⟨TITLE⟩(DRAFT December 11, 2017)

BY ANDREW LOHR

A dissertation submitted to the

Graduate School—New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Mathematics

Written under the direction of

Dr. Zeilberger

and approved by

New Brunswick, New Jersey

May, 2018

ABSTRACT OF THE DISSERTATION

⟨Title⟩(DRAFT December 11, 2017)

by Andrew Lohr

Dissertation Director: Dr. Zeilberger

This thesis deals with applications of experimental mathematics to a number of problems. The First is random graph statistics. In that, a symbolic-computational algorithm, fully implemented in Maple, is described, that computes explicit expressions for generating functions that enable the efficient computations of the expectation, variance, and higher moments, of the random variable 'sum of distances to the root', defined on any given family of rooted ordered trees (defined by degree restrictions). Taking limits, we confirm, via elementary methods, the fact, due to David Aldous, and expanded by Svante Janson and others, that the limiting (scaled) distributions are all the same, and coincide with the limiting distribution of the same random variable, when it is defined on labeled rooted trees.

We also examine generalizations of Sister Celine's method and Gosper's algorithm for evaluating summations. For both, we greatly extend the classes of applicable functions. For the generalization of Sister Celine's method, we allow summations of arbitrary products of hypergeomtric terms and linear recurrent sequences with rational coefficients. For the extension of Gosper's algorithm, we extend it from solely hypergeometric sequences to any multi-basic sequence. For both, we have numerous applications to proving, or reproving in an automated way, interesting combinatorial problems.

Acknowledgements

Still Under Construction

Table of Contents

Chapter 1

Introduction

In Experimental Mathematics, it's often hard to distinguish whether the results, or the methods to find the results are where the real content lies. The techniques used here a heavily based in using computer assistance to solve problems that might otherwise fall to a person to figure out. It is a matter of opinion whether this is a reward in and of itself, showing that something that we thought needed a lot of creativity to determine might actually be just as mechanical as the "lower" mathematics that we teach to our Calculus students, say. There are more concrete cases to be made for the benefits of using a computer to either assist in a result, or find it all on its own. Computers are faster, cheaper, and less prone to errors or boredom than people. Perhaps a less interesting, though not less important use for computers that shows up in this thesis is to compute many, many quantities, and then perform statistical calculations on this data in order to suggest possible conjectures for a person to come along an notice. This is primarily the content of Chapter 3. In Chapter 2 we are able to convert a combinatorial problem involving trees into a mechanical problem involving multi-variable Calculus.

In chapters 4 and 5 we will addressing problems in summation. Recurrence relations will be showing up in abundance. That is, we will have some quantity, either an integer sequence or some expression sequence x_n, and will show that it satisfies some

$$\sum_{j=0}^{N}(Q(n)R^j)x_n = 0$$

where R is the so called "shift operator" which is to say that for and x_n, Rx_n stands for x_{n+1}. We then call N the order of the recurrence. We will sometimes call this whole expression the recurrence, and sometimes will refer to $\sum_{j=0}^{N}(Q(n)R^j)$ as the recurrence that x_n satisfies.

There is a wealth of information on how to analyze something once you know such a recurrence that it satisfies. So, for our cases, if we can analyze something to the point that we know some such recurrence that is satisfies, we will consider it solved. Since we are often starting these summation problems with some undetermined number of terms that is allowed to grow arbitrarily larger, anytime that such a finite description exists, it is a cause for joy. All of the code used for the results here, as well as results of the computations that are too bulky to fit in this document can be found at `http://sites.math.rutgers.edu/~ajl213/DrZ/`

Chapter 2

Limiting Total Height Distributions for Galton Watson Trees

2.1 Background

While many natural families of combinatorial random variables, X_n, indexed by a positive integer n, (for example, tossing a coin n times and noting the number of Heads, or counting the number of occurrences of a specific pattern in an n-permutation) have different expectations, μ_n, and different standard deviations, σ_n, and (usually) largely different asymptotic expressions for these, yet the centralized and scaled versions, $Z_n := \frac{X_n - \mu_n}{\sigma_n}$, very often, converge (in distribution) to the standard normal distribution whose probability density function is famously $\frac{1}{\sqrt{2\pi}} exp(-\frac{x^2}{2})$, and whose moments are $0, 1, 0, 3, 0, 5, 0, 15, 0, 105, \ldots$. Such sequences of random variables are called *asymptotically normal*. Whenever this is **not** the case, it is a cause for excitement [Of course, excitement is in the eyes of the beholder]. One celebrated case (see [21] for an engaging and detailed description) is the random variable 'largest increasing subsequence', defined on the set of permutations, where the intriguing *Tracy-Widom distribution* shows up.

Other, more recent, examples of *abnormal* limiting distributions are described in [29], [6],[7], and [9].

In this chapter we consider, from an elementary, explicit, *symbolic-computational*, viewpoint, the random variable 'sum of distances to the root', defined over an *arbitrary* family of ordered rooted trees defined by degree restrictions. For analysis of this statistic over uniformly chosen random rooted trees, see [25] and [26]. The asymptotic behavior of this statistic for that uniform distribution of random rooted trees is given in [27].

It turns out that the families of trees considered in this paper are special cases of Galton-Watson trees. These have been studied extensively by continuous probability theorists for many years, with a nice, comprehensive introduction given by Janson in [16]. For an analysis of unlabelled Galton-Watson trees, see the work Wagner[28]. In particular, they are trees that are determined by determining the number of children that every node has by independently sampling some fixed distribution with expected value at most 1. Like the trees considered here (described below), they are also types of Galton-Watson trees. It was shown in [1], [2], and [18] that all Galton-Watson generated from a finite variance distribution of vertex degrees followed the same distribution as the area under a Brownian excursion, also a topic well studied in advanced probability theory. In particular, Janson, in section 14 of [14], presents a complicated infinite sum which converges to this distribution originally discovered by Darling (1983). Asymptotic analysis of mean, variance, and higher moments for Galton-Watson trees can be found in [17].

