.
Brendan Nagle, Ph.D.
Masahico Saito, Ph.D.

Date of Approval:
July 07, 2014

Keywords: Prime-Producing Polynomials, Quadratic Diophantine Equations, Double Occurrence Words, Braid Index, Ratios of Successive Moments, Complete Monotonicity

Dedication

To my mom, Nomi, and my sisters, Dottye and Rekha.

My dad, Robert, and brother, Louis, will have to settle for their own dissertations.

Acknowledgments

I am forever indebted to both of my major advisers, Arcadii Grinshpan and Nataša Jonoska, who have supported me throughout my graduate studies, fostered my research interest, and helped to kick-started my academic career. It has been a pleasure working with Masahico Saito, who has been more than generous with his time and assistance, and I owe Brendan Nagle a huge favor for returning from his sabbatical to take part in my defense.

I would be remiss not to thank my fellow members of the USF Math-Bio research group, both past and present, and the numerous friends and colleagues that I have had the good fortune to meet while braving the graduate experience.

This research was made possible by the NSF DUE-0756847, DMS-0900671, and NIH R01 GM109459-01 grants.

Table of Contents

List of Tables

List of Figures

Abstract

Recursion is a fundamental tool of mathematics used to define, construct, and analyze mathematical objects. This work employs induction, sieving, inversion, and other recursive methods to solve a variety of problems in the areas of algebraic number theory, topological and combinatorial graph theory, and analytic probability and statistics. A common theme of recursively defined functions, weighted sums, and cross-referencing sequences arises in all three contexts, and supplemented by sieving methods, generating functions, asymptotics, and heuristic algorithms.

In the area of number theory, this work generalizes the sieve of Eratosthenes to a sequence of polynomial values called polynomial-value sieving. In the case of quadratics, the method of polynomial-value sieving may be characterized briefly as a product presentation of two binary quadratic forms. Polynomials for which the polynomial-value sieving yields all possible integer factorizations of the polynomial values are called recursively-factorable. The Euler and Legendre prime producing polynomials of the form $n^2 + n + p$ and $2n^2 + p$, respectively, and Landau's $n^2 + 1$ are shown to be recursively-factorable. Integer factorizations realized by the polynomial-value sieving method, applied to quadratic functions, are in direct correspondence with the lattice point solutions (X, Y) of the conic sections $aX^2 + bXY + cY^2 + X - nY = 0$. The factorization structure of the underlying quadratic polynomial is shown to have geometric properties in the space of the associated lattice point solutions of these conic sections.

In the area of combinatorial graph theory, this work considers two topological structures that are used to model the process of homologous genetic recombination: assembly graphs and chord diagrams. The result of a homologous recombination can be recorded as a sequence of signed permutations called a micronuclear arrangement. In the assembly graph model, each micronuclear arrangement corresponds to a directed Hamiltonian polygonal path within a directed assembly graph. Starting from a given assembly graph, we construct all the associated micronuclear arrangements. Another way of modeling genetic rearrangement is to represent precursor and product genes as a sequence of blocks which form arcs of a circle. Associating

matching blocks in the precursor and product gene with chords produces a chord diagram. The braid index of a chord diagram can be used to measure the scope of interaction between the crossings of the chords. We augment the brute force algorithm for computing the braid index to utilize a divide and conquer strategy. Both assembly graphs and chord diagrams are closely associated with double occurrence words, so we classify and enumerate the double occurrence words based on several notions of irreducibility.

In the area of analytic probability, moments abstractly describe the shape of a probability distribution. Over the years, numerous varieties of moments such as central moments, factorial moments, and cumulants have been developed to assist in statistical analysis. We use inversion formulas to compute high order moments of various types for common probability distributions, and show how the successive ratios of moments can be used for distribution and parameter fitting. We consider examples for both simulated binomial data and the probability distribution affiliated with the braid index counting sequence. Finally we consider a sequence of multiparameter binomial sums which shares similar properties with the moment sequences generated by the binomial and beta-binomial distributions. This sequence of sums behaves asymptotically like the high order moments of the beta distribution, and has completely monotonic properties.

Chapter 1

Introduction

1.1 Methods of Recursion

Recursion is the process of repetition (repetition), and a recursive method is the use of recursion in a finite or infinite way to define, construct, or analyze a mathematical object. Often times a recursive structure will have a closed form couterpart. For instance, the Fibonacci numbers are expressed through the recursive definition $F_n = F_{n-1} + F_{n-2}$ for $n \geq 2$ such that $F_0 = 0$ and $F_1 = 1$, with Benit's formula

$$F_n = \frac{(1+\sqrt{5})^n - (1-\sqrt{5})^n}{2^n\sqrt{5}}$$

for $n \in \mathbb{N}_0$ as the closed form counterpart. The closed form allows for rapid computation of the n^{th} Fibonacci number, but the recursive definition is more convenient for proving identities and a more natural approach to the sequence. Generating functions such as the formal power series

$$G(t) = \sum_{n=0}^{\infty} a_n\, t^n \tag{1.1}$$

provide a straightforward connection between the recursively defined a_n and the closed $G(t)$. Interplay between a recursive definition and its closed form can be a powerful tool, but many modern problems lack either the recursive relation or the closed form or their definitions are impractical for computational purposes.

Constructing a set from other sets is also another common method of recursion. Consider the following case. Given initial set $K_1 = \{11\}$, the rule to obtain K_{i+1} is to first add a 1 to all the integers in all the words of K_i. Second, form $2n-1$ new words by placing exactly one 1 in all possible positions between the integers in each word. To finally obtain K_{i+1}, concatenate a 1 to the front of each word. The word 11 becomes 22 which generates 122, 212, and 221, leading to the new set $K_2 = \{1122, 1212, 1221\}$. Continuing in this fashion $(n-2)$ more times will yield a set of all double occurrence words with n symbols. It is clear from this construction that $|K_{n+1}| > |K_n|$ for all $n \in \mathbb{N}$. By contrast, the cardinality of the set formed by taking successive differences of $\{1, 4, \ldots, n^2\}$ decreases with each iteration. However, starting from an infinite set $\{\frac{1}{n}\}_{n=1}^{\infty}$ and taking successive differences will not change the cardinality.

Set constructions that increase in cardinality are common in enumeration problems, and give an upper bound to questions of algorithmic complexity and asymptotics. By contrast, set constructions which decrease the cardinality are useful in sieving and optimization problems, where a solution or set of solutions are refined with each iteration.

This study exhibits some of these recursive methods, as applied to particular topics in number theory, combinatorial graph theory, and probability. Organizationally, the chapters are largely self-contained, each chapter providing its own preliminaries and background. An attempt has been made to use the established notation for each field, and as a consequence may conflict from chapter to chapter, e.g., Γ_a is a set of matrices in Chapter 2, Γ_k is an assembly graph in Chapter 3, and $\Gamma(x)$ is the Euler integral of the second kind in Chapter 4. Chapter notations are provided in a table a the end of the chapter. Likewise, the references are sorted alphabetically according to chapter.

The topics covered in each chapter are as follows.

1.1.1 Recursive Methods in Number Theory

We identify a recursive structure among factorizations of polynomial values into two integer factors. Polynomials for which this recursive structure characterizes all non-trivial representations of integer factorizations of the polynomial values into two parts are here called recursively-factorable polynomials. In particular, we prove that n^2+1 and the prime-producing polynomials n^2+n+41 and $2n^2+29$ are recursively-factorable.

For quadratics, we prove that this recursive structure is equivalent to a Diophantine identity involving the product of two binary quadratic forms. We show that this identity may be transformed into geometric terms, relating each integer factorization $an^2 + bn + c = p\,q$ to a lattice point of the conic section $aX^2 + bXY + cY^2 + X - nY = 0$, and vice versa. Surprisingly, the set of all (X, Y) lattice point solutions for these conic sections do not depend on b or c, and may be defined independently of this Diophantine form. Further, interger factorizations for the values of $an^2 + bn + c$ manifest themselves as a line in the geometry of the lattice point solutions.

1.1.2 Recursive Methods in Combinatorial Graph Theory

A double occurrence word over a finite alphabet is a word in which each alphabet letter appears exactly twice. Such words arise naturally in the study of topology, graph theory, and combinatorics. Recently, double occurrence words have been used for studying DNA recombination events. We develop formulas for counting and enumerating several elementary classes of double occurrence words such as palindromic, irreducible, and strongly-irreducible words.

Genome rearrangement and homological recombination processes have been modeled by [27] as 4-regular spacial graphs with rigid vertices, called assembly graphs. These graphs can also be represented by double occurrence words. The rearranged DNA segments are modeled by certain types of paths in the assembly graphs called polygonal paths. We provide a sharp upper bound for the number of polygonal paths in Hamiltonian sets of polygonal paths, and present a family of graphs that achieves this bound. The recombination event can be described through a sequence of signed permutations called a micronuclear arrangements, which in the assembly graph model, correspond to directed Hamiltonian polygonal paths. We characterize all possible micronuclear arrangements associated with a given assembly graph.

Chord diagrams are another visualization of genetic recombination. The braid index of a chord diagram gives a notion of the complexity among the crossings of the chords, and the algorithm given in [29] which computes the braid index is $\mathcal{O}(n!)$. We show that the braid index of a chord diagram can be computed from the closure of its connected linear sub-chord diagrams, producing a divide-and-conquer type algorithm and the classification of chord diagrams with high braid index.

1.1.3 Recursive Methods in Probability

A moment is a property of a probability distribution which abstractly defines its shape. Various types of moments, such as central moments, factorial moments, and cumulants, have been developed to exploit different shape features of the distribution. Converting between the various types of moments is performed by inversion formulas derived from manipulations of the moment generating functions. We introduce a method for fitting a sample to a probability distribution through rational regression of the ratios of various types of moments, and give some examples using sample data and the probability distribution generated by the braid index countings.

Bernstein's theorem shows that for a continuous function f on $[0,1]$, if X has a binomial distribution with parameters n and p, then

$$\lim_{n\to\infty} \mathbb{E}\left[f\left(\frac{X}{n}\right)\right] = f(p), \tag{1.2}$$

and a similar statement holds true for the beta-binomial distribution. Hausdorff showed that moment sequences of distributions defined on the interval $[0,1]$ are completely monotonic, i.e., first differences of the values are non-positive, second differences are non-negative, etc. Taking $f(x) = x^\lambda$ in (1.2) yields a sequence of binomial moments in λ divided by n^λ. We show that this sequence is not completely monotonic for certain $p \in (0,1)$, and provide a similar result for the beta-binomial distribution. However, we exhibit a sequence of multiparameter binomial sums which has characteristics of both the binomial and beta-binomial moment sequences, and completely monotonic properties.

Chapter 2

Polynomial-Value Sieving and Recursively-Factorable Polynomials

The sieve of Eratosthenes is the oldest and most well-known of the integer sieves, and is used to find all the primes up to a given limit N. The sieve begins with the list of integers $L = (2, 3, \ldots, N)$ and proceeds iteratively by marking the smallest number on the list as prime and removing it along with its multiples from the list. The smallest number still left on the list is marked as prime and the procedure continues until the list is empty.

Algorithmically, the sieve of Eratosthenes both identifies the prime numbers in the list and yields a unique prime factorization for the composite numbers through multiple presentations of each polynomial value as product of two integers. In other words, each value $F(n) = n$ in the sequence $L = (F(2), F(3), \ldots, F(N))$ is presented as the factorization presentation $F(n) = p\,q$ for each $p \mid F(n)$. If however F is an arbitrary polynomial with integer coefficients and $p \mid F(n)$, then $p \mid F(n + k\,p)$ for each $k \in \mathbb{Z}$ too. Hence, the algorithm can be generalized to include other polynomials at the cost of missing some of the factorization presentations. Fortunately, the situation can be improved by taking both factors of each composite $F(n)$ into consideration, i.e., if $F(n) = p\,q$ is marked as being divisible by p then all $F(n + kq)$ where $k \in \mathbb{Z}$ can be marked as being divisible by q as well.

To keep track of all the factorization presentations, it suffices to record the initial value along with the sequence of quotients for the multiples of the factors, e.g., if $F(x_1) = 1 \cdot p_1$, $F(x_1 + x_2 p_1) = p_1\,p_2$ and $F(x_1 + x_2 p_1 + x_3 p_2) = p_2\,p_3$ then the factorization presentation can be reconstructed from the sequence (x_1, x_2, x_3). This method of sieving the polynomial values for integer factorizations is expressed in Theorem 2.1, and holds in the context of multivariate polynomials as well. Section 2.2 introduces a family of polynomials called *recursively-factorable polynomials* for which the collection of factorization presentations corresponding to the sequences $\{(x_1, \ldots, x_m) \in \mathbb{Z}^m\}_{m=1}^{\infty}$ yield the unique prime factorization for each value of F via presentations $F(n) = p\,q$ for each $p \mid F(n)$.

In general, recursively-factorable polynomials are rare, but there are some noteworthy instances. Particularly, the *Euler-like* and *Legendre-like* prime producing polynomials of the form $n^2 + n + c$ for $c \in$

$\{2, 3, 5, 11, 17, 41\}$ and $2n^2 + c$ for $c = \{3, 5, 11, 29\}$, respectively, and Landau's $n^2 + 1$ are recursively-factorable. The sieve of Eratosthenes verifies that the line n is also recursively-factorable, but we presently focus on recursively-factorable quadratic equations.

In Section 2.3, we introduce an identity which presents the factorization of a quadratic polynomial value as the product of two binary quadratic forms (Theorem 2.3) and show that this identity associates all the factorization presentations of the aforementioned polynomial-value sieving integer sequences with the set $\Gamma_a := \left\{ \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in M_2(\mathbb{Z}) \,\middle|\, \alpha\delta - a\,\beta\gamma = 1 \right\}$. For monic quadratics, $a = 1$ and the factorization presentations correspond to the transvection generators of $\Gamma_1 = \mathrm{SL}_2(\mathbb{Z})$ (Corollary 2.2).

In Section 2.4, a bijection is established (Theorem 2.5) between Γ_a and the set $\mathcal{L}_a$ of lattice point solutions $(X, Y) \in \mathbb{Z}^2$ for the conic sections $a\,X^2 + b\,XY + c\,Y^2 + X - nY = 0$ with $a, b, c, n \in \mathbb{Z}$, showing that $\mathcal{L}_a$ does not depend on b, c, or n. Following the mappings in Figure 1, each lattice point (X, Y) of the conic section is associated with an element of Γ_a and gives a factorization presentation for $F(n) = an^2 + bn + c$. If a factorization presentation $F(n) = p\,q$ has a corresponding integer sequence $(x_1, \ldots, x_m)$ then there is a matching element of Γ_a which corresponds to a lattice point solution of the conic section.

$$\begin{tikzcd}
F(n) = p\,q \arrow[dd, dashed, bend left, "\text{Thm. 2.2}"] & & \\
 & \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in \Gamma_a \arrow[ul, bend right, "\text{Thm. 2.3}"'] \arrow[r, bend left, "\text{Thm. 2.5}"] & (X, Y) \in \mathcal{L}_a \arrow[l, bend left, "\text{Thm. 2.5}"] \\
(x_1, \ldots, x_m) \arrow[uu, bend left, "\text{Thm. 2.1}"] \arrow[ur, bend right, "\text{Thm. 2.4}"'] & &
\end{tikzcd}$$

Figure 1.: Relationships between factorization presentations $an^2 + bn + c = p\,q$, the polynomial-value sieving sequence $(x_1, \ldots, x_m)$, the set of 2×2 integers matrices Γ_a, and the set of lattice point solutions $\mathcal{L}_a$ to the conic section $aX^2 + bXY + cY^2 + X - nY = 0$.

2.1 Polynomial-Value Sieving

THEOREM 2.1. Let $\mathcal{R}$ be a commutative ring with identity. For any polynomial $F \in \mathcal{R}[x]$ of degree d, there exists a sequence of multivariate polynomials $\{f_m(x_1, \ldots, x_m)\}_{m=0}^{\infty}$ such that $f_m(x_1, \ldots, x_m) \in \mathcal{R}[x_1, \ldots, x_m]$ and

$$F\left(\sum_{k=1}^{m} x_k\, f_{k-1}(x_1, \ldots, x_{k-1})\right) = f_{m-1}(x_1, \ldots, x_{m-1})\, f_m(x_1, \ldots, x_m) \tag{2.1}$$

where $f_0 = 1$, $f_1(x_1) = F(x_1)$, and

$$f_m = f_{m-2} + x_m \sum_{j=1}^{d} \frac{1}{j!} \left(x_m \, \frac{f_{m-1}}{f_{m-2}} \right)^{j-1} \frac{\partial^j f_{m-1}}{\partial x_{m-1}^j} \tag{2.2}$$

for $m \geq 2$ with the convention that f_m is shorthand for $f_m(x_1, \ldots, x_m)$.

Proof. Since $F(x_1 f_0) = f_0 \, f_1(x_1)$ represents the trivial factorization, the statement is initially true and we proceed by induction on m. Let $D^{(j)}$ be the jth order Hasse derivative and $D_x^{(j)} = \frac{1}{j!} \frac{d^j}{dx^j}$ be the jth order Hasse derivative with respect to the intermediate x. Applying $D_{x_{m-1}}^{(j)}$ to both sides of $F\left(\sum_{k=1}^{m-1} x_k f_{k-1}\right) = f_{m-2} \, f_{m-1}$ gives

$$(D^{(j)} F) \left(\sum_{k=1}^{m-1} x_k f_{k-1} \right) \cdot f_{m-2}^j = f_{m-2} \cdot D_{x_{m-1}}^{(j)} f_{m-1}. \tag{2.3}$$

Using the Taylor series expansion for F,

$$\begin{aligned} F\left(\sum_{k=1}^{m} x_k f_{k-1}\right) &= \sum_{j=0}^{d} (D^{(j)} F) \left(\sum_{k=1}^{m-1} x_k f_{k-1}\right) \cdot (x_m f_{m-1})^j \\ &= F\left(\sum_{k=1}^{m-1} x_k f_{k-1}\right) + (x_m f_{m-1}) \sum_{j=1}^{d} (D^{(j)} F) \left(\sum_{k=1}^{m-1} x_k f_{k-1}\right) \cdot (x_m f_{m-1})^{j-1} \\ &= f_{m-1} \cdot \left(f_{m-2} + x_m \sum_{j=1}^{d} (D^{(j)} F) \left(\sum_{k=1}^{m-1} x_k f_{k-1}\right) \cdot (x_m f_{m-1})^{j-1} \right) \end{aligned} \tag{2.4}$$

which gives a definition for $f_m(x_1, \ldots, x_m) \in \mathcal{R}[x_1, \ldots, x_m]$. Substituting (2.3) into (2.4) yields

$$f_m = f_{m-2} + x_m \sum_{j=1}^{d} \left(x_m \frac{f_{m-1}}{f_{m-2}} \right)^{j-1} D_{x_{m-1}}^{(j)} f_{m-1} \, . \qquad \square$$

REMARK 2.1.1. For $F(z) = \sum_{i=0}^{d} a_i z^i$, taking $j = d$ in equation (2.3) gives

$$\frac{D_{x_{m-1}}^{(d)} f_{m-1}}{(f_{m-2})^{d-1}} = a_d$$

for all $d \geq 1$. So for $d = 2$, Theorem 2.1 expresses f_m as

$$f_m = f_{m-2} + x_m \frac{\partial f_{m-1}}{\partial x_{m-1}} + a_2 \, x_m^2 f_{m-1} \, . \tag{2.5}$$

REMARK 2.1.2. For each sequence $(x_1, \ldots, x_m)$, if $x_i = x_{i_a} + x_{i_a}$ then

$$f_m(x_1, x_2, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_m) = f_{m+2}(x_1, x_2, \ldots, x_{i-1}, x_{i_a}, 0, x_{i_b}, x_{i+1} \ldots, x_m) \, . \tag{2.6}$$

Moreover if $(x_1, \ldots, x_m) \in \mathbb{Z}^m$, then there exists f_M such that

$$f_m(x_1, \ldots, x_m) = f_M(z_1, \ldots, z_M)$$

where $z_i \in \{-1, 0, 1\}$ and $M = \sum_{j=1}^{m} 2|x_j| - 1$.

EXAMPLE 2.1.1. Let $F(x) = 3x^2 + 5x + 11$. We compute $f_3(2, -1, 4)$ as follows:

$$\begin{aligned}
f_0 &= 1 \\
f_1(2) &= \frac{F(2 \cdot 1)}{1} = \frac{F(2)}{1} = 33 \\
f_2(2, -1) &= \frac{F(2 + (-1) \cdot 33))}{33} = \frac{F(-31)}{33} = 83 \\
f_3(2, -1, 4) &= \frac{F(-31 + 4 \cdot 83)}{83} = \frac{F(301)}{83} = 3293
\end{aligned}$$

This gives $F(301) = 273319 = 83 \times 3293$. One can also verify that

$$f_3(2, -1, 4) = f_{11}(1, 0, 1, -1, 1, 0, 1, 0, 1, 0, 1).$$

2.2 Recursively-Factorable Polynomials

Theorem 2.1 provides a means of factoring the values of a polynomial F into two integers, but these presentations may not represent the full solution set $\{(n, p, q) \in \mathbb{Z}^3 : F(n) = p\,q\}$. For example when $F(n) = n^2 + n + 7$, the integer factorization $F(1) = 3 \cdot 3$ cannot be presented via Theorem 2.1, i.e., there does not exist a finite sequence of integers $(x_1, x_2, \ldots, x_m)$ for which $f_m = 3$, $f_{m-1} = 3$, and $\sum_{k=1}^{m} x_k f_{k-1} = 1$. Proof of this fact is shown in Remark 2.3.2.

By contrast, Lemma 2.2 provides the existence of a family of polynomials $\mathcal{F}$ for which the prime integer factorization of each value of $F \in \mathcal{F}$ can be reconstructed from the presentations of Theorem 2.1. Theorem 2.2 shows that this family of polynomials contains the recursively-factorable polynomials characterized by the following property.

DEFINITION 2.2.1. Let F be a polynomial with integer coefficients. If for each integer factorization presentation $F(n) = p\,q$ there exists an $r \in \mathbb{Z}$ such that $|F(r)| < |F(n)|$ and $r \equiv n \pmod{|p|}$ or $r \equiv n \pmod{|q|}$, then n is said to satisfy the *recursively-factorable criterion* for F. If each $n \in \mathbb{Z}$ satisfies the recursively-factorable criterion for F, then the polynomial F is said to be *recursively-factorable*.

REMARK 2.2.1. Recursively-factorable polynomials are irreducible over $\mathbb{Z}$. If not then $F(n) = 0$ for some $n \in \mathbb{Z}$, but the non-trivial factorization $0 = 0 \cdot p_0$ has no associated $r \equiv n \pmod{|p_0|}$ such that $|F(r)| < |F(n)| = 0$ for any $p_0 \in \mathbb{Z}$.

LEMMA 2.1. Let F be a polynomial and $G(n) = \pm F(n-h)$ for some $h \in \mathbb{Z}$. If F is recursively-factorable, then so is G.

Proof. Suppose that $G(n) = \pm F(n-h) = p_0\, p_1$ is a non-trivial factorization. Since F is recursively-factorable, we may assume without loss of generality that there exists $q \in \mathbb{Z}$ such that $|F(r)| < |F(n-h)|$ where $r = (n-h) - q\, p_0$. Thus $|G(r+h)| < |G(n)|$ and $r + h = n - q\, p_0 \equiv n \pmod{|p_0|}$, so we may conclude that G is recursively-factorable. □

THEOREM 2.2. If F is recursively-factorable then, for each $n \in \mathbb{Z}$ and $p \in \mathbb{N}$ such that $p \mid F(n)$, there exists a finite sequence of integers $(x_1, x_2, \ldots, x_m)$ such that

$$n = \sum_{k=1}^{m} x_k f_{k-1}(x_1, \ldots, x_{k-1}) \qquad \text{and} \qquad p = |f_m(x_1, \ldots, x_{m-1}, x_m)|. \tag{2.7}$$

Proof. Fix $n \in \mathbb{Z}$. If $p = 1$ or $|F(n)|$ then the sequence (n) gives the presentation $F(n) = F(n \cdot f_0) = f_0\, f_1(n) = 1 \cdot F(n)$. Thus it is sufficient to consider the case where $F(n)$ is a composite integer with a non-trivial factorization $F(n) = p_1\, p_0$ such that $p = |p_0|$.

Let $R = \{r \in \mathbb{Z} : r \equiv n \pmod{|p_0|} \text{ or } r \equiv n \pmod{|p_1|}\}$. Since F is recursively-factorable, there exists an $r \in R$ such that $|F(r)| < |F(n)|$. Moreover there is an $r_1 \in R$ such that $|F(r_1)| \leq |F(r)|$ for all $r \in R$. Set $p_* = p_0$ or p_1 so that $r_1 \equiv n \pmod{|p_*|}$. It follows that $n = q_1\, p_* + r_1$ and $F(r_1) = p_2\, p_*$ for some $q_1, p_2 \in \mathbb{Z}$. If $|p_2| = 1$, then $F(r_1) = p_2\, p_*$ represents a trivial factorization and the sequence (r_1, q_1) yields the presentation

$$F(n) = F(r_1\, p_2 + q_1\, p_*) = f_1(r_1)\, f_2(r_1, q_1). \tag{2.8}$$

If $|p_2| \neq 1$, then $F(r_1) = p_*\, p_2$ represents a non-trivial factorization, and by the minimality of our choice of r_1 relative to all other $r \in R$ there exists an r_2 which minimizes $|F(r_2)| < |F(r_1)|$ over all $r_2 \equiv r_1 \pmod{|p_2|}$, i.e., $r_1 = q_2\, p_2 + r_2$ for some $q_2 \in \mathbb{Z}$.

We may continue in this fashion until we obtain the trivial integer factorization $F(r_{m-1}) = p_{m-1}\, p_m$ where $|p_m| = 1$, produced from a finite sequence of factors $(p^*, p_2, \ldots, p_{m-1}, p_m)$, quotients $(q_1, q_2, \ldots, q_{m-1})$ and remainders $(r_1, r_2, \ldots, r_{m-1})$ such that $r_k = q_{k+1}\, p_{k+1} + r_{k+1}$ and $F(r_k) = p_k\, p_{k+1}$ for each

$2 \le k \le m-1$. Starting with $p_m = 1$ and $F(r_{m-1}) = p_{m-1}\, p_m$ we may reverse this sequence to obtain n and p as follows:

$$p_{m-1} = \frac{F(r_{m-1})}{p_m} = \frac{F(r_{m-1})}{f_0} = f_1(r_{m-1}),$$

$$p_{m-2} = \frac{F(r_{m-2})}{p_{m-1}} = \frac{F(r_{m-1} + q_{m-1}\, p_{m-1})}{p_{m-1}} = \frac{F(r_{m-1}\, f_0 + q_{m-1}\, f_1(r_{m-1}))}{f_1(r_{m-1})} = f_2(r_{m-1}, q_{m-1}).$$

More generally

$$p_k = f_{m-k}(r_{m-1}, q_{m-1}, q_{m-2}, \ldots, q_{k+1})$$

for $2 \le k \le m-2$ and $p_* = f_{m-1}(r_{m-1}, q_{m-1}, \ldots, q_2)$.

Therefore the integer sequence $(r_{m-1}, q_{m-1}, \ldots, q_1)$ gives the presentation

$$\begin{aligned} F(n) &= F\left(r_{m-1}\, f_0 + q_{m-1} f_1(r_{m-1}) + \sum_{k=3}^{m} q_{m-k+1} f_{k-1}(r_{m-1}, q_{m-1}, \ldots, q_{m-k+2}) \right) \\ &= f_{m-1}(r_{m-1}, q_{m-1}, \ldots, q_2)\, f_m(r_{m-1}, q_{m-1}, \ldots, q_2, q_1), \end{aligned}$$

and $p = |f_m(r_{m-1}, q_{m-1}, \ldots, q_1)|$. $\square$

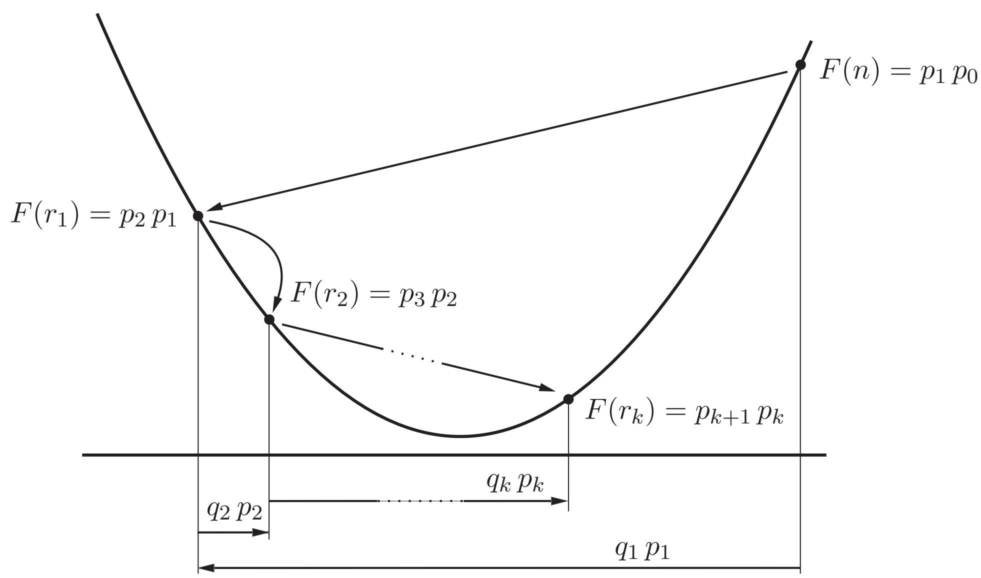


Figure 2.: Sequence of decreasing values of F used to compute $x_1, x_2, \ldots, x_m$ in $f_m(x_1, x_2, \ldots, x_m)$.

The proof of Theorem 2.2 starts with an integer factorization $F(n_0) = p_1\, p_0$ and constructs a sequence of factorizations $F(n_1) = p_1\, p_2$, $F(n_2) = p_2\, p_3, \ldots$ such that $|F(n_0)| > |F(n_1)| > |F(n_2)| \ldots$ until a prime number $F(n_m)$ with the trivial factorization $F(n_m) \cdot 1$ is reached. In this way prime-producing polynomials, which contain a large interval of consecutive prime values, make good candidates for having the recursively-factorable property.

2.2.1 Prime-producing Polynomials

In 1772, Euler [6] discovered that the polynomial $n^2 - n + 41$ produces prime numbers for $n \in [-39, 40]$, and later Legendre [15] noted that both $n^2 + n + 17$ and $n^2 + n + 41$ are prime for $n \in [-16, 15]$ and $n = [-40, 39]$, respectively. Le Lionnais considered polynomials of the type $n^2 + n + \varepsilon$ in general, which he called Euler-like polynomials [16], and integers ε for which $n^2 + n + \varepsilon$ is prime for $n = 0, 1, \ldots, \varepsilon - 2$ have come to be known as *lucky numbers of Euler*.

Rabinowitz [21] proved that ε is a lucky number of Euler if and only if the field $\mathbb{Q}(\sqrt{1 - 4\varepsilon})$ has class number 1. From this, Heegner [13] and Stark [23] showed that there are exactly six lucky numbers of Euler, namely 2, 3, 5, 11, 17, and 41.

Legendre [15] explored other types of prime-producing quadratics such as $2n^2 + \lambda$ which is prime when $\lambda = 29$ for $n = 0, 1, \ldots, 28$. Akin to the Euler-like polynomials, these quadratics give primes for $n = 0, 1, \ldots, \lambda - 1$ for prime λ if and only if $\mathbb{Q}(\sqrt{-2\lambda})$ has class number 2 [8, 17]. Baker [1] and Stark [24] found that the only such λ are 3, 5, 11, and 29.

As seen in Lemma 2.2, Euler-like and Legendre-like prime-producing quadratics are indeed recursively-factorable. Further discussion of prime-producing quadratics can be found in [18, 22].

LEMMA 2.2. The following quadratics (and their horizontal shifts) are recursively-factorable:

(i) $n^2 + c$ where $c \in \{1, 2\}$,

(ii) $n^2 + n + c$ where $c \in \{1, 2, 3, 5, 11, 17, 41\}$

(iii) $2n^2 + c$ where $c \in \{1, 3, 5, 11, 29\}$,

(iv) $2n^2 + 2n + c$ where $c \in \{1, 2, 3, 7, 19\}$,

(v) $3n^2 + c$ where $c = 2$,

(vi) $3n^2 + 3n + c$ where $c \in \{1, 2, 5, 11, 23\}$,

(vii) $4n^2 + c$ where $c \in \{1, 3, 7\}$, and

(viii) $4n^2 + 4n + c$ where $c \in \{2, 3, 5\}$.

Proof. We claim that if F is one of these polynomials and all the values within a suitably large interval $I_{\hat{n}}$ are known to satisfy the recursively-factorable criterion for F, then the remaining values outside of $I_{\hat{n}}$ also satisfy the recursively-factorable criterion.

Supposing $F(n) = an^2 + bn + c$ is one of the polynomials in cases (i)-(viii), F is a positive parabola having a minimum at either $n = 0$ or $n = -\frac{1}{2}$. Furthermore the values $F(n) = F\left(-n - \frac{b}{a}\right)$ for all $n \in \mathbb{Z}$, so if n satisfies the recursively-factorable criterion then so does $-n - \frac{b}{a}$. Also note that $|F(m)| < |F(n)|$ for $m \in I_n = \left(\min\{-n - \frac{b}{a}, n\}, \max\{-n - \frac{b}{a}, n\}\right)$.

For cases (i)-(vi), define $\hat{n}$ such that $|2\,n + \frac{b}{a}| > \lfloor\sqrt{F(n)}\rfloor$ for each $n \geq \hat{n}$. Given that for each factorization presentation $F(n) = p\,q$ either $p \leq \lfloor\sqrt{F(n)}\rfloor$ or $q \leq \lfloor\sqrt{F(n)}\rfloor$, for $n \geq \hat{n}$ there exists a $k \in \mathbb{Z}$ such that either $n - k\,p \in I_{\hat{n}}$ or $n - k\,q \in I_{\hat{n}}$. Thus if we can verify that the values within $I_{\hat{n}}$ satisfy the recursively-factorable criterion, then so do the values greater than $\hat{n}$ (and symmetrically the values less than $-\hat{n} - \frac{b}{a}$), i.e., F is recursively-factorable. In cases (vii) and (viii) we use a sharper approximation of $\min\{p, q\}$ than $\lfloor\sqrt{F(n)}\rfloor$ to determine $\hat{n}$, but the idea is the same.

In cases (i), (iii), (v), and (vii), $F(n)$ is prime (or 1) for $n \in [1 - c, c - 1]$ and $c \mid F(\pm c)$ which means $c \mid F(0) = c$, so the recursively-factorable condition is satisfied for $n \in [-c, c]$. Similarly, $F(n)$ is prime (or 1) for $n \in [1 - c, c - 2]$ in cases (ii), (iv), (vi), and (viii). The recursively-factorable condition is satisfied for $-c$, $c - 1$, and c since $c \mid F(-c), F(c - 1), F(c)$ and $F(0) = F(-1) = c$. Hence for all cases (i)-(viii) the recursively-factorable criterion is satisfied for $n \in [-c, c]$.

Case (i): For $F(n) = n^2 + c$ with $c \in \{1, 2\}$, $\hat{n} = \lfloor\sqrt{\frac{c}{3}}\rfloor = 0$ and $I_{\hat{n}} = [0] \subset [-c, c]$.

Case (ii): For $n^2 + n + c$ with $c \in \{1, 2, 3, 5, 11, 17, 41\}$, $I_{\hat{n}} = \left[-\left\lfloor -\frac{1}{2} + \sqrt{\frac{4c-1}{12}}\right\rfloor - 1, \left\lfloor -\frac{1}{2} + \sqrt{\frac{4c-1}{12}}\right\rfloor\right]$ and gives the respective $I_{\hat{n}}$ intervals corresponding to each c: $[-1, 0] \subseteq [-1, 1]$, $[-1, 0] \subseteq [-2, 2]$, $[-1, 0] \subseteq [-3, 3]$, $[-1, 0] \subseteq [-5, 5]$, $[-2, 1] \subseteq [-11, 11]$, $[-2, 1] \subseteq [-17, 17]$, and $[-4, 3] \subseteq [-41, 41]$.

Case (iii): For $F(n) = 2n^2 + c$ with $c \in \{1, 3, 5, 11, 29\}$, $I_{\hat{n}} = [-\lfloor\sqrt{\frac{c}{2}}\rfloor, \lfloor\sqrt{\frac{c}{2}}\rfloor]$ which gives the respective intervals: $[0] \subseteq [-1, 1]$, $[-1, 1] \subseteq [-3, 3]$, $[-1, 1] \subseteq [-5, 5]$, $[-2, 2] \subseteq [-11, 11]$, and $[-3, 3] \subseteq [-29, 29]$.

Case (iv): Let $F(n) = 2n^2 + 2n + c$ with $c \in \{1, 2, 3, 7, 19\}$, $I_{\hat{n}} = \left[-\lfloor\frac{\sqrt{2c-1}-1}{2}\rfloor - 1, \lfloor\frac{\sqrt{2c-1}-1}{2}\rfloor\right]$ which gives the respective intervals: $[0] \subseteq [-1, 1]$, $[0] \subseteq [-2, 2]$, $[0] \subseteq [-3, 3]$, $[-1, 1] \subseteq [-7, 7]$, and $[-2, 2] \subseteq [-19, 19]$.

Case (v): Let $F(n) = 3n^2 + 2$, then $I_{\hat{n}} = [0] \subseteq [-2, 2]$.

Case (vi): Let $F(n) = 3n^2 + 3n + c$ with $c \in \{1, 2, 5, 11, 23\}$, $I_{\hat{n}} = \left[-\lfloor \frac{\sqrt{4c-3}-1}{2} \rfloor - 1, \lfloor \frac{\sqrt{4c-3}-1}{2} \rfloor\right]$ which gives the respective intervals: $[-1, 0] \subseteq [-1, 1]$, $[-1, 0] \subseteq [-2, 2]$, $[-2, 1] \subseteq [-5, 5]$, $[-3, 2] \subseteq [-11, 11]$, and $[-5, 4] \subseteq [-23, 23]$.

Case (vii): Let F be of the form $4n^2 + c$ with $c \in \{1, 3, 7\}$. We claim that if $F(n) = p\,q$ where $p \leq q$ is an integer factorization presentation, then $p < 2n$. Observe that $p = 2n + 1$ implies that $q \geq 2n + 1$ and

$$4n^2 + 4n + 1 = (2n + 1)^2 \leq p\,q = F(n) = 4n^2 + c \quad \Longrightarrow \quad 4n + 1 \leq c$$

and is a contradiction for $n > c$. Similarly, for $p = 2n$ and $q \geq 2n + 1$,

$$4n^2 + 2n = 2n\,(2n + 1) \leq p\,q = F(n) = 4n^2 + c \quad \Longrightarrow \quad 2n \leq c$$

and is also contradiction for $n > c$. Clearly $q \neq 2n$ since $4n^2 + c = F(n) \neq p\,q = (2n)^2 = 4n^2$. Thus we are guaranteed that $2n > p$ and there exists an $r \in (1 - n, n - 1)$ such that $r \equiv n \pmod{p}$.

Case (viii): Let F be of the form $4n^2 + 4n + c$ with $c \in \{2, 3, 5\}$. As in case (vii), we show that $p < 2n$ for each integer factorization presentation $F(n) = p\,q$ where $p \leq q$. First notice that taking $p = 2n + 2$ and $q \geq 2n + 2$ leads to

$$4n^2 + 8n + 4 = (2n + 2)^2 \leq p\,q = F(n) = 4n^2 + 4n + c \quad \Longrightarrow \quad 4n + 4 \leq c$$

and is a contradiction for $n > c$. Likewise, taking $p = 2n + 1$ and $q \geq 2n + 2$ gives

$$4n^2 + 6n + 2 = (2n + 1)(2n + 2) \leq p\,q = F(n) = 4n^2 + 4n + c \quad \Longrightarrow \quad 2n + 2 \leq c$$

and again is a contradiction for $n > c$. With $p = 2n + 1$ and $q = 2n + 1$, $4n^2 + 4n + c = F(n) \neq p\,q = 4n^2 + 4n + 1$ as $c \neq 1$. Finally assume that $p = 2n$ and $q \geq 2n + 3$,

$$4n^2 + 6n = (2n)(2n + 3) \leq p\,q = F(n) = 4n^2 + 4n + c \quad \Longrightarrow \quad 2n \leq c$$

and is a contradiction for $n > c$. Finally take $q = 2n+2$ to get the contradiction $4n^2+4n = (2n)(2n+2) = p\,q \neq F(n) = 4n^2 + 4n + c$. Therefore if the recursively factorable criterion holds for the values in the interval $[-c, c]$, then $2n > p$ and the criterion holds for the values outside of the interval also. $\square$

REMARK 2.2.2. With some additional casework to show that the values over a suitably large interval satisfy the recursively-factorable criterion, it can also be shown that the polynomials in Table 1 are recursively-factorable. Some of these quadratics are prime-producing polynomials, or a horizontal shift of one, listed in [18] and [26].

Table 1: Recursively-factorable polynomials with real roots.

	$c \leq 5000$
$n^2 - c$	2, 3, 6, 7, 11, 14, 23, 38, 47, 62, 83, 167, 227, 398
$n^2 + n - c$	1, 3, 4, 5, 7, 8, 9, 10, 13, 14, 15, 17, 18, 19, 22, 23, 25, 27, 28, 33, 37, 39, 43, 45, 49, 53, 59, 67, 69, 73, 75, 79, 85, 87, 93, 103, 109, 113, 115, 127, 129, 139, 153, 163, 169, 179, 193, 199, 205, 213, 235, 269, 283, 313, 337, 349, 373, 385, 409, 469, 499, 619, 643, 655, 763, 829, 865, 883, 997, 1063, 1555
$2n^2 - c$	1, 3, 5, 7, 11, 13, 15, 19, 21, 29, 31, 35, 37, 47, 55, 61, 67, 69, 79, 91, 101, 103, 133, 139, 157, 159, 181, 199, 229, 283, 439, 571, 643, 661, 1069
$2n^2 + 2n - c$	1, 2, 3, 5, 6, 7, 9, 10, 11, 14, 15, 17, 21, 23, 26, 27, 29, 35, 38, 41, 43, 53, 63, 65, 71, 81, 83, 86, 107, 113, 146, 149, 173, 185, 191, 215, 218, 223, 251, 317, 323, 371, 413, 491, 743, 833
$3n^2 - c$	1, 2, 5, 10, 14, 29, 46, 106, 149
$3n^2 + 3n - c$	1, 2, 3, 4, 5, 7, 8, 11, 13, 17, 19, 23, 29, 31, 37, 41, 47, 55, 59, 65, 67, 79, 89, 95, 97, 107, 119, 131, 157, 163, 173, 199, 229, 257, 275, 317, 325 457, 479, 635, 637, 1379
$4n^2 - c$	1, 2, 3, 5, 7, 11, 13, 17, 19, 23, 33, 41, 47, 59, 83 107, 167, 227, 563
$4n^2 + 4n - c$	1, 2, 3, 5, 6, 7, 10, 11, 13, 19, 21, 22, 27, 31, 37, 43, 46, 51, 61, 67, 82, 85, 115, 127, 163, 166, 226, 277, 397

For these real-root quadratics, the condition $|F(m)| < |F(n)|$ for $m \in [2-n, n-1]$ no longer holds as it did in Lemma 2.2. However for $n > \max\left\{\frac{-b-\sqrt{b^2+8ac}}{2a}, \frac{-b+\sqrt{b^2+8ac}}{2a}\right\}$, $|F(m)| < |F(n)|$ for all $m \in \left(-n - \frac{b}{a}, n\right)$. Hence $\hat{n}$ can be chosen to be sufficiently large so that, for all $n > \hat{n}$, both $|F(m)| < |F(n)|$ for $m \in \left(-n - \frac{b}{a}, n\right)$ and $\lfloor\sqrt{|F(n)|}\rfloor < |2n + \frac{b}{a}|$.

2.3 Presentation as a Product of Binary Quadratic Forms

We show in this section that, for quadratic polynomials, the factorization presentations of Theorem 2.1, defined recursively as $F\left(\sum_{k=1}^{m} x_k f_{k-1}\right) = f_{m-1} f_m$, may be expressed in a closed form as the product of two binary quadratic forms. Theorem 2.4 establishes that, in this context, each factorization presentation sequence $(x_1, \ldots, x_m)$ corresponds with a particular $A_m \in M_2(\mathbb{Z})$.