All these authors used *continuous*, advanced, probability theory, that while very powerful, only gives you the limit. We are interested in *explicit* expressions for the first few moments themselves, or failing this, for explicit expressions for the generating functions, for *any* family of rooted ordered trees given by degree restrictions. In particular, we study in detail the case of *complete binary trees*, famously counted by the Catalan numbers.

We proceed in the same vein as in [7]. In that article, the random variable 'sum of the distances from the root', defined on the set of *labelled rooted trees* on n vertices, was considered, and it was shown how to find explicit expressions for any given moment, and the first 12 moments were derived, extending the pioneering work of John Riordan and Neil Sloane ([20]), who derived an explicit formula for the expectation. The exact and approximate values for the *limits*, as $n \to \infty$, of α_3 (the *skewness*), α_4 (the *kurtosis*), and the higher moments through the ninth turn out to be as follows.

$$\alpha_3 = \frac{\left(6\,\pi - \frac{75}{4}\right)\sqrt{3}\sqrt{\frac{\pi}{10-3\,\pi}}}{10-3\,\pi} = 0.7005665293596503\ldots \quad ,$$

$$\alpha_4 = \frac{-189\,\pi^2 + 315\,\pi + 884}{7\,(10 - 3\,\pi)^2} = 3.560394897132889\ldots \quad ,$$

$$\alpha_5 = \frac{\left(36\,\pi^2 + \frac{75}{2}\,\pi - \frac{105845}{224}\right)\sqrt{3}\sqrt{\frac{\pi}{10-3\,\pi}}}{(10 - 3\,\pi)^2} = 7.2563753582799571\ldots \quad ,$$

$$\alpha_6 = \frac{15}{16016}\,\frac{-144144\,\pi^3 - 720720\,\pi^2 + 3013725\,\pi + 2120320}{(10 - 3\,\pi)^3} = 27.685525695770609\ldots \quad ,$$

$$\alpha_7 = \frac{\left(162\,\pi^3 + \frac{6615}{4}\,\pi^2 - \frac{103965}{32}\,\pi - \frac{101897475}{9152}\right)\sqrt{3}\sqrt{\frac{\pi}{10-3\,\pi}}}{(10 - 3\,\pi)^3} = 90.0171829093603301\ldots \quad ,$$

$$\alpha_8 = \frac{3}{2586584}\,\frac{-488864376\,\pi^4 - 8147739600\,\pi^3 - 455885430\,\pi^2 + 86568885375\,\pi + 32820007040}{(10 - 3\,\pi)^4}$$
$$= 358.80904151261251\ldots \quad ,$$

$$\alpha_9 = \frac{\left(648\,\pi^4 + 15795\,\pi^3 + \frac{591867}{16}\,\pi^2 - \frac{461286225}{2288}\,\pi - \frac{188411947088175}{662165504}\right)\sqrt{3}\sqrt{\frac{\pi}{10-3\,\pi}}}{(10 - 3\,\pi)^4} = 1460.7011342971821\ldots$$

[Note that when the moments are not centralized, the expressions are simpler, but we prefer it this way].

2.2 Overview

In this chapter we extend the work of [7] and treat infinitely many other families of trees. For any given set of positive integers, S, we will have a 'sample space' of all ordered rooted trees where a vertex may have no children (i.e. be a *leaf*) or it **must** have a number of children that belongs to S. If $S = \{2\}$ we have the case of *complete binary trees.*

For each such family, defined by S, we will show how to derive explicit expressions for the generating functions of the numerators of the straight moments, from which one can easily get many values, and very efficiently find the numerical values for the moments-about-the-mean and hence the scaled moments. For the special case of complete binary trees, we will derive explicit expressions for the first nine moments (that may be extended indefinitely), as well as explicit expressions for the asymptotics of the scaled moments, and indeed (as predicted by the above-mentioned authors) they coincide *exactly* with those found in [7] for the case of labeled rooted trees. This is a specific example of a more general statement about Galton Watson trees given in [17].

2.3 Rooted Ordered Trees

Recall that an *ordered rooted tree* is an unlabeled graph with the root drawn at the top, and each vertex has a certain number (possibly zero) of children, drawn from left to right. For any finite set of positive integers, S, let $\mathcal{T}(S)$ be the set of all rooted labelled trees where each vertex either has no children, or else has a number of children that belongs to S. The set $\mathcal{T}(S)$ has the following structure ("grammar")

$$\mathcal{T}(S) = \{\cdot\} \bigcup_{i \in S} \{\cdot\} \times \mathcal{T}(S)^i \quad .$$

Fix S, Let f_n be number of rooted ordered trees in $\mathcal{T}(S)$ with exactly n vertices. It follows immediately, by elementary generatingfunctionology, that the ordinary generating function

$$f(x) := \sum_{n=0}^{\infty} f_n \, x^n \quad ,$$

(that is the sum of the weights of *all* members of $\mathcal{T}(S)$ with the weight $x^{NumberOfVertices}$ assigned to each tree) satisfies the **algebraic** equation

$$f(x) = x \left(1 + \sum_{i \in S} f(x)^i \right) \quad .$$

Given an ordered tree, t, define the random variable $H(t)$ to be the sum of the distances to the root of all vertices. Let H_n be its restriction to the subset of $\mathcal{T}(S)$, let's call it $\mathcal{T}_n(S)$, of members of $\mathcal{T}(S)$ with exactly n vertices. Our goal in this chapter is to describe a symbolic-computational algorithm that, for *any* finite set S of positive integers, *automatically* finds generating functions that enable the fast computation of the average, variance, and as many higher moments as desired. We will be particularly interested in the limit, as $n \to \infty$, of the centralized-scaled distribution, and we confirm that it is always the same as the one for rooted labelled trees found in [7] as we'd expect by [17].

Let $P_n(y)$ be the generating polynomial defined over $\mathcal{T}_n(S)$, of the random variable, 'sum of distances from the root'. Define the *grand generating function*

$$F(x, y) = \sum_{n=0}^{\infty} P_n(y) x^n \quad .$$

Consider a typical tree, t, in $\mathcal{T}_n(S)$, and now define the more general *weight* by $x^{NumberOfVertices}\, y^{H(t)} = x^n\, y^{H(t)}$. If t is a singleton, then its weight is simply $x^1 y^0 = x$, but if its sub-trees (the trees whose roots are the children of the original root) are $t_1, t_2, \ldots t_i$ (where $i \in S$), then

$$H(t) = H(t_1) + \cdots + H(t_i) + n - 1 \quad ,$$

since when you make the tree t, out of subtrees $t_1, \ldots, t_i$ by placing them from left to right and then attaching them to the root, each vertex gets its 'distance to the root' increased by 1, so altogether the sum of the vertices' heights gets increased by the total number of vertices in $t_1, \ldots, t_i$ (i.e. $n - 1$). Hence $F(x, y)$ satisfies the **functional equation**

$$F(x,y) = x \cdot \left(1 + \sum_{i \in S} F(xy, y)^i\right) \quad ,$$

that can be used to generate many terms of the sequence of generating polynomials $\{P_n(y)\}$.

Note that when $y = 1$, $F(x, 1) = f(x)$, and we get back the algebraic equation satisfied by $f(x)$.

2.4 From Enumeration to Statistics in General

Suppose that we have a finite set, A, on which a certain numerical attribute, called *random variable*, X, (using the probability/statistics lingo), is defined.

For any non-negative integer i, let's define

$$N_i := \sum_{a \in A} X(a)^i \quad .$$

In particular, $N_0(X)$ is the number of elements of A.

The expectation of X, $E[X]$, denoted by μ, is, of course,

$$\mu = \frac{N_1}{N_0} \quad .$$

For $i > 1$, the i-th straight moment is

$$E[X^i] = \frac{N_i}{N_0} \quad .$$

The i-th *moment about the mean* is

$$m_i := E[(X-\mu)^i] = E[\sum_{r=0}^{i} \binom{i}{r} (-1)^r \mu^r X^{i-r}] = \sum_{r=0}^{i} (-1)^r \binom{i}{r} \mu^r E[X^{i-r}]$$

$$= \sum_{r=0}^{i} (-1)^r \binom{i}{r} \left(\frac{N_1}{N_0}\right)^r \frac{N_{i-r}}{N_0}$$

$$= \frac{1}{N_0^i} \sum_{r=0}^{i} (-1)^r \binom{i}{r} N_1^r N_0^{i-r-1} N_{i-r} \quad .$$

Finally, the most interesting quantities, statistically speaking, apart from the mean μ and variance m_2 are the **scaled-moments**, also known as, *alpha coefficients*, defined by

$$\alpha_i := \frac{m_i}{m_2^{i/2}} \quad .$$

2.5 Using Generating functions

In our case X is H_n (the sum of the vertices' distances to the root, defined over rooted ordered trees in our family, with n vertices), and we have

$$N_1(n) = P_n'(1)$$

$$N_i(n) = (y\frac{d}{dy})^i P_n(y)\big|_{y=1}.$$

It is more convenient to first find the numerators of the factorial moments

$$F_i(n) = (\frac{d}{dy})^i P_n(y)|_{y=1} \quad ,$$

from which $N_i(n)$ can be easily found, using the Stirling numbers of the second kind.

2.6 Automatic Generation of Generating functions for the (Numerators of the) Factorial Moments

Let's define

$$P(X) = 1 + \sum_{i \in S} X^i \quad ,$$

then our functional equation for the grand-generating function, $F(x,y)$ can be written

$$F(x,y) = xP(F(xy,y)) \quad .$$

If we want to get generating functions for the first k factorial moments of our random variable H_n, we need the first k coefficients of the Taylor expansion, about $y = 1$, of $F(x, y)$. Writing $y = 1 + z$, and

$$G(x, z) = F(x, 1 + z) \quad ,$$

we get the functional equation for $G(x, z)$

$$G(x, z) = x\, P(G(x + xz, z)) \quad . \qquad (FE)$$

Let's write the Taylor expansion of $G(x, z)$ around $z = 0$ to order k

$$G(x, z) = \sum_{r=0}^{k} g_r(x) \frac{z^r}{r!} + O(z^{k+1}) \quad .$$

It follows that

$$G(x + xz, z) = \sum_{r=0}^{k} g_r(x + xz) \frac{z^r}{r!} + O(z^{k+1}) \quad .$$

We now do the Taylor expansion of $g_r(x + xz)$ around x, getting

$$g_r(x + xz) = g_r(x) \;+\; g_r'(x)(xz) \;+\; g_r''(x)\frac{(xz)^2}{2!} \;+\; \ldots \;+\; g_r^{(k)}(x)\frac{(xz)^k}{k!} \;+\; O(z^{k+1}) \quad .$$

Plugging all this into (FE), and comparing coefficients of respective terms of z^r for r from 0 to k we get $k+1$ equations relating $g_r^{(j)}(x)$ to each other. It is easy to see that one can express $g_r(x)$ in terms of $g_s^{(j)}(x)$ with $s < r$ (and $0 \le j \le k$) .

Using implicit differentiation, the derivatives of $g_0(x)$, $g_0^{(j)}(x)$ (where $g_0(x)$ is the same as $f(x)$), can be expressed as rational functions of x and $g_0(x)$. As soon as we get an expression for $g_r(x)$ in terms of x and $g_0(x)$, we can use calculus to get expressions for the derivatives $g_r^{(j)}(x)$ in terms of x and $g_0(x)$. At the end of the day, we get expressions for each $g_r(x)$ in terms of x and $g_0(x)$ (alias $f(x)$), and since it is easy to find the first ten thousand (or whatever) Taylor coefficients of $g_0(x)$, we can get the first ten thousand coefficients of $g_r(x)$, for all $0 \le r \le k$, and get the numerical sequences that will enable us to get very good approximations for the alpha coefficients.

The beauty is that this is all done by the computer! Maple knows calculus.

We can do even better. Using the methods described in [13], one should be able to get, *automatically*, asymptotic formulas for the expectation, variance, and as many moments as desired. Using these techniques, it may be possible to obtain expressions for the leading terms of all moments, and so show weak convergence of this distribution to a particular limiting distribution. This should be an interesting future project.

For the special case of complete binary trees, everything can be expressed in terms of Catalan numbers, and hence the asymptotic is easy, and our beloved computer, running the Maple package `TREES.txt` (mentioned above), obtained the results in the next section.