2.3.1 Definitions for Δ_a, η, and ϕ_i

DEFINITION 2.3.1. Fix $F(n) = a\,n^2 + b\,n + c$. Let Δ_F, η_F, $\phi_{F,0}$, and $\phi_{F,1}$ be functions from $M_2(\mathbb{Z}) \to \mathbb{Z}$ defined such that for $A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$,

$$\begin{aligned} \Delta_F[A] &= \alpha\,\delta - a\,\beta\,\gamma, \\ \eta_F[A] &= \alpha\,\gamma + b\,\beta\,\gamma + c\,\beta\,\delta, \\ \phi_{F,0}[A] &= \alpha^2 + b\,\alpha\,\beta + a\,c\,\beta^2, \\ \phi_{F,1}[A] &= a\gamma^2 + b\,\gamma\,\delta + c\,\delta^2, \end{aligned} \tag{2.9}$$

and for natural m,

$$\phi_{F,m}[A] = \begin{cases} \phi_{F,0}[A] & \text{for even } m \\ \phi_{F,1}[A] & \text{for odd } m \end{cases}. \tag{2.10}$$

We suppress the F when it is clear by the context, favoring the notation $\Delta[A]$, $\eta[A]$, $\phi_0[A]$, $\phi_1[A]$, and $\phi_m[A]$.

DEFINITION 2.3.2. For $a \in \mathbb{Z}$, let

$$\Gamma_a := \{A \in M_2(\mathbb{Z}) : \Delta[A] = 1\}. \tag{2.11}$$

In general, the set Γ_a is not closed under matrix multiplication and does not contain inverses for each of its elements. However the case when $a = 1$ is particularly noteworthy as $\Gamma_1 = \mathrm{SL}_2(\mathbb{Z})$ is the special linear group.

2.3.2 Closed form parameterization of $aX^2 + bX + c = YZ$

THEOREM 2.3. Let $F : \mathbb{Z} \to \mathbb{Z}$ such that $F(x) = a\,x^2 + b\,x + c$. For $\alpha, \beta, \gamma, \delta \in \mathbb{Z}$,

$$F(\alpha\,\gamma + b\,\beta\,\gamma + c\,\beta\,\delta) = (\alpha^2 + b\,\alpha\,\beta + a\,c\,\beta^2)(a\,\gamma^2 + b\,\gamma\,\delta + c\,\delta^2) \tag{2.12}$$

if and only if $\alpha\,\delta - a\,\beta\,\gamma = 1$ or $-1 - \frac{b\,(\alpha\,\gamma + b\,\beta\,\gamma + c\,\beta\,\delta)}{c}$, i.e., for $A \in M_2(\mathbb{Z})$,

$$F(\eta[A]) = \phi_0[A]\,\phi_1[A] \tag{2.13}$$

if and only if $\Delta[A] = 1$ or $-1 - \frac{b}{c}\,\eta[A]$.

Proof. By expanding both sides, one can verify that:

$$\begin{aligned} &F(\alpha\,\gamma + b\,\beta\,\gamma + c\,\beta\,\delta) - (\alpha^2 + b\,\alpha\,\beta + a\,c\,\beta^2)\,(a\,\gamma^2 + b\,\gamma\,\delta + c\,\delta^2) \\ &= (1 - (\alpha\,\delta - a\,\beta\,\gamma))\,(c\,(\alpha\,\delta - a\,\beta\,\gamma) + (c + b\,(\alpha\,\gamma + b\,\beta\,\gamma + c\,\beta\,\delta))). \end{aligned}$$

□

REMARK 2.3.1. The set of matrices $\mathcal{K} = \mathcal{K}_1 \cup \mathcal{K}_2 \cup \mathcal{K}_3 \subset \Gamma_a$ where

$$\mathcal{K}_1 = \left\{ \begin{pmatrix} \pm 1 & 0 \\ s & \pm 1 \end{pmatrix} \middle|\, s \in \mathbb{Z} \right\}, \tag{2.14}$$

$$\mathcal{K}_2 = \left\{ \begin{pmatrix} s & \pm 1 \\ \mp 1 & 0 \end{pmatrix} \middle|\, s \in \mathbb{Z} \right\}, \quad \text{and} \quad \mathcal{K}_3 = \left\{ \begin{pmatrix} s & \pm 1 \\ \pm 1 & 0 \end{pmatrix} \middle|\, s \in \mathbb{Z} \right\}, \tag{2.15}$$

respectively, each correspond to the trivial factorization in Theorem 2.3 for each $s \in \mathbb{Z}$.

The Fibonacci-Brahmagupta identity has a long history in mathematics beginning with its first appearance in Diophantus' *Arithmetica* (III, 19) [4,25] c.250 in the form of $(p^2+q^2)(r^2+s^2) = (pr+qs)^2+(ps-qr)^2$. Later in c.628, Brahmagupta generalized Diophantus' identity by showing that numbers of the form $p^2+c\,q^2$ are closed under multiplication. Brahmagupta's identity was popularized in 1225 upon its reprinting in Fibonacci's *Liber Quadratorum* [7] where the first rigorous proof of the identity appeared. Finally in 1770, Euler [5] further generalized Brahmagupta's identity by providing the parametric solution

$$(ad\,p^2 + ce\,q^2)(de\,r^2 + ac\,s^2) = ae(d\,pr \pm c\,qs)^2 + cd(a\,ps \mp e\,qr)^2 \tag{2.16}$$

for the Diophantine equation $Ax^2 + By^2 = C$ with composite C. In Corollary 2.1 we show that the case $b = 0$ in Theorem 2.3 corresponds to the case $d = e = 1$ in Euler's Identity (2.16).

COROLLARY 2.1.

$$a\,(\alpha\,\gamma + c\,\beta\,\delta)^2 + c\,(\alpha\,\delta - a\,\beta\,\gamma)^2 = (\alpha^2 + a\,c\,\beta^2)\,(a\,\gamma^2 + c\,\delta^2) \tag{2.17}$$

Proof. When $b = 0$, $F(x) = a\,x^2 + c$ and

$$\begin{aligned} a\,(\alpha\,\gamma + c\,\beta\,\delta)^2 + c\cdot 1^2 &= F(\alpha\,\gamma + c\,\beta\,\delta) \\ &= (\alpha^2 + a\,c\,\beta^2)\,(a\,\gamma^2 + c\,\delta^2) \end{aligned}$$

where $\alpha\,\delta - a\,\beta\,\gamma = 1$. Hence

$$a\,(\alpha\,\gamma + c\,\beta\,\delta)^2 + c\,(\alpha\,\delta - a\,\beta\,\gamma)^2 = (\alpha^2 + a\,c\,\beta^2)\,(a\,\gamma^2 + c\,\delta^2)\,. \qquad \square$$

2.3.3 Equivalence to Polynomial-Value Sieving

THEOREM 2.4. For $F(n) = a\,n^2 + b\,n + c$ and $m \geq 0$,

$$f_m(x_1, \ldots, x_m) = \phi_m[A_m] \tag{2.18}$$

where $A_m \in \Gamma_a$ defined recursively by

$$A_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \text{and} \qquad A_{k+1} = \begin{pmatrix} \alpha_{k+1} & \beta_{k+1} \\ \gamma_{k+1} & \delta_{k+1} \end{pmatrix} = A_k + x_{k+1} B_k \tag{2.19}$$

for $1 \leq k \leq m-1$ such that

$$B_k = \begin{cases} \begin{pmatrix} a\,\gamma_k & \delta_k \\ 0 & 0 \end{pmatrix} & \text{for odd } k \\ \\ \begin{pmatrix} 0 & 0 \\ \alpha_k & a\,\beta_k \end{pmatrix} & \text{for even } k \end{cases}. \tag{2.20}$$

Proof. We shall proceed by induction on m. For each $1 \leq k \leq m$, define $A_k \in \Gamma_a$ and B_k recursively as stated in the hypothesis. Initially we see that $f_0 = 1 = \phi_0[A_0]$ and $f_1 = F(x_1) = \phi_1[A_1]$ satisfies the hypothesis. Now assume $f_{2j} = \phi_0[A_{2j}]$ and $f_{2j+1} = \phi_1[A_{2j+1}]$ for each $0 \leq j \leq \lceil \frac{m}{2} \rceil$. Suppose $m = 2j$ for some $j \geq 1$. Remark 2.1.1 gives

$$f_{2j} = f_{2j-2} + x_{2j} \frac{\partial}{\partial x_{2j-1}} [f_{2j-1}] + a\, x_{2j}^2 f_{2j-1}. \tag{2.21}$$

By the induction hypothesis

$$f_{2j-2} = \phi_0[A_{2j-2}] = \alpha_{2j-2}^2 + b\, \alpha_{2j-2}\, \beta_{2j-2} + ac\, \beta_{2j-2}^2 \tag{2.22}$$

and

$$f_{2j-1} = \phi_1[A_{2j-1}] = a\, \gamma_{2j-1}^2 + b\, \gamma_{2j-1}\, \delta_{2j-1} + c\, \delta_{2j-1}^2. \tag{2.23}$$

The partial derivative $\frac{\partial}{\partial x_{2j-1}} [\phi_1[A_{2j-1}]]$ may be evaluated through the equation $A_{2j-1} = A_{2j-2} + x_{2j-1} B_{2j-2}$. In particular

$$\frac{\partial}{\partial x_{2j-1}} [\gamma_{2j-1}] = \alpha_{2j-2} \quad \text{and} \quad \frac{\partial}{\partial x_{2j-1}} [\beta_{2j-1}] = a\, \beta_{2j-2} \tag{2.24}$$

which yields

$$\begin{aligned} \frac{\partial}{\partial x_{2j-1}} [f_{2j-1}] &= \frac{\partial}{\partial x_{2j-1}} [\phi_1[A_{2j-1}]] \\ &= \frac{\partial}{\partial x_{2j-1}} \left[a\, \gamma_{2j-1}^2 + b\, \gamma_{2j-1} \delta_{2j-1} + c\, \delta_{2j-1}^2\right] \\ &= 2\, a\, \gamma_{2j-1}\, \alpha_{2j-2} + b\, (a\, \gamma_{2j-1}\, \beta_{2j-2} + \delta_{2j-1}\, \alpha_{2j-2}) + 2\, ac\, \delta_{2j-1}\, \beta_{2j-2}\,. \end{aligned} \tag{2.25}$$

Substituting (2.22), (2.23), and (2.25) into (2.21) gives

$$\begin{aligned} f_{2j} = {} & (\alpha_{2j-2}^2 + b\, \alpha_{2j-2}\, \beta_{2j-2} + ac\, \beta_{2j-2}^2) \\ & + x_{2j} (2\, a\, \gamma_{2j-1}\, \alpha_{2j-2} + ab\, \gamma_{2j-1}\, \beta_{2j-2} + b\, \delta_{2j-1}\, \alpha_{2j-2} \\ & + 2\, ac\, \delta_{2j-1}\, \beta_{2j-2}) + x_{2j}^2\, (a) (a\, \gamma_{2j-1}^2 + b\, \gamma_{2j-1}\, \delta_{2j-1} + c\, \delta_{2j-1}^2). \end{aligned} \tag{2.26}$$

As defined in the hypothesis,

$$\begin{aligned} A_{2j} &= A_{2j-1} + x_{2j} B_{2j-1} \\ &= (A_{2j-2} + x_{2j-1} B_{2j-2}) + x_{2j} B_{2j-1} \\ &= \begin{pmatrix} \alpha_{2j-2} & \beta_{2j-2} \\ \gamma_{2j-2} & \delta_{2j-2} \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ x_{2j-1} \alpha_{2j-2} & a\, x_{2j-1} \beta_{2j-2} \end{pmatrix} + \begin{pmatrix} a\, x_{2j} \gamma_{2j-1} & x_{2j} \delta_{2j-1} \\ 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} \alpha_{2j-2} + a\, x_{2j}\, \gamma_{2j-1} & \beta_{2j-2} + x_{2j}\, \delta_{2j-1} \\ \gamma_{2j-2} + x_{2j-1}\, \alpha_{2j-2} & \delta_{2j-2} + a\, x_{2j-1}\, \beta_{2j-2} \end{pmatrix} \end{aligned} \tag{2.27}$$

so

$$\begin{aligned}
\phi_{2j}[A_{2j}] &= \phi_0[A_{2j}] \\
&= (\alpha_{2j-2} + a\, x_{2j}\, \gamma_{2j-1})^2 + ac\, (\beta_{2j-2} + x_{2j}\, \delta_{2j-1})^2 \\
&\quad + b\, (\alpha_{2j-2} + a\, x_{2j}\, \gamma_{2j-1})(\beta_{2j-2} + x_{2j}\, \delta_{2j-1}).
\end{aligned} \tag{2.28}$$

Comparing (2.26) and (2.28) shows that $f_{2j} = \phi_{2j}[A_{2j}]$.

Initially $\Delta[A_0] = 1$ and by the induction hypothesis $\Delta[A_k] = 1$ for $1 \le k \le m-1$, so we check that $A_m \in \Gamma_a$:

$$\begin{aligned}
\Delta[A_m] &= \Delta\left[\begin{pmatrix} \alpha_{m-1} + x_m\, a\, \gamma_{m-1} & \beta_{m-1} + x_m\, \delta_{m-1} \\ \gamma_{m-1} & \delta_{m-1} \end{pmatrix}\right] \\
&= (\alpha_{m-1} + x_m\, a\, \gamma_{m-1})\, \delta_{m-1} - a\, (\beta_{m-1} + x_m\, \delta_{m-1})\, \gamma_{m-1} \\
&= (\alpha_{m-1}\, \delta_{m-1} - a\, \beta_{m-1}\, \gamma_{m-1}) = \Delta[A_{m-1}] = 1\,.
\end{aligned}$$

Similarly when $m = 2j+1$, Remark 2.1.1 says that

$$f_{2j+1} = f_{2j-1} + x_{2j+1} \frac{\partial}{\partial x_{2j}}\big[f_{2j}\big] + a\, x_{2j+1}^2 f_{2j}. \tag{2.29}$$

whose partial derivative $\frac{\partial}{\partial x_{2j}}\big[f_{2j}\big] = \frac{\partial}{\partial x_{2j}}\big[\phi_{2j}[A_{2j}]\big]$ may be computed through (2.28) as

$$\frac{\partial}{\partial x_{2j}}\big[f_{2j}\big] = 2\, a\, \alpha_{2j}\, \gamma_{2j-1} + b\, \alpha_{2j}\, \delta_{2j-1} + a\, b\, \gamma_{2j-1}\, \beta_{2j} + 2\, a\, c\, \beta_{2j}\, \delta_{2j-1} \tag{2.30}$$

since $\alpha_{2j} = \alpha_{2j-2} + a\, x_{2j}\, \gamma_{2j-1}$ and $\beta_{2j} = \beta_{2j-2} + x_{2j}\, \delta_{2j-1}$. Putting (2.23), (2.29), and (2.30) together with the fact that $f_{2j} = \phi_0[A_{2j}]$ gives

$$\begin{aligned}
f_{2j+1} &= (a\, \gamma_{2j-1}^2 + b\, \gamma_{2j-1}\delta_{2j-1} + c\, \delta_{2j-1}^2) \\
&\quad + x_{2j+1}(2\, a\, \alpha_{2j}\, \gamma_{2j-1} + b\, \alpha_{2j}\, \delta_{2j-1} + a\, b\, \gamma_{2j-1}\, \beta_{2j} \\
&\quad + 2\, a\, c\, \beta_{2j}\, \delta_{2j-1}) + x_{2j+1}^2\, (a)(\alpha_{2j}^2 + b\, \alpha_{2j}\, \beta_{2j} + a\, c\, \beta_{2j}^2)
\end{aligned} \tag{2.31}$$

and may be compared with $\phi_{2j+1}[A_{2j+1}]$ which is computed thusly:

$$\begin{aligned}
\phi_{2j+1}\big[A_{2j+1}\big] &= \phi_1\left[\begin{pmatrix} \alpha_{2j-1} + a\, x_{2j}\, \gamma_{2j-1} & \beta_{2j-1} + x_{2j}\, \delta_{2j-1} \\ \gamma_{2j-1} + x_{2j+1}\, \alpha_{2j} & \delta_{2j-1} + a\, x_{2j+1}\, \beta_{2j} \end{pmatrix}\right] \\
&= a\, (\gamma_{2j-1} + x_{2j+1}\, \alpha_{2j})^2 + c\, (\delta_{2j-1} + a\, x_{2j+1}\, \beta_{2j})^2 \\
&\quad + b\, (\gamma_{2j-1} + x_{2j+1}\, \alpha_{2j})(\delta_{2j-1} + a\, x_{2j+1}\, \beta_{2k}).
\end{aligned} \tag{2.32}$$

Checking that (2.31) is equal to (2.32) shows $f_m = \phi_m[A_m]$.

We have that $\Delta[A_k] = 1$ for $1 \le k \le m-1$, so

$$\begin{aligned}\Delta[A_m] &= \Delta\left[\begin{pmatrix}\alpha_{m-1} & \beta_{m-1} \\ \gamma_{m-1} + x_m\,\alpha_{m-1} & \delta_{m-1} + x_m\,a\,\beta_{m-1}\end{pmatrix}\right] \\ &= \alpha_{m-1}\,(\delta_{m-1} + x_m\,a\,\beta_{m-1}) - a\,\beta_{m-1}(\gamma_{m-1} + x_m\,\alpha_{m-1}) \\ &= (\alpha_{m-1}\,\delta_{m-1} - a\,\beta_{m-1}\,\gamma_{m-1}) = \Delta[A_{m-1}] = 1\,.\end{aligned}$$

which completes the proof. $\square$

Combining Theorems 2.2 and 2.4 implies that for a recursively-factorable polynomial F, each non-trivial factorization presentation $(n, p, q \in \mathbb{Z} : |F(n)| = p\,q)$ is represented by some $A_m \in \Gamma_a$ via the identity $F(\eta[A_m]) = \phi_0[A_m]\,\phi_1[A_m]$ from Theorem 2.3.

EXAMPLE 2.3.1. Returning to Example 2.1.1, for $F(n) = 3n^2 + 5n + 11$ we can compute $f_3(2, -1, 4)$ using Theorem 2.4:

$$\begin{aligned} A_1 &= \begin{pmatrix}1 & 0 \\ 2 & 1\end{pmatrix} \\ A_2 &= \begin{pmatrix}1 & 0 \\ 2 & 1\end{pmatrix} + (-1)\begin{pmatrix}3\cdot 2 & 1 \\ 0 & 0\end{pmatrix} = \begin{pmatrix}-5 & -1 \\ 2 & 1\end{pmatrix} \\ A_3 &= \begin{pmatrix}-5 & -1 \\ 2 & 1\end{pmatrix} + (4)\begin{pmatrix}0 & 0 \\ -5 & 3\cdot(-1)\end{pmatrix} = \begin{pmatrix}-5 & -1 \\ -18 & -11\end{pmatrix}\end{aligned}$$

and

$$f_3(2, -1, 4) = \phi_1[A_3] = 3\,(-18)^2 + 5\,(-18)(-11) + 11\,(-11)^2 = 3293.$$

It is readily checked that $\Delta[A_3] = 1$ and meets the conditions of Theorem 2.3. Since $\eta[A_3] = 301$ and $\phi_2[A_3] = 83$, it follows that

$$F(301) = 3293 \times 83.$$

REMARK 2.3.2. The non-trivial factorization $F(1) = 3 \cdot 3$, but $F(0) = 7$ is the only value less than $F(1)$ and $1 \not\equiv 0 \pmod 3$. Likewise $F(1) = 3 \cdot 3$ cannot be represented by Theorem 2.3, since 3 cannot be represented by the binary form $\phi_0[A] = \alpha^2 + \alpha\,\beta + 7\,\beta^2$, see [3] for more details.

REMARK 2.3.3. Recall that the special linear group may be generated by its transvections [11]. In particular, $\mathrm{SL}_2(\mathbb{Z}) = \langle T, U\rangle$ where $T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $U = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$. It follows that

$$T^i = \begin{pmatrix} 1 & i \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad U^i = \begin{pmatrix} 1 & 0 \\ i & 1 \end{pmatrix} \tag{2.33}$$

for all $i \in \mathbb{Z}$.

COROLLARY 2.2. For $F(n) = n^2 + b\,n + c$,

$$f_m(x_1, x_2, \ldots, x_{2i-1}, x_{2i}, \ldots, x_m) = \phi_m[W^{x_m} \ldots T^{x_{2i}} U^{x_{2i-1}} \ldots T^{x_2} U^{x_1}] \tag{2.34}$$

where $W = \begin{cases} U, & \text{if } m \text{ is odd} \\ T, & \text{if } m \text{ is even.} \end{cases}$

Proof. From Theorem 2.4, $f_m = \phi_m[A_m]$ where $A_0 = I$ and

$$A_k = \begin{cases} \begin{pmatrix} \alpha_{k-1} & \beta_{k-1} \\ \gamma_{k-1} + x_k\alpha_{k-1} & \delta_{k-1} + x_k\beta_{k-1} \end{pmatrix} = U^{x_k} A_{k-1} & \text{for odd } k \\ \begin{pmatrix} \alpha_{k-1} + x_k\gamma_{k-1} & \beta_{k-1} + x_k\delta_{k-1} \\ \gamma_{k-1} & \delta_{k-1} \end{pmatrix} = T^{x_k} A_{k-1} & \text{for even } k \end{cases} \tag{2.35}$$

for each $1 \leq k \leq m$. □

It stands to reason that shifting a polynomial horizontally does not change the integer factorization of its values. In the case of quadratics, the specific correspondence between a parabola and its shift is expressed by the following proposition.

PROPOSITION 2.1. Let $F(n) = a\,n^2 + b\,n + c$ and set $G(n) = F(n - h)$ for some $h \in \mathbb{Z}$. For each $A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in \Gamma_a$ there is a corresponding

$$B = A + h \begin{pmatrix} a\,\beta & 0 \\ \delta & 0 \end{pmatrix} \tag{2.36}$$

for which the following conditions hold:

(i) $B \in \Gamma_a$,

(ii) $\eta_G[B] = \eta_F[A] + h$,

(iii) $\phi_{G,0}[B] = \phi_{F,0}[A]$, and

(iv) $\phi_{G,1}[B] = \phi_{F,1}[A]$.

Proof. Let $B = \begin{pmatrix} \alpha + h\,a\beta & \beta \\ \gamma + h\,\delta & \delta \end{pmatrix}$ such that $\alpha\delta - a\beta\gamma = 1$. Noting that

$$G(n) = F(n-h) = a\,n^2 + (b - 2ah)\,n + (c - bh + ah^2) : \tag{2.37}$$

$$\Delta_G[B] = (\alpha + h\,a\beta)\,\delta - a\,\beta(\gamma + h\,\delta) = \alpha\delta - a\,\beta\gamma = 1. \tag{i}$$

$$\begin{aligned} \eta_G[B] &= (\alpha + h\,a\beta)(\gamma + h\,\delta) + (b - 2ah)\,\beta(\gamma + h\,\delta) + (c - bh + ah^2)\,\beta\delta \\ &= (\alpha\gamma + b\,\beta\gamma + c\,\beta\delta) + h\,(\alpha\delta - a\beta\gamma) = \eta_F[A] + h. \end{aligned} \tag{ii}$$

$$\begin{aligned} \phi_{G,1}[B] &= a\,(\gamma + h\,\delta)^2 + (b - 2ah)(\gamma + h\,\delta)\delta + (c - bh + ah^2)\,\delta^2 \\ &= a\,\gamma^2 + b\,\gamma\delta + c\,\delta^2. \end{aligned} \tag{iii}$$

$$\begin{aligned} \phi_{G,2}[B] &= (\alpha + h\,a\beta)^2 + (b - 2ah)(\alpha + h\,a\beta)\beta + a(c - bh + ah^2)\,\beta^2 \\ &= \alpha^2 + b\,\alpha\beta + ac\,\beta^2. \end{aligned} \tag{iv}$$

$\square$

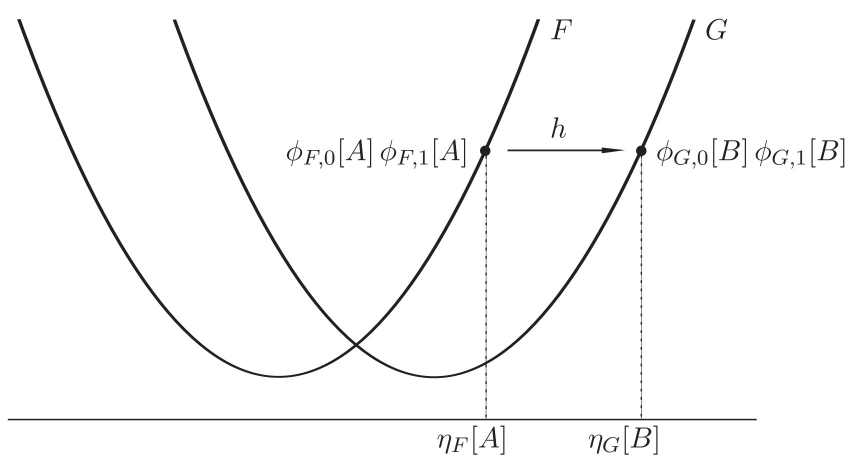


Figure 3.: Correspondence between factorization presentations for shifted parabolas.

2.4 Lattice Points on the Conic Section $aX^2 + bXY + cY^2 + X - nY = 0$

Lastly, Theorem 2.5 relates the set Γ_a with the lattice point solutions of the conic sections $aX^2 + bXY + cY^2 + X - nY = 0$. From Theorem 2.3, each $A_m \in \Gamma_a$ corresponds to an integer factorization presentation of a value of $F(n) = an^2 + bn + c$, i.e., the problem of finding lattice point solutions to these conic sections is equivalent to factoring the value of an associated quadratic polynomial.

THEOREM 2.5. For $a, b, c \in \mathbb{Z}$, let

$$\mathcal{L}_a = \left\{(X, Y) \in \mathbb{Z}^2 \mid aX^2 + bXY + cY^2 + X - nY = 0 \text{ for any } n \in \mathbb{N}\right\}. \tag{2.38}$$

The map $\psi : \Gamma_a/\mathcal{K} \to \mathcal{L}_a/\{(0,0), (-1,0), (1,0)\}$ defined by

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \mapsto \begin{pmatrix} \beta\gamma \\ \beta\delta \end{pmatrix} \tag{2.39}$$

is a bijection.

Proof. Fix $a, b, c \in \mathbb{Z}$ and consider $A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in \Gamma_a$. Set $n = \eta[A]$, $X = \beta\gamma$, $Y = \beta\delta$, and $Z = \alpha\gamma$. Direct substitution shows that

$$Z + b\,X + c\,Y = \alpha\gamma + b\,\beta\gamma + c\,\beta\delta = \eta[A] = n. \tag{2.40}$$

Since $A \in \Gamma_a$, it follows that $\Delta[A] = 1$ and $\beta\gamma\,(\alpha\delta - a\,\beta\gamma) = \beta\gamma(1)$, i.e.,

$$ZY = X + aX^2. \tag{2.41}$$

Solving for Z in (2.40) and substituting it into (2.41) shows that (X, Y) is a solution to

$$aX^2 + bXY + cY^2 + X - nY = 0. \tag{2.42}$$

Now consider the inverse map $\psi^{-1} : \mathcal{L}_a/\{(0,0), (-1,0), (1,0)\} \to \Gamma_a/\mathcal{K}_1 \bigcup \mathcal{K}_2 \bigcup \mathcal{K}_3$ defined by

$$\begin{pmatrix} X \\ Y \end{pmatrix} \mapsto \begin{pmatrix} \frac{\gcd(X,Y)}{Y}(1 + aX) & \gcd(X, Y) \\ \frac{X}{\gcd(X,Y)} & \frac{Y}{\gcd(X,Y)} \end{pmatrix}. \tag{2.43}$$

For each $L = (X, Y) \in \mathcal{L}_a$, $\Delta\left[\psi^{-1}(L)\right] = 1$ and from (2.42)

$$X(1 + aX) = Y(n - bX - cY)$$

so $\frac{\gcd(X,Y)}{Y}(1+aX) \in \mathbb{Z}$. Hence $\psi^{-1}(L) \in \Gamma_a$.

We show that ψ is injective by verifying that $\psi^{-1} \circ \psi(A) = A$ for each $A \in \Gamma_a$. Indeed, since $\Delta[A] = 1$ the $\gcd(\alpha\delta, a\,\beta\gamma) = 1$ implying that $\gcd(\gamma, \delta) = 1$, i.e., $\gcd(\beta\gamma, \beta\delta) = \beta$. Thus,

$$\psi^{-1}\psi[A] = \psi^{-1}\left[\begin{pmatrix} \beta\gamma \\ \beta\delta \end{pmatrix}\right] = \begin{pmatrix} \frac{\beta}{\beta\delta}(1+a\beta\gamma) & \beta \\ \frac{\beta\gamma}{\beta} & \frac{\beta\delta}{\beta} \end{pmatrix} = A$$

since $\Delta[A] = 1$ implies that $\alpha = \frac{1}{\delta}(1 + a\beta\gamma)$.

Likewise, for each $(X, Y) \in \mathcal{L}_a$,

$$\psi \circ \psi^{-1}\left[\begin{pmatrix} X \\ Y \end{pmatrix}\right] = \psi\left[\begin{pmatrix} \frac{G}{Y}(1+aX) & G \\ \frac{X}{G} & \frac{Y}{G} \end{pmatrix}\right] = \begin{pmatrix} X \\ Y \end{pmatrix}$$

meaning ψ is surjective. $\square$

The mapping $\psi : \mathcal{K}_1 \mapsto \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ defined by $\psi\left[\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}\right] = \begin{pmatrix} \beta\gamma \\ \beta\delta \end{pmatrix}$ is well-defined and onto, but is not one-to-one. Similarly, when $a = 1$ or -1 the respective mappings $\psi : \mathcal{K}_2 \mapsto \begin{pmatrix} -1 \\ 0 \end{pmatrix}$ and $\psi : \mathcal{K}_3 \mapsto \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ are onto but not one-to-one. Therefore the image of ψ under Γ_a is $\mathcal{L}_a$.

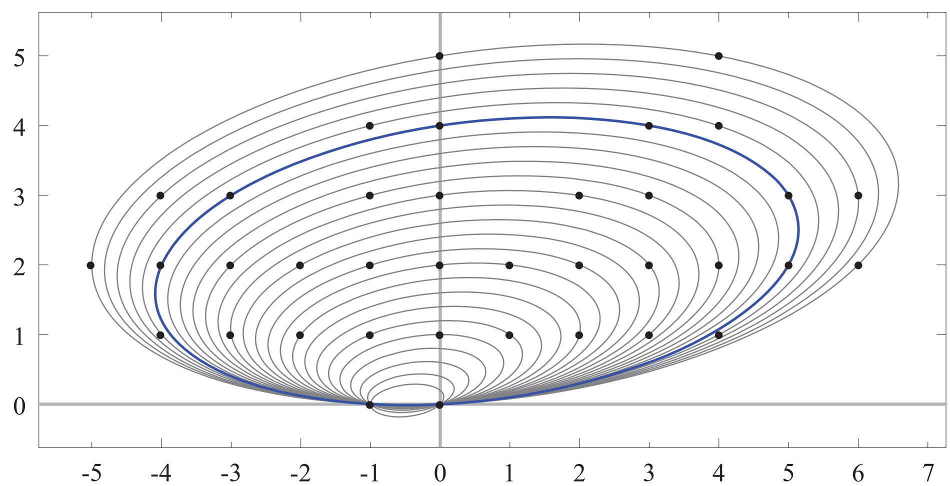


Figure 4.: Plot of $X^2 - XY + 5Y^2 + X - nY = 0$ for $n = 0, \ldots, 25$. The case $n = 20$ is highlighted in blue and lattice points $(X, Y) \in \mathcal{L}_1$ intersecting the ellipses are indicated.

EXAMPLE 2.4.1. Consider the Euler-like polynomial $F(n) = n^2 - n + 5$. It is easy to verify that $(X, Y) = (3, 4)$ is a solution of

$$X^2 - XY + 5Y^2 + X - 20Y = 0. \tag{2.44}$$

By Theorem 2.5, the point $(3, 4)$ corresponds to the element $A \in \Gamma_1$ given by

$$A = \psi^{-1}\left[\begin{pmatrix} 3 \\ 4 \end{pmatrix}\right] = \begin{pmatrix} 1 & 1 \\ 3 & 4 \end{pmatrix}$$

Thus $F(\eta[A]) = F(20) = 5 \cdot 77 = \phi_1[A]\phi_2[A]$. Similarly $(0, 0)$, $(5, 2)$, $(5, 3)$, $(0, 4)$, $(-3, 3)$, $(-4, 2)$ and $(-1, 0)$ are also lattice point solutions (see Figure 4) to (2.44) corresponding to the integer factorizations $1 \cdot 385$, $11 \cdot 35$, $7 \cdot 55$, $77 \cdot 5$, $55 \cdot 7$, $35 \cdot 11$, and $385 \cdot 1$, respectively.

REMARK 2.4.1. Gauss [9,19] showed that the general binary quadratic Diophantine equation can be reduced to a special case of the Pell equation. In particular, (2.42) can be reduced to

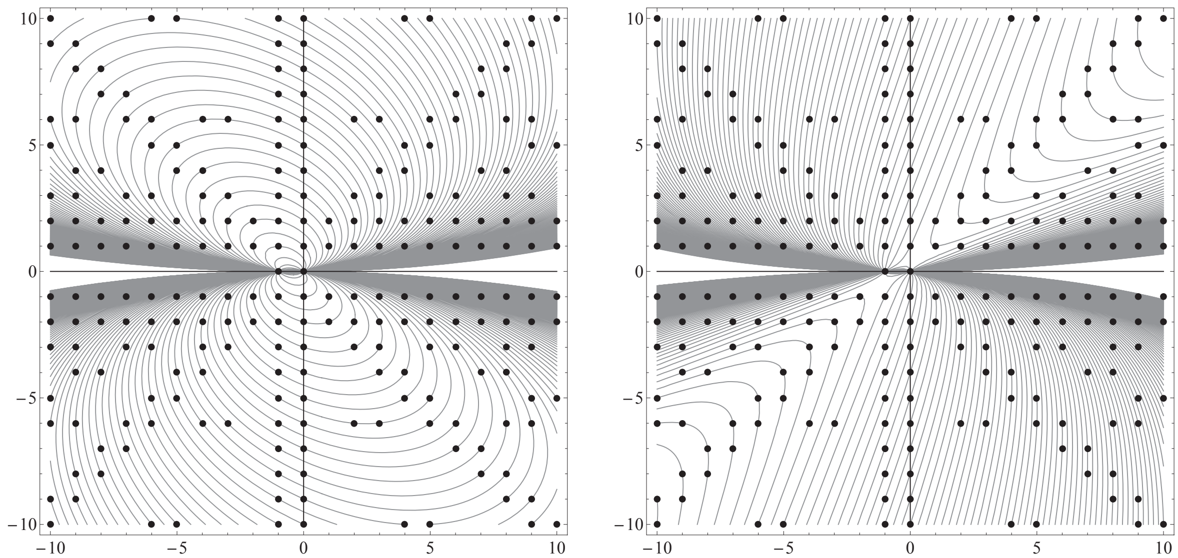
$$U^2 - (b^2 - 4ac)V^2 = 4a(an^2 + bn + c) \tag{2.45}$$

where $U = (b^2 - 4ac)Y + (b + 2an)$ and $V = 2aX + bY + 1$ provided that $b^2 - 4ac \neq 0$. The trivial factorization $F(n) = 1 \cdot F(n)$ corresponds to the solution $U = \pm(2an + b)$ and $V = \pm 1$.

2.5 Geometry of $\mathcal{L}_a$

REMARK 2.5.1. Theorem 2.5 shows that the set $\mathcal{L}_a$ can be found through the image of ψ on Γ_a which does not depend on b, c, or n, i.e., $\mathcal{L}_a$ is determined solely by $a \in \mathbb{Z}$ (see Figure 5). In particular, $(x, y) \in \mathcal{L}_a/\mathcal{K}$ if and only if $\frac{\gcd(x,y)(ax+1)}{y} \in \mathbb{Z}$, and the lattice points $(0, 0) \in L_a$, $(0, -1) \in \mathcal{L}_1$, and $(0, 1) \in \mathcal{L}_{-1}$ also depend only on a.

In this context, if $(x, y) \in \mathcal{L}_a$ then there exists a unique $A_m = \psi^{-1}(x, y) \in \Gamma_a$, and Theorem 2.3 connects this A_m with the factorization presentation $F(\eta[A_m]) = \phi_0[A_m]\phi_1[A_m]$. Theorem 2.1 shows that if $\phi_m[A_m] \mid F(\eta[A_m])$ then $\phi_m[A] \mid F(\eta[A_m] + x_{m+1}\phi_m[A_m]) = F(\eta[A_{m+1}])$ for all $x_{m+1} \in \mathbb{Z}$. Geometrically in terms of $\mathcal{L}_a$, Theorem 2.6 shows this property is equivalent to the fact that $\psi[A_m]$ and $\psi[A_{m+1}]$ lie on the same line that passes through either $(-\frac{1}{a}, 0)$ or $(0, 0)$ depending on whether m is even or odd, respectively (Figure 6).



(a) Plot of $X^2 + XY + 2Y^2 + X - nY = 0$ for $n \in \mathbb{Z}$.

(b) Plot of $X^2 - 3XY + Y^2 + X - nY = 0$ for $n \in \mathbb{Z}$.

Figure 5.: The lattice point solutions $\mathcal{L}_a$ to the conic section $a\,X^2 + b\,XY + c\,Y^2 + X - NY = 0$ are invariant under changes to b and c for fixed a.

2.5.1 Lines in $\mathcal{L}_a$ through the Origin and $(-1/a, 0)$

LEMMA 2.3. Fix $(u, v) \in \mathcal{L}_a$ with $v \neq 0$. If $(x, y) \in \mathbb{Z}^2$ is on the line

$$(-av)\,X + (1 + au)\,Y = v, \tag{2.46}$$

then $\gcd(x, y) \mid \gcd(u, v)$.

Proof. Put $d = \gcd(v, au+1)$ and note that $d \geq 1$ since $v \neq 0$. From (2.46), $v(ax+1) \equiv 0 \pmod{1+au}$ which implies that $ax + 1 \equiv 0 \pmod{(1 + au)/d}$. Since $\gcd(a, 1 + au) = 1$, there is a unique solution $x \equiv -a^{-1} \equiv u \pmod{(1 + au)/d}$ and for some $k \in \mathbb{Z}$,

$$x = \frac{1 + au}{d}k + u \quad \text{and} \quad y = a\frac{v}{d}k + v. \tag{2.47}$$

Since $(u, v) \in \mathcal{L}_a$, we have that $\frac{\gcd(u,v)(1+au)}{v} \in \mathbb{Z}$ and

$$(-a \gcd(u, v))\left(\frac{1 + au}{d}k + u\right) + \left(\frac{\gcd(u, v)(1 + au)}{v}\right)\left(a\frac{v}{d}k + v\right) = \gcd(u, v). \tag{2.48}$$

Thus the $\gcd(u, v)$ is a linear combination of x and y, which means $\gcd(x, y) \mid \gcd(u, v)$. $\square$

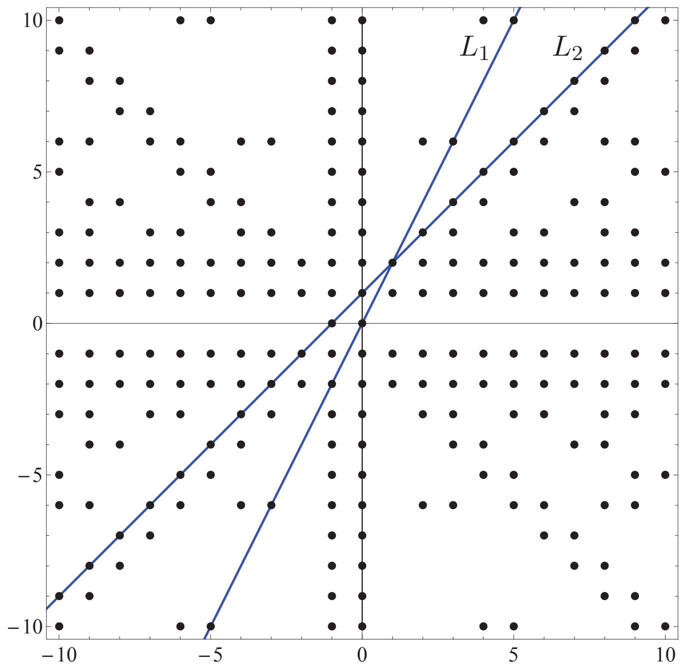


Figure 6.: Two lines $L_1, L_2 \subset \mathcal{L}_1$ intersecting at the point $(1, 2)$ and the image of ψ on Γ_1. If $A = \psi^{-1}[(1, 2)]$, then $\phi_0[A] \mid F(\eta[C_1])$ for each $C_1 \in L_1 \cap \mathcal{L}_1$ and $\phi_1 \mid F(\eta[C_2])$ for each $C_2 \in L_2 \cap \mathcal{L}_1$, or vice versa.

THEOREM 2.6. Let $(u, v) \in \mathcal{L}_a$ such that $v \neq 0$.

i. If $(x, y) \in \mathcal{L}_a$ with $y \neq 0$ is on the line

$$- v\, X + u\, Y = 0, \tag{2.49}$$

then $(\phi_1 \circ \psi^{-1})(x, y) = (\phi_1 \circ \psi^{-1})(u, v)$.

ii. If $(x, y) \in \mathcal{L}_a$ with $y \neq 0$ is on the line

$$(-av)\, X + (1 + au)\, Y = v, \tag{2.50}$$

then $(\phi_0 \circ \psi^{-1})(x, y) = (\phi_0 \circ \psi^{-1})(u, v)$.

Proof. Fix $(u, v) \in \mathcal{L}_a$ with $v \neq 0$, and consider some $(x, y) \in \mathcal{L}_a$ with $y \neq 0$ such that $uy = vx$. We claim that $\gcd(u, v)\, y = \gcd(x, y)\, v$. The $\gcd(u, v), \gcd(x, y) \geq 1$ because $v, y \neq 0$, so

$$\frac{u/\gcd(u, v)}{v/\gcd(u, v)} = \frac{u}{v} = \frac{x}{y} = \frac{x/\gcd(x, y)}{y/\gcd(x, y)}. \tag{2.51}$$

Since the fractions on the left and right hand sides of (2.51) are reduced to their lowest terms, both their numerators and demonimators are equal, respectively. Thus, $v/\gcd(u,v) = y/\gcd(x,y)$ which implies that $v\gcd(x,y) = y\gcd(u,v)$. By Theorem 2.5, $\psi^{-1}(x,y)$ is unique, and

$$\begin{aligned}
(\phi_1 \circ \psi^{-1})(x,y) &= \left(\frac{a\,x^2 + b\,xy + c\,y^2}{\gcd{(x,y)}^2}\right) \cdot \left(\frac{v^2}{v^2}\right) \\
&= \frac{a\,u^2 + b\,uv + c\,v^2}{\left(\frac{v}{y}\gcd{(x,y)}\right)^2} \\
&= \frac{a\,u^2 + b\,uv + c\,v^2}{\gcd{(u,v)}^2} \\
&= (\phi_1 \circ \psi^{-1})(u,v)\,.
\end{aligned}$$

If instead, $(au+1)\,y = v\,(ax+1)$, Lemma 2.3 gives that $\gcd{(x,y)} \mid \gcd{(u,v)}$ and symmetrically $\gcd{(u,v)} \mid \gcd{(x,y)}$, so $\gcd(x,y) = \gcd(u,v)$. Hence

$$\begin{aligned}
(\phi_0 \circ \psi^{-1})(x,y) &= \left(\gcd{(x,y)}\,\frac{ax+1}{y}\right)^2 + b\left(\gcd{(x,y)}^2\,\frac{ax+1}{y}\right) + ac\,(\gcd{(x,y)})^2 \\
&= \left(\gcd{(u,v)}\,\frac{au+1}{v}\right)^2 + b\left(\gcd{(u,v)}^2\,\frac{au+1}{v}\right) + ac\,(\gcd{(u,v)})^2 \\
&= (\phi_0 \circ \psi^{-1})(u,v)\,.
\end{aligned}$$

□

2.5.2 Horizontal and Vertical lines in $\mathcal{L}_a$

PROPOSITION 2.2. For $x, a, b, c \in \mathbb{Z}$, there are exactly $2\,\tau(x^2+x)$ solutions (Y, N) to

$$ax^2 + bxY + cY^2 + x - NY = 0 \tag{2.52}$$

where $\tau(x)$ is the number of positive divisors of x.