Computer-Generated Theorems About the Expectation, Variance, and First Nine Moments for the Total Height on Complete Binary Trees on n **Leaves**

See the output file

`http://www.math.rutgers.edu/~zeilberg/tokhniot/oTREES3.txt` .

2.7 Universality

The computer output, given in the above webpage, proved that for this case, of complete binary trees, the limits of the first nine scaled moments coincide *exactly* with those found in [7], and given above. Confirming, by *purely elementary, finitistic methods*, the universality property mentioned above. We do it for one family at a time, and only for finitely many moments, but on the other hand, we derived *explicit* expressions for the first twelve moments in the case of complete binary trees, and explicit expressions for the generating functions for the moments for other families.

Chapter 3

Rational Sloped Paths

3.1 Background

There is a rich study of Dyck paths in combinatorics. Some of the most ubiquitous results are for the case that the slope of the line is 1. In particular that the number of paths from (0,0) to (n,n) is counted by the Catalan numbers. When we change it from 1 to another rational number a/b, we enter the realm of appropriately named Rational Catalan Combinatorics. For notational convenience, we'll let $A_{a,b,n}$ denote the number of paths from $(0,0)$ to (b,a) staying on or below the line $y = a/bx$.

It was shown by Duchon in 2000 that for any slope $\frac{a}{b}$, the number of paths below a line of that slope is asymptotically $\Theta\left(\frac{1}{n}\binom{(a+b)n}{an}\right)$ [5]. However, it's still unknown what the constant out front is. To show the asymptotics, Duchon showed that the content is somewhere between $\frac{1}{a+b}$ and $\frac{1}{a}$. This upper bound on the number of paths was known at least as far back as 1950 to Grossman. Grossman also had an interesting result, the first proof of which is given by Bizley in 1954 in a now defunt actuarial journal [4]. It gives that gives an exact formula for every $A_{a,b,n}$.Of course, this precision comes at a cost, The formula is given as a sum over a large set of weighted integer partitions. There is no good way to extract estimates from this formula that we know, but it seems powerful and may be useful for this problem in the future. It would be great to have a simpler explanation of the simper problem of determining this value up to a $(1 + o(1))$ factor.

It's clear that $A_{a,b,n} = A_{b,a,n}$, so in this paper, we'll assume that we always have $a > b$.

For the case $b = 1$, there is an exact solution known, using Fuss-Catalan numbers

$\frac{1}{1+an}\binom{(1+a)n}{an}$.

3.2 Approach

We would like to try and find the coefficient out front, that is, α so that the number of paths is $(1+o(1))\,\frac{\alpha}{n}\binom{(a+b)n}{an}$. In an effort to do this, we first are tasked with computing many terms of the sequence

Through a simple dynamic programming algorithm, we are able to compute the number of paths from (0,0) to (bn,an) for n around a thousand. This gives us enough data to try and find a recurrence relation that it satisfies using the maple package available at `http://www.math.rutgers.edu/ zeilberg/tokhniot/FindRec.txt` However, we were only able to successfully find recurrences for the slopes 3/2 and 5/2. Armed with these recurrences we are able to blindingly fast crank out many thousands more terms of this sequence. The recurrence in the data file for slope 3/2 is order 4, whereas the one for 5/2 is order 8 and monstrously long. Though we can't guarantee this is the minimal recurrence, it still gives a massively faster of counting the paths for these two unknown slopes, and potentially for many more slopes that our computer wasn't keen enough to find this time.

Then, once we have exact numbers, we do a statistical fit of the data for many values of n against the model $\left(\frac{\alpha}{n}+\frac{\beta}{n^2}+\frac{\gamma}{n^3}+\frac{\delta}{n^4}\right)\binom{(a+b)n}{an}$ to get our estimate of α. Adding more error terms didn't affect the value of α much. We estimate how close this is to the truth by running it for the first 100 values of n, and the first 200 values of n, and seeing how much our estimate stays the same. We can get quite good estimates of these numbers!

3.3 Data and Figures

There are many more slopes for which we have very exact estimates of α, and are availiable online at

`http://www.math.rutgers.edu/ãjl213/DrZ/RSP.html` in the extra data file.

Though we knew it already with Duchon's result that the value of the coefficient is

Table 3.1: Estimates for α part 1

a\b	2	3
2	1	
3	0.240706636	1
4	0.50000000	0.15972479544
5	0.1613399969	0.1372518253
6	0.333333333	0.50000000
7	0.1216701970	0.1073342967
8	0.250000000	0.09683505915

Table 3.2: Estimates for α part 2

a\b	4	5	6	7
4	1			
5	0.119952918	1		
6	0.240706636	0.09621264003	1	
7	0.09639805178	0.08763172133	0.08039623916	1
8	0.5000000	0.08048157890	0.159724795	0.06908631788

at most $1/a$, it still seems surprising that the value of the coefficient is not monotone in the value of the slope, that is the actual number a/b. We can notice a few simple patterns here, in particular, except in the cases that a and b are not in lowest terms, the value of coefficient decreases as you increase either a or b.

We can investigate this second observation a little further. A similar, but interesting and distinct problem is to try letting something else go to infinity in $A_{a,b,n}$ other than n, as we had before.

Suppose instead that we were to let a go to infinity while b is fixed. There's a nice pattern that appears. In particular, we have the conjecture that α is asymptotically equal to $\frac{gcd(a,b)}{a}$. The $gcd(a,b)$ factor is expected because it makes the expression only depend on the value of a/b, as it should.

Figure 3.1: $a\alpha$ as a function of a for $b = 2$

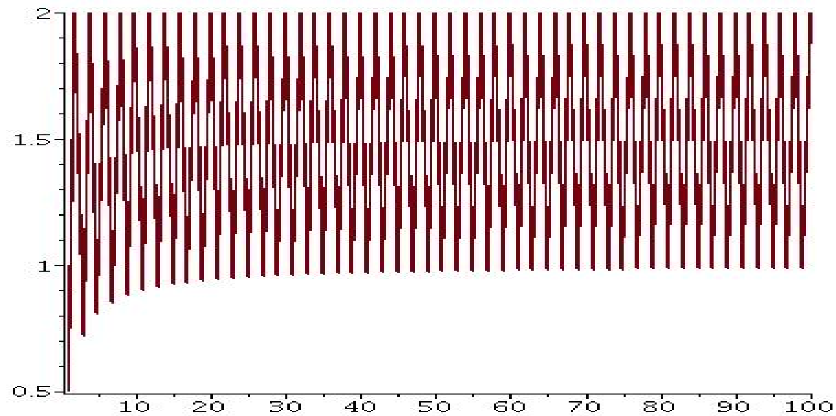
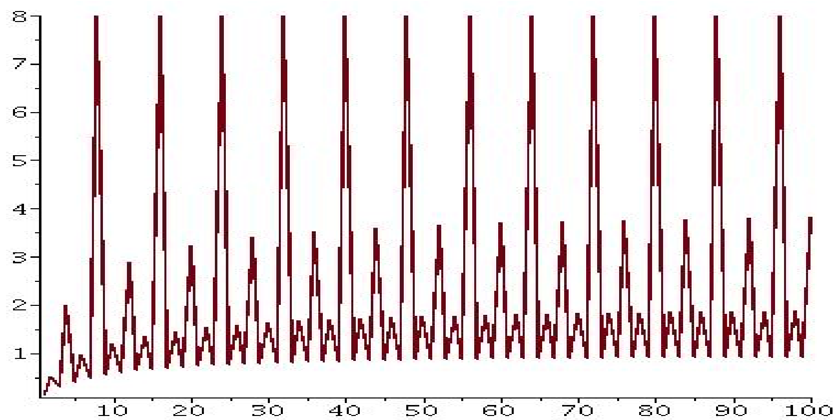


Figure 3.2: $a\alpha$ as a function of a for $b = 8$



3.4 Three Dimensional Lattice Walks

A much less studied area is to consider paths in a three dimensional lattice($\mathbb{Z}^3$) that have to stay to stay in a region bounded by planes. It is simple to extend the dynamic programming solution to this situation. However, since there many more lattice points, the runtime goes up from $\Theta(n^2)$ to $\Theta(n^3)$. This keeps us from getting anywhere near as much data as we did in the previous section. With the data we do have, we have the suggestion that something much more interesting than in the 2D case is happening!

The way that we set up the three dimensional problems, is that we take the number of paths with steps in $\{\langle 1,0,0\rangle, \langle 0,1,0\rangle, \langle 0,0,1\rangle\}$. Instead of the 2D problem of requiring $ax < by$, we define and instance of the counting problem to be indexed by three numbers,

Table 3.3: a=1

b\c	1	2	3	4	5	6	7
1	3.0	2.7	2.6	2.5	2.5	2.4	2.4
2	3.7	3.3	3.0	2.9	2.8	2.7	2.6
3	4.3	3.7	3.4	3.2	3.1	3.0	2.9
4	4.8	4.1	3.8	3.5	3.4	3.2	3.1
5	5.2	4.5	4.1	3.8	3.6	3.4	3.3
6	5.7	4.8	4.4	4.1	3.8	3.6	3.5
7	6.0	5.2	4.6	4.3	4.0	3.8	3.7

Table 3.4: a=2

b\c	1	2	3	4	5	6	7
1	2.7	2.5	2.5	2.4	2.4	2.3	2.3
2	3.3	3.0	2.8	2.7	2.6	2.6	2.5
3	3.7	3.4	3.2	3.0	2.9	2.8	2.8
4	4.1	3.7	3.5	3.3	3.1	3.0	3.0
5	4.5	4.0	3.7	3.5	3.4	3.2	3.2
6	4.8	4.3	4.0	3.7	3.6	3.4	3.3
7	5.2	4.6	4.2	4.0	3.8	3.6	3.5

a, b, c, and require of our paths that they satisfy $ax \le by \le cz$ that end at $x = bc, y = ac, z = ab$. If we have $a = b = c = 1$, this has the precise formula of the 3D Catalan numbers (A005789 in [22]) $\frac{2}{(n+1)^2(n+2)}\binom{3n}{n,n,n}$. However, there is a lot left to understand in this problem, and some things that are distinctly different than the 2D case. In particular, for 2D, it was always $\Theta\left(\frac{1}{n}\binom{(a+b)n}{an}\right)$. That is, the slope of the line didn't affect the fact that you always had a $\Theta(\frac{1}{n})$ fraction of all paths. For the already known $a = b = c = 1$, it is a $\Theta(\frac{1}{n^3})$ fraction of all paths, however this appears to change for different choices of a, b, c. We have some data on the value of this coefficient in tables 3.3 and 3.4.

Chapter 4

Generalization of Sister Celine's Method

4.1 Background

One of the earliest steps in automatically proving identities dates back to Sister Mary Celine Fasenmyer's 1945 Ph.D. thesis [11]. She gave a technique for computing sums of hypergeometric terms, also see [12]. Very briefly, in order to determine if there is an order I recurrence for the sequence $x_n = \sum_k H(n,k)$, it first considers

$$0 = \sum_{i=0}^{I} \sum_{j=0}^{J} y_{i,j}(n) H(n+i, k+j)$$

Where $y_{i,j}(n)$ is an as yet unknown rational function of n. Then, by H being hypergeometric, it is able to reduce all of the $H(n+i, k+j) = G_{i,j}(n,k)H(n,k)$ where $G_{i,j}$ is some rational function of n and k. From there, divide everything through by $H(n,k)$. Now, we have something of the form

$$0 = \sum_{i=0}^{I} \sum_{j=0}^{J} G_{i,j}(n,k) y_{i,j}(n)$$

Combining denominators on the right hand side, and multiplying through by the common denominator, we get that the right hand side becomes a polynomial in n and k, with $\{y_{i,j}(n)\}$ thrown in as well. Collect terms by what power of k appears, and then solve for what the $\{y_{i,j}(n)\}$ have to be in order to make all of the coefficients of powers of k equal to zero. We may get unlucky and have no solution, then, we would need to try a larger I to begin with. If however, we find a solution, we plug that into where we first introduced $y_{i,j}(n)$. Since these have no k's in them, and x_n is obtained by summing over all values of k that make the summand nonzero, we have

$$0 = \sum_{i=0}^{I}\sum_{j=0}^{J} y_{i,j}(n)H(n+i,k+j) = \sum_{i=0}\left(\sum_{j=0}^{J} y_{i,j}(n)\right) x_{n+i}$$

Which, setting $z_i = \sum_{j=0}^{J} y_{i,j}(n)$, we may write in shift operator notation as

$$0 = \left(\sum_{i}^{I} z_i(n)N^i\right) x_n$$

At this point we say that we are done. First, having a recurrence allows you to compute the sequence out to very large values very quickly, storing only a constant number of terms. Also, once you have a rational recurrence like this for x_n then you can extract as good asymptotics as desired like using techniques by Birkhoff-Trjizinski which has been nicely summarized in [24]. Though sometimes you may be able to take these recurrences and recover a really nice formula, there are more sequences to describe than there are nice formulas, so we have to deal with the fact that we can only go so far in making it prettier.