Proof. Fix Y as a divisor of ax^2+x, i.e., there exists a $k \in \mathbb{Z}$ such that $ax^2+x = kY$. Take $N = k+cY+bx$ and observe that

$$ax^2 + x = Yk = Y(N - c\,Y - b\,x).$$

Since Y may be either positive or negative, there are exactly $2\,\tau(ax^2+x)$ solutions to (2.52). □

PROPOSITION 2.3. For $y \in \mathbb{N}$ and $a, b, c \in \mathbb{Z}$:

(i) There are exactly $2^{\omega(y)-\omega(\gcd(y,a))}$ solutions (X, N) to

$$aX^2 + byX + cy^2 + X - Ny = 0 \tag{2.53}$$

where $X \in \mathbb{Z}_y$ and $N \in \mathbb{Z}$ given that $\omega(y)$ is the number of distinct prime divisors of y.

(ii) If (X, N) is a solution to (2.53) then for each $k \in \mathbb{Z}$,

$$(X + k\,y, N + k(1 + 2aX + by + aky))$$

is also a solution.

Proof. Fix $y \in \mathbb{N}$ with prime factorization $y = {p_1}^{q_1}{p_2}^{q_2}\ldots{p_k}^{q_k}$. Since we seek solutions $X \in \mathbb{Z}_y$, it suffices to solve

$$g(X) = aX^2 + X \equiv 0 \pmod{y}$$

which is equivalent to finding simultaneous solutions to the system of equations $X(aX+1) \equiv 0 \pmod{{p_i}^{q_i}}$ for $1 \leq i \leq k$. First note that for each i, if

$$X(aX + 1) \equiv 0 \pmod{p_i} \tag{2.54}$$

then

$$X \equiv \begin{cases} 0, -a^{-1} & \text{if } \gcd(a, p_i) = 1 \\ 0 & \text{if } \gcd(a, p_i) \neq 1 \end{cases} \pmod{p_i}.$$

are the unique solutions modulo ${p_i}^{q_i}$. By inductively applying Hensel's Lemma, we have that

$$X \equiv \begin{cases} 0, -a^{-1} & \text{if } \gcd(a, p_i) = 1 \\ 0 & \text{if } \gcd(a, p_i) \neq 1 \end{cases} \pmod{{p_i}^{q_i}}$$

are the only solutions modulo ${p_i}^{q_i}$. Thus there are 2^{k-l} distinct choices of solutions to the system (2.54) such that l is the number of prime divisors p_i which divide a. The Chinese Remainder Theorem then gives that each of these choices leads to a unique solution of $g(X) \equiv 0 \pmod{y}$ characterized by

$$X \equiv \sum_{i=1}^{k} \chi(i) \left(\frac{y}{{p_i}^{q_i}}\right) \left[\left(-\frac{a\,y}{{p_i}^{q_i}}\right)^{-1} \pmod{{p_i}^{q_i}}\right] \pmod{y}$$

such that

$$\chi(i) = \begin{cases} 0, 1 & \text{if } \gcd(a, p_i) = 1 \\ 0 & \text{if } \gcd(a, p_i) \neq 1 \end{cases}.$$

Now $g(X) \equiv 0 \pmod{y}$ implies there exists $r \in \mathbb{Z}$ such that $aX^2 + X = ry$, so

$$Ny = aX^2 + byX + cy^2 + X \quad \Longrightarrow \quad N = r + bX + cy,$$

i.e., $(X, r + bX + cy)$ is a solution to (2.53). Finally, direct substitution verifies that if (X, N) is a solution to (2.53) then

$$(X + k\,y, N + k(1 + 2aX + by + aky))$$

is also a solution. $\square$

REMARK 2.5.2. Proposition 2.1 says that for a fixed F and one of its horizontal shifts G, each $A \in \Gamma_a$ has a corresponding $B \in \Gamma_a$ such that $F(\eta_F[A]) = G(\eta_G[B])$. This property is mirrored geometrically as a horizontal skewing of (2.42) and algebraically by the fact that if (X, Y) is a solution to $aX^2 + bXY + cY^2 + X - nY = 0$, then $(X + hY, Y)$ is a solution to $aX^2 + (b + 2ah)XY + (ah^2 + bh + c)\,Y^2 + X - (n - h)Y = 0$.

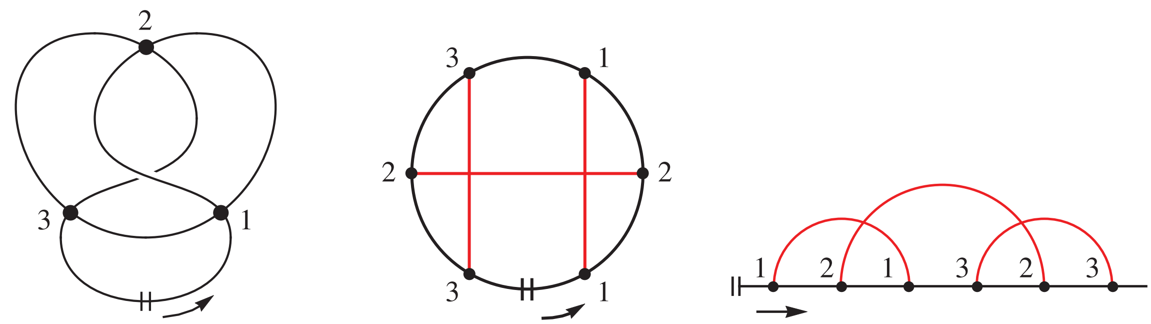
2.6 Chapter Notations

Symbol	Description	Page Reference
A	$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in M_2(\mathbb{Z})$	
A_m	$\begin{pmatrix} \alpha_m & \beta_m \\ \gamma_m & \delta_m \end{pmatrix} \in M_2(\mathbb{Z})$	17
$D^{(j)}$	j^{th} order Hasse derivative	7
$D_x^{(j)}$	j^{th} order Hasse derivative with respect to x	7
$\Delta_F[A]$	$\{A \mid \alpha\delta - a\,\beta\gamma = 1\}$	15
$\eta_F[A]$	$\alpha\gamma + b\,\beta\gamma + c\,\beta\delta$	15
$f_k(x_1, \ldots, x_k)$	a factor in the polynomial-value sieving product	7
F	$F(x) = ax^2 + bx + c$ such that $a, b, c \in \mathbb{Z}$	
Γ_a	$\{A \in M_2(\mathbb{Z}) \mid \Delta_F[A] = 1\}$	15
I_n	$\big(\min\{-n - \frac{b}{a}, n\}, \{\max\{-n - \frac{b}{a}, n\}\}\big)$	12
$\mathcal{K}$	$\mathcal{K} = \mathcal{K}_1 \cup \mathcal{K}_2 \cup \mathcal{K}_3$	16
$\mathcal{K}_1$	$\big\{\left(\begin{smallmatrix} \pm 1 & 0 \\ s & \pm 1 \end{smallmatrix}\right) \,\big|\, s \in \mathbb{Z}\big\}$	16
$\mathcal{K}_2$	$\big\{\left(\begin{smallmatrix} s & \pm 1 \\ \mp 1 & 0 \end{smallmatrix}\right) \,\big|\, s \in \mathbb{Z}\big\}$	16
$\mathcal{K}_3$	$\big\{\left(\begin{smallmatrix} s & \mp 1 \\ \pm 1 & 0 \end{smallmatrix}\right) \,\big|\, s \in \mathbb{Z}\big\}$	16
$\mathcal{L}_a$	$\{(X, Y) \in \mathbb{Z}^2 \mid a\,X^2 + b\,XY + c\,Y^2 + X - n\,Y = 0\}$	23
$M_2(\mathbb{Z})$	set of 2×2 matrices with integer entries	
$\phi_{F,0}[A]$	$\alpha^2 + b\,\alpha\beta + c\,\beta^2$	15
$\phi_{F,1}[A]$	$a\,\gamma^2 + b\,\gamma\delta + c\,\delta^2$	15
$\phi_{F,m}[A]$	$\phi_{F,0}[A]$ when m is even and $\phi_{F,1}[A]$ when m is odd	15
$\psi[A]$	$\psi\,[A] = (\beta\gamma, \beta\delta)$	23
$\mathcal{R}[x_1, \ldots, x_m]$	ring of polynomials with indeterminates $x_1, \ldots, x_m$	6
T	$\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$, one of two transvections that generator $\mathrm{SL}_2(\mathbb{Z})$	21
U	$\left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right)$, one of two transvections that generator $\mathrm{SL}_2(\mathbb{Z})$	21

Chapter 3

Recursive Properties of Double Occurrence Words and Assembly Graphs

A *double occurrence word* (DOW) w of size $n > 0$ is a word containing n distinct letters in any order which appear exactly twice, i.e., the *length* of w is $2n$. There are three common pictorial representations of double occurrence words: self-intersecting closed curves in $\mathbb{R}^3$, chord diagrams, and linear chord diagrams (or linked diagrams) as depicted in Figure 7.



(a) Self-intersecting closed curve (b) Chord Diagram (c) Linear Chord Diagram

Figure 7.: Three graphical representations of the double occurrence word 121323. Base points, indicating the starting point for reading the word, are marked by $\|$.

Topologically, a double occurrence word with n distinct letters can be interpreted as a closed curve traversing n fixed points in $\mathbb{R}^3$ twice. Such a curve (also called an assembly graph [30]) is self-intersecting and may contain over and under crossings when projected into the plane. Each curve of this type can be characterized through the double occurrence word corresponding to a path following the direction of the curve in relation to a fixed *base point*. Self-intersecting closed curves are closely related to Gauss words, knot diagrams, and their shadows [32, 38].

Chord diagrams are defined in the following way. Start with a circle, called a *backbone*, and place n distinctly labeled chords with distinct endpoints in any arrangement (possibly crossing) around the circle. Label the endpoints of each chord with the chord label. Fix a *base point* on the backbone between any two chord endpoints on the circle. The resulting diagram is called a *chord diagram*. Each chord diagram has an associated double occurrence word formed by reading the labels of the endpoints, from the base point back to base point, clockwise around the circle. See [37, 40] for more information on chord diagrams.

A *linear chord* (or *linked*) *diagram* is a pairing of $2n$ distinct ordered points [50]). Graphically, the ordered points are positioned on a line and their pairing is illustrated by an arc connecting them. Such a diagram can be specified by listing the pairs defined by the n arcs. See [49, 51, 52]. A linked diagram can be obtained from a chord diagram by cutting the outer circle at the base point. Conversely, if we arrange the points of the link diagram in a circle and mark a base point between the first and last point, the corresponding representation is a chord diagram.

Since double occurrence words naturally arise in a variety of contexts, insight into their combinatorial structure enriches several fields simultaneously. In this paper, we explore several classifications of double occurrence words based on separating larger double occurrence words into smaller double occurrence words. Further, we count and enumerate members of these classes.

Some of these formulas have been derived in completely different contexts using a variety of approaches. Moreover none of the papers we came across seemed to contain a compilation of the known formulas. In this paper we give a unified approach to deriving these formulas and provide a new formula, giving what appears to be an unobserved integer sequence.

We note that applications of double occurrence words extend to other disciplines. In Section 3.2.2, we observe that certain double occurrence words are related to particular Feynman diagrams in physics, and in Section 3.5 we establish a connection between double occurrence words and DNA recombination events.

3.1 A Note to Reader

Sections 3.2, 3.3, and 3.4 are reprinted from "Four-regular graphs with rigid vertices associated to DNA recombination, **161** (10-11), Jonathan Burns, Egor Dolzhenko, Nataša Jonoska, Tilahun Muche, and Masahico Saito, *Discrete Applied Mathematics*, 1378-1394, Copyright 2013," with permission from Elsevier, and the license agreement appears in the Appendix.

These sections reflect the work of this author with the following execptions: Theorem 3.2 was proved in collaboration with Egor Dolzhenko, and Lemma 3.4 and Theorem 3.3 were joint work with Tilahun Muche and also appear in [31].

3.2 Double Occurrence Words

3.2.1 Types of Equivalences

For convenience, we let $\Sigma = \{1, 2, \ldots, n\}$ and relabel each double occurrence word such that when i appears for the first time in the word, it is preceded by $1, 2, \ldots, i-1$. Double occurrence words labeled by this convention are said to be in *ascending order*. Two double occurrence words are said to be *equivalent* if they are equal after being relabeled in ascending order. If two double occurrence words are not equivalent, they are said to be *distinct*. Throughout this paper, we shall assume that all double occurrence words are in ascending order unless stated otherwise.

The following has been observed over 50 years ago in a different context [51] and is a part of folklore, but we provide a constructive proof for completeness.

LEMMA 3.1. There is a one-to-one correspondence between the full set of DOWs on n letters and the set $[2n-1] \times \cdots \times [3] \times [1]$. The cardinality K_n of this set is $(2n-1)!!$.

Proof. The proof follows by induction on n. As noted above, there is only one double occurrence word with one symbol.

Let w_n be an arbitrary assembly word with n distinct letters. The second occurrence of the letter 1 can appear without restriction in any of the remaining $2n-1$ positions. If $n \geq 2$, we may remove both 1's to form a word w_{n-1} with $n-1$ letters and decrease by 1 the remaining symbols. Note that w_{n-1} is also a double occurrence word in ascending order. By the inductive hypothesis there are $1 \cdot 3 \cdot \cdots \cdot (2n-3) = (2n-3)!!$ possible choices for w_{n-1}. Since there are $(2n-1)$ choices to place the 1s to obtain w_n, the number of possible double occurrence words of n symbols is $(2n-3)!!\,(2n-1) = (2n-1)!!$. $\square$

For example, 122313 is a double occurrence word in ascending order. Its *reverse* with the same letters is 313221, which is not in ascending order. By relabeling 313221 in ascending order we obtain 121332. In this example 122313 is distinct from its reverse 121332. However it is easily checked that 123312 is equivalent to its reverse which motivates the following classification.

DEFINITION 3.2.1. A double occurrence word is *palindromic* (or *symmetric*) if it is equivalent to its reverse. A double occurrence word that is palindromic is called a *palindrome*.

In all three interpretations of double occurrence words (topological, graph theoretic, and linked diagrams), the reverse word induces a diagram, isomorphic to the original, with the orientation reversed. In the topological sense, the orientation refers to the orientation of the closed curve. While the reverse of a linked diagram may be interpreted as reading the diagram right-to-left rather than left-to-right. Finally, the reverse chord diagram may be achieved by reading the letters of the circle in a counter-clockwise fashion rather than clockwise.

If we wish to count the non-isomorphic diagrams generated from double occurrence words, we observe that each diagram can have exactly two orientations. Thus, no more than two distinct double occurrence words can correspond to the same diagram with regard to a starting base point.

If a diagram corresponds to a palindrome, only one distinct double occurrence word is associated with the diagram. Therefore we may count the number of non-isomorphic diagrams with regard to a base point as

$$\begin{aligned}\text{Total Diagrams } &= (\text{\# of Palindromes}) + \frac{1}{2}(\text{\# of Non-Palindromes}) \\ &= \frac{(\text{\# of D.O. Words}) + (\text{\# of Palindromes})}{2}. \qquad (3.1)\end{aligned}$$

We will make use of this formula extensively throughout Section 3.3 to count the number of distinct diagrams corresponding to double occurrence words with each separation property.

It should be noted that omitting the base point in the closed curve or chord diagram makes it possible for more than two double occurrence words to be associated with the same diagram. For instance, rotating the base point around the circle in Figure 7 would lead to 121323, 213231, and 132312 which is 121323, 123132, and 123213 in ascending order, respectively. We do not consider isomorphisms of this type in this paper.

3.2.2 Types of Separations

As mentioned in the introduction, double occurrence words regularly appear in various fields of mathematics. Unfortunately as a result, there are several different, and sometimes conflicting, definitions used to express identical properties. We shall make note of these discrepancies in notation as they come up.

Jacques Touchard was one of the first researchers to comprehensively consider the counting of double occurrence words. In his paper [51], he classified several types of linked diagrams and enumerated the number of diagrams containing a fixed number of crossings. He introduced the classification of "unique systems" and "proper unique systems" which coincide with the following two definitions for irreducible and strongly-irreducible words.

DEFINITION 3.2.2. If a double occurrence word w can be written as a product $w = uv$ of two non-empty double occurrence words u, v, then w is called *reducible*; otherwise, it is called *irreducible*.

The number of irreducible double occurrence words has a close connection with the number of non-isomorphic unlabeled connected Feynman diagrams (also called *irreducible Feynman diagrams* [47]) arising in a simplified model of quantum electrodynamics [34, 41].

This definition for irreducibility agrees with [27] and [30] yet conflicts with [49] where "irreducible" is used for our notion of strongly-irreducible as defined below.

DEFINITION 3.2.3. A non-empty double occurrence word is *strongly-irreducible* if it does not contain a proper sub-word that is also a double occurrence word.

The double occurrence word 12213434 is reducible because it can be written as the product of the two double occurrence words 1221 and 3434, but 12344123 is irreducible. However, since 44 is a proper sub-word of 12344123 it is not strongly-irreducible. The word 12132434 is strongly-irreducible. By definition, strongly-irreducible words are also irreducible, so 12132434 is irreducible as well. In particular 11 is strongly-irreducible.

Strongly-irreducible double occurrence words are also called *connected* words [40]. This terminology is motivated by the *circle graph* associated with a chord diagram. The circle graph is formed by representing the chords as vertices and the intersection of those chords as edges in the graph. In the topological convention, a circle graph is also called an *interlinking graph* [32]. Without too many difficulties it can be proven that a double occurrence word is strongly-irreducible if and only if the circle graph of the corresponding chord diagram, or interlinking graph of the corresponding closed curve, is connected.

LEMMA 3.2. Every double occurrence word contains a strongly-irreducible sub-word.

Proof. If a double occurrence word w is strongly-irreducible, then w itself is a strongly-irreducible sub-word of w. Double occurrence words which are not strongly-irreducible, by definition, contain a proper sub-word w_1 which is a double occurrence word and is either strongly-irreducible or not. If the sub-word is not strongly-irreducible we check the reducibility of its proper sub-word w_2. Since w has finite length, we must reach a double occurrence word w_i, which is a strongly-irreducible proper sub-word of w_{i-1}, through finite recursion. Since w_i must be a proper sub-word of w, this completes the proof. □

3.3 Counting

It is well known [30, 40, 49, 52] and straightforward to show that the total number of double occurrence words K_n of size n is $(2n-1)!!$. Formula (3.1) motivates us to enumerate the number of double occurrence words which correspond to palindromes.

3.3.1 Palindromes

THEOREM 3.1. The number L_n of palindromic double occurrence words of length $2n$, is given by

$$L_n = \sum_{k=0}^{\lfloor n/2 \rfloor} \frac{n!}{(n-2k)!\,k!} = {}_2F_0\left[\begin{matrix} \frac{-n}{2}, & \frac{-n+1}{2} \\ & - \end{matrix}\ ;\ 4\right] \quad \text{for } n \geq 1. \tag{3.2}$$

See section 4.1.2 for the definition of ${}_2F_0$.

Proof. Observe that $L_1 = 1$ since there is a unique one letter palindrome, and $L_2 = 3$ because 1122, 1212, and 1221 are all the two letter palindromes.

If a double occurrence word w of size $n \geq 2$ is a palindrome beginning and ending with 1, then the word formed by removing both 1s is also a palindrome. Hence there are L_{n-1} palindromes with n letters that start and end with 1.

Now consider a word w of size $n \geq 3$ where the second symbol 1 is at the position $j \neq 2n$. Note that there are $2n-2$ possible positions for j. Then the word w is a palindrome if and only if w contains the same symbol s at the positions $2n$ and $n-j+1$. Removing symbols 1 and s from w, and relabeling the resulting word accordingly, produces a palindrome of length $n-2$. Hence there are L_{n-2} palindromes that have a symbol 1 at the jth position for $2 \leq j \leq 2n-1$.

According to the above argument,

$$L_n = L_{n-1} + (2n-2)L_{n-2} \text{ for } n \geq 3, \quad L_1 = 1 \text{ and } L_2 = 3$$

is a recurrence relation for L_n. It is known [46] that the closed formula for this recursive relations is as stated. □

This formula was expressed in 2008 by Ross Drewe in $A047974$ of the OEIS [46], in connection with higher dimensional, invertible, Boolean-like logic functions [48]. Similar results, such as the number of palindromic chord diagrams without a base point, were known in 2000 [50]. The above proof reprinted here is found in [30].

3.3.2 Irreducibles

Though Touchard introduced the classification of irreducible words in 1952, there seems to be little continuation of his efforts. In 2000, Martin and Kearney [41] expressed the number of irreducible words in the broader context of solutions to generating functions. Here, we address the count and construction of both the irreducible double occurrence words and irreducible palindromes directly.

LEMMA 3.3. The number of irreducible double occurrence words I_n with length $2n$ satisfies the recurrence formula $I_1 = 1$ and

$$I_n = (2n-1)!! - \sum_{k=1}^{n-1} I_{n-k}\,(2k-1)!! \quad \text{for } n \geq 2. \tag{3.3}$$

Proof. We shall count the number of irreducible double occurrence words by subtracting the number of reducible double occurrence words from the total number of double occurrence words of length $2n$ and show that each reducible word may be written as the product of an irreducible word and a non-empty double occurrence word.

Without loss of generality, let $w = uv$ be a reducible double occurrence word of length $2n$ such that u is also an irreducible double occurrence word. Note that every proper prefix of an irreducible word is not necessarily a double occurrence word. If the length of v is $2k$, for some $1 \leq k \leq n-1$, then the length of u is $2(n-k)$. By construction, u is irreducible and is counted among I_{n-k} and v is counted among the $(2k-1)!!$ possible double occurrence words of length $2k$.

Summing over the possible symbols in v yields the desired count. Since u is irreducible and v is non-empty, this ensures that each reducible double occurrence word w is counted exactly once. □

THEOREM 3.2. The number of irreducible palindromes J_n with length $2n$ satisfies the recurrence formula $J_1 = 1$ and

$$J_n = L_n - \sum_{k=1}^{\lfloor n/2 \rfloor} (2k-1)!!\; J_{n-2k} \quad \text{for } n \geq 2. \tag{3.4}$$

where L_n is the total number of palindromes with length $2n$.

Proof. Similar to the above argument, we first count the reducible palindromes and subtract them from the total number of palindromic words.

Suppose w is a reducible double occurrence word with length $2n$. Then w can be written as $w = uvu'$ where u is an arbitrary double occurrence word with length $2k$ $(1 \leq k \leq \lfloor n/2 \rfloor)$, u' is the double occurrence word corresponding to u by reversing the orientation, and v is an irreducible palindrome with length $2(n - 2k)$. $\square$

Though the number of irreducible double occurrence words appears in the OEIS (A000698), we note that the number of irreducible palindromes is the only sequence discussed in this paper which is not currently listed in the OEIS [46]. See Table 5, Table 6, and Table 3 for the number of irreducibles, strong-irreducibles, and the number of non-isomorphic diagrams as defined according to (3.1), respectively.

3.3.3 Strong-Irreducibles

The classification of strongly-irreducible double occurrence words was introduced in [52] and the first counting of the strong-irreducibles was done by Stein in [49]. Stein was the first to count both the strongly-irreducible double occurrence words and the strongly-irreducible palindromes, but his counting methods and recursive formulas were simplified in [42] and later by Klazar in [40]. In Theorem 3.3, we present a proof similar to [40] expressed in terms of language theory.

Using language theory to count double occurrence words led directly to a characterization of the strongly-irreducible double occurrence words, which we express in Lemma **??**, and Theorem 3.3 follows as a natural consequence.

LEMMA 3.4. Every strongly-irreducible double occurrence word w in ascending order may be written in a unique form as $w = 1u_1v_11v_2u_2$ where $1u_11u_2$ and v_1v_2 are both strongly-irreducible.

Proof. Let w be strongly-irreducible. Every double occurrence word w in ascending order must be of the form $w = 1p_11p_2$. Delete both 1's. Then we have a double occurrence word $p_1p_2 = u_1xu_2$ where x is the first strongly-irreducible double occurrence word of smallest positive length. Thus u_1 and u_2 are uniquely defined. Note that u_1 and u_2 may be empty words.

Let v_1 be the prefix of x which is a suffix of p_1 and let v_2 be the suffix of x which is the prefix of p_2. This means that $x = v_1v_2$. Neither v_1 nor v_2 is empty as it would imply that x is a sub-word of either p_1 or p_2 which would constitute a proper sub-word of w. Since w is taken to be strongly-irreducible, this cannot be the case.

We show that $1u_11u_2$ is strongly-irreducible. Suppose not. Then there exists a non-empty double occurrence sub-word z in either u_1 or u_2 which implies that w contains z and is not strongly-irreducible. This is a contradiction. Hence $1u_11u_2$ and v_1v_2 are strongly-irreducible. $\square$

THEOREM 3.3. The number of strongly-irreducible double occurrence words S_n with length $2n$ satisfies the recurrence formula

$$S_n = (n-1)\sum_{k=1}^{n-1} S_k\, S_{n-k}, \tag{3.5}$$

where $S_1 = 1$ and $n \geq 2$.

Proof. Note that the only strongly-irreducible double occurrence word of length 2 is 11, i.e., $S_1 = 1$.

Let u and v be strongly-irreducible double occurrence words such that the length of v is $2k$, the length of u is $2(n-k)$, $u = 1u_1 1u_2$, and $v = v_1 v_2$. Since the length of v is $2k$, there are $2k-1$ ways to write $v = v_1 v_2$ with v_1, v_2 not empty. By Lemma **??**, each strongly-irreducible double occurrence word w of length $2n$ can be uniquely represented as $w = 1u_1 v_1 1 v_2 u_2$. Hence there are $2k-1$ possibilities for such w's to be formed from each u and v.

Since there are S_{n-k} choices for u and S_k choices for v the total counting for S_n when $n \geq 2$ is given by

$$S_n = \sum_{k=1}^{n-1} (2k-1)\, S_k\, S_{n-k} = (n-1)\sum_{k=1}^{n-1} S_k\, S_{n-k}. \qquad \square$$

For completeness, we state Klazar's counting formula of the strongly-irreducible palindromes. See [40] for the proof.

THEOREM 3.4. Let S_n and T_n be the number of strongly-irreducible double occurrence words and strongly-irreducible palindromes of length $2n$, respectively. Then

$$T_n = \sum_{i=1}^{n-2} T_i\, T_{n-i} + \sum_{i=1}^{\lfloor n/2 \rfloor} (2n-4i-1)\, S_i\, T_{n-2i} \tag{3.6}$$

for $n \geq 2$ where $T_0 = -1$ and $T_1 = 1$.

Theorem 3.3 and Theorem 3.4 correspond to the sequences A000699 and A004300 listed in the OEIS. For the first few values of these sequences, see Table 5 and Table 6.

3.4 Assembly Graphs

An *assembly curve* k is a smooth map $k : [0,1] \to \mathbb{R}^3$ with no singularities except a finite number of transversal self-intersections, called *rigid vertices* of k. If $k(0) = k(1)$, then k is called a *circular assembly curve*. Each assembly curve k induces a graph structure called an *assembly graph* Γ_k. Consider an assembly curve with self-intersections at $k(t_1), k(t_2), \ldots, k(t_n)$ where $t_i \in (0,1)$ for all $i \in [n]$. The assembly graph Γ_k induced by k consists of the vertex set $V(\Gamma_k) = \{k(0), k(t_1), \ldots, k(t_n), k(1)\}$ and the edge set $E(\Gamma_k) = \{k(t_s, t_{s+1}) : s \in [0, n]$, and $t_0 = 0, t_{n+1} = 1\}$ corresponding to the intervals between rigid vertices of k. By this construction, an assembly graph Γ_k is directed, inheriting the edge direction from the orientation of k, but we use the convention that Γ_k is undirected unless stated otherwise.

If an assembly curve k intersects itself at most once at each rigid vertex, $k(0) \neq k(i)$ for $i \in (0, 1]$, and $k(1) \neq k(j)$ for $j \in [0, 1)$, then Γ_k will consist of 4-valent vertices corresponding to the rigid vertices and possibly 1-valent vertices corresponding to the initial and terminal ends of k.

Assembly graphs were introduced to model genome rearrangement and homological recombination processes, whereby DNA segments are modeled by certain types of paths in the assembly graphs called *polygonal paths* [27]. The minimum number of polygonal paths visiting all vertices in a graph is called an assembly number for the graph and discussed in Section 3.4.3. The biological model is considered in great detail in Section 3.5.1, but we also refer the reader to [27, 28].

3.4.1 Preliminaries

Definitions and notations of assembly graphs and related concepts are listed below.

Definitions of assembly graphs and related concepts.

A graph consists of a set of vertices V and a set of edges E. The endpoints of every edge is either a pair of vertices or a single vertex. In the latter case, the edge is called a *loop*. Multiple edges are also allowed. A *degree* of a vertex v is the number of edges incident to v such that each loop is counted twice. A path in a graph is an alternating sequence of vertices and edges (starting and ending with a vertex) such that consecutive vertices are incident to the edge between them. A single vertex is a path called *singleton*.

A *4-valent rigid vertex* is a vertex of degree 4 for which a cyclic order of edges is specified. For a 4-valent rigid vertex v, if its incident edges appear in order e_1, e_2, e_3, e_4, we say that e_2 and e_4 are *neighbors with respect to* v to e_1 (or e_3). Vice versa, e_1 and e_3 are neighbors to e_2 (or e_4). We note that we allow two of the edges incident to a vertex v to be equal. This does not change the degree of the vertex v, nor the definition of a neighbor. For example, if edges e_1 and e_2 are equal, then this edge is a neighbor to itself, and it is a loop. Similarly if e_1 and e_3 are equal, then this edge is not a neighbor to itself.

An *assembly graph* Γ is a finite connected graph where all vertices are rigid vertices of valency 1 or 4. A vertex of valency 1 is called an *endpoint*. Note that the definition of assembly graph implies that the number of endpoints is always even. The number of 4-valent vertices in Γ is called the *size* of Γ and is denoted by $|\Gamma|$. The assembly graph is called *trivial* if $|\Gamma| = 0$. Two assembly graphs are *isomorphic* if they are isomorphic as graphs and the graph isomorphism preserves the cyclic order of the edges incident to a vertex.

Two types of paths are of interest: (A) paths in which consecutive edges are never neighbors with respect to their common incident vertex and (B) paths in which every pair of consecutive edges are neighbors with respect to their common incident vertex. A path of type (A) where no edge is repeated is called a *transverse* path, or simply a *transversal*. A path of type (B) where no vertex is repeated is called a *polygonal* path. Graphs that have an Eulerian transversal, a transverse path that visits each edge exactly once, are called *simple assembly graphs*. We note that in a simple assembly graph, if a vertex v is an endpoint of a loop e, then e must be a neighbor of itself.

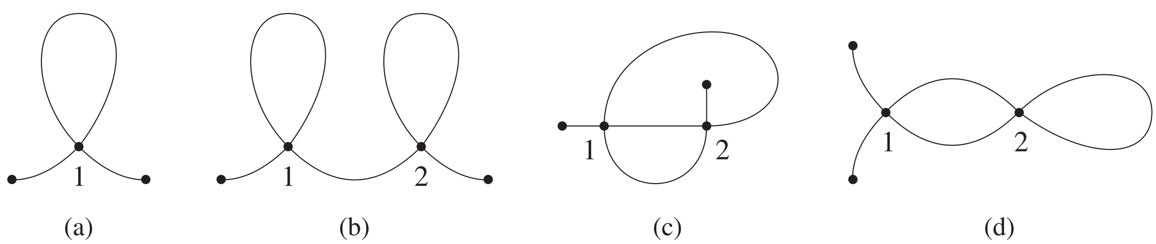


Figure 8.: List of assembly graphs of sizes 1 and 2.

Both transversals and polygonal paths model specific parts of the DNA rearrangement phenomenon. In particular, the transversal represents the micronuclear DNA segment prior to the recombination, the rigid vertices indicate the recombination sites, and polygonal paths are DNA segments after the rearrangements. We are interested to see what types of rearrangements and how many distinct genes can be encoded in a single micronuclear DNA sequence. Therefore, we are interested in graphs with Eulerian transversals, i.e., simple assembly graphs.

Convention: In the rest of the exposition, unless otherwise stated, all graphs are simple assembly graphs with two endpoints. In all planar diagrams representing assembly graphs, the cyclic order of the edges incident to a given vertex is specified by the order depicted within a neighborhood of the vertex in the diagram.

Figure 9 depicts two examples of assembly graphs. Eulerian transverse paths for (a) and (b) are respec-

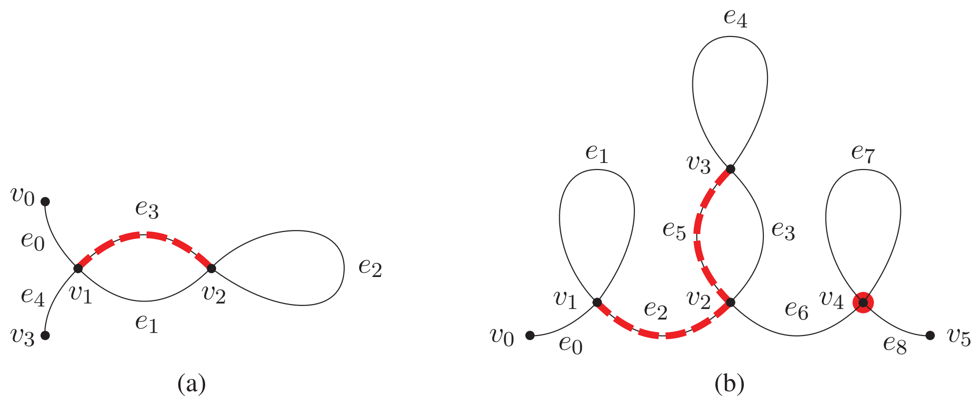


Figure 9.: Simple assembly graphs and polygonal paths in dotted lines. In (b) there are two polygonal paths, one of which is a singleton containing the vertex v_4.

tively given by, given by

$$(v_0, e_0, v_1, e_1, v_2, e_2, v_2, e_3, v_1, e_4, v_3), \quad \text{and}$$

$$(v_0, e_0, v_1, e_1, v_1, e_2, v_2, e_3, v_3, e_4, v_3, e_5, v_2, e_6, v_4, e_7, v_4, e_8, v_5).$$

A path may be given as a sequence of edges only (omitting the vertices) when the sequence of edges uniquely determines the vertices of the path. With this convention the transverse paths for the graphs in (a) and (b) in Figure 9 are $(e_0, e_1, \ldots, e_4)$ and $(e_0, e_1, \ldots, e_8)$, respectively.

Two transverse paths with endpoints are *equivalent* if they are either identical, or one is the reverse of the other. In [27] it was shown that two simple assembly graphs with two endpoints are isomorphic if and only if their Eulerian transversals are equivalent.

Given a simple assembly graph Γ, designate one of the endpoints as initial (i) and the other endpoint as terminal (t). We call such Γ an *oriented* (or *directed*) simple assembly graph with direction from i to t. We consider the transverse path of a directed simple assembly graph as a path starting at the vertex i and terminating at the vertex t.

Let Γ be an oriented simple assembly graph with an initial vertex i and a terminal vertex t. Let the set of 4-valent vertices of Γ be $V = \{v_1, \ldots, v_n\}$ where $n = |\Gamma|$. Starting from i, write down the sequence of vertices in the order they are encountered along the transversal. This is an assembly word over alphabet V. Thus an oriented assembly graph gives rise to an assembly word, and it is known that equivalence classes of assembly words are in one-to-one correspondence with isomorphism classes of assembly graphs [27]. In particular if w is a double occurrence word, we write Γ_w for the simple assembly graph defined by w.

The relationship between assembly words and assembly graphs is similar to unsigned Gauss codes and knot diagrams, with the exceptions that assembly graphs need be neither closed nor planar. Rather the biological motivation of the model implies that the rigid vertices and edges of an assembly graph are embedded in $\mathbb{R}^3$.

Table 3: Non-isomorphic diagrams in (3.1) are obtained by summing all words with the palindromes of each class and halving the total.

Symbols	All	Irreducible	Strongly Irreducible
1	1	1	1
2	3	2	1
3	11	8	3
4	65	47	17
5	513	389	135
6	5,363	4,226	1,463
7	68,219	55,804	19,306
OEIS	A132101 $\left(\frac{K_n+L_n}{2}\right)$	A199136 $\left(\frac{I_n+J_n}{2}\right)$	A199138 $\left(\frac{S_n+T_n}{2}\right)$

3.4.2 Hamiltonian Sets of Polygonal Paths

Two paths are *disjoint* if they do not have a vertex in common. We are interested in disjoint polygonal paths that visit every vertex in an assembly graph. A pairwise disjoint set $\{\gamma_1, \ldots, \gamma_k\}$ of polygonal paths in Γ is called *Hamiltonian* if their union contains all 4-valent vertices of Γ. In particular the set of vertices $V(\Gamma)$ is a Hamiltonian set of singletons. A polygonal path γ with no repeating vertices is called *Hamiltonian* if the set $\{\gamma\}$ is Hamiltonian.

Let Γ be a non-trivial assembly graph. The *assembly number* of Γ that is denoted by $\mathrm{An}(\Gamma)$, is defined by $\mathrm{An}(\Gamma) = \min\{\, k \mid$ there exists a Hamiltonian set of polygonal paths $\{\gamma_1, \ldots, \gamma_k\}$ in $\Gamma\}$. Graphs Γ with $\mathrm{An}(\Gamma) = 1$ are called *realizable*. Otherwise they are called *unrealizable*. These names reflect whether a given assembly graph corresponds to a single scrambled gene or not. In particular the assembly number of a graph gives the minimal number of genes that can be encoded by a corresponding DNA sequence.

A Hamiltonian polygonal path for the assembly graph depicted in Figure 9 (a) is (v_1, e_3, v_2), depicted by a thick dotted curve. Note that for the graph Γ depicted in (b), there is no Hamiltonian polygonal path. Therefore, Γ is unrealizable with $\mathrm{An}(\Gamma) = 2$.

Recall that a polygonal path in an assembly graph is a path in which every two consecutive edges are neighbors with respect to their common vertex. This can be seen as if the path makes a "90-degree" turn at every rigid vertex. A set of paths $\gamma = \{\gamma_1, \ldots, \gamma_s\}$ is called Hamiltonian if every vertex is visited by exactly one γ_i for some $i = 1, \ldots, s$.

In this section, we investigate the number (of sets) of Hamiltonian polygonal paths in assembly graphs. First we consider the number of all sets of such paths.

THEOREM 3.5. If Γ is a simple assembly graph with $|\Gamma| = n$ and $\mathcal{C}$ is the collection of all sets of Hamiltonian polygonal paths of Γ, then

$$|\mathcal{C}| \leq F_{2n+1} - 1, \tag{3.7}$$

where F_n is the n^{th} Fibonacci number.

Proof. Orient Γ and enumerate the edges between the 4-valent vertices successively along the transversal as $\{1, \ldots, 2n-1\}$ (the edges incident to the endpoints are not enumerated). For a set $\gamma = \{\gamma_1, \ldots, \gamma_s\}$ of Hamiltonian polygonal paths, let $||\gamma||$ be the number of edges visited by paths in γ. Vertices not incident to edges in γ are considered to be singleton paths. Let $\mathcal{C}_k$ be the collection of all γ such that $||\gamma|| = k$. Clearly, $||\gamma|| \leq n-1$ for all γ and $|\mathcal{C}| = |\mathcal{C}_{n-1}| + \cdots + |\mathcal{C}_1| + |\mathcal{C}_0|$.

To obtain $|\mathcal{C}_k|$, note that each $\gamma \in \mathcal{C}_k$ corresponds to a unique subset $S_\gamma \in \binom{[2n-1]}{k}$ formed by the edges belonging to γ. Since γ is a set of polygonal paths, S_γ contains no consecutive integers, and we can count the elements in $\mathcal{C}_k$ by the classic *stars and bars* argument.

Take any S_γ derived from $\gamma \in \mathcal{C}_k$ and represent the elements of S_γ by stars "$\star$" and the elements of the complement of S_γ within set $[2n-1]$ with $2n-1-k$ vertical bars "$|$". Note that there is at least one bar between any two stars. Every position before, between, and after the vertical bars represents a possible slot for k different stars "$\star$" for a total of $2n-k$ positions. Therefore, there are a total of $\binom{2n-k}{k}$ arrangements of stars and bars such that no arrangement contains two consecutive stars. Thus, for each k such that $0 \leq k \leq n-1$, $|\mathcal{C}_k| \leq \binom{2n-k}{k}$ and

$$|\mathcal{C}| = \sum_{k=0}^{n-1} |\mathcal{C}_k| \leq \sum_{k=0}^{n-1} \binom{2n-k}{k} = F_{2n+1} - 1.$$

The last equality is straightforward by induction (see also [46]). $\square$

DEFINITION 3.4.1. The *tangled cord* $\mathcal{T}_n$ of size n, for a positive integer n, is an assembly graph with assembly word

$$1213243\cdots(n-1)(n-2)n(n-1)n. \tag{3.8}$$

Specifically, $\mathcal{T}_1 = 11$, $\mathcal{T}_2 = 1212$, $\mathcal{T}_3 = 121323$, and $\mathcal{T}_n$ is obtained from $\mathcal{T}_{n-1}$ by replacing the last letter $(n-1)$ by the subword $n(n-1)n$. Figure 10 shows the structure of the tangled cord.

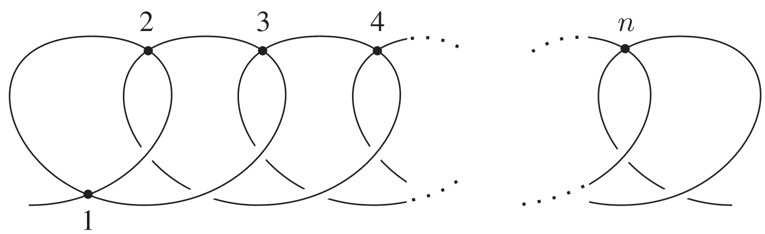


Figure 10.: Assembly graph corresponding to the tangled cord $\mathcal{T}_n$.

THEOREM 3.6. The tangled cord $\mathcal{T}_n$ has $\binom{n+1}{2}$ distinct Hamiltonian polygonal paths.

Proof. Observe that $\binom{n+1}{2} = \binom{2n-n+1}{n-1}$ which is precisely the number of ways one can choose $n-1$ non-consecutive numbers from $1, 2, \ldots, 2n-1$ (see proof of Theorem 3.5). We prove the theorem by induction on n.

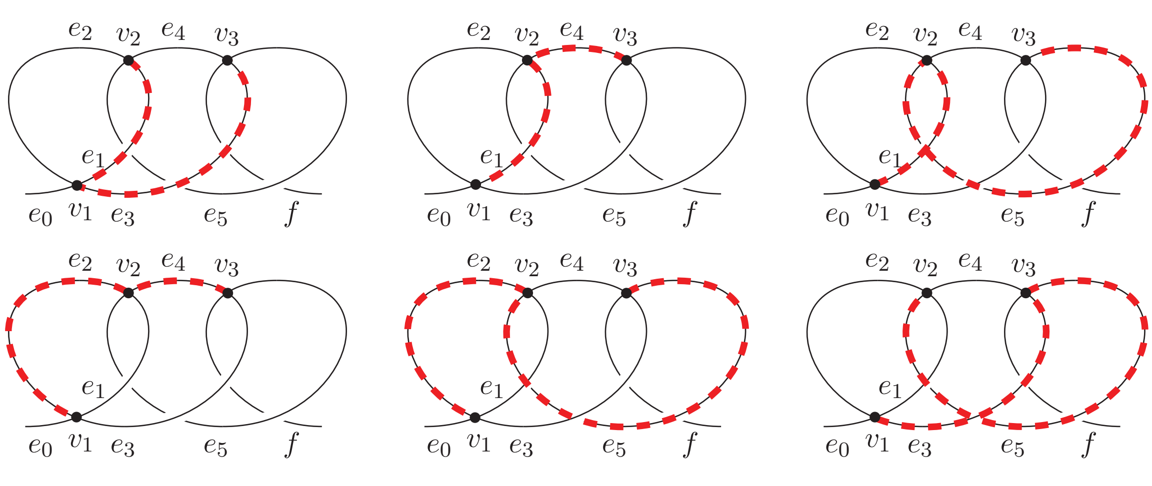


Figure 11.: Six Hamiltonian polygonal paths for $\mathcal{T}_3$.

Trivially, the singleton vertex of $\mathcal{T}_1$ constitutes a Hamiltonian polygonal path. Further, all three edges between the two 4-valent rigid vertices of $\mathcal{T}_2$ constitute Hamiltonian polygonal paths (see Figure 8(c)). For

$n = 3$, $\mathcal{T}_3$ contains 7 edges labeled $e_0, e_1, \ldots, e_5, f$. Out of these, the following six form polygonal paths: e_1e_3, e_1e_4, e_1e_5, e_2e_4, e_2e_5 and e_3e_5. These paths are indicated in Figure 11. Hence, the claim holds for $n = 1, 2, 3$. Let $n \geq 4$ and consider $\mathcal{T}_{n-1}$.