For a more complete explanation of Sister Celine's method, look at chapter 4 of [19]. There are some generalizations of Sister Celine's method given in [30], in particular to certain classes of multiple summations and to a continuous analog.

Some of our applications of the expanded method presented in this paper relate to binomial transforms of functions. There are nice treatments of binomial transforms of Fibonacci like sequences given in [23].

4.2 Overview

We will be taking this technique of Sister Celine and extending it to allow many more kinds of summands. In particular, it can be of the form $x_n = \sum_k^n a_k^d H(k,n)$ where d is any number, H is hypergeometric, and a_k is some sequence defined by a rational recurrence relation. Since so many sequences can be so described by rational recurrence relations, this is a significant extension in scope.

It works very similarly to Sister Celine, in that we will consider ratios of successive terms. That is, to find a recurrence with order at most I, start with

$$\sum_{i=0}^{I}\sum_{j=0}^{J}\frac{F(n+i,k+j)}{F(n,k)}a_{j+k}^{d}y_{i,j}(n)$$

Let D be the order of the recurrence describing $\{a_k\}$. Then, we use that relation to rewrite all of the $\{a_{k+j}\}_{j=D}^{J}$ in terms of $\{a_{k+j}\}_{j=0}^{D-1}$. That is, by repeatedly applying the relation, we can write each a_{j+k} as a linear combination:

$$a_{j+k} = \sum_{m=0}^{D-1} c_{k,j,m}a_{k+m}$$

where for the $j < D$, we just let $c_{k,j,m} = \begin{cases} 1 & j = m \\ 0 & j \neq m \end{cases}$. Then, since we have an expression with D terms to the d, we can expand that out to get at most Dd terms. Then, unlike in Sister Celine, where we have a polynomial in k, we now have a polynomial in $\{k, a_k, a_{k+1}, \ldots a_{k+D-1}\}$. But, once we have collected the coefficients of each of the combinations of those variables, we set all of them equal to zero, and then try to solve for the $y_{i,j}(n)$. As in Sister Celine, we are not guaranteed that we can find such a solution for our particular choice of I and J. We are guaranteed by WZ theory that for a large enough choice of I and J, it gives us a recurrence relation that looks like

$$0 = \left(\sum_{i=0}^{I}\left(\sum_{j=0}^{I} y_{i,j}(n)\right)N^{i}\right)x_n$$

4.3 Application to Enumerating Chess King Walks

Suppose that there is a king wandering around on an infinite d-dimensional chess board, we want to know how many of the $(3^d-1)^n$ walks of length n that the king could take would end up bringing him back to where he started. Given a polynomial p, we will use the notation $Ct(p)$ to denote the constant term of p. Then, by using the powers of

z_i to keep track of our total displacement in the i dimension, we have:

$$
\begin{aligned}
x_n &= Ct\left(\left(\left(\prod_{i=1}^{d} z_i + z_i^{-1} + 1\right) - 1\right)^n\right)\\
&= Ct\left(\sum_{k=0}^{n}\left(\prod_{i=1}^{d} z_i + z_i^{-1} + 1\right)^k \binom{n}{k}(-1)^{n-k}\right)\\
&= \sum_{k=0}^{n} Ct\left(\left(\prod_{i=1}^{d} z_i + z_i^{-1} + 1\right)^k\right)\binom{n}{k}(-1)^{n-k}\\
&= \sum_{k=0}^{n} Ct\left(\left(z + z^{-1} + 1\right)^k\right)^d \binom{n}{k}(-1)^{n-k}
\end{aligned}
$$

Luckily for us, $Ct\left(\left(z + z^{-1} + 1\right)^k\right)$ is already well understood. It is the central trinomial coefficients (A002426 [22]). Also luckily, it is a known that this sequence satisfies the recurrence.

$$0 = \left(N^2 - \frac{2n-1}{n}N - \frac{3n-3}{n}\right)x_n$$

So, we are in exactly the set up of this extended method. In which case you can describe the number of d dimensional king walks which end at the origin after taking n steps by

$$\sum_{k=0}^{n} a_k^d \binom{n}{k}(-1)^{n-k}$$

This clearly falls into the scope of this modified algorithm, and using it you are able to find rational recurrences (effectively solve) for all dimensions up to 4. Here is the one for a two dimensional king walking around

$$
\begin{aligned}
g(n,N) =& (3n^3 + 40n^2 + 175n + 250)N^3\\
&+ (9n^3 + 138n^2 + 703n + 1190)N^2\\
&+ (108n^3 + 1548n^2 + 7364n + 11632)N\\
&+ 96n^3 + 1280n^2 + 5632n + 8192
\end{aligned}
$$

then

$$0 = g(n,N)x_n$$

Although this already does not look super nice, at least it is short, which is more than can be said of those describing higher dimensions, but they are included in an appendix. Also important is that they were found by a computer.