Label the 4-valent vertices of $\mathcal{T}_{n-1}$ and edges between them sequentially along a transversal as $v_1, v_2, \ldots, v_{n-1}$ and $e_0, e_1, \ldots, e_{2n-3}, f$ respectively (see Figure 12(a)). In this ordering, the edges e_0 and f are incident to the endpoints. Suppose that every choice of $n-2$ non-consecutive edges from $e_1, e_2, \ldots, e_{2n-3}$ forms a Hamiltonian polygonal path. The number of such paths is $\binom{n}{2}$.

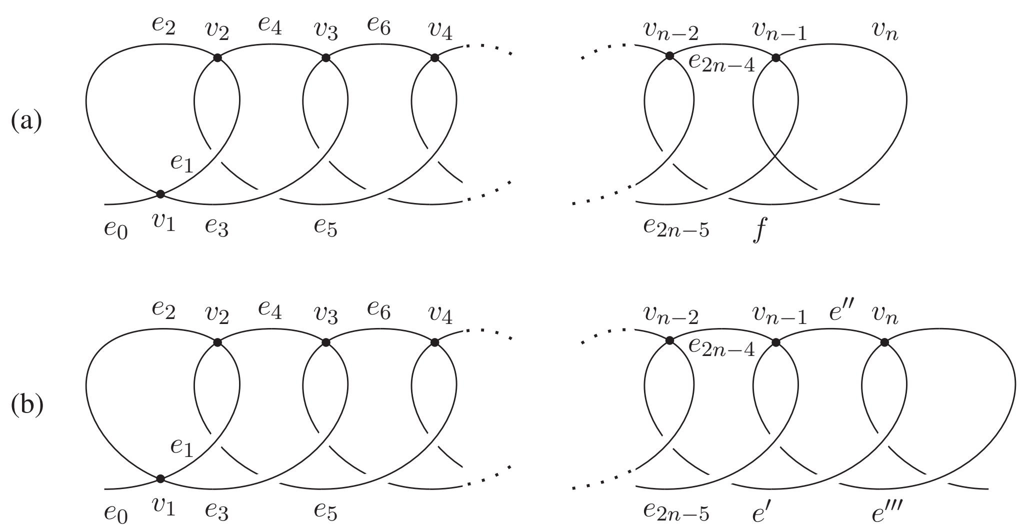


Figure 12.: Addition of a new vertex to $\mathcal{T}_{n-1}$ to obtain $\mathcal{T}_n$.

Observe that $\mathcal{T}_n$ is formed from $\mathcal{T}_{n-1}$ by extending f and intersecting it with edge e_{2n-3}. Such a construction makes the following changes to $\mathcal{T}_{n-1}$ (see Figure 12(b)):

i. The intersection of f with e_{2n-3} becomes a new vertex which we call v_n,

ii. The edge e_{2n-3} is split into two edges by v_n which we call e' and e'' in the direction of the original transversal, and

iii. Another edge is created between v_{n-1} and v_n following e'' which we call e'''.

By the induction hypothesis, every choice of $n-2$ non-consecutive edges in $\mathcal{T}_{n-1}$ forms a polygonal path and there are $\binom{n}{2}$ distinct Hamiltonian polygonal paths of $\mathcal{T}_{n-1}$. Let γ be a Hamiltonian polygonal path in $\mathcal{T}_{n-1}$. We have the following possibilities for reaching v_{n-1} with a Hamiltonian polygonal path in $\mathcal{T}_n$.

Case 1. Suppose $e_{2n-3} \in \gamma$. Let g_1 be the number of such paths in $\mathcal{T}_{n-1}$. Note that v_{n-1} and v_{n-2} are ends of e_{2n-3}, and the neighbors of e_{2n-3} are e_{2n-4} and e_{2n-5}. The path $e'e'''$ also ends at v_{n-1} and v_{n-2}, and the neighboring edges are e_{2n-4} and e_{2n-5}. Therefore, γ may be extended to a Hamiltonian polygonal path $\gamma' = (\gamma \setminus e_{2n-3}) \cup \{e', e'''\}$ of $\mathcal{T}_n$. Hence with this extension, the number of paths in $\mathcal{T}_n$ that contain both e' and e''' is the same as the number of paths in $\mathcal{T}_{n-1}$ that contain e_{2n-3}. We have g_1 such paths in $\mathcal{T}_n$.

Case 2. Now suppose $e_{2n-3} \notin \gamma$, but $e_{2n-5} \in \gamma$. Let g_2 be the number of such paths in $\mathcal{T}_{n-1}$. Then v_{n-1} is an end of γ. Since e_{2n-5} is in γ, the vertex v_{n-2} must be visited by e_{2n-7}, so v_{n-2} must be an end of γ as well. Hence, γ can extend to a Hamiltonian polygonal path in $\mathcal{T}_n$ in three possible ways: $\gamma_1 = \gamma \cup \{e''\}$, or $\gamma_2 = \gamma \cup \{e'''\}$ or $\gamma_3 = \gamma \cup \{e'\}$. Thus there are $3g_2$ such paths in $\mathcal{T}_n$. Note that there is only one path in $\mathcal{T}_{n-1}$ that contains e_{2n-5} but not e_{2n-3} (the path with $n-2$ edges: $e_1 e_3 \cdots e_{2n-5}$). Hence $g_2 = 1$.

Case 3. Suppose $e_{2n-4} \in \gamma$. Let g_3 be the number of such paths in $\mathcal{T}_{n-1}$. Since γ is polygonal and Hamiltonian, it visits every vertex in $\mathcal{T}_{n-1}$ and neither e_{2n-5} nor e_{2n-3} are in γ. In this case γ ends at vertex v_{n-1}. Hence, γ may be extended to one of two different Hamiltonian polygonal paths $\gamma_1' = \gamma \cup \{e''\}$ or $\gamma_2' = \gamma \cup \{e'''\}$ in $\mathcal{T}_n$. Thus there are $2g_3$ such paths in $\mathcal{T}_n$. Now γ has $n-2$ edges, and if the last edge is fixed to be e_{2n-4} then the rest of $n-3$ non-consecutive edges are chosen from $e_1, e_2, \ldots, e_{2n-6}$. Note that there are $\binom{(2n-6)-(n-3)+1}{n-3} = \binom{n-2}{n-3} = n-2$ such paths, i.e., $g_3 = n-2$.

We observe that all Hamiltonian polygonal paths in $\mathcal{T}_n$ are obtained as described in the three cases above. Let β be a Hamiltonian polygonal path of $\mathcal{T}_n$. Then β contains $(n-1)$ edges. Since β is Hamiltonian, for the edges e', e'', e''' we have four mutually exclusive cases: (*a*) only $e' \in \beta$, (*b*) only $e'' \in \beta$, (*c*) only $e''' \in \beta$, or (*d*) $e', e''' \in \beta$, but $e'' \notin \beta$. In the first three cases (*a*)–(*c*), $(n-2)$ edges of β correspond to non-consecutive edges in $\mathcal{T}_{n-1}$, and these cases correspond to extensions of Hamiltonian polygonal paths in $\mathcal{T}_{n-1}$. Hence, those are the cases when the polygonal paths in $\mathcal{T}_{n-1}$ end at vertices v_{n-1} or v_{n-2}. Such paths are described in Case 2 and Case 3. In the case of (*d*), we observe that for every path in $\mathcal{T}_n$ that contains e' and e''' the neighboring edge of e''', the edge e_{2n-4} cannot be part of β. This means that the rest of $n-3$ edges of β are non-consecutive edges in $\mathcal{T}_{n-1}$ chosen from $e_1, \ldots, e_{2n-5}$. Hence we can obtain β from a Hamiltonian polygonal path γ in $\mathcal{T}_{n-1}$ containing e_{2n-3} as described in Case 1.

Now we observe that there are exactly $\binom{n+1}{2}$ Hamiltonian paths in $\mathcal{T}_n$. From Case 1 we have g_1 paths in $\mathcal{T}_n$, from Case 2 there are $3g_2$ paths in $\mathcal{T}_n$ and from Case 3 there are $2g_3$ paths in $\mathcal{T}_n$. Considering that $g_1 + g_2 + g_3$ is the number of Hamiltonian polygonal paths in $\mathcal{T}_{n-1}$ and the fact that $g_2 = 1$ and $g_3 = n-2$

we have total of

$$g_1 + 3g_2 + 2g_3 = \binom{n}{2} + 2g_2 + g_3 = \binom{n}{2} + 2 + n - 2 = \binom{n+1}{2}$$

Hamiltonian polygonal paths of $\mathcal{T}_n$. □

COROLLARY 3.1. The upper bound in Theorem 3.5 is achieved for every positive integer n, and it is achieved by a tangled cord $\mathcal{T}_n$.

Proof. A set K of non-consecutive edges in an assembly graph is not a subset of a Hamiltonian set of polygonal paths only if the edges in K form a cycle. Label the edges of $\mathcal{T}_n$ sequentially along a transversal as $e_0, e_1, \ldots, e_{2n-1}, f$ respectively. We observe that the only cycle that can be formed by non-consecutive edges in $\mathcal{T}_n$ is the cycle that contains all odd numbered edges (there are n of them). Hence any set of k non-consecutive edges in $\mathcal{T}_n$ where $k < n$ does not form a cycle. This implies that every set of k non-consecutive edges corresponds to a Hamiltonian set of polygonal paths (see proof of Theorem 3.5). The vertices not visited by these k edges form paths that are singletons. Thus the inequality in Theorem 3.5 is tight. □

CONJECTURE 3.1. The upper bound in Theorem 3.5 is achieved for every positive integer n *only* by the tangled cord $\mathcal{T}_n$.

3.4.3 Assembly Numbers

We recall that the assembly number of an assembly graph is the minimum number of disjoint polygonal paths required to cover each rigid vertex of the assembly graph. The proof of Theorem 3.5 gives a means of constructing potential Hamiltonian sets of polygonal paths from sets of non-adjacent edges. Hence the assembly number may be calculated through a brute force search to determine if the potential edge sets form disjoint connected paths in the assembly graph.

Table 4 shows the assembly numbers for assembly graph isomorphism classes corresponding to graphs with a small number of rigid vertices. These numbers are obtained by computer calculations. In particular, we notice that the assembly numbers 2 and 3 appear for the first time for graphs with 4 and 7 vertices, respectively. In the case of assembly number 2 there is only one realization graph with 4 vertices out of the total 65 isomorphism classes, and in the case of assembly number 3, there are 3 realization graphs (with 7

Table 4: Assembly numbers.

Assembly number / # of rigid vertices	2	3	4	5	6	7	OEIS
1	3	11	64	504	5241	66515	A199135
2	0	0	1	9	122	1701	A199139
3	0	0	0	0	0	3	-

vertices) among over 67,000 isomorphism classes. The last column in Table 4 indicates the newly inserted sequences in the OEIS.

The assembly number would increase if the polygonal paths were "forced" to visit certain edges. This enforcement can be obtained by introducing loops on the edges, therefore introducing the necessity for the polygonal paths to visit vertices of the loops. Hence, here we study the assembly number of graphs obtained by adding loops on edges, and investigate possible lengths of polygonal paths for such graphs.

We recall that the assembly number of an assembly graph is the minimum number of scrambled genes that can be represented by the graph. In particular, we say that an assembly graph is realizable if it can encode a single gene, in other words, if the graph has a Hamiltonian polygonal path.

3.5 Micronuclear Arrangements

Several genera of ciliates, such as *Oxytricha* and *Stylonychia*, undergo massive genome rearrangement during sexual reproduction. These massively occurring recombination processes make them ideal model organisms to study gene rearrangements. See [35] and references therein for further details of the following biological description.

Ciliates contain two types of nuclei: a somatic macronucleus which controls the bulk of the cell functions and a germline micronucleus which gives rise to a new macronucleus during sexual reproduction. During sexual reproduction, up to 95% of the micronuclear genetic sequence can be excised and the remaining fragments reorganize to form a new macronucleus. Sequences which are retained during this recombination process are called *macronuclear destined sequences* or *MDSs*, and those that are excised are called *internal eliminated sequences* or *IESs*. The micronuclear fragments unscramble to form macronuclear genes by fusing the matching segments at the ends of the MDSs called *pointers* (see Figures 14 and 16).

In relation to the unscrambled macronuclear gene, a scrambled micronuclear gene may have permuted or inverted MDS segments separated by IESs. Formation of the macronuclear genes in these ciliates thus requires any combination of the following three events: unscrambling of segment order, DNA inversion, and IES removal.

Since the micronuclear IESs are not retained in the macronuclei, the recombination event may be characterized solely through the order and orientation of the MDSs in the scrambled micronuclear sequence relative to the unscrambled macronuclear sequence. A *micronuclear arrangement* (cf. [35]) is a sequence of permuted and inverted MDSs. In particular, if a micronuclear sequence contains d scrambled macronuclear genes $g_1, g_2, \ldots, g_d$ having $n_1, n_2, \ldots, n_d$ MDSs per macronuclear gene respectively, then the micronuclear arrangement α will contain $n = \sum_{i=1}^{d} n_i$ MDSs and may be presented as

$$\alpha(\sigma, \tau, \epsilon) = M_{\tau(1),\sigma(1)}^{\epsilon(1)} M_{\tau(2),\sigma(2)}^{\epsilon(2)} \ldots M_{\tau(n),\sigma(n)}^{\epsilon(n)}$$

where σ_i indicates the order that the i^{th} MDS appears in the unscrambled macronuclear gene τ_i with orientation ϵ_i relative to macronuclear gene. Said another way, $\tau : [n] \to [d]$ assigns each MDS to a macronuclear gene, $\epsilon : [n] \to \{-1, 1\}$ is a signing function, and $\sigma_i = \sigma(i, \tau(i))$ such that if $\tau(i) = t$ then $\sigma_i : [n_t] \to [n_t]$ is a permutation.

EXAMPLE 3.5.1. The micronuclear arrangement

$$\alpha = M_{3,3}^{+1} M_{1,2}^{-1} M_{2,1}^{+1} M_{3,1}^{+1} M_{1,3}^{-1} M_{3,2}^{+1} M_{1,1}^{+1} M_{2,2}^{+1}$$

characterizes the genetic recombination taking place in Figure 13.

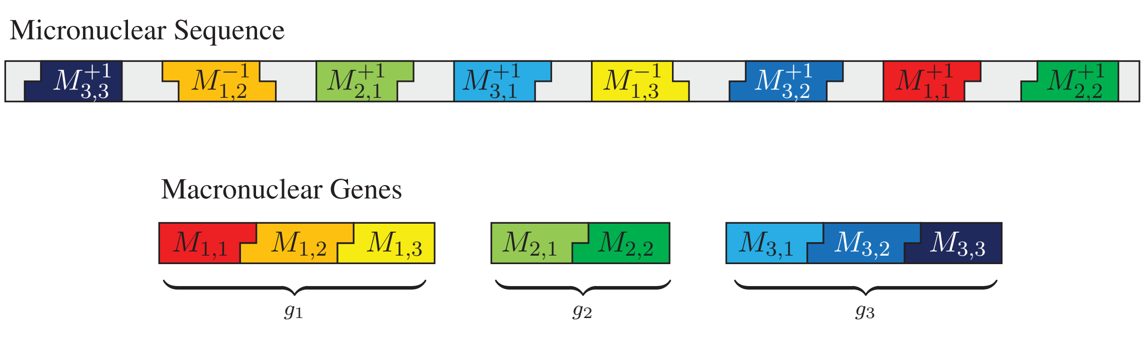


Figure 13.: A scrambled micronuclear sequence and the corresponding unscrambled macronuclear sequences. MDSs (colored segments) are separated by IES (gray segments) in the micronucleus. Macronuclear genes are formed when micronuclear MDS pointers align and fuse.

When it is contextually clear that the micronuclear sequence contains only one macronuclear gene, we omit τ. In this case σ becomes a permutation of $[n]$. Further, it is conventional to notate the orientation of the MDS using a bar (or no bar) rather than a sign of -1 (or 1), e.g., $\overline{M}_2$ is the same as M_2^{-1} and $M_{2,1}^{-1}$ and indicates that the MDS M_2 is inverted in the scrambled gene relative to the single unscrambled gene. For example, the micronuclear arrangement of the Actin I gene in Figure 14 is

$$M_{1,3}^{+1}M_{1,4}^{+1}M_{1,6}^{+1}M_{1,5}^{+1}M_{1,7}^{+1}M_{1,9}^{+1}M_{1,2}^{-1}M_{1,1}^{+1}M_{1,8}^{+1}$$

and conventionally denoted as $M_3M_4M_6M_5M_7M_9\overline{M}_2M_1M_8$.

Micronuclear arrangements give a means of recording which MDSs are effected during the unscrambling process, but do not provide an explanation for how the recombination actually occurs. There are several theoretical models attempting to describe these DNA recombination processes [35, 36, 39, 44]. It has been conjectured that an additional molecule (called a template) takes part in the recombination process [28, 44] and experimental support for this model was obtained in [43].

3.5.1 Connection with Assembly Graphs

A connection between micronuclear arrangements and the assembly graph model was established in [28]. However to further emphasize the topological nature of the genetic sequence, we adopt the following convention. For a genetic alphabet $\Sigma = \{G, A, T, C\}$, let $w = a_1a_2 \dots a_n \in \Sigma^+$ be a finite genetic sequence, and consider the coloring χ such that $\chi\left(\left[\frac{i-1}{n}, \frac{i}{n}\right)\right) = a_i$ for each $i \in [n]$ and $\chi(1) = n$. Recall that an assembly curve is a smooth map $k : [0, 1] \to \mathbb{R}^3$ with no singularities except a finite number of transversal self-intersections. The assembly graph model associates a micronuclear sequence with an assembly curve, and represents the recombination process where the MDS pointers align as rigid vertices.

DEFINITION 3.5.1. Let w be a micronuclear sequence with well-defined MDS, IES, and pointer sequences. For each pointer sequence $p = a_i \dots a_{i+j} \subseteq w$, the *center of the pointer* is taken as

$$c(p) = \frac{2i + j - 1}{2n} \in [0, 1] .$$

We can now construct an assembly curve k from a micronuclear sequence by identifying the pointers which overlap in the unscrambled macronuclear gene sequences, i.e., $k(t_1) = k(t_2)$ for some $t_1, t_2 \in [0, 1]$ if and only if t_1 and t_2 are the centers of pointers which overlap in the macronuclear gene sequences (or $t_1 = 0$ and $t_2 = 1$ when k is a circular assembly curve). As a consequence, the macronuclear genes now

Figure 14.: Scrambled Actin I micronuclear gene in *Oxytricha nova* comprised of MDSs (colored segments), IESs (gray segments), and pointers (black segments) [45].

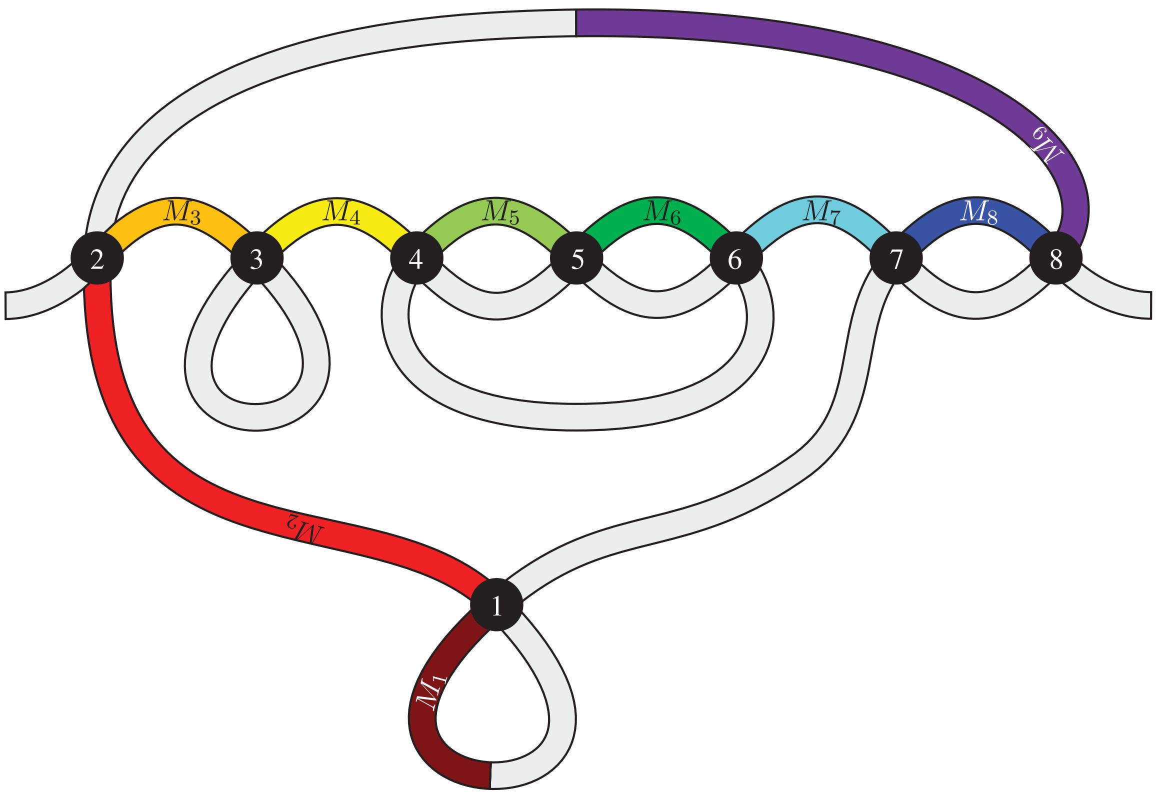


Figure 15.: Assembly graph of the *Oxytricha nova* Actin I gene [45] representing intermediary steps in the unscrambling of the micronuclear gene in Figure 14 into the macronuclear gene in Figure 16.

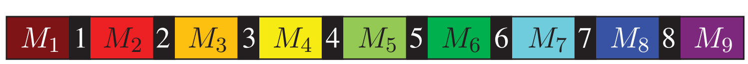


Figure 16.: Unscrambled Actin I macronuclear gene in *Oxytricha nova* [45]. All MDSs are sequenced and oriented correctly, the MDS pointers overlap, and the IESs have been removed.

correspond to a set of intervals $[i_1, i_2) \cup [i_3, i_4) \cup \cdots \cup [i_{j-1}, i_j)$ which are now connected in the image of k. Since there is always an MDS between two IESs in the micronuclear arrangement, the MDSs will be connected as polygonal paths in the assembly curve. Thus, the full set of macronuclear genes appear as a Hamiltonian set of polygonal paths in the assembly graph induced by k. See Figures 14, 15, and 16.

Note that neither the length of the micronuclear sequence nor the colorings of the nucleotide intervals were necessary to generate the induced assembly graph. Indeed an assembly graph can be generated directly from a micronuclear arrangement, and the transversal of the assembly graph corresponds to a double occurrence word called an *assembly word*. Details of this procedure are provided in the definition of the MicArrToArrWord function, $\varrho(\alpha)$. For the micronuclear arrangement in Example 3.5.1,

$$\begin{aligned}\varrho(\alpha) &= \varrho(M_{3,3}^{+1})\,\varrho(M_{1,2}^{-1})\,\varrho(M_{2,1}^{+1})\,\varrho(M_{3,1}^{+1})\,\varrho(M_{1,3}^{-1})\,\varrho(M_{3,2}^{+1})\,\varrho(M_{1,1}^{+1})\,\varrho(M_{2,2}^{+1}) \\ &= \;(\,[5]\,)\;\;(\,[2]\;[1]\,)\;\;(\,[3]\,)\;\;\;(\,[4]\,)\;\;\;(\,[2]\,)\;\;(\,[4]\;[5]\,)\;\;(\,[1]\,)\;\;\;(\,[3]\,)\end{aligned}$$

which corresponds to the double occurrence word 1234525134 in ascending order.

The function MicArrToArrWord establishes that each micronuclear arrangement corresponds to exactly one assembly word, however the converse is not true as many micronuclear arrangements will correspond with the same assembly word. For example, $\rho(M_1M_2M_3M_4)$ and $\rho(M_4\overline{M}_3\overline{M}_2M_1)$ both correspond to the word 112233 in ascending order. This leads to the question: What are the possible micronuclear arrangements which can induce a particular assembly graph?

Observe the following features of the assembly graph construction:

- each micronuclear arrangement has an associated assembly curve k;
- each assembly curve k induces a directed assembly graph Γ_k;
- the order and orientations of the MDSs in the micronuclear arrangement correspond with directed edges of a Hamiltonian set of polygonal paths;
- the first and last MDS of each macronuclear gene correspond to partial edges in Γ_k.

Theorem 3.5 provides a means to enumerate all potential Hamiltonian sets of unoriented polygonal paths, but does not include the placement of partial edges at the path ends. Since the paths in the Hamiltonian set are disjoint, any path which ends at a rigid vertex may be extended to include a partial edge from any of the remaining three adjacent edges. However, only two of these placements extend as polygonal paths.

Function MicArrToArrWord(α) : $\varrho(\alpha)$

Input: A micronuclear arrangement $\alpha = M_{\tau(1),\sigma(1)}^{\epsilon(1)} M_{\tau(2),\sigma(2)}^{\epsilon(2)} \dots M_{\tau(n),\sigma(n)}^{\epsilon(n)}$ with scrambled macronuclear genes of sizes $n_1, n_2, \dots, n_d$.

Output: A double occurrence arrangement word W.

$W \leftarrow \emptyset$;

for $i \leftarrow 1$ **to** n **do**

  if $\sigma(i) = 1$ **then**

    $W := W\left(1 + \sum_{j=1}^{\tau(i)-1} (n_j - 1)\right)$; /* First MDS of MAC gene */

  else if $\sigma(i) = n_j$ **then**

    $W := W\left(\sum_{j=1}^{\tau(i)} (n_j - 1)\right)$; /* Last MDS of MAC gene */

  else

    /* Interior MDS of MAC gene */

    $W := W\left(\sigma(i) - \frac{1}{2}(1+\epsilon(i)) + \sum_{j=1}^{\tau(i)-1} (n_j - 1)\right)\left(\sigma(i) - \frac{1}{2}(1-\epsilon(i)) + \sum_{j=1}^{\tau(i)-1} (n_j - 1)\right)$;

  end

end

return W ;

The orientation of a polygonal path within an oriented assembly graph determines the MDS order and orientation, i.e., the first edge of the path will correspond to M_1, the second to M_2, etc. If the orientations of the polygonal path and assembly graph coincide for the edge corresponding to the i^{th} MDS, then mark the edge as M_i. If the orientations disagree, then we mark the edge as $\overline{M}_i$. Recording the MDS sequence as the assembly graph is transversed recovers the associated micronuclear arrangement; see Figure 17. The proceeding discussion summarizes as the following theorem.

THEOREM 3.7. All micronuclear arrangements which correspond to a given assembly graph are generated from the combinations of Hamiltonian sets of polygonal paths, extensions of the paths to polygonal partial edges, and orientations of the extended polygonal paths.

Tables 7, 8, and 9 give all possible micronuclear arrangements for assembly graphs of size two.

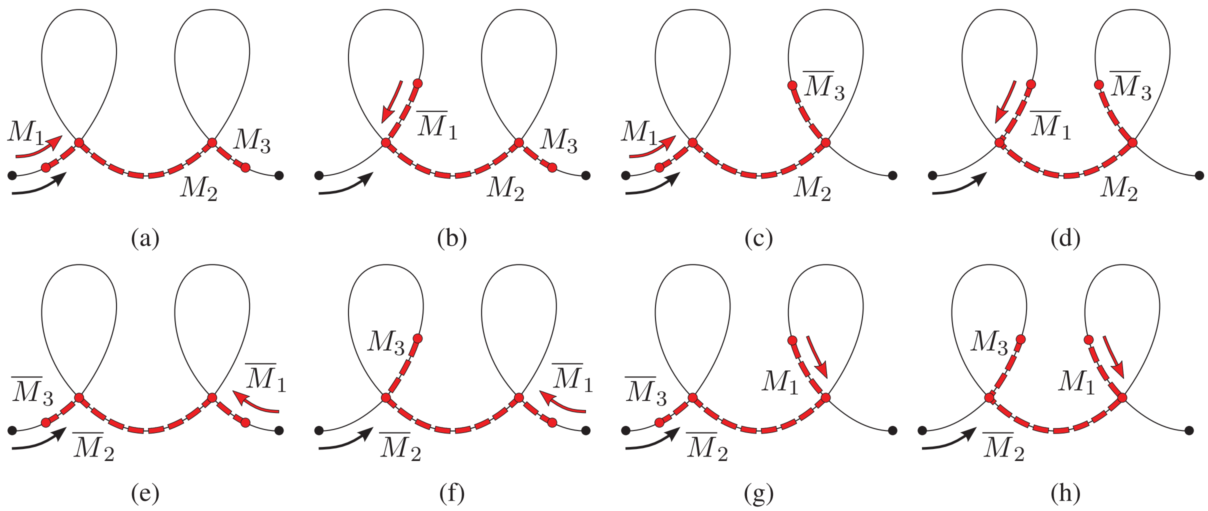


Figure 17.: Possible micronuclear arrangements corresponding to the assembly graph 1122 and the Hamiltonian polygonal path $\{v_1, e_1, v_2\}$.

3.6 Braidings of Chord Diagrams

Recall from the chapter introduction that a *chord diagram* (CD) consists of a finite number of *chords* intersecting a circle, called the *backbone*, at distinct points called the *chord endpoints*. Homologous genetic recombinations can be readily visualized via chord diagrams by representing the precursor gene and product gene as disjoint arcs on a backbone. Chords are drawn between the two arcs to represent matching genetic sequences. See Figure 18 for an example.

For this model, it is clear that non-scrambled genetic regions present as parallel chords and scrambled regions present as intersecting chords, relative to the arc orientations on the CD backbone. If the orientation of one of the backbone arcs is reversed, the chords which were previously intersecting become parallel (and vice versa). Hence, a good measure for the complexity of a genetic scrambling should either weight a section of parallel chords equally with a section of chords that all cross each other, or weight them as antipodal extremities.

In this section, we consider a property of chord diagrams called the braid index to quantify the degree of complication in the crossing structure of the chords in a CD. In 1998, Birman and Trapp [29] provided an algorithm to compute the braid index, and characterized properties of CDs with braid index 3. In Section 3.6.4, we improve the Birman-Trapp algorithm with a divide and conquer strategy, characterize and count the CDs with high braid index, and identify the types of CDs for which the braid index remains com-

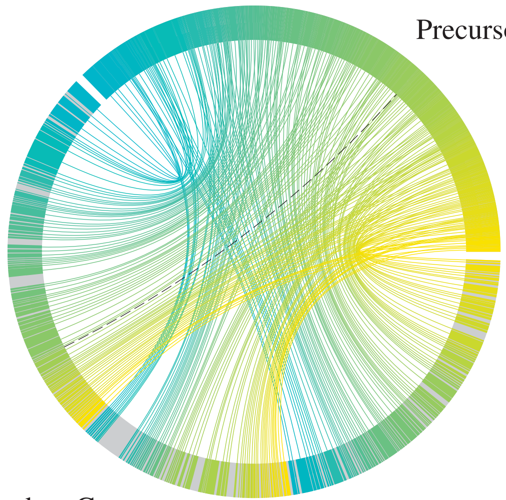


Figure 18.: Chord diagram modeling the genetic recombination which occurs in *Oxytricha trifallax* during sexual reproduction. Figure from [33].

putationally intractable. Finally in Section 3.6.5 we find the braid index for some special CDs, namely the tangled cord and the waffle graph.

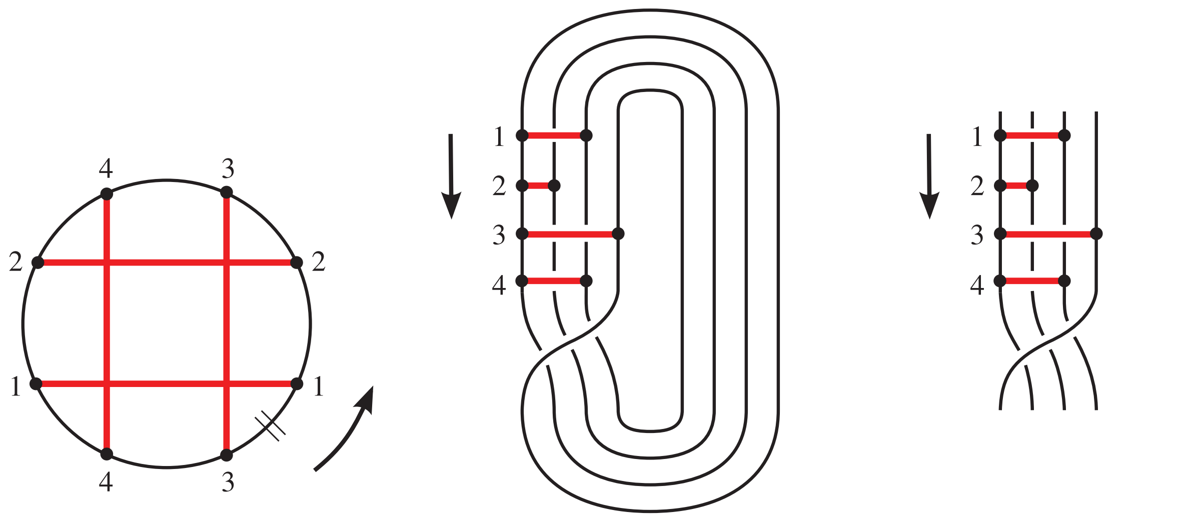
3.6.1 Preliminaries

All definitions in this section come from either [29] or [50]. Namely, the definitions for chord diagrams, linear chord diagrams, and isolated chords are adapted from [50], and the preliminaries to discuss braided chord diagrams are covered in [29].

A *linear chord diagram* (LCD) is an oriented chord digram such that the backbone is "cut" at a *base point* on the backbone, that is distinct from the chord endpoints. Cutting the backbone of the CD forms a linear backbone with an *inital* and *terminal backbone endpoint*, designated by the orientation of the CD. The *closure* of a LCD L, written $\overline{L}$, is the CD formed by identifying the initial and terminal backbone endpoints of L. The CD and LCD representing the DOW 121323 are pictured in Figure 7.

Both CDs and LCDs containing n chords are said to have *size* n. A CD (or LCD) is *degenerate* if it has an *isolated chord*, i.e., one not crossed by any other chord. A CD (or LCD) having only isolated chords is called *fully-degenerate*.

We represent a CD as a *braided chord diagram* by replacing the backbone by the standard m-braid repre-



(a) Chord diagram (b) Braided chord diagram (closed) (c) Braided chord diagram (open)

Figure 19.: The braided chord diagram (b) is generated by the word $A(1,3)A(1,2)A(1,4)A(1,3) \in C_4$ and represents a braiding of the chord diagram 12342143. Similar to figures in [29].

sentative $\sigma_{m-1}\sigma_{m-2}\ldots\sigma_1$ of the unknot, and express the chords of D as horizontal segments in the m-braid that connect two distinct strings Figure 19 (b). Such a braided chord diagram is called a *braid representative* of the CD D. Since the structure of the m-braid is implicit in the construction of the braided chord diagram, the "tops" and "bottoms" of the m-braid may be suppressed without losing any information forming an *open braided chord diagram*, as in Figure 19 (c). Likewise a closed braid can be recovered from an open braid by connecting the top and bottom of each string.

DEFINITION 3.6.1. The *chord monoid* $\mathcal{C}_m$ is defined as having *generators* $A(i,j)$, $i \leq j \leq m$ and relations

$$A(i,j)A(k,l) = A(k,l)A(i,j)\,, \tag{3.9}$$

for distinct i, j, k, l.

The geometric interpretation of a generator $A(i,j)$ is a chord joining the i^{th} and j^{th} strand of the identity m-braid, and multiplication in $\mathcal{C}_m$ is concatenation. Relation (3.9) implies that chords with ends on distinct strings can commute in the m-braid, in the sense of Figure 21 (a), while preserving the underlying CD association.

For $W \in \mathcal{C}_m$, the *closure of W* is the braided chord diagram $\mathbf{W}$ obtained by adjoining the *standard* m-braid representative of the unknot to W to form an open braided chord diagram which is then closed.

We can obtain a name (DOW) for D by labeling the generators in W in the order that they occur and *reading the name* for D from W by starting in the upper left-most strand and transversing the backbone of the braid.

The operation of replacing the word $W'A(i,j) \in \mathcal{C}_m$ with $A(i+1, j+1)W'$ (with the appropriate modification of $j+1 = 1$ when $j = m$) is called *cyclic permutation* and preserves the CD that a braid represents without effecting the number of strings in the m-braid.

DEFINITION 3.6.2. The *braid index* of a chord diagram D, written as $br(D)$, is the least number of strands used in any braid representative for D, i.e.,

$$br(D) = \min\{m \in \mathbb{N} \mid W \in C_m \text{ closes to the braid representative } \mathbf{W} \text{ of } D\}. \tag{3.10}$$

Similarly, the braid index of a linear chord diagram L is the braid index of the chord diagram D obtained by connecting the initial and terminal backbone endpoints of L.

3.6.2 Braid Index

LEMMA 3.5. If D' is a sub-chord diagram of D, then $br(D') \leq br(D)$.

Proof. Suppose not, i.e., there exists a minimal braiding $\mathbf{W}$ of D with fewer strings than the minimal braiding of D'. Since D' is a sub-chord diagram of D the chords of D which do not correspond to chords of D' can be removed from W to form a braiding of D' containing fewer strings than the minimal braiding of D'. Clearly this is a contradiction. □

LEMMA 3.6. If D_n is a sub-diagram of D_{n+1}, with n and $n+1$ chords respectively, then the braid index of D_{n+1} is either $br(D_n)$ or $br(D_n)+1$.

Proof. Let D_{n+1} be the chord diagram corresponding to the DOW $a_1 a_2 \ldots a_i \ldots a_j \ldots a_{2(n+1)}$ such that $a_i \in [n+1]$. Construct D_n from D_{n+1} by removing a chord (a_i, a_j), where a_i and a_j represent the same letter in the DOW, appearing at locations i and j in the name of D. From Lemma 3.5 we get that $br(D_n) \leq br(D_{n+1})$. Assume $br(D_n) < br(D_{n+1})$ and let $\mathbf{W}$ be a minimal braiding of D_n. If the chord (a_i, a_j) is added to $\mathbf{W}$ such that a_i intersects the braid immediately after a_{i-1}, and a_j intersects immediately before a_{j-1}, then the we must have that a_i and a_j belong to the same string. Otherwise $br(D_n) = br(D_{n+1})$, and we are done. Performing an increasing stabilization (adding a trivial loop) anywhere between a_i and a_j will increase the string count by one, but provides a valid braiding of D_{n+1} with $br(D_n)+1$ strings. □

COROLLARY 3.2 ([29]). If D_n is a CD with n chords, then $br(D_n) \leq n + 1$.

3.6.3 Birman-Trapp Algorithm

A *canonical braiding of a diagram* D_n, with n chords, is a $2n$-braiding in which each chord end is placed on a string by itself. Details of constructing the conical braiding of D_n are given by the function CanonicalBraid and illustrated by Figure 20. A *decreasing stabilization* of a braiding removes the trivial loops in the braid between two strings; see Figure 21. A precise definition is given by the function DecreasingStablilization.

The Birman-Trapp algorithm for computing the braid index is as follows [29]:

Step 1. Fix any chord of D_n, label it 1, and choose one endpoint of chord 1, labeling it $*$.

Step 2. For each labeling of the remaining chords:

i. Form the name resulting from beginning to read the labels at $*$.

ii. Construct the canonical braidings of D_n relative to that name.

iii. Perform all possible decreasing stabilizations to the canonical braiding, and record the braid index of the resulting braid.

This method amounts to a brute-force algorithm as there are $(n-1)!$ relabelings of the chords.

3.6.4 Divide and Conquer Algorithm

We show that if the CD contains clusters of isolated chords, then Birman-Trapp algorithm can be run on the individual clusters rather than the entire diagram.

LEMMA 3.7. If $W = A(i_1, j_1)A(i_2, j_2) \ldots A(i_n, j_n) \in \mathcal{C}_m$ represents the chord diagram D_n, then the chords that correspond to the generators $A(i_s, j_s)$ and $A(i_t, j_t)$ with $s < t$ intersect in D_n if and only if $i_s \leq i_t < j_s \leq j_t$ or $i_t < i_s \leq j_t < j_s$.

Proof. This is proved through case-by-case analysis; see Figure 22. □

In some cases the order of the generators is necessary to determine whether the corresponding chords intersect in the CD. For instance the chords related to the closure of $W_1 = A(1, 2)A(1, 3)$ intersect in the CD, but the chords of $W_2 = A(1, 3)A(1, 2)$ do not intersect. However if $i_s < j_s < i_t < j_t$, then the chords related to $A(i_s, j_s)$ and $A(i_t, j_t)$ are parallel. This also implies that i_s, j_s, i_t, and j_t are distinct, in which case $A(i_s, j_s)$ and $A(i_t, j_t)$ commute as in Figure 21(a).

```
Function CanonicalBraid(N)
  Input: A chord diagram D_n with name N
  Output: W = A(i_1, j_1) ... A(i_n, j_n)

W ← ∅;
for k ← 1 to n do
    W ← W A(position(N,k,1), position(N,k,2));
    // where position(N,k,j) returns the position of the j^th letter
       k in N
end
return W;
```

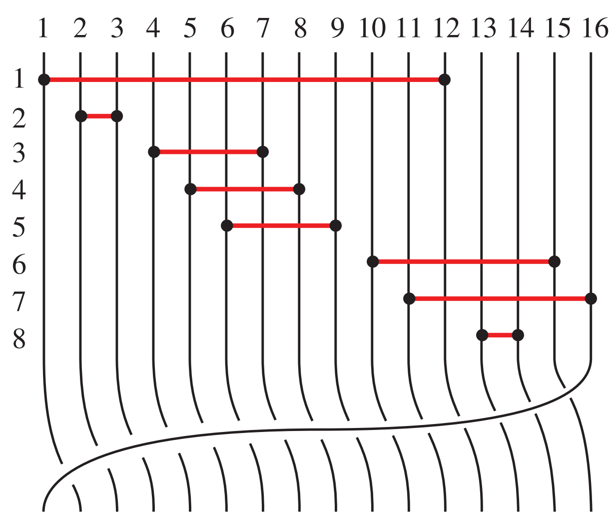


Figure 20.: The canonical braiding of the chord diagram represented by the DOW 1223453456718867.

Function DecreasingStablilization(W)

Input: $W = A(i_1, j_1) \dots A(i_n, j_n)$ corresponding to a chord diagram D_n with name N

Output: W which closes to form a braid containing no trivial loops.

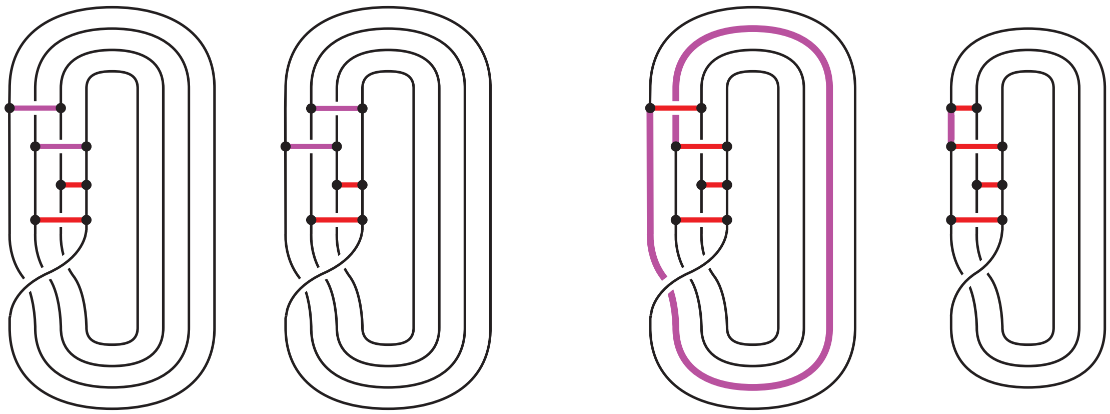
for $k \leftarrow 1$ **to** n **do**
    if $i_k - 1 \notin \{i_t, j_t \mid k \leq t \leq n\}$ *and* $i_k \notin \{i_t, j_t \mid 1 \leq t \leq k-1\}$ **then**
        $W \leftarrow$ `Decrease(`W, i_k`)` ;
        `// where Decrease(`W,c`) decreases` $i_t, j_t \leq c$ `by 1 for all generators of W`
    end
    if $j_k - 1 \notin \{i_t, j_t \mid k \leq t \leq n\}$ *and* $j_k \notin \{i_t, j_t \mid 1 \leq t \leq k-1\}$ **then**
        $W \leftarrow$ `Decrease(`W, j_k`)` ;
    end
end
$W \leftarrow$ `Increase(`W`)` ;
`// where Increase(`W`) increases` i_t, j_t `by 1 for all generators of W`
return W;

Algorithm 1: Computing Braid Index of a Chord Diagram

Input: A chord diagram D_n with name N

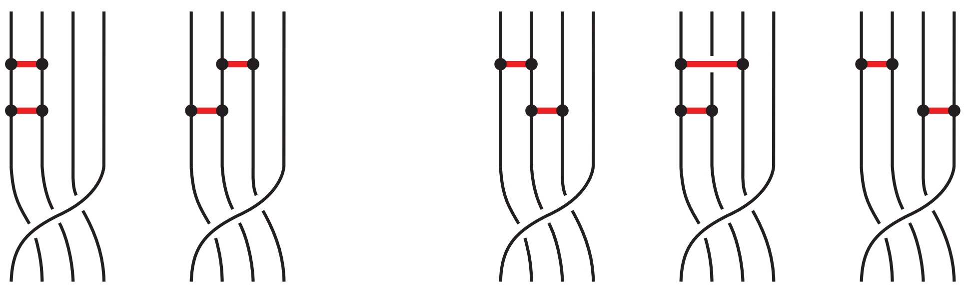
Output: $br(D_n)$

$m \leftarrow n + 1$;
$W^* \leftarrow$ `CanonicalBraid(`N`)` ;
for $k \leftarrow 1$ **to** $(n-1)!$ **do**
    $W_k \leftarrow \sigma^{(1)}_{n,k}(W^*)$;   `//` k^{th} `permutation of` S_n `fixing 1`
    $A(i_1, j_1) \dots A(i_n, j_n) \leftarrow$ `DecreasingStablilization(`W_k`)` ;
    $m \leftarrow \min\{m, \max\{i_t, j_t \mid 1 \leq t \leq n\}\}$;
end
return m ;



(a) Chord Commutation (b) Decreasing Stabilization (Removing trivial loops)

Figure 21.: Chord diagram preserving braid moves used in the Birman-Trapp algorithm. [29]



(a) Intersecting (b) Parallel

Figure 22.: Chords in the m-braid that correspond to either intersecting or parallel (non-intersecting) chords in the CD.

THEOREM 3.8. If the chord diagram D is the composition of linear chord diagrams $L_1, L_2, \ldots, L_n$, then

$$br(D) = \sum_{i=1}^{n} br(D_i) - (n-1), \tag{3.11}$$

where each chord diagram D_i corresponds to the closure of L_i, respectively.

Proof. Let D be composed of LCDs $L_1, L_2, \ldots, L_n$ having $l_1, l_2, \ldots, l_n$ chords, respectively, and fix a base point for D between L_n and L_1. Set $N = \sum_{i=1}^{n} l_i$.

We claim that there exists a $W \in \mathcal{C}_m$ which closes to a minimal braiding $\mathbf{W}$ for D, and has the form $W = W_1 W_2 \ldots W_n$ where each $W_i \in C_{m_i}$ (and $\sum_{i=1}^{n} m_i = m - n + 1$) closes to a braiding of D_i, such that, for $s < t$, all generators $A(i_{s_u}, j_{s_u}) \in W_s$ and $A(i_{t_v}, j_{t_v}) \in W_t$, satisfy $i_{s_u} < j_{s_u} \leq i_{t_v} < j_{t_v}$ for each $u \in [l_s]$ and $v \in [l_t]$. This braiding is pictured in Figure 23.

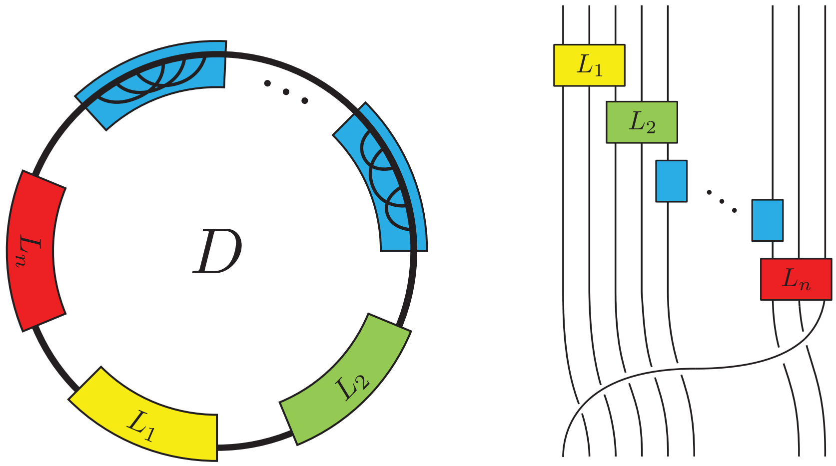


Figure 23.: A minimal braiding of a chord diagram composed of isolated linear chord diagrams.

Suppose $U = A(i_1, j_1)A(i_2, j_2)\dots A(i_k, j_k)\dots A(i_N, j_N) \in \mathcal{C}_m$ also represents D. Without loss of generality, we may assume that $A(i_1, j_1)$ corresponds to the first chord of L_1 and that i_1 and j_1 correspond to the first and second intersections in the CD, respectively. This assumption is justified, since we may cyclically permute $A(i_1, j_1)$ until it is the first generator and $i_1 = 1$. If L_1 contains more than one chord, take $A(i_k, j_k)$ to be the next generator representing a chord of L_1. If $k > 2$, then $A(i_2, j_2)$ belongs to one of $L_2, \dots, L_n$ and corresponds to a parallel chord in D. Since $A(i_k, j_k)$ belongs to L_1 and $A(i_2, j_2)$ belongs to L_i for some $i > 1$, it follows that $i_2 \geq j_k$. From Lemma 3.7, $A(i_2, j_2)$ must commute with $A(i_k, j_k)$, as only the right-most case of the parallel chords in Figure 22 can occur. Indeed, all the chords $A(i_3, j_3), \dots, A(i_{k-1}, j_{k-1})$ must also be parallel with $A(i_k, j_k)$ and commute. Thus, $U' = A(i_1, j_1)A(i_k, j_k)A(i_2, j_2)\dots A(i_{k-1}, j_{k-1})A(i_{k+1}, j_{k+1})\dots A(i_N, j_N)$ also reads as D. Continuing in this fashion, we may construct U'' such that the first l_1 generators of U'' correspond to the chords in L_1.

Similarly, if the first generator to follow $A(i_{l_1}, j_{l_1})$ in U'' does not belong to L_2, then it commutes with the chords from L_2. This follows from the facts that both ends from any L_2 chord must precede all other chord ends from L_i where $i > 2$, and all such chords are parallel to the L_2 chords. Thus the generators of U can be organized so that the first l_1 chords correspond to L_1, the next l_2 chords correspond to L_2, etc., as in Figure 23. Note that the only two operations used to rearrange U into W were cyclic permutation and commutation. Therefore if U was a minimal braiding of D, then so is W and the claim is proven.

Note that the closure $\mathbf{W}_i$ of each W_i is a minimal braiding of D_i. Suppose not, i.e., there exists a

$W_i^\star \in \mathcal{C}_m$ that closes to $\mathbf{W}_i^\star$ which also represents D_i and contains fewer strings than $\mathbf{W}_i$. Cyclically permute $W_i^\star$ so that the first chord of L_i is the first generator of $W_i^\star$, and the first chord end is on the first string. Thus $W^\star = W_1 \dots W_i^\star \dots W_n$ would close to a braid with fewer strings than $\mathbf{W}$, but this is a contradiction as $\mathbf{W}$ is taken to be a minimal braiding.

Finally observe that the generators of L_k and L_{k+1} intersect on one exactly one common string for each $1 \leq k \leq n-1$. If they intersected on more than one string, the chords of L_k and L_{k+1} would intersect. However if the generators did not intersect on a common chord, there would be a decreasing stabilization and $\mathbf{W}$ would not be minimal. $\square$

LEMMA 3.8. [29] If the set of chords of the diagram D has a subset of p isolated chords, then $br(D) \geq p+1$.

THEOREM 3.9. Let D_n be a chord diagram with n chords. Then $br(D_n) = n+1$ iff D_n is fully-degenerate.

Proof. ($\Longrightarrow$) It is straightforward to verify that 1122 and 1221 have braid index 3 and are fully-degenerate, but 1212 has braid index 2.

Now inductively assume that all chord diagrams with $n \geq 3$ chords and braid index $n+1$ are fully-degenerate, and consider any chord diagram D_{n+1} with $n+1$ chords such that $br(D_{n+1}) = n+2$. If D_{n+1} is fully-degenerate then we are done, so assume to the contrary that D_{n+1} has at least one pair, $\{a, b\}$ of intersecting chords. At least one of these chords, a, does not intersect the remaining $n-2$ chords, as removing its paired counterpart, b, would decrease the braid index by exactly one by Corollary 3.2. By our induction hypothesis, the resulting diagram would have braid index n, so it must be fully-degenerate. However if we remove any one of the remaining $n-2$ non-intersecting chords, we decrease the braid index by exactly one and have a chord diagram that is not fully-degenerate. This is a contradiction, so no such intersecting pair of chords exist and D_{n+1} is fully-degenerate.

($\Longleftarrow$) Since D_n is fully degenerate, it contains n isolated chords so $br(D_n) > n$, by Lemma 3.8. Therefore $br(D_n) = n+1$ because $br(D_n) \leq n+1$ for all chord diagrams by Corollary 3.2. $\square$

COROLLARY 3.3. The number of linear chord diagrams with n chords that have braid index $n+1$ is

$$\frac{1}{n+1}\binom{2n}{n},$$

i.e., the n^{th} Catalan number C_n.

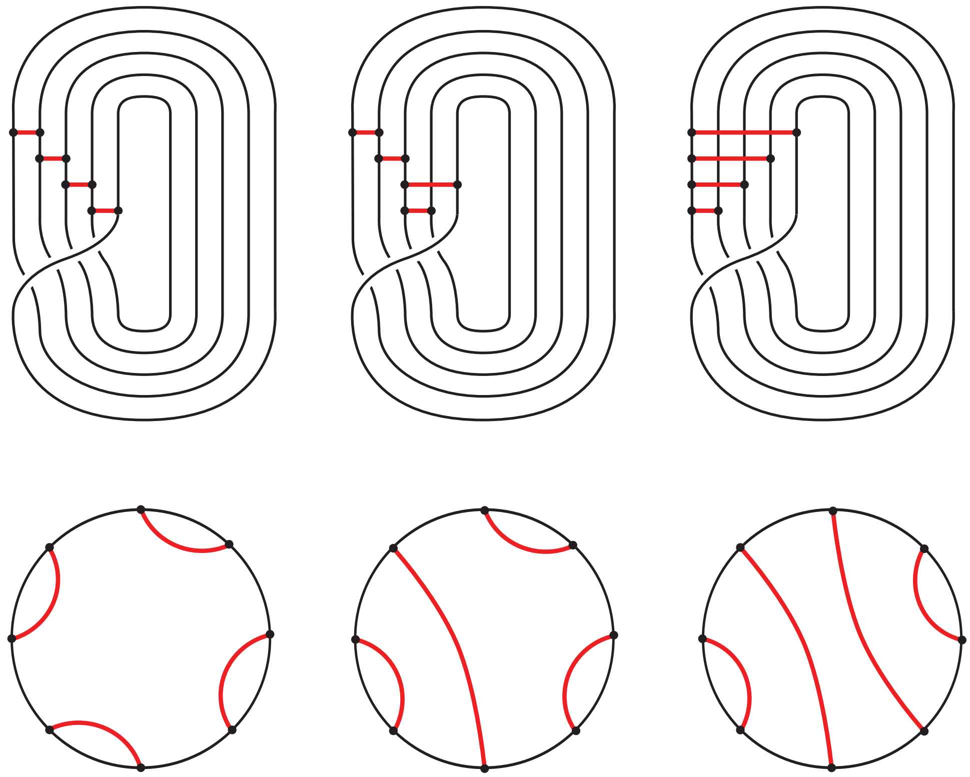


Figure 24.: Minimal braidings for fully-degenerate chord diagrams with 4 chords.

The total number of fully-degenerate linear chord diagrams with n chords is C_n [50], and the number of symmetric linear chord diagrams with n chords that have braid index $(n+1)$ is $\binom{2n}{n}$.

CONJECTURE 3.2. The number of linear chord diagrams with n chords that have braid index n is

$$\frac{n^2(n-1)}{2(n+2)}\left[\frac{1}{n+1}\binom{2n}{n}\right]. \tag{3.12}$$

3.6.5 Minimal Braidings of Select Chord Diagrams

DEFINITION 3.6.3. The *waffle* $\mathcal{W}_{m,n}$ is the chord diagram named

$$12\ldots(m+n)(m)(m-1)\ldots 21(m+n)(m+n-1)\ldots(m+1). \tag{3.13}$$

See Figure 25.

LEMMA 3.9. For $m \geq 1$, $br(\mathcal{W}_{m,1}) = m+1$.

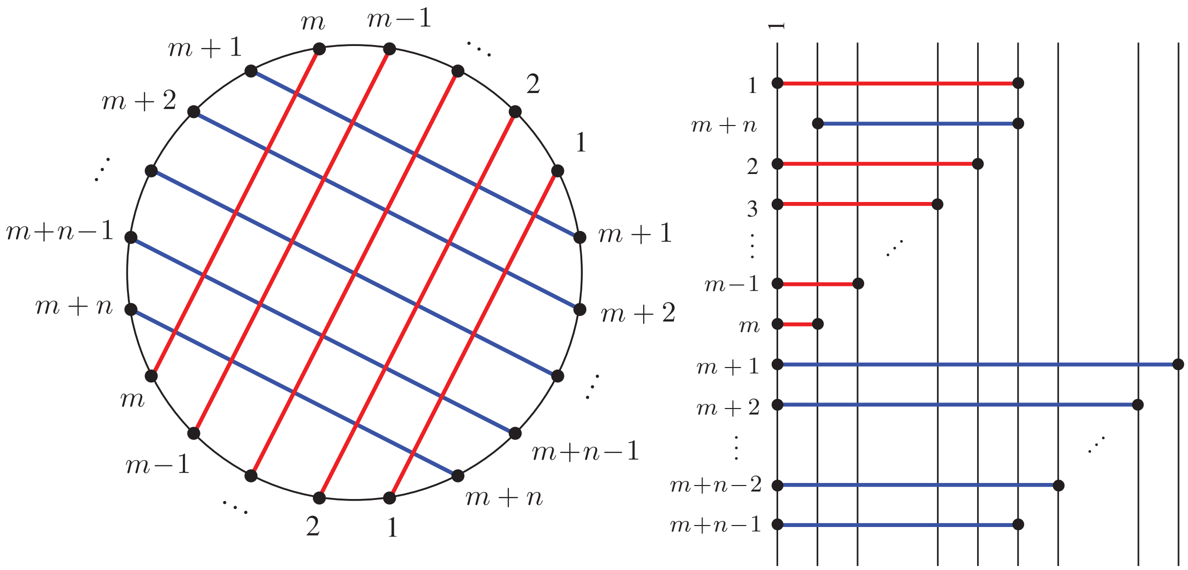


Figure 25.: The waffle chord diagram $\mathcal{W}_{m,n}$ and one of its minimal braidings.

Proof. By construction, the waffle $\mathcal{W}_{m,1}$ contains m non-intersecting edges so $br(\mathcal{W}_{m,1}) \geq m+1$. It is short work to verify that $W = A(1, m+1)A(1, m) \ldots A(1, 2)A(1, m+1)$ reads the name of $\mathcal{W}_{m,1}$. $\square$

PROPOSITION 3.1. *For* $m, n \geq 2$,

$$br(\mathcal{W}_{m,n}) \leq m + n - 1\,. \tag{3.14}$$

Proof. Without loss of generality, let $m \geq n \geq 1$. The sequence of generators

$$\underbrace{A(1,m)A(2,m)A(1,m-1)A(1,m-2)\ldots A(1,1)}_{m+1 \text{ generators}}\underbrace{A(1,m+n-1)A(1,m+n-2)\ldots A(1,m+1)}_{n-1 \text{ generators}}$$

uses $m + n - 1$ strings to generate a braid corresponding to $\mathcal{W}_{m,n}$, i.e., $br(\mathcal{W}_{m,n}) \leq n + m - 1$ (see Figure 25). $\square$

CONJECTURE 3.3. *For* $m, n \geq 2$, $br(\mathcal{W}_{m,n}) \leq m + n - 1$.

PROPOSITION 3.2. The braid index of the tangled cord $\mathcal{T}_n$ is $\lceil n/2 \rceil + 1$.

Proof. Let $\mathcal{T}_n = t_1 t_2 t_3 \ldots t_{2n}$ where $t_i \in [n]$ for each $i \in [2n]$. Then sequence of generators

$$A(t_1, t_1 + 1)A(t_2, t_2 + 1)A(t_3, t_3 + 1) \ldots A(t_n, t_n + 1)$$

generates a braid corresponding to the tangled cord (see Figure 26). Further,

$$\max\{t_1, t_1+1, t_2, t_2+1, \ldots, t_n, t_n+1\} = \lceil n/2 \rceil + 1. \tag{3.15}$$

Thus $br(\mathcal{T}_n) \leq \lceil n/2 \rceil + 1$.

Note that the odd numbered chords in the chord diagram of the tangled cord are non-intersecting, so $br(\mathcal{T}_n) \geq \lceil n/2 + 1 \rceil$. □

n	Generators
1	A(1,2)
2	A(1,2)**A(1,2)**
3	A(1,2)**A(2,3)**A(1,2)
4	A(1,2)A(2,3)A(1,2)**A(2,3)**
5	A(1,2)A(2,3)A(1,2)**A(3,4)**A(2,3)
6	A(1,2)A(2,3)A(1,2)A(3,4)A(2,3)**A(3,4)**
7	A(1,2)A(2,3)A(1,2)A(3,4)A(2,3)**A(4,5)**A(3,4)
$\vdots$	$\vdots$

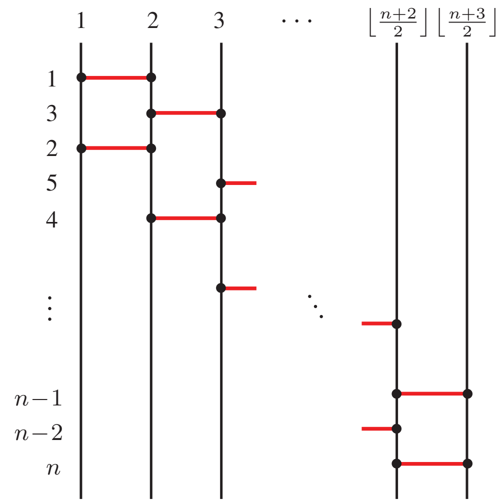


Figure 26.: A minimum braiding of the tangled cord $\mathcal{T}_n$.

3.7 Tables & Figures

Table 5: All Double Occurrence Words.

Symbols	All	Irreducible	Strongly Irreducible
1	1	1	1
2	3	2	1
3	15	10	4
4	105	74	27
5	945	706	248
6	10,395	8,162	2,830
7	135,135	110,410	38,232
8	2,027,025	1,708,394	593,859
9	34,459,425	29,752,066	10,401,712
10	654,729,075	576,037,442	202,601,898
11	13,749,310,575	12,277,827,850	4,342,263,000
12	316,234,143,225	285,764,591,114	101,551,822,350
13	7,905,853,580,625	7,213,364,729,026	2,573,779,506,192
14	213,458,046,676,875	196,316,804,255,522	70,282,204,726,396
15	6,190,283,353,629,375	5,731,249,477,826,890	2,057,490,936,366,320
OEIS	A001147 (K_n)	A000698 (I_n)	A000699 (S_n)

Table 6: Palindromic Double Occurrence Words.

Symbols	All	Irreducible	Strongly Irreducible
1	1	1	1
2	3	2	1
3	7	6	2
4	25	20	7
5	81	72	22
6	331	290	96
7	1,303	1,198	380
8	5937	5,452	1,853
9	26,785	25,176	8,510
10	133,651	125,874	44,940
11	669,351	637,926	229,836
12	3,609,673	3,448,708	1,296,410
13	19,674,097	18,919,048	7,211,116
14	113,525,595	109,412,210	43,096,912
15	664,400,311	642,798,510	256,874,200
OEIS	A047974 (L_n)	A195186 (J_n)	A004300 (T_n)

Table 7: Micronuclear arrangements corresponding to the assembly graph 1122.

Assembly Word	Hamiltonian Set of Polygonal Paths	α	α^{R_1}	α^{R_2}	$\alpha^{R_{1,2}}$
1122	$\{v_1, e_2, v_2\}$	$M_1M_2M_3$	$M_3\overline{M}_2M_1$	–	–
		$\overline{M}_1M_2M_3$	$\overline{M}_3\overline{M}_2M_1$		
		$M_1M_2\overline{M}_3$	$M_3\overline{M}_2\overline{M}_1$		
		$\overline{M}_1M_2\overline{M}_3$	$\overline{M}_3\overline{M}_2\overline{M}_1$		
	$\{v_1\}, \{v_2\}$	$M_{1,1}M_{1,2}M_{2,1}M_{2,2}$	$M_{1,2}M_{1,1}M_{2,1}M_{2,2}$	$M_{1,1}M_{1,2}M_{2,2}M_{2,1}$	$M_{1,2}M_{1,1}M_{2,2}M_{2,1}$
		$\overline{M}_{1,1}M_{1,2}M_{2,1}M_{2,2}$	$\overline{M}_{1,2}M_{1,1}M_{2,1}M_{2,2}$	$\overline{M}_{1,1}M_{1,2}M_{2,2}M_{2,1}$	$\overline{M}_{1,2}M_{1,1}M_{2,2}M_{2,1}$
		$M_{1,1}\overline{M}_{1,2}M_{2,1}M_{2,2}$	$M_{1,2}\overline{M}_{1,1}M_{2,1}M_{2,2}$	$M_{1,1}\overline{M}_{1,2}M_{2,2}M_{2,1}$	$M_{1,2}\overline{M}_{1,1}M_{2,2}M_{2,1}$
		$\overline{M}_{1,1}\overline{M}_{1,2}M_{2,1}M_{2,2}$	$\overline{M}_{1,2}\overline{M}_{1,1}M_{2,1}M_{2,2}$	$\overline{M}_{1,1}\overline{M}_{1,2}M_{2,2}M_{2,1}$	$\overline{M}_{1,2}\overline{M}_{1,1}M_{2,2}M_{2,1}$
		$M_{1,1}M_{1,2}\overline{M}_{2,1}M_{2,2}$	$M_{1,2}M_{1,1}\overline{M}_{2,1}M_{2,2}$	$M_{1,1}M_{1,2}\overline{M}_{2,2}M_{2,1}$	$M_{1,2}M_{1,1}\overline{M}_{2,2}M_{2,1}$
		$\overline{M}_{1,1}M_{1,2}\overline{M}_{2,1}M_{2,2}$	$\overline{M}_{1,2}M_{1,1}\overline{M}_{2,1}M_{2,2}$	$\overline{M}_{1,1}M_{1,2}\overline{M}_{2,2}M_{2,1}$	$\overline{M}_{1,2}M_{1,1}\overline{M}_{2,2}M_{2,1}$
		$M_{1,1}\overline{M}_{1,2}\overline{M}_{2,1}M_{2,2}$	$M_{1,2}\overline{M}_{1,1}\overline{M}_{2,1}M_{2,2}$	$M_{1,1}\overline{M}_{1,2}\overline{M}_{2,2}M_{2,1}$	$M_{1,2}\overline{M}_{1,1}\overline{M}_{2,2}M_{2,1}$
		$\overline{M}_{1,1}\overline{M}_{1,2}\overline{M}_{2,1}M_{2,2}$	$\overline{M}_{1,2}\overline{M}_{1,1}\overline{M}_{2,1}M_{2,2}$	$\overline{M}_{1,1}\overline{M}_{1,2}\overline{M}_{2,2}M_{2,1}$	$\overline{M}_{1,2}\overline{M}_{1,1}\overline{M}_{2,2}M_{2,1}$
		$M_{1,1}M_{1,2}M_{2,1}\overline{M}_{2,2}$	$M_{1,2}M_{1,1}M_{2,1}\overline{M}_{2,2}$	$M_{1,1}M_{1,2}M_{2,2}\overline{M}_{2,1}$	$M_{1,2}M_{1,1}M_{2,2}\overline{M}_{2,1}$
		$\overline{M}_{1,1}M_{1,2}M_{2,1}\overline{M}_{2,2}$	$\overline{M}_{1,2}M_{1,1}M_{2,1}\overline{M}_{2,2}$	$\overline{M}_{1,1}M_{1,2}M_{2,2}\overline{M}_{2,1}$	$\overline{M}_{1,2}M_{1,1}M_{2,2}\overline{M}_{2,1}$
		$M_{1,1}\overline{M}_{1,2}M_{2,1}\overline{M}_{2,2}$	$M_{1,2}\overline{M}_{1,1}M_{2,1}\overline{M}_{2,2}$	$M_{1,1}\overline{M}_{1,2}M_{2,2}\overline{M}_{2,1}$	$M_{1,2}\overline{M}_{1,1}M_{2,2}\overline{M}_{2,1}$
		$\overline{M}_{1,1}\overline{M}_{1,2}M_{2,1}\overline{M}_{2,2}$	$\overline{M}_{1,2}\overline{M}_{1,1}M_{2,1}\overline{M}_{2,2}$	$\overline{M}_{1,1}\overline{M}_{1,2}M_{2,2}\overline{M}_{2,1}$	$\overline{M}_{1,2}\overline{M}_{1,1}M_{2,2}\overline{M}_{2,1}$
		$M_{1,1}M_{1,2}\overline{M}_{2,1}\overline{M}_{2,2}$	$M_{1,2}M_{1,1}\overline{M}_{2,1}\overline{M}_{2,2}$	$M_{1,1}M_{1,2}\overline{M}_{2,2}\overline{M}_{2,1}$	$M_{1,2}M_{1,1}\overline{M}_{2,2}\overline{M}_{2,1}$
		$\overline{M}_{1,1}M_{1,2}\overline{M}_{2,1}\overline{M}_{2,2}$	$\overline{M}_{1,2}M_{1,1}\overline{M}_{2,1}\overline{M}_{2,2}$	$\overline{M}_{1,1}M_{1,2}\overline{M}_{2,2}\overline{M}_{2,1}$	$\overline{M}_{1,2}M_{1,1}\overline{M}_{2,2}\overline{M}_{2,1}$
		$M_{1,1}\overline{M}_{1,2}\overline{M}_{2,1}\overline{M}_{2,2}$	$M_{1,2}\overline{M}_{1,1}\overline{M}_{2,1}\overline{M}_{2,2}$	$M_{1,1}\overline{M}_{1,2}\overline{M}_{2,2}\overline{M}_{2,1}$	$M_{1,2}\overline{M}_{1,1}\overline{M}_{2,2}\overline{M}_{2,1}$
		$\overline{M}_{1,1}\overline{M}_{1,2}\overline{M}_{2,1}\overline{M}_{2,2}$	$\overline{M}_{1,2}\overline{M}_{1,1}\overline{M}_{2,1}\overline{M}_{2,2}$	$\overline{M}_{1,1}\overline{M}_{1,2}\overline{M}_{2,2}\overline{M}_{2,1}$	$\overline{M}_{1,2}\overline{M}_{1,1}\overline{M}_{2,2}\overline{M}_{2,1}$

Table 8: Micronuclear arrangements corresponding to the assembly graph 1212.

Assembly Word	Hamiltonian Set of Polygonal Paths	α	α^{R_1}	α^{R_2}	$\alpha^{R_{1,2}}$
1212	$\{v_1, e_1, v_2\}$	$M_2M_1M_3$	$M_3M_1\overline{M}_2$	–	–
		$M_2\overline{M}_1M_3$	$\overline{M}_3M_1\overline{M}_2$		
		$M_2M_1\overline{M}_3$	$M_3\overline{M}_1\overline{M}_2$		
		$M_2\overline{M}_1\overline{M}_3$	$\overline{M}_3\overline{M}_1\overline{M}_2$		
	$\{v_1, e_2, v_2\}$	$M_3M_2M_1$	$M_1\overline{M}_2M_3$	–	–
		$\overline{M}_3M_2M_1$	$\overline{M}_1\overline{M}_2M_3$		
		$M_3M_2\overline{M}_1$	$M_1\overline{M}_2\overline{M}_3$		
		$\overline{M}_3M_2\overline{M}_1$	$\overline{M}_1\overline{M}_2\overline{M}_3$		
	$\{v_1, e_3, v_2\}$	$M_1M_3M_2$	$\overline{M}_2M_3M_1$	–	–
		$\overline{M}_1M_3M_2$	$\overline{M}_2\overline{M}_3M_1$		
		$M_1\overline{M}_3M_2$	$\overline{M}_2M_3\overline{M}_1$		
		$\overline{M}_1\overline{M}_3M_2$	$\overline{M}_2\overline{M}_3\overline{M}_1$		
	$\{v_1\}, \{v_2\}$	$M_{1,1}M_{2,1}M_{1,2}M_{2,2}$	$M_{1,2}M_{2,1}M_{1,1}M_{2,2}$	$M_{1,1}M_{2,2}M_{1,2}M_{2,1}$	$M_{1,2}M_{2,2}M_{1,1}M_{2,1}$
		$\overline{M}_{1,1}M_{2,1}M_{1,2}M_{2,2}$	$\overline{M}_{1,2}M_{2,1}M_{1,1}M_{2,2}$	$\overline{M}_{1,1}M_{2,2}M_{1,2}M_{2,1}$	$\overline{M}_{1,2}M_{2,2}M_{1,1}M_{2,1}$
		$M_{1,1}\overline{M}_{2,1}M_{1,2}M_{2,2}$	$M_{1,2}\overline{M}_{2,1}M_{1,1}M_{2,2}$	$M_{1,1}\overline{M}_{2,2}M_{1,2}M_{2,1}$	$M_{1,2}\overline{M}_{2,2}M_{1,1}M_{2,1}$
		$\overline{M}_{1,1}\overline{M}_{2,1}M_{1,2}M_{2,2}$	$\overline{M}_{1,2}\overline{M}_{2,1}M_{1,1}M_{2,2}$	$\overline{M}_{1,1}\overline{M}_{2,2}M_{1,2}M_{2,1}$	$\overline{M}_{1,2}\overline{M}_{2,2}M_{1,1}M_{2,1}$
		$M_{1,1}M_{2,1}\overline{M}_{1,2}M_{2,2}$	$M_{1,2}M_{2,1}\overline{M}_{1,1}M_{2,2}$	$M_{1,1}M_{2,2}\overline{M}_{1,2}M_{2,1}$	$M_{1,2}M_{2,2}\overline{M}_{1,1}M_{2,1}$
		$\overline{M}_{1,1}M_{2,1}\overline{M}_{1,2}M_{2,2}$	$\overline{M}_{1,2}M_{2,1}\overline{M}_{1,1}M_{2,2}$	$\overline{M}_{1,1}M_{2,2}\overline{M}_{1,2}M_{2,1}$	$\overline{M}_{1,2}M_{2,2}\overline{M}_{1,1}M_{2,1}$
		$M_{1,1}\overline{M}_{2,1}\overline{M}_{1,2}M_{2,2}$	$M_{1,2}\overline{M}_{2,1}\overline{M}_{1,1}M_{2,2}$	$M_{1,1}\overline{M}_{2,2}\overline{M}_{1,2}M_{2,1}$	$M_{1,2}\overline{M}_{2,2}\overline{M}_{1,1}M_{2,1}$
		$\overline{M}_{1,1}\overline{M}_{2,1}\overline{M}_{1,2}M_{2,2}$	$\overline{M}_{1,2}\overline{M}_{2,1}\overline{M}_{1,1}M_{2,2}$	$\overline{M}_{1,1}\overline{M}_{2,2}\overline{M}_{1,2}M_{2,1}$	$\overline{M}_{1,2}\overline{M}_{2,2}\overline{M}_{1,1}M_{2,1}$
		$M_{1,1}M_{2,1}M_{1,2}\overline{M}_{2,2}$	$M_{1,2}M_{2,1}M_{1,1}\overline{M}_{2,2}$	$M_{1,1}M_{2,2}M_{1,2}\overline{M}_{2,1}$	$M_{1,2}M_{2,2}M_{1,1}\overline{M}_{2,1}$
		$\overline{M}_{1,1}M_{2,1}M_{1,2}\overline{M}_{2,2}$	$\overline{M}_{1,2}M_{2,1}M_{1,1}\overline{M}_{2,2}$	$\overline{M}_{1,1}M_{2,2}M_{1,2}\overline{M}_{2,1}$	$\overline{M}_{1,2}M_{2,2}M_{1,1}\overline{M}_{2,1}$
		$M_{1,1}\overline{M}_{2,1}M_{1,2}\overline{M}_{2,2}$	$M_{1,2}\overline{M}_{2,1}M_{1,1}\overline{M}_{2,2}$	$M_{1,1}\overline{M}_{2,2}M_{1,2}\overline{M}_{2,1}$	$M_{1,2}\overline{M}_{2,2}M_{1,1}\overline{M}_{2,1}$
		$\overline{M}_{1,1}\overline{M}_{2,1}M_{1,2}\overline{M}_{2,2}$	$\overline{M}_{1,2}\overline{M}_{2,1}M_{1,1}\overline{M}_{2,2}$	$\overline{M}_{1,1}\overline{M}_{2,2}M_{1,2}\overline{M}_{2,1}$	$\overline{M}_{1,2}\overline{M}_{2,2}M_{1,1}\overline{M}_{2,1}$
		$M_{1,1}M_{2,1}\overline{M}_{1,2}\overline{M}_{2,2}$	$M_{1,2}M_{2,1}\overline{M}_{1,1}\overline{M}_{2,2}$	$M_{1,1}M_{2,2}\overline{M}_{1,2}\overline{M}_{2,1}$	$M_{1,2}M_{2,2}\overline{M}_{1,1}\overline{M}_{2,1}$
		$\overline{M}_{1,1}M_{2,1}\overline{M}_{1,2}\overline{M}_{2,2}$	$\overline{M}_{1,2}M_{2,1}\overline{M}_{1,1}\overline{M}_{2,2}$	$\overline{M}_{1,1}M_{2,2}\overline{M}_{1,2}\overline{M}_{2,1}$	$\overline{M}_{1,2}M_{2,2}\overline{M}_{1,1}\overline{M}_{2,1}$
		$M_{1,1}\overline{M}_{2,1}\overline{M}_{1,2}\overline{M}_{2,2}$	$M_{1,2}\overline{M}_{2,1}\overline{M}_{1,1}\overline{M}_{2,2}$	$M_{1,1}\overline{M}_{2,2}\overline{M}_{1,2}\overline{M}_{2,1}$	$M_{1,2}\overline{M}_{2,2}\overline{M}_{1,1}\overline{M}_{2,1}$
		$\overline{M}_{1,1}\overline{M}_{2,1}\overline{M}_{1,2}\overline{M}_{2,2}$	$\overline{M}_{1,2}\overline{M}_{2,1}\overline{M}_{1,1}\overline{M}_{2,2}$	$\overline{M}_{1,1}\overline{M}_{2,2}\overline{M}_{1,2}\overline{M}_{2,1}$	$\overline{M}_{1,2}\overline{M}_{2,2}\overline{M}_{1,1}\overline{M}_{2,1}$

Table 9: Micronuclear arrangements corresponding to the assembly graph 1221.

Assembly Word	Hamiltonian Set of Polygonal Paths	α	α^{R_1}	α^{R_2}	$\alpha^{R_{1,2}}$
1221	$\{v_1, e_1, v_2\}$	$M_2M_3M_1$	$M_1M_3\overline{M}_2$	–	–
		$M_2\overline{M}_3M_1$	$\overline{M}_1M_3\overline{M}_2$		
		$M_2M_3\overline{M}_1$	$M_1\overline{M}_3\overline{M}_2$		
		$M_2\overline{M}_3\overline{M}_1$	$\overline{M}_1\overline{M}_3\overline{M}_2$		
	$\{v_1, e_3, v_2\}$	$M_3M_1M_2$	$\overline{M}_2M_1M_3$	–	–
		$\overline{M}_3M_1M_2$	$\overline{M}_2\overline{M}_1M_3$		
		$M_3\overline{M}_1M_2$	$\overline{M}_2M_1\overline{M}_3$		
		$\overline{M}_3\overline{M}_1M_2$	$\overline{M}_2\overline{M}_1\overline{M}_3$		
	$\{v_1\}, \{v_2\}$	$M_{1,1}M_{2,1}M_{2,2}M_{1,2}$	$M_{1,2}M_{2,1}M_{2,2}M_{1,1}$	$M_{1,1}M_{2,2}M_{2,1}M_{1,2}$	$M_{1,2}M_{2,2}M_{2,1}M_{1,1}$
		$\overline{M}_{1,1}M_{2,1}M_{2,2}M_{1,2}$	$\overline{M}_{1,2}M_{2,1}M_{2,2}M_{1,1}$	$\overline{M}_{1,1}M_{2,2}M_{2,1}M_{1,2}$	$\overline{M}_{1,2}M_{2,2}M_{2,1}M_{1,1}$
		$M_{1,1}\overline{M}_{2,1}M_{2,2}M_{1,2}$	$M_{1,2}\overline{M}_{2,1}M_{2,2}M_{1,1}$	$M_{1,1}\overline{M}_{2,2}M_{2,1}M_{1,2}$	$M_{1,2}\overline{M}_{2,2}M_{2,1}M_{1,1}$
		$\overline{M}_{1,1}\overline{M}_{2,1}M_{2,2}M_{1,2}$	$\overline{M}_{1,2}\overline{M}_{2,1}M_{2,2}M_{1,1}$	$\overline{M}_{1,1}\overline{M}_{2,2}M_{2,1}M_{1,2}$	$\overline{M}_{1,2}\overline{M}_{2,2}M_{2,1}M_{1,1}$
		$M_{1,1}M_{2,1}\overline{M}_{2,2}M_{1,2}$	$M_{1,2}M_{2,1}\overline{M}_{2,2}M_{1,1}$	$M_{1,1}M_{2,2}\overline{M}_{2,1}M_{1,2}$	$M_{1,2}M_{2,2}\overline{M}_{2,1}M_{1,1}$
		$\overline{M}_{1,1}M_{2,1}\overline{M}_{2,2}M_{1,2}$	$\overline{M}_{1,2}M_{2,1}\overline{M}_{2,2}M_{1,1}$	$\overline{M}_{1,1}M_{2,2}\overline{M}_{2,1}M_{1,2}$	$\overline{M}_{1,2}M_{2,2}\overline{M}_{2,1}M_{1,1}$
		$M_{1,1}\overline{M}_{2,1}\overline{M}_{2,2}M_{1,2}$	$M_{1,2}\overline{M}_{2,1}\overline{M}_{2,2}M_{1,1}$	$M_{1,1}\overline{M}_{2,2}\overline{M}_{2,1}M_{1,2}$	$M_{1,2}\overline{M}_{2,2}\overline{M}_{2,1}M_{1,1}$
		$\overline{M}_{1,1}\overline{M}_{2,1}\overline{M}_{2,2}M_{1,2}$	$\overline{M}_{1,2}\overline{M}_{2,1}\overline{M}_{2,2}M_{1,1}$	$\overline{M}_{1,1}\overline{M}_{2,2}\overline{M}_{2,1}M_{1,2}$	$\overline{M}_{1,2}\overline{M}_{2,2}\overline{M}_{2,1}M_{1,1}$
		$M_{1,1}M_{2,1}M_{2,2}\overline{M}_{1,2}$	$M_{1,2}M_{2,1}M_{2,2}\overline{M}_{1,1}$	$M_{1,1}M_{2,2}M_{2,1}\overline{M}_{1,2}$	$M_{1,2}M_{2,2}M_{2,1}\overline{M}_{1,1}$
		$\overline{M}_{1,1}M_{2,1}M_{2,2}\overline{M}_{1,2}$	$\overline{M}_{1,2}M_{2,1}M_{2,2}\overline{M}_{1,1}$	$\overline{M}_{1,1}M_{2,2}M_{2,1}\overline{M}_{1,2}$	$\overline{M}_{1,2}M_{2,2}M_{2,1}\overline{M}_{1,1}$
		$M_{1,1}\overline{M}_{2,1}M_{2,2}\overline{M}_{1,2}$	$M_{1,2}\overline{M}_{2,1}M_{2,2}\overline{M}_{1,1}$	$M_{1,1}\overline{M}_{2,2}M_{2,1}\overline{M}_{1,2}$	$M_{1,2}\overline{M}_{2,2}M_{2,1}\overline{M}_{1,1}$
		$\overline{M}_{1,1}\overline{M}_{2,1}M_{2,2}\overline{M}_{1,2}$	$\overline{M}_{1,2}\overline{M}_{2,1}M_{2,2}\overline{M}_{1,1}$	$\overline{M}_{1,1}\overline{M}_{2,2}M_{2,1}\overline{M}_{1,2}$	$\overline{M}_{1,2}\overline{M}_{2,2}M_{2,1}\overline{M}_{1,1}$
		$M_{1,1}M_{2,1}\overline{M}_{2,2}\overline{M}_{1,2}$	$M_{1,2}M_{2,1}\overline{M}_{2,2}\overline{M}_{1,1}$	$M_{1,1}M_{2,2}\overline{M}_{2,1}\overline{M}_{1,2}$	$M_{1,2}M_{2,2}\overline{M}_{2,1}\overline{M}_{1,1}$
		$\overline{M}_{1,1}M_{2,1}\overline{M}_{2,2}\overline{M}_{1,2}$	$\overline{M}_{1,2}M_{2,1}\overline{M}_{2,2}\overline{M}_{1,1}$	$\overline{M}_{1,1}M_{2,2}\overline{M}_{2,1}\overline{M}_{1,2}$	$\overline{M}_{1,2}M_{2,2}\overline{M}_{2,1}\overline{M}_{1,1}$
		$M_{1,1}\overline{M}_{2,1}\overline{M}_{2,2}\overline{M}_{1,2}$	$M_{1,2}\overline{M}_{2,1}\overline{M}_{2,2}\overline{M}_{1,1}$	$M_{1,1}\overline{M}_{2,2}\overline{M}_{2,1}\overline{M}_{1,2}$	$M_{1,2}\overline{M}_{2,2}\overline{M}_{2,1}\overline{M}_{1,1}$
		$\overline{M}_{1,1}\overline{M}_{2,1}\overline{M}_{2,2}\overline{M}_{1,2}$	$\overline{M}_{1,2}\overline{M}_{2,1}\overline{M}_{2,2}\overline{M}_{1,1}$	$\overline{M}_{1,1}\overline{M}_{2,2}\overline{M}_{2,1}\overline{M}_{1,2}$	$\overline{M}_{1,2}\overline{M}_{2,2}\overline{M}_{2,1}\overline{M}_{1,1}$

Table 10: Number of chord diagrams and linear chord diagrams by braid index.

	CD Braid Index						
n	2	3	4	5	6	7	8
1	1						
2	1	1					
3	1	2	2				
4	1	6	7	3			
5	1	12	38	22	6		
6	1	29	198	235	79	12	
7	1	65	1,107	2,390	1,409	284	27

	LCD Braid Index							
n	2	3	4	5	6	7	8	9
1	1							
2	1	2						
3	1	9	5					
4	1	34	56	14				
5	1	122	480	300	42			
6	1	414	3,743	4,620	1,485	132		
7	1	1,339	27,554	62,769	36,036	7,007	429	
8	1	4,190	193,084	797,760	752,100	246,428	32,032	1,430

3.8 Chapter Notations

Symbol	Description	Page Reference
$A(i,j)$	a generator of the chord algebra	58
$\mathrm{An}(\Gamma)$	assembly number of the assembly graph Γ	44
$\alpha(\sigma,\tau,\epsilon)$	a micronuclear arrangement	51
$br(D)$	braid index of a chord diagram D	59
C_m	the chord algebra	58
C_n	the Catalan numbers	65
CD	chord diagram	56
D_w	the chord diagram represented by the DOW w	56
DOW	double occurrence word	32
e_i	an edge from E	41
E	set of edges of an assembly graph	41
ϵ_i	sign function of a micronuclear arrangement	51
F_n	the Fibonacci numbers	45
γ	a Hamiltonian set of polygonal paths	44
Γ	an assembly graph	41
$\lvert\Gamma\rvert$	size of Γ	42
I_n	number of irreducible DOWS of size n	38
J_n	number of irreducible, palindromic DOWs of size n	38
K_n	number of DOWs of size n	37
L_n	number of palindromic DOWs of size n	37
LCD	linear chord diagram	57
$M_{i,j}^k$	j^{th} MDS of macronuclear gene i with orientation k	51
ρ	function assigning a DOW to a micronuclear arrangement	54
S_n	number of strongly-irreducible DOWs of size n	40
σ_i	permutation for MDSs of a micronuclear arrangement	51
T_n	number of strongly-irreducible, palindromic DOWs of size n	40
$\mathcal{T}_n$	tangled cord of size n	46
$\mathbf{W}_w$	a braiding of the a chord diagram D_w	59

Symbol	Description	Page Reference
W_w	a word in the braid algebra C_m corresponding to the chord diagram D_w for a DOW w	60
$\mathcal{W}_{n,m}$	waffle chord diagram with groupings of n and m chords	66
v_i	a vertex from V	41
V	set of verticies of an assembly graph	41

Chapter 4

High Order Moments and Complete Monotonicity

Moments are a central feature of probability theory, and give an abstract quantification for the sparseness of points in a set. For instance, the first moment indicates where "middle" of the points occurs and the second central moment says how spread out the points are relative to that middle. As the study of moments developed, new type of moments were considered, e.g., absolute moments, standard moments, factorial moments, cumulants, etc. Often, one type of moment can be expressed recursively as the weighted sum of another type of moment using an inversion formula.