Something probably more insightful than these walls of text that exactly describes these sequences is their asymptotics:

For the two dimensional king, the number of paths of length n is

$$c_2 \frac{8^n}{n} \left(1 - \frac{4}{9n} + \frac{1}{18n^2} + O\left(\frac{1}{n^3}\right)\right)$$

For three dimensions:

$$c_3 \frac{26^n}{n^{\frac{3}{2}}} \left(1 - \frac{11}{18n} + \frac{683}{5832n^2} + O\left(\frac{1}{n^3}\right)\right)$$

and for four dimensions:

$$c_4 \frac{80^n}{n^2} \left(1 - \frac{25}{9n} + \frac{36439}{6561n^2} + O\left(\frac{1}{n^3}\right)\right)$$

The dominant asymptotics are somewhat unsurprising. The exponential part is all possible paths. The dominant power of n is $\left(\frac{1}{\sqrt{n}}\right)^d$, and it is well known that the central binomial coefficient is asymptotically $\frac{2^n}{\sqrt{n}}$, and we are doing something somewhat like that in d dimensions. The value of c_2 is approximately equal to $\frac{2}{3\pi}$. This value for c_2 can be proven in a rigorous way using classical analysis. For c_3 and c_4, we are not so lucky, instead, all we can say from non-rigorous observation is that $c_3 \approx .110225343716$ and $c_4 \approx .068412392872$. There might be some way using a more traditional approach that would get us the true value of these constants.

The $d = 2$ case was first worked out by a computer using a different approach. For more information on this, see [8], or for information on the techniques, see [3].

4.4 Application to other sequences

This also allows for computing binomial transforms of other sequences. An example of this is if you were to let F_k be the k-th Fibonacci number and consider the sequence

$$x_n = \sum_{k=0}^{n} F_k \binom{n}{k}$$

You immediately receive the recurrence that defines x_n is $0 = (-N^2 + 3N - 1)x_n$, and this is identical to the recurrence given for (A001906 [22]) which is the sequence describing the sum. Though this is already a known fact, if you just bump the power up on F_k to F_k^2, you still get a rather nice recurrence relation for the sum, in particular it is described by $0 = (-N^3 + 5N^2 - 5)x_n$. This integer sequence is as yet unnamed in the OEIS, but has both a simple definition in terms of Fibonacci, and a lovely formula where it is just 5 times the difference of two earlier terms. All powers of Fibonacci seem to follow this nice pattern that a linear recurrence where the terms do not depend on n suffices, instead of in general, where the recurrence may need rational functions of n showing up to describe the next term. These C-finite sequences are discussed in greater detail in [31]. The techniques given in that paper can also be applied to some of the problems considered here.

Also of interest, suppose that you are considering a_k to be the m-Fibonacci sequence, that is, $a_{k+2} = ma_{k+1} + a_k$ then, it is simple enough to plug in this recurrence, and lo and behold, an answer is found. For

$$x_n = \sum_{k=0}^{n} a_k \binom{n}{k}$$

we have

$$x_{n+2} = (2+m)x_{n+1} - x_n$$

Which is also known, but is a main theorem of a twelve page paper by Falcon and Plaza [10] instead of a two second calculation by the computer.

4.5 Application to multiple summations

Another promising application of this technique is to evaluating multiple sums over hypergeometric terms. A toy example of this would be if you wanted to compute

$$\sum_{i=0}^{n} \sum_{k=0}^{i} \binom{i}{k} \binom{n}{i}$$

To do this, pick out any of the factors which contain k, and run some automated process to evaluate single summation such as the Zeilberger Algorithm [19]. Often, this sum will not have a nice formula, so you are left with a possibly high order recurrence describing it. However, that is precisely what the techniques here are made to handle, so you can feed this partial evaluation into the procedure. Given enough computing this allows any number of summation signs to be dealt with. For each summation, we have the usual requirements of the original Sister Celine's method, namely that for each summation, the boundaries extend as far as the terms can be without becoming zero. In this particular case, evaluating the inner sum, you get $0 = (N-2)x_n$, and plugging that recurrence in, we get that the whole sum satisfies $0 = (N-3)z_n$. Which is to say, the sum evaluates to 3^n. Though this has a nice combinatorial proof where you count the number of assignments from $\{1, \ldots, n\}$ to $\{1, 2, 3\}$ by first picking the k elements that map to either 1 or 2, and then, from those k elements, picking the i elements that map to 2, That requires a moment of thought where such a simple recurrence for the computer only requires less than a second of thought. Or, suppose the harder problem, where we would want to compute

$$\sum_{i=0}^{n} \sum_{k=0}^{n} \binom{i-k}{k}^2 \binom{n}{i}$$

It **may** be possible as a person to figure this out in a more human way, but for the computer it is just a few seconds away from spitting out the the solution is described by the recurrence

$$0 = \left(-(n+9)N^5 + (7n+54)N^4 - (17n+103)N^3 + (21n+97)N^2 - (15n+50)N + 5n + 5\right) x_n$$

A maple package for multiple summations has already been described in [3] and is available at:

`http://sites.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/multiZ.html`

However this package is roughly the same time on the simple first example given, and faster than their package on the second example. Their package, however, gives a 'better' analysis of the summation, in that it does indefinite summation, and does not require that on the bounds of summation, the summand is zero. That is, theirs generalizes Zeilberger's algorithm, instead of Sister Celine's.

4.6 Example Usage of this Maple Package

Hopefully by this point, you are asking yourself, how to use these powerful tools. Though there is more detailed documentation in the maple package itself. The first step is to figure out the recurrence that is satisfied by your a_k, called rec1. Then, call $findrec(I, J, timeout, rec1, F, d, n, N)$ where both rec1 and the output will be in shift operator notation, with N the shift operator. This call will attempt to find the recurrence for the sum:

$$x_n = \sum_{k=0}^{n} a_k^d F(n, k)$$

Where the recurrence is of order at most I, and degree at most J. *timeout* is the most time (in seconds) that you are willing to wait on a particular attempt, if it exceeds that time, the procedure exits.

Chapter 5

A Practical Variation on Gosper's Algorithm

This chapter will be filled in once work on the article has been completed separately, so as to only have one incomplete version of it floating around at a time.

References

[1] David Aldous, *The continuum random tree II, The continuum random tree II: an overview*, Stochastic analysis **167** (1991), 23-70.

[2] David Aldous, *The continuum random tree III*, Ann. Probab. **21** (1993), 248-289,.