Urn models are used in combinatorics to describe the recursive behavior of placing colored balls into containers called urns. The balls may iteratively be inserted, removed, or replaced and generate a stochastic process. For example, suppose a red and blue ball are placed into a single urn and, during each round, a ball is drawn at random from the urn, the color is recorded, and placed back into the urn. After n rounds the probability of drawing a total of k red balls follows the binomial distribution Bin$(n, \frac{1}{2})$. See [72] for more urn models and their associated probability distributions.

Distributions derived from urn models, having a parameter n corresponding to the number of rounds, can be scaled so that the support of the distribution belongs to the unit interval $[0, 1]$. In the binomial case, the probabilities $P[X = k] = \binom{n}{k}p^k(1-p)^{n-k}$ defined on $k \in \{0, 1, \ldots, n-1, n\}$ can be scaled to $P[Y = \frac{k}{n}] = \binom{n}{k}p^k(1-p)^{n-k}$ for $k \in \left\{0, \frac{1}{n}, \ldots, \frac{n-1}{n}, 1\right\}$, which we call the *scaled distribution*.

In 1912, Bernstein [56] used a sequence of scaled-binomial distributions to give a constructive proof of the Stone-Weierstrass approximation theorem: given a function f which is continuous on $[0, 1]$, there is a sequence of polynomials which converge uniformly to f on $[0, 1]$. In 1921, Hausdorff showed that a sequence of moments, derived from a probability distribution defined on the interval $[0, 1]$, has the property that the sequential differences will all be negative. Indeed the n^{th} iterative differences will all be nonnegative if n is even and nonpositive if n is odd. Sequences of this type are called *completely monotonic*. Indeed, Hausdorff showed that a sequence is completely monotonic if and only if it corresponds with such a moment sequence [69, 70]. If, however, the order of the moment for a scaled distribution is fixed, then sequence c_n, formed by increasing the round size n, may no longer be completely monotonic (see Section 4.6).

This chapter considers two moment problems:

i. Given sample points from an unknown discrete random variable, determine the form of the underlying distribution and approximate the corresponding parameters to best-fit the data.

ii. Construct a completely monotonic sequence which behaves asymptotically like a moment sequence of a scaled random variable.

Sections 4.1 contains a general combinatorial, probabilistic, analytic, and statistical background to contextualize the later sections. A review of the various types of moments and their corresponding generating functions are provided in Section 4.2, and the techniques needed to convert one type of a moment into another are covered in Section 4.3.

Section 4.4 begins to address moment problem i. and contains a new method of distribution fitting by means of regression on the successive ratios of moments. We apply this new method to both simulated data from a binomial distribution and the braid index counts from Table 10. Application of the moment ratio method, shows that the sequence of probability distributions generated by the braid index counts appears to fit a sequence of Gamma distributions, indicating an asymptotic growth rate for the sequence.

Section 4.5 discuss Bernstein's Theorem and completely monotonic sequences in greater depth, preparing the discussion of moment problem ii. (above). Finally, in Section 4.6 we consider a sequence $S_n(\alpha, \beta, \gamma, \lambda, \mu)$ of multiparameter negative binomial sums with the desired monotonic properties and consider the asymptotics of this sequence when various parameters are fixed. In one case, S_n tends towards the moment sequence of a scaled Beta-Binomial distribution with a limiting case of the moments of the Beta distribution, and in another case, S_n tends towards the moment sequence of the scaled-Binomial distribution, which are the Bernstein polynomials for the exponential function.

4.1 Preliminaries

For completeness, we provide a summary of the concepts covered in this chapter, stemming from a variety of fields. Subsection 4.1.1 provides the necessary combinatorial background for Sections 4.2 and 4.3. Supporting backgrounds for the analysis and special functions topics are provided in Subsection 4.1.2, in preparation for Sections 4.4, 4.5, and 4.6. The probability content in Subsection 4.1.3 is used throughout the entire chapter, but the statistical applications are confined to Subsection 4.2.4 and 4.4.

Readers familiar with some or all of these fields may wish to skip this section or use it as reference if the need arises.

4.1.1 Counting Functions

We define the factorial function in the usual recursive way, namely

$$n! := \begin{cases} 1, & n = 0 \\ n\,(n-1)!, & n \in \mathbb{N} \end{cases} . \tag{4.1}$$

Analogously, $n^{\underline{0}} = 1$ and

$$n^{\underline{k}} := n\,(n-1)\ldots(n-k+1) \tag{4.2}$$

is called the *falling factorial* (of n of length k), and the related $n^{\overline{0}} = 1$ and

$$n^{\overline{k}} := n\,(n+1)\ldots(n+k-1) \tag{4.3}$$

is called the *rising factorial* (of n of length k). Note that the rising and falling factorial functions are related through $n^{\overline{k}} = (n+k-1)^{\underline{k}}$, e.g., $1^{\overline{k}} = k^{\underline{k}} = k!$.

The symbols $\begin{bmatrix} n \\ k \end{bmatrix}$ and $\begin{Bmatrix} n \\ k \end{Bmatrix}$ denote the Stirling numbers of the first and second kind, respectively. The Stirling numbers of the first and second kind may be defined recursive in the following way:

i. $\begin{bmatrix} n \\ k \end{bmatrix} = \begin{bmatrix} n-1 \\ k-1 \end{bmatrix} + (n-1)\begin{bmatrix} n-1 \\ k \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 0 \end{bmatrix} = 1, \quad \begin{bmatrix} 0 \\ k \end{bmatrix} = 0,$ and $\begin{bmatrix} n \\ 0 \end{bmatrix} = 0.$

ii. $\begin{Bmatrix} n \\ k \end{Bmatrix} = \begin{Bmatrix} n-1 \\ k-1 \end{Bmatrix} + k\begin{Bmatrix} n-1 \\ k \end{Bmatrix}, \quad \begin{Bmatrix} 0 \\ 0 \end{Bmatrix} = 1, \quad \begin{Bmatrix} 0 \\ k \end{Bmatrix} = 0,$ and $\begin{Bmatrix} n \\ 0 \end{Bmatrix} = 0.$

The $B_{n,k}(x_1, \ldots, x_{n-k+1})$ represents the k^{th} partial Bell polynomial given by

$$B_{n,k}(x_1, \ldots, x_{n-k+1}) = \sum_{P \in \mathcal{P}_n} \frac{n!}{j_1!\, j_2! \,\cdots\, j_{n-k+1}!} \left(\frac{x_1}{1!}\right)^{j_1} \left(\frac{x_2}{2!}\right)^{j_2} \cdots \left(\frac{x_{n-k+1}}{(n-k+1)!}\right)^{j_{n-k+1}} \tag{4.4}$$

such that $P = \{j_1, \ldots, j_{n-k+1}\}$ is an integer partion of n into k parts, i.e.,

$$\sum_{i=1}^{n-k+1} i\, j_i = n \quad \text{and} \quad \sum_{i=1}^{n-k+1} j_i = k.$$

For instance, 5 can be partitioned as

$$\begin{aligned}
5 &\implies \{0,0,0,0,1\}\\
4+1 &\implies \{1,0,0,1\}\\
3+2 &\implies \{0,1,1,0\}\\
3+1+1 &\implies \{2,0,1\}\\
2+2+1 &\implies \{1,2,0\}\\
2+1+1+1 &\implies \{3,1\}\\
1+1+1+1+1 &\implies \{5\}
\end{aligned}$$

of which two of these partitions are composed of three parts, namely $3+1+1$ and $2+2+1$. Thus,

$$\begin{aligned}
B_{5,3}(x_1,x_2,x_3) &= \frac{5!}{2!\,0!\,1!}\left(\frac{x_1}{1!}\right)^2\left(\frac{x_2}{2!}\right)^0\left(\frac{x_3}{3!}\right)^1 + \frac{5!}{1!\,2!\,0!}\left(\frac{x_1}{1!}\right)^1\left(\frac{x_2}{2!}\right)^2\left(\frac{x_3}{3!}\right)^0\\
&= 10\,{x_1}^2 x_3 + 15\,x_1{x_2}^2.
\end{aligned}$$

4.1.2 Binomial Coefficients and Special Functions

The binomial coefficient

$$\binom{n}{k} = \frac{n!}{(n-k)!\,k!} = \frac{n!/(n-k)!}{k!} = \frac{n^{\underline{k}}}{k!} \qquad (n \in \mathbb{N}, k \in [0,n]) \tag{4.5}$$

is a fundamental tool in counting, and has been a cornerstone of combinatorics, probability, and statistics for hundreds of years. Most notably, the coefficients naturally arise in the expansion of $(a+b)^n$ for $n \in \mathbb{N}$ in the aptly named Binomial Theorem:

$$(a+b)^n = \sum_{k=0}^{n}\binom{n}{k}a^k b^{n-k}\,, \qquad (n \in \mathbb{N}). \tag{4.6}$$

There are many generalizations of factorials, binomial coefficients, and the binomial theorem. For instance, the *Gamma function*

$$\Gamma(z) = \int_0^\infty t^{z-1}e^{-t}dt\,, \qquad (\Re\, z > 0), \tag{4.7}$$

also known as the *Euler integral of the second kind*, extends the factorial function in that $n! = \Gamma(n+1)$ for each $n \in \mathbb{N}$. In this fashion, the binomial coefficient generalizes to

$$\binom{x}{y} = \frac{\Gamma(x+1)}{\Gamma(x-y+1)\,\Gamma(y+1)}\,. \tag{4.8}$$

The *Beta function*, also know as the *Euler integral of the first kind*, is given by

$$B(x,y) = \int_0^1 t^{x-1}(1-t)^{y-1}\,dt, \qquad (\Re x, \Re y > 0), \tag{4.9}$$

and related to the Gamma function and binomial coefficients via

$$B(x,y) = \frac{\Gamma(x)\,\Gamma(y)}{\Gamma(x+y)} = \frac{1}{(x+y-1)\binom{x+y-2}{x-1}} \tag{4.10}$$

when $x+y \neq 1$. One generalization of the Binomial Theorem is the *Negative Binomial Theorem* [60, Ch. 1]:

$$(1-z)^{-\alpha} = \sum_{n=0}^{\infty} d_n(\alpha)\, z^n \qquad (\alpha > 0), \tag{4.11}$$

where $d_0(\alpha) = 1$ and $d_n(\alpha) = \frac{\alpha^{\overline{n}}}{n!}$ for each $n \in \mathbb{N}$ are called the *negative binomial coefficients*. The negative binomial coefficients may be expressed in terms of the Gamma and Beta functions as

$$d_n(\alpha) = \frac{\Gamma(\alpha+n)}{\Gamma(\alpha)\, n!} = \frac{1}{n\, B(\alpha, n)} \tag{4.12}$$

and inherit some of the properties of binomial coefficients, i.e., for $n \in \mathbb{N}$, $d_1(\alpha) = \alpha$, $d_n(1) = 1$,

$$n\, d_n(\alpha) = \alpha\, d_{n-1}(\alpha+1), \qquad d_n(\alpha) = d_n(\alpha+1) - d_{n-1}(\alpha+1), \tag{4.13}$$

$$\text{and} \quad d_n(\alpha+\beta) = \sum_{k=0}^{n} d_k(\alpha)\, d_{n-k}(\beta)\,. \tag{4.14}$$

Stirling's approximation for the Gamma function [60, p. 47],

$$\Gamma(z) = \sqrt{2\pi} e^{-z} z^{z-1/2} \left(1 + \mathcal{O}\left(\frac{1}{|z|}\right)\right) \quad \text{as} \quad |z| \to \infty \tag{4.15}$$

for $|\arg(z)| < \pi$, used for large $|z|$ and bounded a, b [83], gives

$$\frac{\Gamma(z+a)}{\Gamma(z+b)} = z^{a-b}\left[1 + \frac{(a-b)(a+b-1)}{2z} + \mathcal{O}(|z|^{-2})\right] \quad \text{as} \quad |z| \to \infty. \tag{4.16}$$

In particular, the asymptotic behavior in n of the negative binomial coefficient is

$$d_n(\alpha) = \frac{n^{\alpha-1}}{\Gamma(\alpha)}\left(1 + \mathcal{O}\left(\frac{1}{n}\right)\right) \quad \text{as} \quad n \to \infty \tag{4.17}$$

where $\alpha > 0$.

A *hypergeometric series* is a series $\sum c_n$ such that c_{n+1}/c_n is a rational function of n. With this convention, we consider the power series

$${}_pF_q\left[\begin{matrix} a_1, \ldots, a_p \\ b_1, \ldots, b_q \end{matrix}\,;\, z\right] := \sum_{n=0}^{\infty} \frac{a_1^{\overline{n}} \ldots a_p^{\overline{n}}}{b_1^{\overline{n}} \ldots b_q^{\overline{n}}} \frac{z^n}{n!}, \tag{4.18}$$

which we call a *generalized hypergeometric function* if the series converges. This means that $b \notin \{0\} \bigcup \mathbb{Z}^-$.

Some notable examples of hypergeometric functions are:

$$ {}_0F_0[z] = \sum_{n=0}^{\infty} \frac{z^n}{n!} = e^z \tag{4.19} $$

$$ {}_1F_0\left[\begin{matrix} a \\ - \end{matrix}\, ; z\right] = \sum_{n=0}^{\infty} \frac{a^{\overline{n}}}{n!} z^n = \sum_{n=0}^{\infty} d_n(a) z^n = (1-z)^{-a} \tag{4.20} $$

$$ {}_0F_1\left[\begin{matrix} - \\ 1/2 \end{matrix}\, ; -\frac{z^2}{4}\right] = \sum_{n=0}^{\infty} \frac{1}{(1/2)^{\overline{n}}} \frac{(-z^2/4)^n}{n!} = \sum_{n=0}^{\infty} \frac{(-1)^n z^{2n}}{(2n)!} = \cos z \tag{4.21} $$

$$ z\,{}_2F_1\left[\begin{matrix} 1,\, 1 \\ 2 \end{matrix}\, ; -z\right] = z \sum_{n=0}^{\infty} \frac{1^{\overline{n}} 1^{\overline{n}}}{2^{\overline{n}}} \frac{z^n}{n!} = \sum_{n=0}^{\infty} \frac{z^{n+1}}{n+1} = \log(1+z)\,. \tag{4.22} $$

4.1.3 Random Variables and Expected Value

In this section we provide the probabilistic and statistical background for the chapter. These definitions are taken from [71, 81], but see [80, 81] for introduction to these topics.

A *σ-field* is a collection $\mathcal{S}$ of subsets of a set Ω that contains the empty set, $\emptyset$, as a member and is closed under countable unions and complements. A *probability measure* $P : \Omega \to [0, 1]$ on a σ-field $\mathcal{S}$ of subsets of Ω is a function such that $P(\Omega) = 1$ and the probability measure of a countable union of disjoint sets $\{E_i\}$ is equal to $\sum P(E_i)$. A *sample space* is the pair $(\Omega, \mathcal{S})$ and a *probability space* is a triple $(\Omega, \mathcal{S}, P)$, where Ω is a set, $\mathcal{S}$ is σ-field, and P is a probability measure on $\mathcal{P}$ on $\mathcal{S}$. Informally, Ω is set of possible outcomes of an event, $\mathcal{S}$ is the collection of all combinations of these outcomes, and $\mathcal{P}$ assigns the likelihood of each outcome.

A *random variable* (RV) for a sample space $(\Omega, \mathcal{S})$ is a finite, single-valued function $X : \Omega \to \mathbb{R}$ if the inverse images under X of all Borel sets in $\mathbb{R}$ are events, i.e.,

$$ X^{-1}(B) = \{\omega : X(\omega) \in B\} \in \mathcal{S} \qquad \forall B \in \mathfrak{B}\,. \tag{4.23} $$

In particular, X is a RV if and only if for each $x \in \mathbb{R}$, $\{\omega : X(\omega) \leq x\} = \{X \leq x\} \in \mathcal{S}$ [81]. The RV X, defined on the probability space $(\Omega, \mathcal{S}, P)$ induces a probability space $(\mathbb{R}, \mathfrak{B}, Q)$ by means of the correspondence

$$ Q(B) = P\{X^{-1}(B)\} = P\{\omega : X(\omega) \in B\} \qquad B \in \mathfrak{B}, \tag{4.24} $$

and Q is called the (*probability*) *distribution* of X. The function

$$ F(x) = Q(-\infty, x] = P\{\omega : X(\omega) \leq x\} \qquad x \in \mathbb{R} \tag{4.25} $$

is called the *cumulative distribution function* (CDF) or *distribution function* (DF) of the RV X.

A RV X defined on $(\Omega, \mathcal{S}, P)$ is said to be *discrete* if there exits a countable set $E \subseteq \mathbb{R}$ such that $P\{X \in E\} = 1$. For a discrete RV X, the collection of numbers $\{p_k\}$ satisfying $P\{X = x_k\} = p_k \geq 0$ for all i and $\sum_{k=1}^{\infty} p_k = 1$, is called the *probability mass function* (PMF) of the RV X, and the corresponding set $\{x_k\}$ is the *support* of X.

A RV X defined on $(\Omega, \mathcal{S}, P)$ with DF F is said to be *continuous* if F is absolutely continuous, i.e., there exists a function $f : \mathbb{R} \to [0, \infty)$ such that for every $x \in \mathbb{R}$,

$$F(x) = \int_{-\infty}^{x} f(t) dt \,. \tag{4.26}$$

Here the function f is called the *probability density function* (PDF) of the RV X.

A sequence $\{X_n\}$ of RVs is said to be *independent, identically distributed* (i.i.d. for short) if for every $n \geq 2$ the RVs $X_1, X_2, \ldots, X_n$ are independent with the common law $\mathcal{L}(X)$.

Let X be a discrete RV X with support $\{x_k\}$ and PMF $\{p_k\}$. If the sum $\sum\limits_{k \in \Omega} |x_k| \, p_k < \infty$, we say that the *expected value* exists and set

$$\mathbb{E}[X] = \sum_{k \in \Omega} x_k \, p_k \,. \tag{4.27}$$

Similarly if X is a continuous RV with PDF f, support Ω, and $\int_{\Omega} |x| f(x) dx < \infty$, then the expectation exists and is given by

$$\mathbb{E}[X] = \int_{x \in \Omega} x \, f(x) \, dx \,. \tag{4.28}$$

Let h be a Borel-measurable function on Ω. Similarly we say that $\mathbb{E}[g(x)]$ exists if $\int_{x \in \Omega} |g(x)| \, f(x) \, dx < \infty$, in which case

$$\mathbb{E}[g(X)] = \int_{x \in \Omega} g(x) \, f(x) \, dx \,. \tag{4.29}$$

For both the discrete and continuous type RVs, the expected value has the following properties:

i. **(Constants)** If c is a constant, then $\mathbb{E}[c] = c$.

ii. **(Linearity)** If X and Y are RVs and c is a constant, then $\mathbb{E}[c\, X + Y] = c\, \mathbb{E}[X] + \mathbb{E}[Y]$.

iii. **(Inequality)** $|\mathbb{E}[X]| \leq \mathbb{E}[|X|]$.

iv. If X and Y are independent, then $\mathbb{E}[X\, Y] = \mathbb{E}[X]\, \mathbb{E}[Y]$.

4.2 Probability Generating Functions and Moments

Given a sequence $\{a_n\}_{n=0}^{\infty}$, the *ordinary generating function*, or simply *generating function* (GF), of a_n is the formal power series

$$G(a_n; x) = \sum_{n=0}^{\infty} a_n x^n \,, \tag{4.30}$$

and if $P\{k = n\} = a_n$ for all $n \in \mathbb{N}_0$ is the PMF of a discrete RV, then $G(a_n; x)$ is called the *probability generating function* (PGF) of a_n. The coefficients of a GF can be extracted by taking successive derivatives,

$$\frac{d^m}{dx^m} \left[G(a_n; x)\right]_{x=0} = \left[\sum_{n=m}^{\infty} n^{\underline{m}}\, a_n\, x^{n-m}\right]_{x=0} = n!\, a_n \,. \tag{4.31}$$

For this reason, it is often more convenient to consider the *exponential generating function* (EGF), expressed as

$$EG(a_n; x) = \sum_{n=0}^{\infty} a_n \frac{x^n}{n!} \tag{4.32}$$

for which $\frac{d^m}{dx^m} \left[EG(a_n; x)\right]_{x=0} = a_m$. GFs can be useful in counting, expressing recursive relations, establishing identities, exploring the growth and asymptotics of sequences [53, 85].

4.2.1 Moments

The *moment generating function* (MGF) for a random variable X is defined as

$$\mathrm{M}_X(t) = \mathbb{E}[e^{Xt}] \tag{4.33}$$

provided this expectation exists. If the moment generating function exists in an open interval containing zero, then

$$\mathrm{M}_X(t) = \mathbb{E}[e^{Xt}] = \mathbb{E}\left[\sum_{i=0}^{\infty} \frac{(Xt)^i}{i\,!}\right] = 1 + \sum_{i=1}^{\infty} \mathbb{E}[X^i]\, \frac{t^i}{i\,!}. \tag{4.34}$$

by the Taylor series expansion of the exponential function and the linearity of expectation. From (4.34) it is clear that the MGF is indeed an exponential generating function. For this reason the i^{th} coefficient of the MGF, $\mathbb{E}[X^i]$, is called the i^{th} *moment* and denoted as μ_i'. It should be noted that moments are sometimes called ordinary, raw, crude [72] or non-central moments.

As is a feature of exponential generating functions, the coefficients of (4.34) may be computed through repeated differentiation and evaluation at $t = 0$. In particular,

$$\frac{d^i}{dt^i} [\mathrm{M}_X(t)]_{t=0} = \mathbb{E}[X^i] = \mu_i'. \tag{4.35}$$

Most notably, the mean and variance of X are of great statistical significance and may be evaluated in terms of moments as

$$\mu = \mathbb{E}[X] = \mu_1' \quad \text{and} \quad \mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \mu_2' - (\mu_1')^2, \tag{4.36}$$

respectively. However, the *variance* of X is defined as $\mathbb{E}[(X-\mu)^2]$ which suggests the notion of central moments.

4.2.2 Central and Standardized Moments

The *central moment generating function* (CMGF) for a random variable X is defined as

$$\mathrm{C}_X(t) = \mathbb{E}[e^{(X-\mu)t}] = e^{-\mu\, t} M_X(t) \tag{4.37}$$

provided μ and the expectation exist. Similar to (4.34), when the central moment generating function exists in an open interval around μ the CMGF may be presented as exponential generating function whose coefficients are called *central moments* and denoted μ_i. In general

$$\mu_i = \mathbb{E}[(X-\mu)^i] = \mathbb{E}\left[\sum_{j=0}^{i} \binom{i}{j} (-\mu)^{i-j} X^j\right] = \sum_{j=0}^{i} \binom{i}{j} (-\mu)^{i-j} \mu_j' \tag{4.38}$$

when such expectation exists. Note that if μ exists, then $\mu_1 = \mathbb{E}[X-\mu] = \mathbb{E}[X] - \mu = 0$. The second central moment is given the distinction of being called the *variance*, usually denoted σ^2, and the square root of the variance is called the *standard deviation*, denoted by σ.

The i^{th} *standard moment*, denoted α_i is defined as $\alpha_i = \mu_i \,/\, \sigma^i$. It is easy to see that $\alpha_1 = \mu_1 = 0$ and $\alpha_2 = 1$. The third and fourth standard moment are less trivial and are called the *skewness* and *kurtosis*, respectively. Standard moments are closely connected with cumulants, and this connection is made explicit in 4.2.3.

Central and standard moments are useful in describing some visual aspects of a probability distribution. Just as the mean describes the "center" of a distribution, the standard deviation and variance quantify how "spread out" it is, the skewness measures how asymmetric or "lopsided" it is, and the kurtosis measures how peaked or "pointy" it is. The abstractness of the corresponding shape parameter increases with the order of the moment in the same way that the graphical properties of the higher order derivatives become increasingly abstract. For this reason, moments greater than order 4, known as *higher (order) moments*, are often omitted for statistical considerations.

Table 12: List of probabilistic generating functions. [71]

Generating Function	Notation	Definition
Probability Generating Function	$G(t)$	$\mathbb{E}[t^X]$
Moment Generating Function	$M(t)$	$\mathbb{E}[e^{tX}]$
Central Moment Generating Function	$C(t)$	$\mathbb{E}[e^{t(X-\mu)}]$
Characteristic Function	$\phi(t)$	$\mathbb{E}[e^{itX}]$
Factorial Moment Generating Function	$\widetilde{M}(t)$	$\mathbb{E}[(t+1)^X]$
Cumulant Generating Function	$K(t)$	$\ln \mathbb{E}[e^{tX}]$
Factorial Cumulant Generating Function	$\widetilde{K}(t)$	$\ln \mathbb{E}[(t+1)^X]$

4.2.3 Factorial Moments, Cumulants, and Factorial Cumulants

The *factorial moment generating function* (FMGF) for a RV is defined as

$$\widetilde{M}_X(t) = \mathbb{E}[(t+1)^X] = 1 + \mathbb{E}\left[\sum_{i=1}^{\infty} X^{\underline{i}} \frac{t^i}{i!}\right] \tag{4.39}$$

and provided the expectation exists, the coefficients $\mu'_{[i]} = \mathbb{E}[X^{\underline{i}}]$ are called *factorial moments*. Similarly, the *cumulant generating function* (CGF) and *factorial cumulant generating function* for a RV are defined as

$$K_X(t) = \ln \mathbb{E}[(t+1)^X] \qquad \text{and} \qquad \widetilde{K}_X(t) = \ln \mathbb{E}[e^{tX}]\,, \tag{4.40}$$

and the i^{th} coefficients of their respective series expansions are called *cumulants*, κ_i, and *factorial cumulants*, $\kappa_{[i]}$, respectively. The connection between cumulants and moments is expressed through the following relationship [82]:

$$\mu'_r = \sum_{i=0}^{r-1} \binom{r-1}{i} \kappa_{r-i}\, \mu'_i \qquad \text{and} \qquad \kappa_r = \mu'_r - \sum_{i=1}^{r-1} \binom{r-1}{i} \kappa_{r-i}\, \mu'_i\,. \tag{4.41}$$

4.2.4 Sample Moments

Let $X_1, X_2, \ldots, X_n$ be an i.i.d. sample from a population DF F. The *sample* (or *empirical*) *distribution function* is defined as

$$F_n^*(t) = \frac{1}{n} \sum_{j=1}^{n} \chi(t; X_j) \tag{4.42}$$

where $\chi(t; X_j)$ is the step function at $X_j \in \mathbb{R}$ given by

$$\chi(t; X_j) = \begin{cases} 1, & t \geq X_j \\ 0, & t < X_j \end{cases}. \tag{4.43}$$

As one might suspect, the expectation of the sample distribution $F_n^*(t)$ is the population DF F, $\mathbb{E}[F_n^*(t)] = F(t)$. In fact the Clivenko-Cantelli Theorem [63, p. 391] gives the stronger statement that $F_n^*(t)$ converges uniformly to $F(x)$, so $\mathbb{E}(X_i) = \mathbb{E}(X)$ for each $i \in [n]$. By the independence of the sample, for each $i, j \in [n]$

$$\mathbb{E}[X_i\, X_j] = \mathbb{E}[X_i]\, \mathbb{E}[X_j] = \mathbb{E}[X^2] \tag{4.44}$$

and motivates the following definition.

DEFINITION 4.2.1. For an i.i.d. sample $X_1, \ldots, X_n$ taken from a population X, define

$$\hat{\mu}_i = \frac{1}{n} \sum_{j=1}^{n} X_j^i \qquad \text{and} \qquad \hat{\mu}_1 = \frac{1}{n} \sum_{j=1}^{n} X_j \tag{4.45}$$

as the *sample moment of order* i and the *sample mean*, respectively.

From (4.44), it is straightforward to see that if $X_1, X_2, \ldots, X_n$ are i.i.d., then $\mathbb{E}[m_i] = \mu_i'$, provided μ_i' exists. Analogously, a *sample factorial moment* is given as

$$m_{[i]} = \frac{1}{n} \sum_{j=1}^{n} X_j^{\underline{i}} \tag{4.46}$$

THEOREM 4.1. [81, p. 334] Let $\{X_j\}_{j=1}^n$ be a sample of n i.i.d. RVs from a population X. The sequence $\{\sum_{j=1}^n {X_j}^i/n\}$ is asymptotically normal with mean $\mu_i' = \mathbb{E}[X^i]$ and variance $\text{Var}(X^i)/n$, i.e., as $n \to \infty$

$$\frac{m_{i,n} - \mu_i'}{\sqrt{\text{Var}(X^i)/n}} \xrightarrow{L} \mathcal{N}(0, 1) \tag{4.47}$$

where $m_{i,n}$ is the sample moment of order i from $\{X_j\}_{j=1}^n$ and $\mathcal{N}(0, 1)$ is the standard normal distribution.

Hence, the moments of a sample distribution approach the moments of the underlying population distribution as the sample size increases. It should be noted that for fixed n, as $i \to \infty$ the variance $\text{Var}(X^i)/n \to \infty$, provided the $\text{Var}(X^i)$ exists for each $i \in \mathbb{N}$.

COROLLARY 4.1. If $\{X_j\}_{j=1}^n$ is an i.i.d. sample of RVs from a population X, the sequence $\{\sum_{j=1}^n X_j^{\underline{i}}/n\}$ is asymptotically normal with mean $\mu_{[i]}'$ and variance $\text{Var}(X^{\underline{i}})/n$.

4.3 Inversion Formulas

In general, a closed formula for the raw moments, factorial moments, or cumulants may not exist, and the existence of a closed form for one does not imply a closed form for the others. However, we show that the full existence of one, implies the full existence of rest. Moreover, one can express each type of moment as a weighted sum of the other types.

For any constant μ,

$$v_n = \sum_{k=0}^{n} \binom{n}{k} \mu^{n-k} u_k \quad (\forall n) \quad \Longleftrightarrow \quad u_n = \sum_{k=0}^{n} (-\mu)^{n-k} \binom{n}{k} v_k \quad (\forall n) \tag{4.48}$$

$$v_n = \sum_{k=0}^{n} \left\{ {n \atop k} \right\} u_k \quad (\forall n) \quad \Longleftrightarrow \quad u_n = \sum_{k=0}^{n} \left[{n \atop k} \right] v_k \quad (\forall n) \tag{4.49}$$

COROLLARY 4.2. For $n \in \mathbb{N}$,

i. $\mu'_n = \sum_{j=1}^{n} \left[{n \atop j} \right] \mu'_{[j]}$ and $\mu'_{[n]} = \sum_{j=1}^{n} \left\{ {n \atop j} \right\} \mu'_j$

ii. $\kappa_n = \sum_{j=1}^{n} \left[{n \atop j} \right] \kappa_{[j]}$ and $\kappa_{[n]} = \sum_{j=1}^{n} \left\{ {n \atop j} \right\} \kappa_j$

$$v_n = \sum_{k=1}^{n} B_{n,k}(u_1, \ldots, u_{n-k+1}) \quad (\forall n) \Longleftrightarrow u_n = \sum_{k=1}^{n} (-1)^{k-1} (k-1)!\, B_{n,k}(v_1, \ldots, v_{n-k+1}) \quad (\forall n) \tag{4.50}$$

THEOREM 4.2. [55] Let $\{u_j : j = 1, 2, \ldots\}$ be a sequence satisfying $\sum_{j=1}^{\infty} |u_j|/j! < \infty$. Then

$$\exp\left(\sum_{j=1}^{\infty} u_j \frac{z^j}{j!} \right) = \sum_{k=0}^{\infty} v_j \frac{z^k}{k!} \quad (\forall\ |z| \leq 1), \tag{4.51}$$

where $v_0 = 1$ and for $k = 1, 2, \ldots,$

$$v_j = \sum_{j=1}^{k} B_{n,k}(u_1, \ldots, u_{k-j+1}).$$

This allows us to establish a direct correspondence between raw moments via cumulants and factorial moments via factorial cumulants, and vice versa.

COROLLARY 4.3. For $n > 0$,

i. $\mu'_n = \sum\limits_{j=1}^{n} B_{n,k}\left(\kappa_1, \ldots, \kappa_{n-k+1}\right)$ and $\kappa_n = \sum\limits_{j=1}^{n} (-1)^{k-1}(k-1)!\, B_{n,k}\left(\mu'_1, \ldots, \mu'_{n-k+1}\right)$

ii. $\mu'_{[n]} = \sum\limits_{j=1}^{n} B_{n,k}\left(\kappa_{[1]}, \ldots, \kappa_{[n-k+1]}\right)$ and $\kappa_{[n]} = \sum\limits_{j=1}^{n} (-1)^{k-1}(k-1)!\, B_{n,k}\left(\mu'_{[1]}, \ldots, \mu'_{[n-k+1]}\right)$

THEOREM 4.3. Let $X \sim \text{Bin}(n,p)$ be a RV with a *binomial* PMF

$$P\{X=k\} = \binom{n}{k} p^k (1-p)^{n-k} \tag{4.52}$$

such that $0 \le p \le 1$, $n \in \mathbb{N}_0$ and $k \in [0,n]$. Then the factorial moments and factorial cumulants are given by

$$\mu'_{[i]} = p^i n^{\underline{i}} \qquad \text{and} \qquad \kappa_{[i]} = (-1)^{i-1} n\, p^i (i-1)!\,, \tag{4.53}$$

respectively.

Proof. Let $\widetilde{M}_{\mathbf{B}}(t;n,p)$ be the FMGF for the binomial distribution. From the binomial theorem,

$$\widetilde{M}_{\mathbf{B}}(t;n,p) = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k}(1+t)^k = [(1+t)\,p + (1-p)]^n,$$

which means that for each $i \in \mathbb{N}$,

$$\mu'_{[i]} = \left\{ \frac{d^i}{dt^i}\left[(p\,t+1)^n\right] \right\}_{t=0} = \left\{ n^{\underline{i}}\,(p\,t+1)^{n-i}\; p^i \right\}_{t=0} = n^{\underline{i}} p^i\,.$$

Now consider that the FCGF, $\widetilde{K}_{\mathbf{B}}(t;n,p) = \ln \mathbb{E}[(t+1)^X]$, i.e.,

$$\widetilde{K}_{\mathbf{B}}(t;n,p) = \ln \widetilde{M}_{\mathbf{B}}(t;n,p) = n\,\ln(p\,t+1),$$

and for each $i \in \mathbb{N}$,

$$\kappa_{[i]} = n\left\{ \frac{d^i}{dt^i}\left[\ln(p\,t+1)\right] \right\}_{t=0} = n\left\{ (-1)^{\underline{i-1}} \left(\frac{p}{p\,t+1}\right)^i \right\}_{t=0} = (-1)^{i-1}(i-1)!\,n\,p^i\,. \qquad \square$$

COROLLARY 4.4. Let $X \sim \text{Bin}(n,p)$. Then for $i \in \mathbb{N}$,

i. $\mu'_i = \sum\limits_{k=1}^{i} \left\{ {i \atop k} \right\} n^{\underline{k}}\, p^k,$

ii. $\kappa_i = \sum\limits_{k=1}^{i} \left\{ {i \atop k} \right\} (-1)^{k-1} n\, p^k (k-1)!,$

iii. $\mu_i = (-n\,p)^i + \sum\limits_{j=1}^{i}\sum\limits_{k=1}^{j} \binom{i}{j} (-n\,p)^{i-j} \left\{ {j \atop k} \right\} n^{\underline{k}}\, p^k,$

iv. $\alpha_i = -\left(\frac{1}{1-p}\right)^{-i/2}\left((-n\,p)^{i-1} + \sum_{j=1}^{i}\sum_{k=1}^{j}\binom{i}{j}(-n\,p)^{i-j-1}\left\{ {j \atop k} \right\} n^{\underline{k}}\,p^k\right),$

Proof. Statements i. and ii. follow immediately by applying Corollary 4.2 to Theorem 4.3. Note that $\mu_0 = 1$ for all distributions and $\mu_1 = \mu_1' = n\,p$. The raw moments μ_i may be expressed as a sum of the weighted factorial moments $\mu_{[j]}'$ with $j \in [i]$ through the binomial theorem (4.38). Finally, the standard moments may be computed from iii. by using the definition in (4.36) for the variance. □

COROLLARY 4.5. Let X be a RV with a *beta-binomial* PMF

$$P\{X = k\} = \binom{n}{k}\frac{B(k+\alpha, n-k+\beta)}{B(\alpha,\beta)} \tag{4.54}$$

such that $\alpha, \beta > 0$, $n \in \mathbb{N}_0$ and $k \in [n]$. Then

$$\mu_{[i]}' = \frac{\alpha^{\overline{i}}}{(\alpha+\beta)^{\overline{i}}}\,n^{\underline{k}} \qquad \text{and} \qquad \mu_i' = \sum_{k=1}^{i}\left\{ {i \atop k} \right\}\frac{\alpha^{\overline{k}}}{(\alpha+\beta)^{\overline{k}}}n^{\underline{k}}. \tag{4.55}$$

Proof. The FMGF for the beta-binomial distribution is

$$\widetilde{M}_{\mathrm{BB}}(t;\alpha,\beta,n) = {}_2F_1\left[\begin{matrix}-n,\ \alpha\\ \alpha+\beta\end{matrix};\,-t\right] = \sum_{k=0}^{\infty}\frac{(-n)^{\overline{k}}\,\alpha^{\overline{k}}}{(\alpha+\beta)^{\overline{k}}}\frac{(-t)^k}{k!} = \sum_{k=0}^{n}\frac{n^{\underline{k}}\,\alpha^{\overline{k}}}{(\alpha+\beta)^{\overline{k}}}\frac{t^k}{k!}$$

which is an EGF as in (4.32), so $\mu_{[i]}' = \frac{\alpha^{\overline{i}}}{(\alpha+\beta)^{\overline{i}}}n^{\underline{k}}$. □

REMARK 4.3.1. The beta-binomial distribution is also called the *negative* (*inverse*) *hypergeometric* distribution or the *hypergeometric waiting-time* distribution [71].

4.4 Fitting of Probability Distributions and Estimation of Parameters

In the proceeding section, several type of moments are defined along with the inversion formulas to transform each type of moment into another. This section highlights the application of moments for distribution fitting and parameter estimation. In particular, we present a new method for distribution fitting which involves properties of the sequential ratios of moments.

4.4.1 General Hypergeometric Distributions

The *general hypergeometric distributions* (GHD) have a PMF of

$$\Pr[X = x] = \frac{\binom{a}{x}\binom{b}{n-x}}{\binom{a+b}{n}}\,, \tag{4.56}$$

a PGF and FMGF of the forms

$$G(z) = \frac{{}_2F_1\left[\begin{matrix}-n,\ -a\\ b-n+1\end{matrix}\,;\,z\right]}{{}_2F_1\left[\begin{matrix}-n,\ -a\\ b-n+1\end{matrix}\,;\,1\right]} \qquad \text{and} \qquad \widetilde{M}_X(t) = {}_2F_1\left[\begin{matrix}-a,\ -n\\ -a-\ b\end{matrix}\,;\,-t\right], \tag{4.57}$$

respectively. Hence, when the factorial moments exist

$$\mu'_{[i]} = \frac{n!\,a!\,(a+b-i)!}{(n-i)!\,(a-i)!\,(a+b)!}\,, \tag{4.58}$$

yielding

$$\xi'_{[i]} = \frac{(a-i)(n-i)}{(a+b)-i} \tag{4.59}$$

4.4.2 Generalized Hypergeometric Factorial Moment Distributions

The *Generalized Hypergeometric Factorial Moment Distributions* (GHFMD) is a class of discrete distributions for which the PGFs have the form

$$G(z) = {}_pF_q\left[\begin{matrix}a_1,\ \ldots,\ a_p\\ b_1,\ \ldots,\ b_q\end{matrix}\,;\,\lambda(z-1)\right] \tag{4.60}$$

[74], which means their factorial moments are generated by

$$G(1+t) = {}_pF_q\left[\begin{matrix}a_1,\ \ldots,\ a_p\\ b_1,\ \ldots,\ b_q\end{matrix}\,;\,\lambda\,t\right] \tag{4.61}$$

and [71]

$$\mu'_{[i]} = \frac{(a_1+i-1)!\ldots(a_p+i-1)!\,(b_1-1)!\ldots(b_q-1)!}{(a_1-1)!\ldots(a_p-1)!\,(b_1+i-1)!\ldots(b_q+i-1)!}\,\lambda^i\,, \tag{4.62}$$

so from (4.60), the PMFs of the GHFMD class are given by

$$P[X=x] = \frac{\lambda^x\,\prod_{k=1}^p \Gamma(a_k+x)/\Gamma(a_k)}{x!\prod_{k=1}^q \Gamma(b_k+x)/\Gamma(b_k)}\ {}_pF_q\left[\begin{matrix}a_1+x,\ \ldots,\ a_p+x\\ b_1+x,\ \ldots,\ b_q+x\end{matrix}\,;\,-\lambda\right]. \tag{4.63}$$

Many classic discrete distribution belong to this class, such as the Binomial, Negative Binomial, Hypergeometric, Negative Hypergeometric, and Poisson (see Table 17).

In view of (4.62), the successive ratios of factorial moments are rational functions in i :

$$\xi'_{[i]} = \frac{\mu'_{[i+1]}}{\mu'_{[i]}} = \lambda\,\frac{(i+a_1)\ldots(i+a_p)}{(i+b_1)\ldots(i+b_q)}\,. \tag{4.64}$$

Consider an i.i.d. sample $\{X_1,\ldots,X_n\}$ from a GHFMD X. According to Corollary 4.1 if n is large and i is small, $\hat{\xi'}_{[i]} = \frac{\hat{\mu}'_{[i+1]}}{\hat{\mu}'_{[i]}}$ is a good estimator for $\xi'_{[i]}$. Hence the model

$$y = \frac{\beta_0+\beta_1\,x+\cdots+\beta_p\,x^p}{1+\beta_{p+1}\,x+\cdots+\beta_{p+q}\,x^q} \tag{4.65}$$

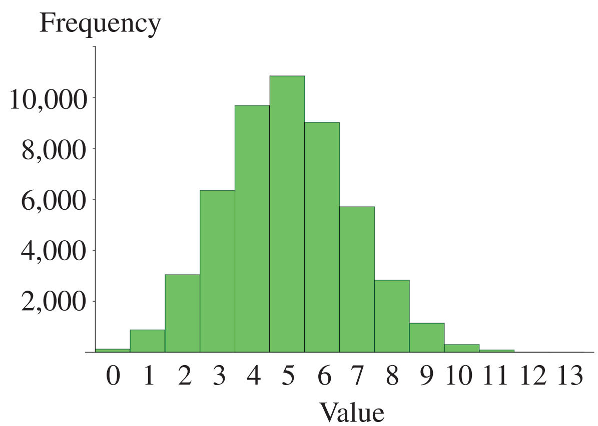
should be a reasonable fit for the points $(1, \xi'_{[1]}), (2, \xi'_{[2]}), \ldots, (m, \xi'_{[m]})$. If $m \geq p + q + 1$, then linear least-squares regression for the model

$$y = \beta_0 + \beta_1\, x + \cdots + \beta_p\, x^p - \beta_{p+1}\, xy - \cdots - \beta_{p+q} x^p y \tag{4.66}$$

returns the best-fit values $\hat{\boldsymbol{\beta}} = \{\hat{\beta}_0, \beta_1, \ldots, \hat{\beta}_{p+q}\}$. Using the values $\hat{\boldsymbol{\beta}}$, the numerator and denominator of (4.64) can be factored, giving approximations for $a_1, \ldots, a_p$ and $b_1, \ldots, b_q$ and λ.

If either $\beta_0 + \beta_1\, x + \cdots + \beta_p\, x^p$ or $1 + \beta_{p+1}\, x + \cdots + \beta_{p+q}\, x^q$ contains imaginary roots, then it is unlikely that the sample is generated from a distribution in the GHFMD class.

EXAMPLE 4.4.1. A sample of 50,000 sample points were generated from a RV $X \sim \mathrm{Bin}(15, \frac{1}{3})$ using Wolfram's Mathematica© version 9.0. Figure 28 shows a histogram of the frequencies for each of the values in the sample data.



Value	Frequency
0	125
1	877
2	3,043
3	6,345
4	9,674
5	10,841
6	9,014

Value	Frequency
7	5,707
8	2,830
9	1,141
10	302
11	90
12	9
13	2

Figure 27.: Frequencies of the values from the 50,000 simulated $\mathrm{Bin}\big(15, \frac{1}{3}\big)$ sample points.

The sample factorial moments can be easily calculated from the frequency table in Figure 28 as

$$\hat{\mu}'_{[i]} = \frac{1}{50,000}\left(125 \cdot 0^{\underline{i}} + 877 \cdot 1^{\underline{i}} + 3,043 \cdot 2^{\underline{i}} + \cdots + 2 \cdot 13^{\underline{i}}\right) , \tag{4.67}$$

from which the ratios of the sample factorial moments may be computed. In this case the few ratios are $\{(1, 4.9997), (2, 4.66584), (3, 4.32844), (4, 3.99174), (5, 3.65845), \ldots\}$, which when plotted (see Figure 28) form a line. Hence (4.65) simply becomes

$$y = \beta_0 + \beta_1\, x, \tag{4.68}$$

which requires at least two distinct sample points for a best-fit estimate. From Corollary (4.1), we know that the variance of $\hat{\mu}'_{[i]}$ increases as i increases, so we find the best-fit line for the ratios of sample factorial moments using the smallest order moments available, namely $\hat{\xi}'_{[1]} = 4.9997$ and $\hat{\xi}'_{[2]} = 4.66584$. The passing through $(1, \hat{\xi}'_{[1]})$ and $(2, \hat{\xi}'_{[2]})$ is

$$y = -0.33386\, x + 5.33356 = -0.33386\,(x - 15.9754) \tag{4.69}$$

and corresponds to the PGF in (4.60) as

$${}_1F_0\left[\begin{matrix}-15.9754 \\ -\end{matrix}\,;\,-0.33386(z-1)\right] \approx {}_1F_0\left[\begin{matrix}-16 \\ -\end{matrix}\,;\,\frac{1}{3}(1-z)\right] = G_X(z) \tag{4.70}$$

which gives $X \sim \mathrm{Bin}\big(16, \frac{1}{3}\big)$.

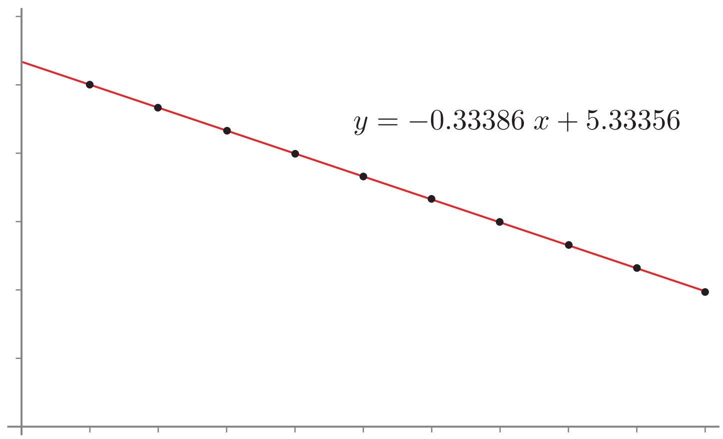


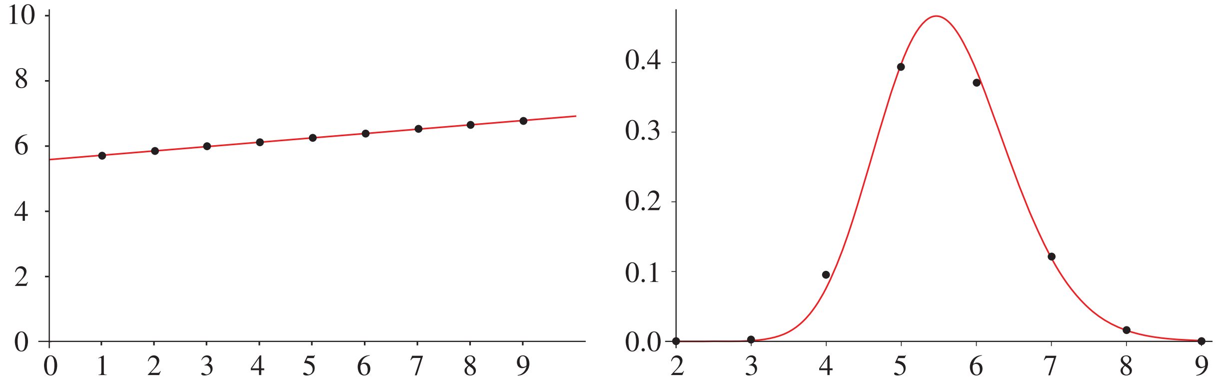
Figure 28.: Plot of i versus the ratios of the sample factorial moments $\xi_{[i]}$ for $i = 1, \ldots, 10$. The best-fit linear least-squares line is shown in red.

EXAMPLE 4.4.2. For a fixed number of chords, the frequency of the linear chord diagrams having with a particular braid index, given in Table 10 of Chapter 2, divided by the total number of linear chord diagrams is a PMF for some RV X. If we apply the same technique used in Example 4.4.1, the numerator and denominator for (4.65) contain complex roots unless $p = q = 1$. Plotting the best-fit distribution with $p = q = 1$ shows that X does not belong to the class of GHFMD.

However, the plot of the successive ratio of raw moments $\hat{\boldsymbol{\xi}}' = \{(1, \hat{\xi}'_1), \ldots, (1, \hat{\xi}'_m)\}$ shows a linear correspondence (see Figure 29a). Noting that

$$\xi'_i = \frac{\mu'_{i+1}}{\mu'_i} = \frac{\beta^{i+1}\,\alpha^{\overline{i+1}}}{\beta^i\,\alpha^{\overline{i}}} = \beta\,(\alpha + i) = \beta\, i + \alpha\beta \tag{4.71}$$

for the distribution Gamma(α, β) for $\alpha, \beta > 0$, it is reasonable to assume that the Gamma distribution will be a good fit for X. In this case, X is representative of the entire population, so there is no need to consider variance problem incurred using high order moments with the central limit theorem as in Corollary (4.1). Hence the full set $\boldsymbol{\xi'}$ should be used for the linear least-squares regression, which gives $\xi'_x = 0.133307x + 5.58579$ implying $\alpha = 41.9016 \approx 42$ and $\beta = 0.133307 \approx \frac{2}{15}$. The plot of X and Gamma $\left(42, \frac{2}{15}\right)$ are pictured in Figure 29b.



(a) Linear relationship for ξ'_i.

(b) Plot of X and Gamma $\left(42, \frac{2}{15}\right)$.

Figure 29.: Fitting the RV X, the braid index of linear chord diagrams with 8 chords, by examining the ratio of the successive moments.

4.5 Complete Monotonicity and Bernstein's Polynomial Theorem

4.5.1 Completely Monotonic Sequences

DEFINITION 4.5.1. A sequence $\{c_k\}_{k=0}^{\infty}$ is called *completely monotonic* if, for all $k, r \in \mathbb{N}_0 := \{0\} \bigcup \mathbb{N}$,

$$(-1)^r \Delta^r c_k \geq 0, \tag{4.72}$$

where Δ^r is the *forward difference operator* defined by the rules $\Delta^0 c_k = c_k$ and $\Delta^{r+1} c_k = \Delta^r c_{k+1} - \Delta^r c_k$.

Similar to Pascal's triangle, iterating the definition of Δ^r leads to the identity

$$\Delta^r c_k = \sum_{j=0}^{r} \binom{r}{j} (-1)^j c_{k+r-j} = \sum_{j=0}^{r} \binom{r}{j} (-1)^{r+j} c_{k+j} \tag{4.73}$$

for $k, r \in \mathbb{N}_0$ (see [84] for more details). Thus the condition that

$$\sum_{j=0}^{r} \binom{r}{j} (-1)^j c_{k+j} \geq 0 \tag{4.74}$$

for all $k, r \in \mathbb{N}_0$ is equivalent to (4.72).

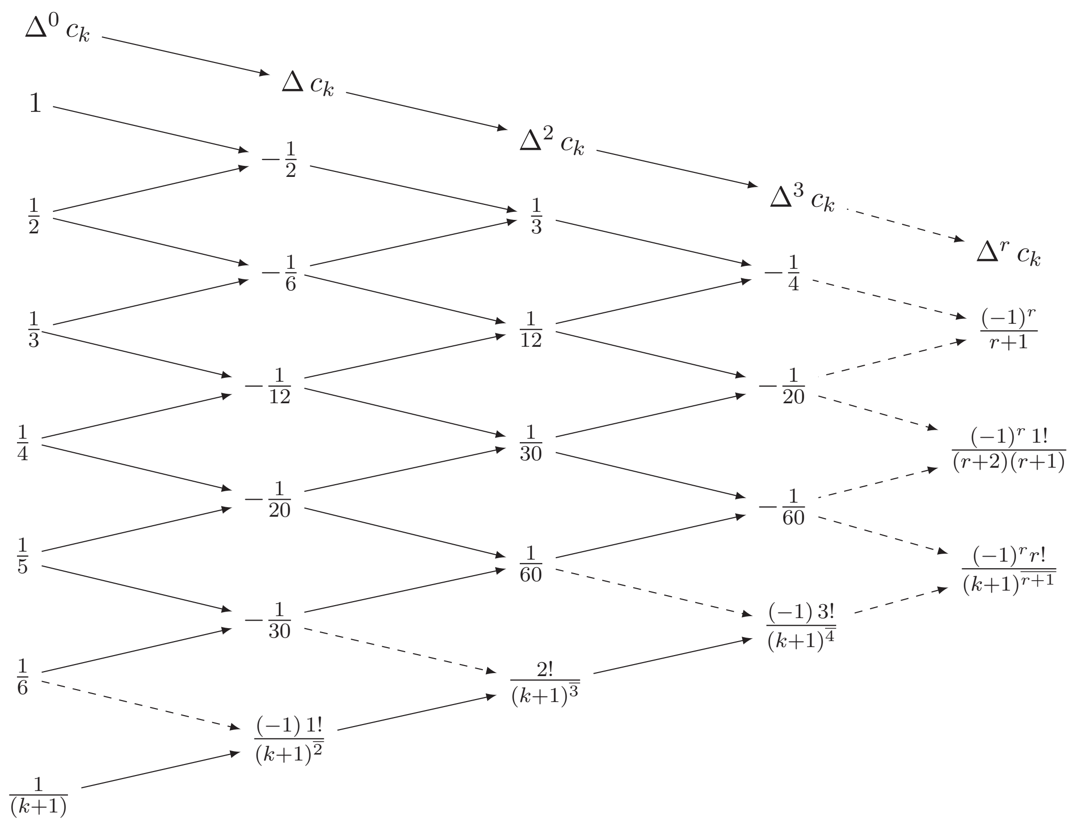


Figure 30.: Sequential differences of completely monotonic sequences alternate signs.

EXAMPLE 4.5.1. Let $\{c_k\} = \left\{\frac{1}{k+1}\right\}_{k=0}^{\infty}$. The consecutive differences between the terms are:

The sequence $\{c_k\}$ is completely monotonic since $(-1)^r \Delta^r c_k = \frac{r!}{(k+1)^{\overline{r+1}}} \geq 0$ for all $k, r \in \mathbb{N}_0$.

EXAMPLE 4.5.2. Fix $x \in (0,1)$ and consider the sequence $c_n = x^n, (n \in \mathbb{N})$. From (4.74),

$$(-1)^r \Delta^r c_n = \sum_{j=0}^{r} \binom{r}{j} (-1)^j x^{k+j} = x^k \left(\sum_{j=0}^{r} \binom{r}{j} (-1)^j x^j \right) = x^k (1-x)^r > 0 \tag{4.75}$$

since $0 < x < 1$.

Likewise, if X be a random variable with support $\Omega \subseteq [0,1]$, then by the linearity of expectation

$$(-1)^r \Delta^r \mu_k' = \sum_{j=0}^{r} \binom{r}{j} (-1)^j \, \mathbb{E}[X^{k+j}] = \mathbb{E}[X^k (1-X)^r] \geq 0\,, \tag{4.76}$$

showing that $\{\mu_k'\}$ is a completely monotonic sequence.

THEOREM 4.4. (Hausdorff's Moment Theorem, [61]) The moments c_r of a probability distribution F form a completely monotonic sequence with $c_0 = 1$. Conversely, an arbitrary completely monotonic sequence $\{c_r\}$ subject to the norming $c_0 = 1$ coincides with the moment sequence of a unique probability distribution.

Completely monotonic sequences may be produced from *completely monotonic functions* for which successive derivatives change the parity, i.e., a function f is completely monotonic if for all n, $(-1)^n f^{(n)}(x) \geq 0$ on $(0, \infty)$. It follows that $f(x) = e^{-h(x)}$ is completely monotonic if $h'(x)$ is completely monotonic, and such functions f are called *logarithmically completely monotonic*.

4.5.2 Bernstein's Relation for Polynomials

DEFINITION 4.5.2. The *Bernstein polynomial of degree* n of a function $h : [0, 1] \to \mathbb{R}$ is defined as

$$\mathbf{B}_n(h; t) = \sum_{k=0}^{n} \binom{n}{k} t^k (1-t)^{n-k}\, h\left(\frac{k}{n}\right). \tag{4.77}$$

Bernstein polynomials were originally developed to provide a constructive proof of the Stone-Weierstrass polynomial approximation theorem, but have shown to be a useful tool in a variety of other applications such as Béizer curves in computer graphics.

THEOREM 4.5. (Bernstein's Theorem, [56]) For a continuous function h on $[0, 1]$, the relation

$$\lim_{n\to\infty} \mathbf{B}_n(h; t) = h(t) \tag{4.78}$$

holds uniformly on $[0, 1]$.

COROLLARY 4.6. [64, 65] Let $h(t)$ be a continuous function on $[0, 1]$ and $\alpha, \beta > 0$, then

$$\lim_{n\to\infty} \sum_{k=0}^{n} \frac{d_k(\alpha)\, d_{n-k}(\beta)}{d_n(\alpha+\beta)}\, h\left(\frac{k}{n}\right) = \int_0^1 t^{\alpha-1}(1-t)^{\beta-1}\, h(t) \frac{dt}{B(\alpha, \beta)}. \tag{4.79}$$

In the 1950s, Bohman [57] and Korovkin [77] were able to identify sufficient properties of the Bernstein basis polynomials to ensure $B_n(h; t) \to h(t)$. In fact, it is enough to check that the Bernstein polynomials converge for the functions $h := 1$, t, and t^2 since they have the following property.

DEFINITION 4.5.3. Let L denote a linear operator that maps a function f on $[a, b]$ to a function Lf defined on $[c, d]$. Then L is said to be a *monotone* (*positive*) operator if

$$f(x) \geq g(x), \quad \forall x \in [a, b] \quad \text{implies} \quad (Lf)(x) \geq (Lg)(x), \quad \forall x \in [c, d], \tag{4.80}$$

where $(Lf)(x)$ denotes the value of Lf at $x \in [a, b]$.

THEOREM 4.6 (Monotone Operator Theorem, [57, 77]). Let $\{L_n\}$ denote a sequence of monotone linear operators that map a function $f \in C[a,b]$ to a function $L_n f \in C[a,b]$, and let $L_n f \to f$ uniformly on $[a,b]$ for $f := 1, x,$ and x^2. Then $L_n f \to f$ uniformly on $[a,b]$ for all $f \in C[a,b]$.

Proof. See [79, p. 263] □

EXAMPLE 4.5.3. Let $h(t) = t^\lambda$ for $\lambda > 0$. Then from (4.77) and (4.79),

$$\lim_{n\to\infty} \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} \left(\frac{k}{n}\right)^\lambda = p^\lambda \tag{4.81}$$

and

$$\lim_{n\to\infty} \sum_{k=0}^{n} \frac{d_k(\alpha)\, d_{n-k}(\beta)}{d_n(\alpha+\beta)} \left(\frac{k}{n}\right)^\lambda = \frac{\int_0^1 t^{\alpha+\lambda-1}(1-t)^{\beta-1} dt}{B(\alpha,\beta)} = \frac{B(\alpha+\lambda,\beta)}{B(\alpha,\beta)}\,. \tag{4.82}$$

4.6 Monotonicity and Asymptotics for a Moment Sequence of a Scaled Random Variable

According to Hausdorff's moment theorem 4.4, moment sequences of continuous distributions on $[0,1]$ are completely monotonic, and Bernstein's theorem 4.5 gives a construction of a sequence of polynomials which converge to a continuous function $h(t)$ on $[0,1]$. Taking $h(t) = t^\lambda$, as in Example 4.5.3, the sequence of Bernstein polynomials is μ'_λ/n^λ, such that μ'_λ is the λ^{th} moment of the binomial distribution with parameters n and p.

A natural question is: "For fixed λ and p, is this sequences of Bernstein polynomials completely monotonic?" However, it is straightforward to check that for $\lambda = 3$ the 10^{th} difference fails for certain p, i.e., in the notation of Definition 4.72

$$(-1)^{10}\Delta^{10} c_1 = \frac{p\,(1-p)}{304,920}\,(83,711 - 84,262\,p)\ , \tag{4.83}$$

is negative, and therefore not completely monotonic, for $p \in \left(\frac{83{,}711}{84{,}262}, 1\right)$. A similar construction for the scaled beta-binomial moment sequences, μ'_λ/n^λ shows that for $\lambda = 3$,

$$(-1)^{10}\Delta^{10} c_1 = \frac{\alpha\,\beta}{304,920(\alpha+\beta)(1+\alpha+\beta)(2+\alpha+\beta)}\,(83,160 - 511\,\alpha + 83,711\,\beta)\ , \tag{4.84}$$

and not completely monotonic when $\alpha > \frac{83{,}160}{551}$ and $0 < \beta < \frac{511\,\alpha - 83{,}160}{83711}$. In this section, we introduce a sequence which behaves asymptotically like the binomial and beta-binomial distributions while having completely monotonic properties.

4.6.1 Multiparameter Binomial Sum and Inequality

THEOREM 4.7. [64] Let $\mathbf{a} = (a_0, \ldots, a_n)$ and $\mathbf{b} = (b_0, \ldots, b_n)$ be nonzero complex vectors $(n = 1, 2, \ldots)$. Then for any numbers $\alpha, \beta > 0$, $\lambda \geq 0$, $p > 1$ $(1/p + 1/q = 1)$, and $\tau \in (0, \min(p, q)]$, the following inequality holds:

$$\begin{aligned}[d_n(\alpha+\beta+\lambda)]^{1-1/\tau} & \left[\sum_{k=0}^{n} \frac{d_{n-k}(\lambda)}{[d_k(\alpha+\beta)]^{\tau-1}} \left|\sum_{l=0}^{k} a_l b_{k-l}\right|^{\tau}\right]^{1/\tau} \\ & \leq \left[\sum_{k=0}^{n} \frac{d_{n-k}(\beta+\lambda)}{[d_k(\alpha)]^{p-1}} |a_k|^p\right]^{1/p} \cdot \left[\sum_{k=0}^{n} \frac{d_{n-k}(\alpha+\lambda)}{[d_k(\beta)]^{q-1}} |b_k|^q\right]^{1/q} . \end{aligned} \tag{4.85}$$

DEFINITION 4.6.1. Define the sequence $\{S_n\}_{n=0}^{\infty}$ such that

$$S_n(\alpha, \beta, \gamma, \lambda, \mu) = \sum_{k=0}^{n} \frac{d_k(\alpha)\, d_{n-k}(\beta)}{d_n(\alpha+\beta)} \left(\frac{d_{n-k}(k+\gamma)}{d_{n-k}(k+\gamma+\lambda)}\right)^{\mu} \tag{4.86}$$

where $\alpha, \beta, \gamma, \lambda \geq 0$, and $\alpha + \beta > 0$, $\gamma + \lambda > 0$ and $\mu \geq 0$. Some other equivalent presentations of S_n are

$$\begin{aligned} & S_n(\alpha, \beta, \gamma, \lambda, \mu) \\ & = \sum_{k=0}^{n} \frac{\binom{\alpha+k-1}{k}\binom{\beta+n-k-1}{n-k}}{\binom{\alpha+\beta+n-1}{n}} \left(\frac{\binom{\gamma+n-1}{n-k}}{\binom{\gamma+\lambda+n-1}{n-k}}\right)^{\mu} && (4.87) \\ & = \sum_{k=0}^{n} \frac{\Gamma(n+1)\,\Gamma(k+\alpha)\,\Gamma(n-k+\beta)\,\Gamma(\alpha+\beta)}{\Gamma(k+1)\,\Gamma(n-k+1)\,\Gamma(n+\alpha+\beta)\,\Gamma(\alpha)\,\Gamma(\beta)} \left(\frac{\Gamma(n+\gamma)\,\Gamma(k+\gamma+\lambda)}{\Gamma(k+\gamma)\,\Gamma(n+\gamma+\lambda)}\right)^{\mu} && (4.88) \\ & = \sum_{k=0}^{n} \binom{n}{k} \frac{B(k+\alpha, n-k+\beta)}{B(\alpha, \beta)} \left(\frac{B(n+\gamma, \lambda)}{B(k+\gamma, \lambda)}\right)^{\mu} . && (4.89) \end{aligned}$$

REMARK 4.6.1. For $\mu \in \mathbb{N}$, the sequence $S_n(\alpha, \beta, \gamma, \lambda, \mu)$ has the following hypergeomtric presentation:

$$\frac{\Gamma(n+\beta)\,\Gamma(\alpha+\beta)}{\Gamma(n+\alpha+\beta)\,\Gamma(\beta)} \left(\frac{\Gamma(n+\alpha)\,\Gamma(\gamma+\lambda)}{\Gamma(n+\gamma+\lambda)\,\Gamma(\gamma)}\right)^{\mu} {}_{\mu+2}F_{\mu+1}\left[\begin{matrix} -n,\ \alpha,\ \gamma+\lambda,\ \ldots,\ \gamma+\lambda \\ 1-(n+\beta),\ \gamma,\ \ldots,\ \gamma \end{matrix} ; 1\right] . \tag{4.90}$$

THEOREM 4.8. For every $n \in \mathbb{N}$,

$$S_n(\alpha, \beta+\lambda, \alpha, -\gamma, \mu)^{\nu} \cdot S_n(\beta, \alpha+\lambda, \beta, \gamma, \nu)^{\mu} \geq 1 , \tag{4.91}$$

where $\alpha, \beta > 0$, $\lambda \geq 0$, and $\gamma \in [\min\{0, \alpha-\beta\}, \max\{0, \alpha-\beta\}]$ and $\mu > 1$ $(1/\mu + 1/\nu = 1)$.

Proof. By definition,

$$\begin{aligned} & S_n(\alpha, \beta+\lambda, \alpha, -\gamma, \mu)^{\nu} \cdot S_n(\beta, \alpha+\lambda, \beta, \gamma, \nu)^{\mu} \\ & = \left[\sum_{k=0}^{n} \frac{d_k(\alpha)\, d_{n-k}(\beta+\lambda)}{d_n(\alpha+\beta+\lambda)} \left(\frac{d_{n-k}(k+\alpha)}{d_{n-k}(k+\alpha-\gamma)}\right)^{\mu}\right]^{\nu} \left[\sum_{k=0}^{n} \frac{d_k(\beta)\, d_{n-k}(\alpha+\lambda)}{d_n(\alpha+\beta+\lambda)} \left(\frac{d_{n-k}(k+\beta)}{d_{n-k}(k+\beta+\gamma)}\right)^{\nu}\right]^{\mu} , \end{aligned}$$

and noting that $\frac{d_{n-k}(k+\alpha)}{d_{n-k}(k+\alpha-\gamma)} = \frac{d_n(\alpha)\, d_k(\alpha-\gamma)}{d_n(\alpha-\gamma)\, d_k(\alpha)}$ we obtain

$$\begin{aligned}
&S_n(\alpha, \beta+\lambda, \alpha, -\gamma, \mu)^\nu \cdot S_n(\beta, \alpha+\lambda, \beta, \gamma, \nu)^\mu \\
&\qquad = \left(\frac{d_n(\alpha)\, d_n(\beta)}{d_n(\alpha-\gamma)\, d_n(\beta+\gamma)}\right) \left[\sum_{k=0}^n \frac{d_k(\alpha)\, d_{n-k}(\beta+\lambda)}{d_n(\alpha+\beta+\lambda)} \left(\frac{d_k(\alpha-\gamma)}{d_k(\alpha)}\right)^\mu\right]^\nu \\
&\qquad\qquad \times \left[\sum_{k=0}^n \frac{d_k(\beta)\, d_{n-k}(\alpha+\lambda)}{d_n(\alpha+\beta+\lambda)} \left(\frac{d_k(\beta+\gamma)}{d_k(\beta)}\right)^\nu\right]^\mu .
\end{aligned} \tag{4.92}$$

We claim that

$$\frac{d_n(\alpha)}{d_n(\alpha-\gamma)} \frac{d_n(\beta)}{d_n(\beta+\gamma)} \geq 1\,, \tag{4.93}$$

for each $n \in \mathbb{N}_0$. Note that equality in (4.93) holds for both $\gamma = 0$ and $\gamma = \alpha - \beta$. Put $g : (-\beta, \alpha) \to \mathbb{R}$ as

$$\begin{aligned}
g(\gamma) &= \log\left[\frac{d_n(\alpha)\, d_n(\beta)}{d_n(\alpha-\gamma)\, d_n(\beta+\gamma)}\right] \qquad (4.94)\\
&= \log\left[\frac{\Gamma(\alpha-\gamma)\,\Gamma(\beta+\gamma)}{\Gamma(n+\alpha-\gamma)\,\Gamma(n+\beta+\gamma)}\right] + \log\left[\frac{\Gamma(n+\alpha)\,\Gamma(n+\beta)}{\Gamma(\alpha)\,\Gamma(\beta)}\right],
\end{aligned}$$

and we seek values γ for which $g(\gamma) \geq 0$. Let ψ be the *digamma function* defined by $\psi(z) := \frac{d}{dz}\log\Gamma(z) = \frac{\Gamma'(z)}{\Gamma(z)}$ with an integral representation of

$$\psi(z) = -c + \int_0^1 \frac{1-t^{z-1}}{1-t}\, dt \qquad (\Re\, z > 0), \tag{4.95}$$

where c is the Euler constant [60]. Hence

$$\begin{aligned}
g'(\gamma) &= \psi(n+\alpha-\gamma) + \psi(\beta+\gamma) - \psi(n+\beta+\gamma) - \psi(\alpha-\gamma) \\
&= \int_0^1 \frac{t^{n+\alpha-\gamma-1} + t^{\beta+\gamma-1} - t^{n-\beta-\gamma-1} - t^{\alpha-\gamma-1}}{1-t}\, dt \\
&= \int_0^1 \frac{1-t^n}{1-t}\left(t^{\beta+\gamma-1} - t^{\alpha-\gamma-1}\right) dt
\end{aligned} \tag{4.96}$$

and $g'(\gamma) > 0$ when $t^{\beta+\gamma-1} - t^{\alpha-\gamma-1} > 0$, i.e., $g(\gamma)$ is increasing for $\gamma \in \left(-\beta, \frac{\alpha-\beta}{2}\right)$ and decreasing for $\gamma \in \left(\frac{\alpha-\beta}{2}, \alpha\right)$. Since $g(0) = g(\alpha-\beta) = 0$, the function $g(\gamma) \geq 0$ for $\gamma \in [0, \alpha-\beta]$ when $\alpha \geq \beta$ and $\gamma \in [\alpha-\beta, 0]$ when $\beta \geq \alpha$. This completes the claim of (4.93), meaning

$$\left[\sum_{k=0}^n \frac{d_k(\alpha)\, d_{n-k}(\beta+\lambda)}{d_n(\alpha+\beta+\lambda)} \left(\frac{d_k(\alpha-\gamma)}{d_k(\alpha)}\right)^\mu\right]^\nu \left[\sum_{k=0}^n \frac{d_k(\beta)\, d_{n-k}(\alpha+\lambda)}{d_n(\alpha+\beta+\lambda)} \left(\frac{d_k(\beta+\gamma)}{d_k(\beta)}\right)^\nu\right]^\mu \tag{4.97}$$

is less than or equal to (4.92). Taking $a_k = d_k(\alpha-\gamma)$ and $b_k = d_k(\beta+\gamma)$ in Theorem 4.7 bounds (4.97) below by 1. $\square$

4.6.2 Asymptotics of S_n

We see that, as in Example 4.5.3, S_n shares asymptotic properties with the scaled binomial distribution.

THEOREM 4.9. For $\alpha, \beta, \gamma > 0$ and $\lambda, \mu \geq 0$,

$$\lim_{n\to\infty} \sum_{k=0}^{n} \binom{n}{k} \frac{B(k+\alpha, n-k+\beta)}{B(\alpha,\beta)} \left(\frac{B(n+\gamma,\lambda)}{B(k+\gamma,\lambda)}\right)^{\mu} = \frac{B(\alpha+\lambda\,\mu,\beta)}{B(\alpha,\beta)} . \tag{4.98}$$

Proof. We claim that

$$\lim_{n\to\infty} \left| \sum_{k=0}^{n} \frac{d_k(\alpha)\, d_{n-k}(\beta)}{d_n(\alpha+\beta)} \left[\left(\frac{d_{n-k}(k+\gamma)}{d_{n-k}(k+\gamma+\lambda)}\right)^{\mu} - \left(\frac{k}{n}\right)^{\lambda\mu} \right] \right| = 0. \tag{4.99}$$

To establish this, fix $\epsilon > 0$ and consider that

$$\begin{aligned} &\left| \sum_{k=0}^{n} \frac{d_k(\alpha)\, d_{n-k}(\beta)}{d_n(\alpha+\beta)} \left[\left(\frac{d_{n-k}(k+\gamma)}{d_{n-k}(k+\gamma+\lambda)}\right)^{\mu} - \left(\frac{k}{n}\right)^{\lambda\mu} \right] \right| \\ &\qquad \leq \sum_{k=0}^{N} \frac{d_k(\alpha)\, d_{n-k}(\beta)}{d_n(\alpha+\beta)} \left| \left(\frac{d_{n-k}(k+\gamma)}{d_{n-k}(k+\gamma+\lambda)}\right)^{\mu} - \left(\frac{k}{n}\right)^{\lambda\mu} \right| \qquad (4.100) \\ &\qquad\quad + \sum_{k=N+1}^{n} \frac{d_k(\alpha)\, d_{n-k}(\beta)}{d_n(\alpha+\beta)} \left| \left(\frac{d_{n-k}(k+\gamma)}{d_{n-k}(k+\gamma+\lambda)}\right)^{\mu} - \left(\frac{k}{n}\right)^{\lambda\mu} \right| \qquad (4.101) \end{aligned}$$

for each $0 \leq N \leq n$. Given the asymptotics for the ratio of Gamma functions (4.16), large n and $k \in [N+1, n]$ yield

$$\left(\frac{d_{n-k}(k+\gamma)}{d_{n-k}(k+\gamma+\lambda)}\right)^{\mu} = \left(\frac{\Gamma(n+\gamma)}{\Gamma(n+\gamma+\lambda)} \frac{\Gamma(k+\gamma+\lambda)}{\Gamma(k+\gamma)}\right)^{\mu} = \left(\frac{k}{n}\right)^{\lambda\mu} \left(1+\mathcal{O}(k^{-1})\right) . \tag{4.102}$$

Hence N can be chosen sufficiently large so that

$$\sum_{k=N+1}^{n} \frac{d_k(\alpha)\, d_{n-k}(\beta)}{d_n(\alpha+\beta)} \left| \left(\frac{d_{n-k}(k+\gamma)}{d_{n-k}(k+\gamma+\lambda)}\right)^{\mu} - \left(\frac{k}{n}\right)^{\lambda\mu} \right| < \frac{\epsilon}{2} . \tag{4.103}$$

Now for this fixed N, consider $k \in [0, N]$ and large n. Given that $k+\gamma+\lambda \geq k+\gamma$, $n \geq k$, and $\lambda\,\mu \geq 0$, it follows that $0 \leq \left(\frac{d_{n-k}(k+\gamma)}{d_{n-k}(k+\gamma+\lambda)}\right)^{\mu} \leq 1$ and $0 \leq \left(\frac{k}{n}\right)^{\lambda\,\mu} \leq 1$. Asymptotically in n,

$$\frac{d_k(\alpha)\, d_{n-k}(\beta)}{d_n(\alpha+\beta)} = \left(\frac{\Gamma(k+\alpha)\,\Gamma(\alpha+\beta)}{\Gamma(k+1)\,\Gamma(\alpha)\,\Gamma(\beta)}\right) \frac{\Gamma(n-k+\beta)}{\Gamma(n-k+1)} \frac{\Gamma(n+1)}{\Gamma(n+\alpha+\beta)} \leq L\, n^{-\alpha} \left(1+\mathcal{O}(n^{-1})\right) \tag{4.104}$$

where $L = \max\limits_{k\in[0,N]} \left\{\frac{\Gamma(k+\alpha)\,\Gamma(\alpha+\beta)}{\Gamma(k+1)\,\Gamma(\alpha)\,\Gamma(\beta)}\right\}$. Thus for large enough n,

$$\sum_{k=0}^{N} \frac{d_k(\alpha)\, d_{n-k}(\beta)}{d_n(\alpha+\beta)} \left| \left(\frac{d_{n-k}(k+\gamma)}{d_{n-k}(k+\gamma+\lambda)}\right)^{\mu} - \left(\frac{k}{n}\right)^{\lambda\mu} \right| \leq N L\, n^{-\alpha-\lambda\mu}(1+\mathcal{O}(n^{-1})) < \frac{\epsilon}{2} . \tag{4.105}$$

Therefore from (4.103) and (4.105), a sufficiently large value of N can be chosen, whereafter a larger value of n may then be selected to ensure that

$$\left|\sum_{k=0}^{n}\frac{d_k(\alpha)\,d_{n-k}(\beta)}{d_n(\alpha+\beta)}\left[\left(\frac{d_{n-k}(k+\gamma)}{d_{n-k}(k+\gamma+\lambda)}\right)^{\mu}-\left(\frac{k}{n}\right)^{\lambda\mu}\right]\right|<\epsilon\,. \tag{4.106}$$

Hence,

$$\begin{aligned}\lim_{n\to\infty}S_n(\alpha,\beta,\gamma,\lambda,\mu) &= \lim_{n\to\infty}\sum_{k=0}^{n}\frac{d_k(\alpha)\,d_{n-k}(\beta)}{d_n(\alpha+\beta)}\left(\frac{k}{n}\right)^{\lambda\mu}\\ &= \lim_{n\to\infty}\sum_{k=0}^{n}\binom{n}{k}\frac{B(k+\alpha,n-k+\beta)}{B(\alpha,\beta)}\left(\frac{k}{n}\right)^{\lambda\mu}\,,\end{aligned} \tag{4.107}$$

and by Corollary 4.79,

$$\begin{aligned}\lim_{n\to\infty}S_n(\alpha,\beta,\gamma,\lambda,\mu) &= \int_0^1 t^{\alpha-1}(1-t)^{\beta-1}t^{\lambda\mu}\frac{dt}{B(\alpha,\beta)} \\ &= \frac{B(\alpha+\lambda\mu)}{B(\alpha,\beta)}\,.\end{aligned} \tag{4.108}$$

□

REMARK 4.6.2. Note that if $X\sim\text{BetaBin}(\alpha,\beta,n)$ then $Y=\frac{X}{n}\sim\text{Beta}(\alpha,\beta)$ as $n\to\infty$ [67, p. 425]. Since $f(x)=x^{\lambda\mu}$ is a bounded and continuous function on $x\in[0,1]$, Theorem 4.9 suggests that

$$\mathbb{E}\left[\left(\frac{X}{n}\right)^{\lambda\mu}\right]\to\mathbb{E}[Y^{\lambda\mu}]=\frac{B(\alpha+\lambda\,\mu,\beta)}{B(\alpha,\beta)} \tag{4.109}$$

which is the $\lambda\mu$ moment of $\text{Beta}(\alpha,\beta)$.

4.6.3 Complete Monotonicity of S_n

THEOREM 4.10. The sequence $S_n(\alpha,\beta,\alpha,\lambda,1)$ is logarithmically completely monotonic.

Proof. Observe that

$$\begin{aligned}S_n(\alpha,\beta,\alpha,\lambda,1) &= \frac{d_n(\alpha)}{d_n(\alpha+\beta)\,d_n(\alpha+\lambda)}\sum_{k=0}^{n}d_k(\beta)\,d_{n-k}(\alpha+\lambda)\\ &= \frac{d_n(\alpha)\,d_n(\alpha+\beta+\lambda)}{d_n(\beta)\,d_n(\alpha+\lambda)}\\ &= \frac{\Gamma(\alpha+\beta)\,\Gamma(\alpha+\lambda)}{\Gamma(\alpha)\,\Gamma(\alpha+\beta+\lambda)}\left(\frac{\Gamma(n+\alpha)\,\Gamma(n+\alpha+\beta+\lambda)}{\Gamma(n+\alpha+\lambda)\,\Gamma(n+\alpha+\beta)}\right)\,.\end{aligned} \tag{4.110}$$

and taking $x=n+\alpha$, $a=\lambda$, and $b=\beta$, the second product of (4.110) becomes

$$\frac{\Gamma(x)\,\Gamma(x+a+b)}{\Gamma(x+a)\,\Gamma(x+b)} \tag{4.111}$$

which is logarithmically completely monotonic in x for any $a,b\geq 0$ [58,66]. □

THEOREM 4.11. For $\alpha, \beta, \gamma > 0$, $\lambda \geq 0$, and $\mu \geq 1$, the sequence

$$\lim_{t\to\infty} S_n(\alpha, \beta, \gamma t, \lambda t, \mu) \tag{4.112}$$

is completely monotonic.

Proof. From (4.16),

$$\lim_{t\to\infty} \frac{B(n+\gamma\, t, \lambda\, t)}{B(k+\gamma\, t, \lambda\, t)} = \lim_{t\to\infty} \frac{\Gamma(n+\gamma\, t)\, \Gamma(k+(\gamma+\lambda)\, t)}{\Gamma(k+\gamma\, t)\, \Gamma(n+(\gamma+\lambda)\, t)} = \left(\frac{\gamma}{\gamma+\lambda}\right)^{n-k}$$

and by (4.89) gives

$$\begin{aligned}
&\lim_{t\to\infty} S_n(\alpha, \beta, \gamma, \lambda, \mu) \\
&\qquad = \frac{1}{B(\alpha,\beta)} \sum_{k=0}^{n} \binom{n}{k} B(k+\alpha, n-k+\beta) \left(\frac{\gamma}{\gamma+\lambda}\right)^{\mu(n-k)} \\
&\qquad = \frac{1}{B(\alpha,\beta)} \sum_{k=0}^{n} \binom{n}{k} \left[\int_0^1 x^{k+\alpha-1}(1-x)^{(n-k)+\beta-1} dx\right] \left(\frac{\gamma}{\gamma+\lambda}\right)^{\mu(n-k)} \\
&\qquad = \frac{1}{B(\alpha,\beta)} \int_0^1 x^{\alpha-1}(1-x)^{\beta-1} \left[\sum_{k=0}^{n} \binom{n}{k} x^k \left(\left(\frac{\gamma}{\gamma+\lambda}\right)^{\mu} (1-x)\right)^{n-k}\right] dx \\
&\qquad = \frac{1}{B(\alpha,\beta)} \int_0^1 x^{\alpha-1}(1-x)^{\beta-1} \left[x + \left(\frac{\gamma}{\gamma+\lambda}\right)^{\mu} (1-x)\right]^{n} dx\,.
\end{aligned} \tag{4.113}$$

Let $T_n(x, \gamma, \lambda, \mu) = \left(x + \left(\frac{\gamma}{\gamma+\lambda}\right)^{\mu} (1-x)\right)^{n}$. Since $\gamma > 0$, $\lambda \geq 0$ and $\mu \geq 1$, it follows that $0 \leq T_n(x, \gamma, \lambda, \mu) \leq 1$ for all $x \in (0,1)$ and $n \in \mathbb{N}$. By Example 4.5.2, $T_n(x, \gamma, \lambda, \mu)$ is completely monotonic so (4.113) is also completely monotonic, since $x^{\alpha-1}(1-x)^{\beta-1} \geq 0$ for all $x \in [0,1]$. $\square$

CONJECTURE 4.1. The sequence $S_n(\alpha, \beta, \alpha, \lambda, \mu)$ is completely monotonic for $\mu \geq 1$.

4.7 Moment Tables

4.7.1 Probability Distributions and Moment Generating Functions

Table 13: Probability mass functions and moment generating functions for some common discrete distributions.

Distribution	Support	PMF $= p(x)$	MFG $= M(t)$
Uni(a, b)	$[a, b]$	$\frac{1}{(b+1)-a}$	$\frac{e^{at}-e^{(b+1)t}}{[(b+1)-a]\,(1-e^{t})}$
Bin(n, p)	$[0, n]$	$\binom{n}{t}p^t(1-p)^{n-t}$	$(1+p\,(e^t-1))^n$
Geo(p)	$\mathbb{Z}^+$	$(1-p)^t\,p$	$\frac{p}{1-e^t\,(1-p)}$
NegBin(r, p)	$\mathbb{Z}^+$	$\binom{t+r-1}{r-1}(1-p)^t p^r$	$\left(\frac{p}{1-e^t(1-p)}\right)^r$
HyperGeo(n, m, N)	$*$[a]	$\frac{\binom{m}{t}\binom{N-m}{n-t}}{\binom{N}{n}}$	${}_2F_1\left[\begin{matrix}-m, -n\\ -N\end{matrix}\,;\,1-e^t\right]$
BetaBin(α, β, n)	$[0, n]$	$\binom{n}{t}\frac{B(t+\alpha,n-t+\beta)}{B(\alpha,\beta)}$	${}_2F_1\left[\begin{matrix}\alpha, -n\\ \alpha+\beta\end{matrix}\,;\,1-e^t\right]$
BetaNegBin(α, β, n)	$\mathbb{Z}^+$	$\binom{n+t-1}{t}\frac{B(n+\alpha,t+\beta)}{B(\alpha,\beta)}$	${}_2F_1\left[\begin{matrix}\beta, n\\ 1-\alpha\end{matrix}\,;\,1-e^t\right]$
Pois(λ)	$\mathbb{Z}^+$	$\frac{\lambda^t}{t!}e^{-\lambda}$	$e^{\lambda\,(e^t-1)}$
Zeta(s)	$\mathbb{N}$	$\frac{t^{-s}}{\zeta(s)}$	$\frac{\mathrm{Li}_s(e^t)}{\zeta(s)}$

[a] $[\max(0, n+m-N), \min(m, n)]$

Table 14: Probability density functions and moment generating functions for some common continuous distributions.