[3] Moa Apagodu and Doron Zeilberger, Multi-Variable Zeilberger and Almkvist-Zeilberger Algorithms and the Sharpening of Wilf-Zeilberger Theory, Advances in Applied Mathematics, Vol 37, Issue 2, August 2006, 139-152

[4] M. T. L. Bizley, *DERIVATION OF A NEW FORMULA FOR THE NUMBER OF MINIMAL LATTICE PATHS FROM (o, o) TO (km, kn) HAVING JUST t CONTACTS WITH THE LINE my=nx AND HAVING NO POINTS ABOVE THIS LINE; AND A PROOF OF GROSSMAN'S FORMULA FOR THE NUMBER OF PATHS WHICH MAY TOUCH BUT DO NOT RISE ABOVE THIS LINE*, Journal of the Institute of Actuaries **80** (1954), 55-62.

[5] Phillipe Duchon, *On the enumeration and generation of generalized Dyck words*, Discrete Mathematics **225** (2000), 121-135

[6] Shalosh B. Ekhad and Doron Zeilberger, *Explicit Expressions for the Variance and Higher Moments of the Size of a Simultaneous Core Partition and its Limiting Distribution* , The Personal Journal of Shalosh B. Ekhad and Doron Zeilberger, `http://www.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/stcore.html` .

[7] Shalosh B. Ekhad and Doron Zeilberger, *Going Back to Neil Sloane's FIRST LOVE (OEIS Sequence A435): On the Total Heights in Rooted Labeled Trees*, The Personal Journal of Shalosh B. Ekhad and Doron Zeilberger, `http://www.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/a435.html` .

[8] FShalosh B. Ekhad, "In How Many Ways Can a Chess King Return Home After n Steps?" `http://www.math.rutgers.edu/~zeilberg/mamarim/mamarimPDF/melech.pdf`

[9] Shalosh B. Ekhad and Doron Zeilberger, *Automatic Proofs of Asymptotic ABNORMALITY (and much more!) of Natural Statistics Defined on Catalan-Counted Combinatorial Families*, The Personal Journal of Shalosh B. Ekhad and Doron Zeilberger, `http://www.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/abnormal.html` .

[10] Sergio Falcon and Angel Plaza, Binomial Transforms of the k-Fibonacci Sequence, International Journal of Nonlinear Sciences & Numerical Simulation 10, 2009, 1527-1535

[11] Fasenmyer, Sister Mary Celine, Some generalized hypergeometric polynomials, Ph.D. dissertation, University of Michigan, November, 1945.

[12] Fasenmyer, Sister Mary Celine, A note on pure recurrence relations, Amer. Math. Monthly 56 (1949), 14-17.

[13] Philippe Flajolet Robert Sedgewick, "*Analytic Combinatorics*", Cambridge University Press, 2009. (Free download from Flajolet's homepage.)

[14] Svante Janson *Brownian excursion area, Wright's constants in graph enumeration, and other Brownian areas*, Probability Surveys 3 (2007), 80-145. Available on line from
`https://arxiv.org/abs/ 0704.2289`

[15] Svante Janson *Patterns in random permutations avoiding the pattern 132*, Available on line from
`https://arxiv.org/abs/1401.5679`

[16] Svante Janson *Simply generated trees, conditioned GaltonWatson trees, random allocations and condensation*, Probability Surveys 9 (2012), 103-252. Available on line from
`http://www2.math.uu.se/ svante/papers/sj264.pdf`

[17] Svante Janson *The Wiener Index of Simply Generated Trees*, Random Structures and algorithms 22, issue 4 (2003), 337-358. Available on line from
`http://www2.math.uu.se/ svante/papers/sj146.pdf`

[18] Jean-Franois Marckert and Abdelkader Mokkadem *The depth first processes of Galton–Watson trees converge to the same Brownian excursion*, The Annals of Probability Volume 31, number 3 (2003) 1655-1678. Available on line from
`http://projecteuclid.org/euclid.aop/1055425793` .

[19] Mario Petkovsek, Herbert Wilf, and Doron Zeilberger, A=B, Taylor & Francis, 1996. `http://www.math.rutgers.edu/~zeilberg/AeqB.pdf`

[20] John Riordan and Neil J. A. Sloane, *The enumeration of rooted trees by total height*, J. Australian Math. Soc. **10**(1969), 278-282. Available on line from:
`http://neilsloane.com/doc/riordan-enum-trees-by-height.pdf` .

[21] Dan Romik, "*The Surprising Mathematics of Longest Increasing Subsequences*", Cambridge University Press, 2015.

[22] N. J. A. Sloane, editor, The On-Line Encyclopedia of Integer Sequences, published electronically at `https://oeis.org`, 2017

[23] Michael Spivey, Combinatorial Sums and Finite Differences, Discrete Mathematics Vol 307 Issue 24, November 2007, 3130-3146

[24] Jet Wimp and Doron Zeilberger Resurrecting the Asymptotics of Linear Recurrences, Journal of Mathematical Analysis and Applications Vol 111, Issue 1, October 1985, 162 - 176

[25] L. Takcs, *On the Total Heights of Random Rooted Trees*, J Appl Prob, 29(3) (1992), 543-556.

[26] L. Takcs, Limit theorems for random trees, Proc Natl Acad Sci USA, 89 (1992), 5011-5014.

[27] L. Takcs, *The asymptotic distribution of the total heights of random rooted trees*, Acta Sci Math (Szeged) 57(14) (1993), 613625.

[28] S. Wagner, *On the Wiener index of random trees*, Disc Math 312(9) (2012), 1502-1511.

[29] Doron Zeilberger, *Doron Gepner's Statistics on Words in* $\{1,2,3\}^*$ *is (Most Probably) Asymptotically Logistic*, The Personal Journal of Shalosh B. Ekhad and Doron Zeilberger,
`http://www.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/gepner.html`
.

[30] Doron Zeilberger, Sister Celine's Technique and Its Generalizations, Journal of Mathematical Analysis and Applications, Volume 85, Issue 1, January 1982, 114-145
`http://www.math.rutgers.edu/~zeilberg/mamarimY/Zeilberger_y1982_p114.pdf`

[31] Doron Zeilberger, The C-Finite Ansatz, Ramanujan Journal 31(2013), 23-32
`http://sites.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/cfinite.html`