Distribution	PDF $= f(x)$	MGF $= M(t)$
Beta(α, β)	$\frac{1}{B(\alpha,\beta)}\, x^{\alpha-1}(1-x)^{\beta-1}$	${}_1F_1(\alpha; \alpha+\beta; t)$
Cauchy(θ, σ)	$\left(\pi\sigma\left[1+\left(\frac{x-\theta}{\sigma}\right)^2\right]\right)^{-1}$	d.n.e.
χ^2_p	$\frac{(p/2)^{1/2}}{\Gamma(p/2)} x^{p/2-1} e^{-(p/2)\,x}$	$(1-2t)^{-p/2}$
Exp(θ)	$\theta\, e^{-\theta\, x}$	$\left(1-\frac{t}{\theta}\right)^{-1}$
$F_{n,m}$	$\frac{1}{B\left(\frac{n}{2},\frac{m}{2}\right)} \left(\frac{n}{m}\right)^{n/2} x^{\frac{n}{2}-1} \left(1+\frac{n}{m}\,x\right)^{-\frac{n+m}{2}}$	d.n.e.
Gamma(α, β)	$\frac{1}{\Gamma(\alpha)\,\beta^{\alpha}} x^{\alpha-1} e^{-x/\beta}$	$(1-\beta t)^{-\alpha}$
Gumbel(μ, σ)	$\frac{1}{\sigma}\exp\left\{-\left(\frac{x-\mu}{\sigma}\right)-e^{-\left(\frac{x-\mu}{\sigma}\right)}\right\}$	$\Gamma\left(1-\sigma\, t\right) e^{\mu\, t}$
Laplace(μ, β)	$\frac{1}{2\beta}\exp\left\{-\frac{\lvert x-\mu\rvert}{\beta}\right\}$	$\frac{e^{t\mu}}{1-t^2\beta^2}$
Logistic(μ, β)	$\dfrac{e^{-(x-\mu)/\beta}}{\beta\,(1+e^{-(x-\mu)/\beta})^2}$	$e^{\mu\, t} B(1-\beta\, t, 1+\beta\, t)$
Lognormal(μ, σ^2)	$\dfrac{1}{x\,\sigma\sqrt{2\pi}} e^{-\frac{(\ln x-\mu)^2}{2\sigma^2}}$	d.n.e.
$\mathcal{N}(\mu, \sigma^2)$	$\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	$e^{t\mu+\frac{1}{2}\sigma^2 t^2}$
Pareto(α, β)	$\frac{\alpha\,\beta^{\alpha}}{x^{\alpha+1}}$	$\alpha(-\beta\, t)^{\alpha}\,\Gamma(-\alpha, -\beta\, t)$
Rayleigh(σ)	$\frac{x}{\sigma^2}\exp\left\{-\frac{x^2}{2\sigma^2}\right\}$	$1+\sqrt{\pi}\sqrt{\frac{t^2\sigma^2}{2}}\, e^{t^2\sigma^2/2}\left(1+\mathrm{erf}\sqrt{\frac{t^2\sigma^2}{2}}\right)$
Student's t_ν	$\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\,\Gamma\left(\frac{\nu}{2}\right)}\left(1+\frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}$	d.n.e.
Uniform(a, b)	$\frac{1}{b-a}$	$\frac{e^{tb}-e^{ta}}{t(b-a)}$
Weibull(λ, k)	$\frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1} e^{-(x/\lambda)^k}$	$\sum_{n=0}^{\infty}\frac{t^n\lambda^n}{n!}\,\Gamma\left(1+\frac{n}{k}\right)$

4.7.2 Moments and Factorial Moments

Table 15: High order moments and factorial moments for some common discrete distributions.

Distribution	Factorial Moments ($\mu'_{[i]}$)	Moments (μ'_i)
Uni(a, b)	$\frac{(b+1)^{\underline{i+1}}-a^{\underline{i+1}}}{(i+1)[(b+1)-a]}$	$\sum_{k=1}^{i} \left\{ \begin{matrix} i \\ k \end{matrix} \right\} \frac{(b+1)^{\underline{k+1}}-a^{\underline{i+1}}}{(k+1)[(b+1)-a]}$
Bin(n, p)	$p^i n^{\underline{i}}$	$\sum_{k=1}^{i} \left\{ \begin{matrix} i \\ k \end{matrix} \right\} p^k n^{\underline{k}}$
Geo(p)	$\left(\frac{1-p}{p}\right)^i i!$	$\sum_{k=1}^{i} \left\{ \begin{matrix} i \\ k \end{matrix} \right\} \left(\frac{1-p}{p}\right)^k k!$
NegBin(r, p)	$\left(\frac{1-p}{p}\right)^i r^{\overline{i}}$	$\sum_{k=1}^{i} \left\{ \begin{matrix} i \\ k \end{matrix} \right\} \left(\frac{1-p}{p}\right)^k r^{\overline{k}}$
HyperGeo(n, m, N)	$\frac{m^{\underline{i}} n^{\underline{i}}}{N^{\underline{i}}}$	$\sum_{k=1}^{i} \left\{ \begin{matrix} i \\ k \end{matrix} \right\} \frac{m^{\underline{k}} n^{\underline{k}}}{N^{\underline{k}}}$
BetaBin(α, β, n)	$\frac{\alpha^{\overline{i}}}{(\alpha+\beta)^{\overline{i}}} n^{\underline{i}}$	$\sum_{k=1}^{i} \left\{ \begin{matrix} i \\ k \end{matrix} \right\} \frac{\alpha^{\overline{k}}}{(\alpha+\beta)^{\overline{k}}} n^{\underline{k}}$
BetaNegBin(α, β, n)	$\frac{\beta^{\overline{i}}}{(\alpha-1)^{\underline{i}}} n^{\overline{i}}$	$\sum_{k=1}^{i} \left\{ \begin{matrix} i \\ k \end{matrix} \right\} \frac{\beta^{\overline{k}}}{(\alpha-1)^{\underline{k}}} n^{\overline{k}}$
Pois(λ)	λ^i	$\sum_{k=1}^{i} \left\{ \begin{matrix} i \\ k \end{matrix} \right\} \lambda^k$
Zeta(s)	$\begin{cases} \sum_{k=1}^{i} \left[\begin{matrix} i \\ k \end{matrix} \right] \frac{\zeta(s-k)}{\zeta(s)}, & i < s-1 \\ \infty, & i \geq s-1 \end{cases}$	$\begin{cases} \frac{\zeta(s-i)}{\zeta(s)}, & i < s-1 \\ \infty, & i \geq s-1 \end{cases}$

Table 16: High order moments of common continuous distributions

Distribution	Moments (μ_r')
Beta(α, β)	$\dfrac{\alpha^{\overline{i}}}{(\alpha+\beta)^{\overline{i}}}$
Cauchy(θ, σ)	$\begin{cases} -, & i \text{ odd} \\ \infty, & i \text{ even} \end{cases}$
χ^2_p	$\prod_{j=1}^{i} (p + 2(j-1))$
Exp(θ)	$\dfrac{i!}{\theta^i}$
$F_{n,m}$	$\dfrac{\Gamma\left(\frac{n+2i}{2}\right)\,\Gamma\left(\frac{m-2i}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\,\Gamma\left(\frac{m}{2}\right)}\left(\frac{m}{n}\right)^i$
Gamma(α, β)	$\beta^i \alpha^{\overline{i}}$
Laplace(μ, β)	$\sum_{k=0}^{\lfloor i/2 \rfloor} \binom{i}{2k} (2k)!\, \mu^{i-2k} \beta^{2k}$
Logistic(μ, β)	$\mu^r + \sum_{k=1}^{n} \binom{n}{k} \mu^{n-k} (\pi s)^k (2^k - 2) \lvert \mathcal{B}_k \rvert$
Lognormal(μ, σ^2)	$e^{i\,\mu + \frac{1}{2}(i\,\sigma)^2}$
$\mathcal{N}(\mu, \sigma^2)$	$\sum_{k=0}^{\lfloor i/2 \rfloor} \binom{i}{2k} (2k-1)!!\, \mu^{i-2k} \sigma^{2k}$
Pareto(α, β)	$\begin{cases} \frac{\alpha^r \beta}{\beta - r}, & \beta > r \\ \infty, & \text{otherwise} \end{cases}$
Rayleigh(σ)	$2^{\frac{i}{2}}\, \Gamma\left(1 + \frac{i}{2}\right)\, \sigma^i$
t_ν	$\dfrac{\Gamma\left(\frac{\nu+1}{2}\right)\, \Gamma(\nu - \frac{i}{2})}{\sqrt{\pi}\, \Gamma\left(\frac{\nu}{2}\right)}\, \nu^{\frac{i}{2}}$
Uniform(a, b)	$\dfrac{b^{i+1} - a^{i+1}}{(i+1)\,(b-a)}$
Weibull(λ, k)	$\lambda^r\, \Gamma\left(1 + \frac{r}{k}\right)$

Table 17: Named GHFMDs with PGFs Expressible as ${}_pF_q[\lambda(z-1)]$. Reprinted from [71, 75].

Name	PGF Form	Reference
	Terminating	
Binomial	${}_1F_0\left[\begin{matrix}-n\\-\end{matrix}\,;\,p\,(1-z)\right]$	
Hypergeometric	${}_2F_1\left[\begin{matrix}-n,\,-Np\\-N\end{matrix}\,;\,1-z\right]$	
Negative hypergeometric	${}_2F_1\left[\begin{matrix}-n,\,a\\a+b\end{matrix}\,;\,1-z\right]$	
Discrete rectangular	${}_2F_1\left[\begin{matrix}-n,\,1\\2\end{matrix}\,;\,1-z\right]$	
Chung-Feller	${}_2F_1\left[\begin{matrix}-n,\,1/2\\1/2\end{matrix}\,;\,1-z\right]$	[59]
Pólya	${}_2F_1\left[\begin{matrix}-n,\,M/c\\(M+N)/c\end{matrix}\,;\,1-z\right]$	
Matching Distribution	${}_1F_1\left[\begin{matrix}-n\\-n\end{matrix}\,;\,1-z\right]$	
Gumbel's matching distribution	${}_1F_1\left[\begin{matrix}-n\\-n\end{matrix}\,;\,\lambda\,(z-1)\right]$	[62]
Laplace–Haag	${}_1F_1\left[\begin{matrix}-n\\-N\end{matrix}\,;\,M(z-1)\right]$	[62]
Anderson's matching distribution	${}_1F_{N-1}\left[\begin{matrix}-n\\-n,\,\ldots,\,-n\end{matrix}\,;\,(-1)^N(z-1)\right]$	[54]
Stevens-Craig (Coupon collecting)	${}_nF_{n-1}\left[\begin{matrix}1-k,\,\ldots,\,1-k\\-k,\,\ldots,\,-k\end{matrix}\,;\,(1-z)\right]$	[20]
STERRED rectangular	${}_3F_2\left[\begin{matrix}-n,\,1,\,1\\2,\,2\end{matrix}\,;\,1-z\right]$	[76]
	Nonterminating	
Poisson	${}_0F_0\left[\begin{matrix}-\\-\end{matrix}\,;\,\lambda(z-1)\right]$	
Negative Binomial	${}_1F_0\left[\begin{matrix}k\\-\end{matrix}\,;\,p(z-1)\right]$	
Poisson $\wedge$ Beta	${}_1F_1\left[\begin{matrix}a\\a+b\end{matrix}\,;\,\lambda(z-1)\right]$	
Type H_2	${}_2F_1\left[\begin{matrix}k,\,a\\a+b\end{matrix}\,;\,\lambda(z-1)\right]$	[68, 73]
STERRED geometric	${}_2F_1\left[\begin{matrix}1,\,1\\2\end{matrix}\,;\,q\,\frac{z-1}{1-q}\right]$	[76]

Table 18: Stirling numbers of the first kind, $\left[{n \atop k}\right]$. (OEIS: A048994)

	0	1	2	3	4	5	6	7
0	1							
1	0	1						
2	0	-1	1					
3	0	2	-3	1				
4	0	-6	11	-6	1			
5	0	24	-50	35	-10	1		
6	0	-120	274	-255	85	-15	1	
7	0	720	-1764	1624	-735	175	-21	1

Table 19: Stirling numbers of the second kind, $\left\{{n \atop k}\right\}$. (OEIS: A008277)

	0	1	2	3	4	5	6	7
0	1							
1	0	1						
2	0	1	1					
3	0	1	3	1				
4	0	1	7	6	1			
5	0	1	15	25	10	1		
6	0	1	31	90	65	15	1	
7	0	1	63	301	350	140	21	1

4.8 Chapter Notations

Symbol	Description	Page Reference
$\begin{bmatrix} n \\ k \end{bmatrix}$	stirling numbers of the first kind	79
$\begin{Bmatrix} n \\ k \end{Bmatrix}$	stirling numbers of the second kind	79
$\binom{n}{k}$	binomial coefficient	80
α_i	standard moment	85
$B(\alpha, \beta)$	beta function	81
$\mathfrak{B}$	the Borel σ-algebra	
$\mathcal{B}_n$	Bernoulli numbers	
$B_{n,k}(x_1, \ldots, x_{n-k+1})$	partial Bell polynomial	79
BetaBin(α, β, n)	beta-binomial distribution	90
Bin(n, p)	binomial distribution	89
$\mathbf{B}_n(h; t)$	Bernstein polynomial	96
CDF	cumulative distribution function	83
Δ_n	forward difference operator	94
DF	distribution function	83
$d_n(\alpha)$	negative binomial coefficient	81
EGF	exponential moment generating function	84
erf	Gauss error fuction	
$\mathbb{E}[X]$	expected value	83
FMGF	factorial moment generating function	86
$\Gamma(z)$	gamma function	80
$\Gamma(z, s)$	incomplete gamma function	
GF	generating function	84
GHFMD	generalized hypergeometric factorial moment distribution	91
${}_pF_q$	generalized hypergeometric function	82
i.i.d.	independent, identically distributed	83
μ'_i	moment	84
$\hat{\mu}'_i$	sample moment	87
μ_i	central moment	85

Symbol	Description	Page Reference
$\mu'_{[i]}$	factorial moment	86
$\hat{\mu}'_{[i]}$	sample factorial moment	87
MGF	moment generating function	84
$M_X(t)$	moment generating function of the RV X	84
$\widetilde{M}_X(t)$	factorial moment generating function of the RV X	86
$\mathbb{N}$	natural numbers $(1, 2, \dots)$	
$\mathbb{N}_0$	natural numbers with 0 $(0, 1, 2, \dots)$	
$[n]$	natural numbers up to n $(1, 2, \dots, n)$	
$n!$	factorial function	79
$n^{\underline{k}}$	falling factorial function	79
$n^{\overline{k}}$	rising factorial function	79
Ω	possible outcomes in a sample space	82
PDF	probability density function	83
PGF	probability generating function	84
PMF	probability mass function	83
ψ	digamma function	99
$\mathbb{R}$	the real numbers	
RV	random variable	82
$\mathcal{S}$	a σ-field of subset from Ω	82
$S_n(\alpha, \beta, \gamma, \lambda, \mu)$	a series	98
Var(X)	variance	85
$\xi'_{[i]}$	ratio of factorial moments	91

Chapter 5

Conclusions

At first glance, the topics considered in Chapters 2,3, and 4 seem unrelated, but further inspection reveals a common theme. In each case, a geometric problem is shown to have a solution which is expressed recursively as a weighted sum. In some of these cases, the weighted sums have equivalent closed forms, which can be exploited.

5.1 Future Research

We have shown that general tools for working with recursive relations, such as inversions, generating functions, and hypergeometric presentations are useful in many mathematical fields. Further development of these general tools can lead to unexpected consequences, such as the counting sequences for the braid index following a Gamma distribution, see Example 4.4.2. Other steps in this direction may include

- Establish additional hypergeometric identities and methods of manipulating generating functions to convert weighted sums into closed forms, along the lines of the WZ algorithm [85].

The following sections include partial summaries of the results from each chapter in order to contextualize the statements of future research topics.

5.1.1 Number Theory

Chapter 2 concerns the integer factorizations of polynomial values. A recursive structure among the integer factorizations is expressed in Theorem 2.1, by making use of the fact that if f divides a polynomial $F(n)$ then f divides $F(n+f)$ as well. For certain prime-producing polynomials, the polynomial-value sieving method characterizes all the positive integer factorizations of the polynomial values. These prime-producing polynomials are deeply connected with many outstanding problems in number theory such as the

k-tuple conjecture [18, p.106], Landau's $n^2 + 1$ problem [14], and the Riemann zeta hypothesis [10], so any development of the study could have a great impact.

Some direct extensions of this number theory work may include:

- Identify sufficient initial conditions for the polynomial-value sieve to characterize all the positive integer factorizations for all quadratic progressions,
- Use Theorem 2.1 to establish closed identities of the form $F(n) = PQ$, where F is a polynomial with integer coefficients, for cubic and higher order progressions,
- Apply the identity for linear progressions to give an alternative proof for Dirichlet's arithmetic progression theorem,
- Develop an integer factorization algorithms based on Theorem 2.1, Theorem 2.3, and Corollary 2.2, and
- Investigate the connection between polynomial-value sieving and the Bunyakovsky [2] and Hardy-Littlewood [12] conjectures about the prime density of polynomial progressions.

5.1.2 Combinatorial Graph Theory

In Chapter 3, double occurrence words are studied through the context of self-intersecting smooth curves, assembly graphs, chord diagrams, and linear chord diagrams. This study was initially motivated by a biological model of genome rearrangement, and recent data [33] shows that copies of the same genetic sequence may unscramble in different ways to form distinct genes. Generalizing the assembly graph model to account for these potentially stochastic rearrangements may involve the use of multiple occurrence words and their generalizations to chord diagrams, leading to the following research topics:

- Classify and enumerate multiple occurrence and arbitrary occurrence words based on notions of irreducibility, and
- Consider assembly curves with multiple self-intersects at a fixed point and extend this generality to the assembly graph model.

Several outstanding problems remain for computing the braid index of an arbitrary chord diagram. The divide-and-conquer algorithm improves up the Birman-Trapp brute-force algorithm by decomposing a chord diagram into sub-linear chord diagrams, but strongly-irreducible chord diagrams cannot be decomposed this way. Moreover, there is no clear connection of how the braid index relates to other properties of chord diagrams. These concerns motive the following research tasks:

- Determine an algorithm for computing the braid index of strongly-irreducible chord diagrams,
- Explore the connection between a chord diagram's braid index and circle graph, and
- See whether the braid index of comparative genomics chord diagrams indicates a notion of evolutionary complexity by cross-referencing the braid index with phylogenetic data.

5.1.3 Probabilty

Inversion formulas make it easy to convert between various type of moments, and computing sample moments is simple and computationally inexpensive. Hence, distribution and parameter fitting of a sample using the ratios of successive moments is feasible for large datasets. The ratio of successive factorial moments of most discrete distributions is a rational function, making them easy to distinguish with this method. Similarly, most continuous distributions have either ratios of successive moments or cumulants that are rational functions. This work introduces the ratios of successive moments method through a discrete example and a continuous example, and the full scope of the method can be investigated immediately. In particular:

- Exhaust the pros and cons of using fractional moments to circumvent the increasing variance and sensitivity towards outliers when using high order moments,
- Build new types of moments from orthogonal sequences, such as the Tanh and Lah numbers, and examine their graphical features,
- Define and study the classes of probability distributions based on the successive ratios of these new types of moments, and
- Use rational regression on the ratios of successive moments for joint and mixture distributions.

The moment sequences of scaled random variables derived from urn models can be used to build sequences of polynomials which approximate continuous function on a bounded interval. However these moment sequence are not necessarily completely monotonic. We construct a multiparameter binomial sum, having a hypergeometric presentation, which is completely monotonic and behaves asymptotically like a scaled random variable. The most common method of proving that a hypergeometric function is completely monotonic is to work with its closed form as the ratio of gamma functions. However not all hypergeometric functions can be expressed in a closed form using modern identities [78], and it is rare that special cases of S_n can be shown to be completely monotonic without the aid of a presentation as the ratio of gamma functions. Additional considerations:

- Consider the moment sequences of scaled random variables derived from other urn models,
- Fully characterize the conditions under which $S_n(\alpha, \beta, \gamma, \lambda, \mu)$ is completely monotonic, and
- Use the monotone operator theorem 4.6 to establish new approximating polynomials based on common probability distributions.

References

[1] A. Baker, *Imaginary Quadratic Fields with Class Number 2*, Annals of Mathematics **94** (1971), no. 1, 139 –152, DOI 10.2307/1970739.

[2] V. Bouniakowsky, *Nouveaux théorèmes relatifs à la distinction des nombres premiers et à la décomposition des entiers en facteurs*, Mém. Acad. Sc. St. Pétersbourg, vol. 6, 1857.

[3] J.H. Conway, *The Sensual (Quadratic) Form*, Carus Mathematical Monographs, vol. 26, Mathematical Association of America, Washington, DC, 1997.

[4] Diophantus, *Arithmetica*, Book III, c. 250. Problem 19.

[5] L. Euler, *Algebra*, Vol. 2, St. Petersburg, 1770.

[6] ———, *Extrait d'une lettre de M. Euler le père à M. Bernoulli concernant le memoire imprimé parmi ceux de 1771*, Nouveaux mémoirs de'Acadé des Sciences de Berlin 1772, 1774.

[7] L. Fibonacci, *The book of squares* (L.E. Sigler, ed.), translated by L.E. Sigler, Academic Press, Boston, MA, 1987.

[8] F.G. Frobenius, *Über quadratische Formen, die viele Primzahlen darstellen: unter Benutzung einer Mitteilung des Hrn. Dr. R. Remak*, Preussische Akademie der Wissenschaften Berlin: Sitzungsberichte der Preußischen Akademie der Wissenschaften zu Berlin, Reichsdr., 1912.

[9] C.F. Gauss, *Disquisitiones Arithmeticae*, Springer, New York, NY, 1986.

[10] D. Goldfeld, *Gauss' class number problem for imaginary quadratic fields*, Bulletin (New Series) of the American Mathematical Society **13** (1985), no. 1, 23 –37, DOI 10.1090/S0273-0979-1985-15352-2.

[11] A.J. Hahn, O.T. O'Meara, and J. Dieudonne, *The Classical Groups and K-Theory*, Grundlehren der mathematischen Wissenschaften, Springer-Verlag, New York, NY, 2010.

[12] G.H. Hardy and J.E. Littlewood, *Some problems of Partitio numerorum; III: On the expression of a number as a sum of primes*, Acta Mathematica **44** (1923), no. 1, 1 –70, DOI 10.1007/BF02403921.

[13] K. Heegner, *Diophantische Analysis und Modulfunktionen*, Mathematische Zeitschrift **56** (1952), no. 3, 227 –253, DOI 10.1007/BF01174749.

[14] E. Landau, *Gelöste und ungelöste Probleme aus der Theorie der Primzahlverteilung und der Riemannschen Zetafunktion*, Proceedings of the Fifth International Congress of Mathematicians, vol. 1, University press, Cambridge, 1913.

[15] A.M. Legendre, *Théorie des nombres*, Libraire Scientifique, A. Hermann, Paris, 1798.

[16] F.L. Lionnais and J. Brette, *Les nombres remarquables*, Actualitâes scientifiques et industrielles, A. Hermann, Paris, 1983.

[17] S. Louboutin, *Extensions du théorème de Frobenius-Rabinovitsch*, C. R. Acad. Sci. Paris **312** (1991), 711 –714.

[18] R.A. Mollin, *Quadratics*, CRC Press Series on Discrete Mathematics and its Applications, CRC Press, Boca Raton, FL, 1996.

[19] L.J. Mordell, *Diophantine equations*, Pure and Applied Mathematics, vol. 30, Academic Press, London, 1969.

[20] G.P. Patil and S.W. Joshi, *A dictionary and bibliography of discrete distributions*, International Statistical Institute, Oliver & Boyd, Edinburgh, 1968.

[21] G. Rabinowitz, *Eindeutigkeit der Zerlegung in Primzahlfaktoren in quadratischen Zahlkörpern*, Proceedings of the Fifth International Congress of Mathematicians, vol. 1, University press, Cambridge, 1913.

[22] P. Ribenboim, *The Little Book of Bigger Primes*, 2nd, Springer-Verlag, New York, NY, 2004.

[23] H.M. Stark, *A complete determination of the complex quadratic fields of class-number one*, The Michigan Mathematical Journal **14** (1967), no. 1, 1 –27, DOI 10.1307/mmj/1028999653.

[24] ______, *A Transcendence Theorem for Class Number Problems*, Annals of Mathematics **94** (1971), no. 1, 153 –173, DOI 10.2307/1970740.

[25] J. Stillwell, *Mathematics and Its History*, Undergraduate Texts in Mathematics, Springer, New York, NY, 2010.

[26] E.W. Weisstein, *Prime-Generating Polynomial* (2014), From MathWorld –A Wolfram Web Resource., `http://mathworld.wolfram.com/Prime-GeneratingPolynomial.html`.

[27] A. Angeleska, N. Jonoska, and M. Saito, *DNA recombinations through assembly graphs*, Discrete Applied Mathematics **157** (2009), 3020 –3037, DOI 10.1016/j.dam.2009.06.011.

[28] A. Angeleska, N. Jonoska, M. Saito, and L.F. Landweber, *RNA-guided DNA assembly*, Journal of Theoretical Biology **248** (2007), no. 4, 706 –720, DOI 10.1016/j.jtbi.2007.06.007.

[29] J.S Birman and R. Trapp, *Braided Chord Diagrams*, Journal of Knot Theory and Its Ramifications **7** (1998), no. 1, 1 –22, DOI 10.1142/S0218216598000024.

[30] J. Burns, E. Dolzhenko, N. Jonoska, T. Muche, and M. Saito, *Four-regular Graphs with Rigid Vertices Associated to DNA Recombination*, Discrete Applied Mathematics **161** (2013), no. 10 –11, 1378 – 1394, DOI 10.1016/j.dam.2013.01.003.

[31] J. Burns and T. Muche, *Counting irreducible double occurrence words*, Congressus Numerantium **207** (2011), 181 –196. arXiv:1105.2926.

[32] G. Cairns and D.M. Elton, *The Planarity Problem for Signed Gauss Words*, Journal of Knot Theory and Its Ramifications **2** (1993), no. 4, 359 –367, DOI 10.1142/S0218216593000209.

[33] X. Chen, J.R. Bracht, A.D. Goldman, E. Dolzhenko, E.C. Swart, D.M. Clay, D.H. Perlman, T.G. Doak, A. Stuart, C.T. Amemiya, and L.F. Landweber, *The architecture of a scrambled genome reveals massive levels of genomic rearrangement during development*, to appear in *Cell*.

[34] P. Cvitanović, B. Lautrup, and R.B. Pearson, *Number and weights of Feynman diagrams*, Phys. Rev. D **18** (1978), 1939 –1949, DOI 10.1103/PhysRevD.18.1939.

[35] A. Ehrenfeucht, T. Harju, I. Petre, D.M. Prescott, and G. Rozenberg, *Computation in Living Cells : Gene Assembly in Ciliates*, Natural Computing Series, Springer-Verlag, Berlin, Heidelberg, 2004.

[36] A Ehrenfeucht, T. Harju, and G. Rozenberg, *Gene assembly through cyclic graph decomposition*, Theoretical Computer Science **281** (2002), no. 1–2, 325 –349, DOI 10.1016/S0304-3975(02)00019-1.

[37] C. Godsil and G. Royle, *Algebraic Graph Theory*, Graduate Texts in Mathematics, vol. 207, Springer-Verlag, New York, 2001.

[38] L.H. Kauffman, *Virtual Knot Theory*, European Journal of Combinatorics **20** (1999), no. 7, 663 –691, DOI 10.1006/eujc.1999.0314.

[39] L. Kari and L.F. Landweber, *Computational Power of Gene Rearrangement* (E. Winfree and D. Gifford, eds.), DNA5: DIMACS Series in Discrete Mathematics and Theoretical Computer Science, vol. 54, Amer. Math. Soc., Providence, RI, 2000.

[40] M. Klazar, *Non-P-recursiveness of numbers of matchings or linear chord diagrams with many crossings*, Advances in Applied Mathematics **30** (2003), no. 1-2, 126 –136, DOI 10.1016/S0196-8858(02)00528-6.

[41] R.J. Martin and M.J. Kearney, *An exactly solvable self-convolutive recurrence*, Aequationes mathematicae **80** (2010), no. 3, 291 –318, DOI 10.1007/s00010-010-0051-0.

[42] A. Nijenhuis and H.S. Wilf, *The enumeration of connected graphs and linked diagrams*, Journal of Combinatorial Theory, Series A **27** (1979), no. 3, 356 –359, DOI 10.1016/0097-3165(79)90023-2.

[43] M. Nowacki, V. Vijayan, Yi Zhou, K. Schotanus, T.G. Doak, and L.F. Landweber, *RNA-mediated epigenetic programming of a genome-rearrangement pathway*, Nature **451**, no. 7175, 153 –158, DOI 10.1038/nature06452.

[44] D.M. Prescott, A. Ehrenfeucht, and G. Rozenberg, *Template-guided recombination for IES elimination and unscrambling of genes in stichotrichous ciliates*, Journal of Theoretical Biology **222** (2003), no. 3, 323 –330, DOI 10.1016/S0022-5193(03)00037-7.

[45] D.M. Prescott and A.F. Greslin, *Scrambled actin I gene in the micronucleus of Oxytricha nova*, Developmental Genetics **13** (1992), no. 1, 66 –74, DOI 10.1002/dvg.1020130111.

[46] N.J.A. Sloane, *Online Encyclopedia of Integer Sequences* (2014), `http://www.oeis.org`.

[47] R.W. Robinson, *Counting irreducible Feynman diagrams exactly and asymptotically*, Abstracts Amer. Math. Soc., 2002. 975-05-270.

[48] D. Ross, 2013. Personal Communication.

[49] P.R. Stein, *On a class of linked diagrams, I. Enumeration*, Journal of Combinatorial Theory, Series A **24** (1978), no. 3, 357 –366, DOI 10.1016/0097-3165(78)90065-1.

[50] A. Stoimenow, *On the number of chord diagrams*, Discrete Mathematics **218** (2000), no. 1 –3, 209 –233, DOI 10.1016/S0012-365X(99)00347-7.

[51] J. Touchard, *Contributions à l'étude du problème des timbres-poste*, Canad. J. Math. **2** (1950), 385 –398, DOI 10.4153/CJM-1950-035-6.

[52] ______, *Sur un problème de configurations et sur les fractions continues*, Canad. J. Math. **4** (1952), 2 –25, DOI 10.4153/CJM-1952-001-8.

[53] M. Aigner, *Discrete Mathematics*, translated by D. Kramer, American Mathematical Society, 2007.

[54] T.W. Anderson, *On card matching*, The Annals of Mathematical Statistics **14** (1943), no. 4, 426 –435, DOI 10.1214/aoms/1177731362.

[55] G.E. Andrews, *The Theory of Partitions* (G.C. Rota and P. Turán, eds.), Encyclopedia of Mathematics and its Applications, vol. 2, Addison-Wesley Publishing Company, 1976.

[56] S.N. Bernstein, *Démonstration du théorém de Weierstrass, fondée sur le calcul des probabilités*, Commun. Soc. Math. Kharkow **2** (1912/1913), no. 13, 1 –2.

[57] H. Bohman, *On approximation of continuous analytic functions*, Arkiv. Mat. **1** (1952), no. 2, 43 –56.

[58] J. Bustoz and M.E.H. Ismail, *On gamma function inequalities*, Mathematics of Computation **47** (1986), no. 176, 659 –667, DOI 10.2307/2008180.

[59] K.L. Chung and W. Feller, *On Fluctuations in Coin Tossing*, Proceedings of the National Academy of Sciences U.S.A. **35** (1949), 605 –608, DOI 10.1073/pnas.35.10.605.

[60] A. Erdélyi, W. Magnus, F. Oberhettinger, and G.F. Tricomi (eds.), *Higher Transcendental Functions. Vol. I.*, McGraw-Hill, New York, 1953.

[61] W. Feller, *An Introduction to Probability Theory and Its Applications*, Second Edition, Wiley Series in Probability and Mathematical Statistics, vol. 2, Wiley, 1971.

[62] M. Fréchet, *Les probabilités associées a un systéme d'evénements compatibles et dépendants. 2. Cas particuliers et applications*, Actualités Sci. Indust. **859** (1943). Hermann, Paris.

[63] M. Fisz, *Probability Theory and Mathematical Statistics*, 3rd, Wiley, New York, 1963.

[64] A.Z. Grinshpan, *Weighted inequalities and negative binomials*, Advances in Applied Mathematics **45** (2010), no. 4, 564 –606, DOI 10.1016/j.aam.2010.04.004.

[65] ______, *General inequalities, consequences and applications*, Advances in Applied Mathematics **34** (2005), no. 1, 71 –100, DOI 10.1016/j.aam.2004.05.001.

[66] A.Z. Grinshpan and M.E.H. Ismail, *Completely Monotonic Functions Involving the Gamma and q-Gamma Functions*, Proceedings of the American Mathematical Society **134**, no. 4, 1153 –1160, DOI 10.1090/S0002-9939-05-08050-0.

[67] A.K. Gupta and S. Nadarajah (eds.), *Handbook of Beta Distribution and Its Applications*, Statistics: A Series of Textbooks and Monographs, CRC Press, 2004.

[68] J. Gurland, *A generalized class of contagious distributions*, Biometrics **14** (1958), 229 –49.

[69] F. Hausdorff, *Summationmethoden und Momentfolgen. I*, Mathematische Zeitschrift **9** (1921), no. 1, 74 –109, DOI 10.1007/BF01378337 (German).

[70] ______, *Summationmethoden und Momentfolgen. II*, Mathematische Zeitschrift **9** (1921), no. 3–4, 280 –299, DOI 10.1007/BF01279032 (German).

[71] N.L. Johnson, S. Kotz, and A.W. Kemp, *Univariate Discrete Distributions*, Third Edition, Wiley Series in Probability and Statistics, Wiley-Interscience, 2005.

[72] N.L. Johnson and S. Kotz, *Urn Models and Their Application: An Approach to Modern Discrete Probability Theory*, Wiley Series in Probability and Statistics, John Wiley & Sons, 1977.

[73] S.K. Katti, *Interrelations among generalized distributions and their components*, Biometrics **22** (1966), 44 –52, DOI 10.2307/2528212.

[74] A.W. Kemp, *Studies in Univariate Discrete Distribution Theory Based on the Generalized Hypergeometric Function and Associated Differential Equations*, Ph.D. Thesis, 1968.

[75] A.W. Kemp and C.D. Kemp, *A family of discrete distributions defined via their factorial moments*, Communications in Statistics - Theory and Methods **3** (1974), no. 12, 1187 –1196, DOI 10.1080/03610927408827220.

[76] ______, *Some distributions arising from an inventory decision problem*, Bulletin of the International Statistical Institute **43** (1969), no. 2, 336 –338.

[77] P.P. Korovkin, *On convergence of linear positive operators in the space of continuous functions*, Doklady Akad. Nauk. SSSR **90** (1953), 961 –964 (Russian).

[78] M. Milgram, *On Hypergeometric* ${}_3F_2(1)$ *- A Review*, ArXiV (2006). arXiv:math/0603096.

[79] G.M. Phillips, *Interpolation and Approximation by Polynomials*, CMS Books in Mathematics, vol. XIV, Springer, New York, 2003.

[80] J.A. Rice, *Mathematical Statistics and Data Analysis*, Second Edition, Duxbury Press, Belmont, CA, 1995.

[81] V.K. Rohatgi and A.K. Md. E. Saleh, *An Introduction to Probability and Statistics*, 2nd Edition, Wiley Series in Probability and Statistics, 2001.

[82] P.J. Smith, *A recursive formulation of the old problem of obtaining moments from cumulants and vice versa*, The American Statistician **49** (1995), no. 2, 217 –218, DOI 10.1080/00031305.1995.10476146.

[83] F.G. Tricomi and A. Erdélyi, *The Asymptotic Expansion of a Ratio of Gamma Functions*, Pacific Journal of Mathematics **1** (1951), no. 1, 133 –142.

[84] D.V. Widder, *The Laplace Transform*, Princeton University Press, Princeton, NJ, 1946.

[85] H.S. Wilf, *generatingfunctionology*, Third Edition, A K Peters, Ltd., 2006.

Appendix A - Copyright Permission

ELSEVIER LICENSE TERMS AND CONDITIONS

This is a License Agreement between Jonathan Burns ("You") and Elsevier ("Elsevier") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Elsevier, and the payment terms and conditions.

Supplier	Elsevier Limited The Boulevard, Langford Lane Kidlington, Oxford, OX5 1GB,UK
Registered Company Number	1982084
Customer name	Jonathan Burns
License number	3410200441935
License date	Jun 15, 2014
Licensed content publisher	Elsevier
Licensed content publication	Discrete Applied Mathematics
Licensed content title	Four-regular graphs with rigid vertices associated to DNA recombination
Licensed content author	Jonathan Burns, Egor Dolzhenko, Nataša Jonoska, Tilahun Muche, Masahico Saito
Licensed content date	July 2013
Licensed content volume	161
Licensed content issue number	10-11
Number of pages	17
Start Page	1378
End Page	1394
Type of Use	reuse in a thesis/dissertation
Portion	full article
Format	electronic
Are you the author of this publication?	Yes
Will you be translating?	No
Title of your thesis/dissertation	Recursive Methods in Number Theory, Combinatorial Graph Theory, and Probability

Expected completion date	Jun 2014
Estimated size (number of pages)	120
Elsevier VAT number	GB 494 6272 12
Permissions price	0.00 USD

INTRODUCTION

1. The publisher for this copyrighted material is Elsevier. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your Rightslink account and that are available at any time at `http://myaccount.copyright.com`).

GENERAL TERMS

2. Elsevier hereby grants you permission to reproduce the aforementioned material subject to the terms and conditions indicated.

3. **Acknowledgement:** If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permission is not obtained then that material may not be included in your publication/copies. Suitable acknowledgement to the source must be made, either as a footnote or in a reference list at the end of your publication, as follows:

   "Reprinted from Publication title, Vol /edition number, Author(s), Title of article / title of chapter, Pages No., Copyright (Year), with permission from Elsevier [OR APPLICABLE SOCIETY COPYRIGHT OWNER]." Also Lancet special credit - "Reprinted from The Lancet, Vol. number, Author(s), Title of article, Pages No., Copyright (Year), with permission from Elsevier."

4. Reproduction of this material is confined to the purpose and/or media for which permission is hereby given.

5. **Altering/Modifying Material:** Not Permitted. However figures and illustrations may be altered/adapted minimally to serve your work. Any other abbreviations, additions, deletions and/or any other alterations

shall be made only with prior written authorization of Elsevier Ltd. (Please contact Elsevier at permissions@elsevier.com)

6. If the permission fee for the requested use of our material is waived in this instance, please be advised that your future requests for Elsevier materials may attract a fee.

7. **Reservation of Rights:** Publisher reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

8. **License Contingent Upon Payment:** While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by publisher or by CCC) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and publisher reserves the right to take any and all action to protect its copyright in the materials.

9. **Warranties:** Publisher makes no representations or warranties with respect to the licensed material.

10. **Indemnity:** You hereby indemnify and agree to hold harmless publisher and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.

11. **No Transfer of License:** This license is personal to you and may not be sublicensed, assigned, or transferred by you to any other person without publisher's written permission.

12. **No Amendment Except in Writing:** This license may not be amended except in a writing signed by both parties (or, in the case of publisher, by CCC on publisher's behalf).

13. **Objection to Contrary Terms:** Publisher hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent

with these terms and conditions or CCC's Billing and Payment terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and publisher (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.

14. **Revocation:** Elsevier or Copyright Clearance Center may deny the permissions described in this License at their sole discretion, for any reason or no reason, with a full refund payable to you. Notice of such denial will be made using the contact information provided by you. Failure to receive such notice will not alter or invalidate the denial. In no event will Elsevier or Copyright Clearance Center be responsible or liable for any costs, expenses or damage incurred by you as a result of a denial of your permission request, other than a refund of the amount(s) paid by you to Elsevier and/or Copyright Clearance Center for denied permissions.

LIMITED LICENSE

The following terms and conditions apply only to specific license types:

15. **Translation:** This permission is granted for non-exclusive world English rights only unless your license was granted for translation rights. If you licensed translation rights you may only translate this content into the languages you requested. A professional translator must perform all translations and reproduce the content word for word preserving the integrity of the article. If this license is to re-use 1 or 2 figures then permission is granted for non-exclusive world rights in all languages.

16. **Posting licensed content on any Website:** The following terms and conditions apply as follows: Licensing material from an Elsevier journal: All content posted to the web site must maintain the copyright information line on the bottom of each image; A hyper-text must be included to the Homepage of the journal from which you are licensing at `http://www.sciencedirect.com/science/journal/xxxxx` or the Elsevier homepage for books at `http://www.elsevier.com`; Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.

Licensing material from an Elsevier book: A hyper-text link must be included to the Elsevier homepage at http://www.elsevier.com. All content posted to the web site must maintain the copyright information line on the bottom of each image.

Posting licensed content on Electronic reserve: In addition to the above the following clauses are applicable: The web site must be password-protected and made available only to bona fide students registered on a relevant course. This permission is granted for 1 year only. You may obtain a new license for future website posting.

For journal authors: the following clauses are applicable in addition to the above: Permission granted is limited to the author accepted manuscript version* of your paper.

*Accepted Author Manuscript (AAM) Definition: An accepted author manuscript (AAM) is the authors version of the manuscript of an article that has been accepted for publication and which may include any author-incorporated changes suggested through the processes of submission processing, peer review, and editor-author communications. AAMs do not include other publisher value-added contributions such as copy-editing, formatting, technical enhancements and (if relevant) pagination.

You are not allowed to download and post the published journal article (whether PDF or HTML, proof or final version), nor may you scan the printed edition to create an electronic version. A hyper-text must be included to the Homepage of the journal from which you are licensing at http://www.sciencedirect.com/science/journal/xxxxx. As part of our normal production process, you will receive an e-mail notice when your article appears on Elseviers online service ScienceDirect (www.sciencedirect.com). That e-mail will include the article's Digital Object Identifier (DOI). This number provides the electronic link to the published article and should be included in the posting of your personal version. We ask that you wait until you receive this e-mail and have the DOI to do any posting.

Posting to a repository: Authors may post their AAM immediately to their employers institutional repository for internal use only and may make their manuscript publically available after the journal-specific embargo period has ended.

Please also refer to Elsevier's Article Posting Policy for further information.

18. **For book authors** the following clauses are applicable in addition to the above: Authors are permitted to place a brief summary of their work online only.. You are not allowed to download and post the

published electronic version of your chapter, nor may you scan the printed edition to create an electronic version. Posting to a repository: Authors are permitted to post a summary of their chapter only in their institutions repository.

20. **Thesis/Dissertation:** If your license is for use in a thesis/dissertation your thesis may be submitted to your institution in either print or electronic form. Should your thesis be published commercially, please reapply for permission. These requirements include permission for the Library and Archives of Canada to supply single copies, on demand, of the complete thesis and include permission for UMI to supply single copies, on demand, of the complete thesis. Should your thesis be published commercially, please reapply for permission.

Elsevier Open Access Terms and Conditions

Elsevier publishes Open Access articles in both its Open Access journals and via its Open Access articles option in subscription journals.

Authors publishing in an Open Access journal or who choose to make their article Open Access in an Elsevier subscription journal select one of the following Creative Commons user licenses, which define how a reader may reuse their work: Creative Commons Attribution License (CC BY), Creative Commons Attribution Non Commercial - ShareAlike (CC BY NC SA) and Creative Commons Attribution Non Commercial No Derivatives (CC BY NC ND)

Terms & Conditions applicable to all Elsevier Open Access articles:

Any reuse of the article must not represent the author as endorsing the adaptation of the article nor should the article be modified in such a way as to damage the authors honour or reputation.

The author(s) must be appropriately credited.

If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source it is the responsibility of the user to ensure their reuse complies with the terms and conditions determined by the rights holder.

Additional Terms & Conditions applicable to each Creative Commons user license:

CC BY: You may distribute and copy the article, create extracts, abstracts, and other revised versions, adaptations or derivative works of or from an article (such as a translation), to include in a collective work (such as an anthology), to text or data mine the article, including for commercial purposes without permission from Elsevier

CC BY NC SA: For non-commercial purposes you may distribute and copy the article, create extracts, abstracts and other revised versions, adaptations or derivative works of or from an article (such as a translation), to include in a collective work (such as an anthology), to text and data mine the article and license new adaptations or creations under identical terms without permission from Elsevier

CC BY NC ND: For non-commercial purposes you may distribute and copy the article and include it in a collective work (such as an anthology), provided you do not alter or modify the article, without permission from Elsevier.

Any commercial reuse of Open Access articles published with a CC BY NC SA or CC BY NC ND license requires permission from Elsevier and will be subject to a fee. Commercial reuse includes:

- Promotional purposes (advertising or marketing)
- Commercial exploitation (e.g. a product for sale or loan)
- Systematic distribution (for a fee or free of charge)

Please refer to Elsevier's Open Access Policy for further information.

21. **Other Conditions:**

v1.7

About the Author

Jonathan Burns graduated from the University of North Texas in 2005, earning a bachelor of science in Pure & Applied Mathematics. In 2010, he received a Masters of Arts in Pure & Applied Mathematics from the University of South Florida, before entering the Ph.D. program.

In the span between his undergraduate and graduate studies, Dr. Burns worked as a plastics extrusion engineer and metal machinist, during which time, he designed, fabricated, and tested parts for the industrial, medical, aerospace, and defense sectors.

During his graduate studies, he was supported by the NSF as research assistant, responsible for further developing the Mathematics Umbrella Model for teaching calculus to STEM students, and supervising undergraduate applied-calculus research projects. Additionally, he was sponsored by the NSF and NIH to study genome rearrangements through topological and combinatorial models. Dr. Burns has published in 2 peer-reviewed journals, given talks at 8 national and international research conferences and 6 USF presentations at seminars & student meetings. He is currently a research fellow in the Center for Industrial and Interdiciplinary Mathematics at the University of South Florida and the managing editor for the *Undergraduate Journal of Mathematical Modeling: One + Two*.